Note 00. Symmetry

Alex Fu

Fall 2022

Symmetry is many things at once: a familiar general concept, a formal mathematical object, and a powerful problem-solving technique. While we will not need a formal definition of symmetry, we can borrow the idea behind its definition in *group theory* to understand symmetry as a problem-solving technique.

Definition 1 (Symmetry).

A **symmetry** of an object X is some action φ such that X is left *invariant*, or unchanged, under the action of φ.

As usual, a highly abstract definition is made clear through examples.

Example 1 (Reflections and rotations).

Reflections and *rotations* can be kinds of symmetry. We say that a geometric object has reflectional or rotational symmetry if it is unchanged by the action of a reflection or rotation, so, for instance, reflections and rotations are symmetries of a circle.

These symmetries are also not limited to geometric objects. A function f is **even** if it has reflectional symmetry about the vertical axis: $f(-x) = f(x)$, and f is **odd** if it has 180-degree rotational symmetry about the origin: $f(-x) = -f(x)$. Using symmetry, what can we say about integrals of even or odd functions over intervals like $[-a, a]$?

We will soon see even more examples in probability: symmetric probability distributions have reflectional symmetry, multivariate normal distributions are (uniquely!) rotationally invariant, and Markov chains can even exhibit both kinds of symmetries. After recognizing symmetry, we can then exploit it in applications like finding expected values, concentration inequalities, or stationary distributions.

Example 2 (Bijections and permutations).

Bijections can be a kind of symmetry. Two sets have the same number of elements, or **cardinality**, iff there exists a bijection between them, so cardinality is something unchanged under the action of a bijection. For instance, if we can show a bijection between some subset $T \subseteq S$ and its complement T^c, then we have just shown that $|T| = |T^c|$, and thus $|T| = |S|/2$; this can be a helpful technique in *combinatorics*, the art of counting.

In particular, a bijection on a finite set of n elements is called a *permutation*. For instance, the **trace** of a matrix is invariant under cyclic permutations: $\mathrm{tr}(ABC) = \mathrm{tr}(BCA) = \mathrm{tr}(CAB)$. A permutation cannot be a symmetry of an ordered list, unless it is the identity function, because permutations do not preserve order, but it can be a symmetry of a (unordered) set.

Example 3 (Linear transformations and transition probabilities).

Linear transformations can be a kind of symmetry. If (λ, v) is an eigenvalue-eigenvector pair of a linear transformation A, where λ is nonzero, then the span of v is invariant under the action of A by $Av = \lambda v$. In particular, if P is a matrix with nonnegative entries and $P\mathbf{1} = \mathbf{1}$, where $\mathbf{1}$ is the vector with all entries 1, then every row of P sums to 1. Such a matrix is called **row-stochastic**, and the transition probabilities of any finite-state Markov chain are uniquely described by a row-stochastic matrix.

Furthermore, a key assumption we will make about Markov chains is **time homogeneity**: the transition probabilities between states are invariant under shifts in time. Distributions over states can also be invariant under shifts in time: a distribution π is **stationary** if $\pi P = \pi$ is a row eigenvector of P with eigenvalue 1. Stationary distributions are an important part of the study of random processes — they often represent some sort of stability, predictability, or long-term behavior of their random process.

Example 4 (In proofs and problem solving).

This is the key example of symmetry for our purposes. If we can reuse a step in a proof, the formula for a calculation, or a particular argument, only needing to exchange certain variable names ($x \leftrightarrow y$), labels (Alice $\leftrightarrow$ Bob), or specific values ($-1 \leftrightarrow 1$), then *this is a kind of symmetry as well*, often indicated by stating "by symmetry" or "without loss of generality."

We might name this a kind of "symmetry under exchanging symbols." For instance, we can find the **density of a transformed random variable** using invariance under change of variables: $|f(x)\,dx| = |f(y)\,dy|$. Or, because the order of intersections does not matter, we can permute our way to an order of variables that might be easier to work with: $\mathbb{P}(A \cap B) = \mathbb{P}(A \mid B)\,\mathbb{P}(B) = \mathbb{P}(B \mid A)\,\mathbb{P}(A)$ — **Bayes' rule**.

Though we cannot always point out this kind of symmetry for you, a key to recognizing it is seeing when an arbitrary choice of a variable is made, or when different objects are nonetheless the same in some aspect. For a short list, consider the sides of a (fair) coin; the numbers on a die; the suits in a deck of cards; the names of two players; m indistinguishable balls; n indistinguishable bins; k points on a circumference; d dimensions of some object identical in any particular dimension; two minimizers that you suspect are equal, etc.

The last example draws an interesting connection between symmetry and probability, a reason why symmetry is in some way fundamental to probability theory: one definition of "arbitrary" is simply "random." If symmetry holds in situations where an arbitrary choice is made, would it also hold in situations with randomness?

For example, consider the set of n outcomes $\Omega = \{\omega_1, \ldots, \omega_n\}$ of a random experiment, where an outcome "happens" if it is arbitrarily chosen. If we want to assign a *probability value* to every outcome, a value between 0 and 1 describing the likelihood that the outcome happens (is chosen), then by symmetry, this *probability function* $p : \Omega \to [0, 1]$ must satisfy

$$p(\omega_i) = p(\omega_j) \qquad \forall \omega_i, \omega_j \in \Omega$$

because there is no reason to assign one outcome a different value over another outcome — the **principle of indifference**. In other words, $p(\cdot)$ is symmetric under exchanging its argument with any another possible argument. Then, it must be the case that

$$p(\omega_i) = \frac{1}{|\{\omega_1, \ldots, \omega_n\}|} = \frac{1}{n} \qquad \forall i : 1 \le i \le n$$

because there are n possible outcomes to choose from in total. If $\Omega = \{H, T\}$, or $\{1, 2, 3, 4, 5, 6\}$, or the cards in a deck, having such a $p(\cdot)$ is called fairness, but in probability, we give another name to this type of symmetry: **uniformity**.

With uniformity, probability is simply proportion. To find the probability of some set of outcomes labelled as "favorable," $A \subseteq \Omega$, we can just count, then divide:

$$p(A) = \sum_{\omega \in A} p(\omega) = \sum_{\omega \in A} \frac{1}{|\Omega|} = \frac{|A|}{|\Omega|} = \frac{\text{number of favorable outcomes}}{\text{total number of outcomes}}.$$

Or, if Ω represents a region in space, with each outcome being a point, then the probability that a randomly chosen point lands in a particular region $A \subseteq \Omega$ is also a proportion: the measure (length, area, or volume) of A, divided by the total measure of Ω.

While the definition above marked the beginning of *probability theory*, the mathematical formalization of probability, the field eventually grew past it for a simple reason: symmetry does not always apply. One of the more well-known jokes from overextending symmetry is that

"Every event has probability $\frac{1}{2}$: it either happens or it doesn't."

So, with symmetry in mind, let us turn to a more powerful definition of probability.

■