Reading sample · Chapter 2, pages 23–30
https://aibook.art

AI in Colors

From Simple Neural Nets to Large Language Models

DIMITRI REISWICH

Chapter 2

Building Functions

This chapter is all about building your own functions, and a bit about Minecraft, but more on that later. Think back to the functions you learned in school, how many do you remember? Most people can name only a handful: maybe the linear function, x^2, the logarithm, and the sine and cosine functions. These are useful examples, but for our purposes we will work with an even simpler set. We will use just three very simple functions as building blocks and learn how to combine them to construct more complex functions, including those you may remember from school. This is similar to Minecraft, you start with a limited set of simple blocks, yet with them you can build complex structures such as buildings, bridges, and castles. Our goal is to combine these building blocks to construct functions that fit the data we saw in the last chapter.

These three building blocks are a line, a step, and a hockey stick. They are shown in Figure 2.1. We will refer to the line as $f_{/}$, to the step as $f_{\_\ulcorner}$, and to the hockey stick as $f_{\_/}$, where their visual appearance is reflected in the notation. If a friend happens to peek over while you are reading, I do not want them to think you are diving into trivial literature, this is AI, after all. At first glance, this notation might seem to represent complex and abstract mathematics. In reality, it is just a set of small icons designed to make it easier to remember which function is which.

At first glance, these may seem like simple toy functions, too limited to be useful beyond this book. But you might be surprised, they form the foundation of many machine learning models, including those running on your phone. These three functions are easy to underestimate because of their apparent simplicity. However, when many of them are combined and chained together, they function like atoms in the real world—individually small and seemingly insignificant, yet together, they can represent everything around us. So, let's introduce each function one by one. It's important to take the time to understand these functions as well as their diagrams, as they will be essential throughout the rest of the book.

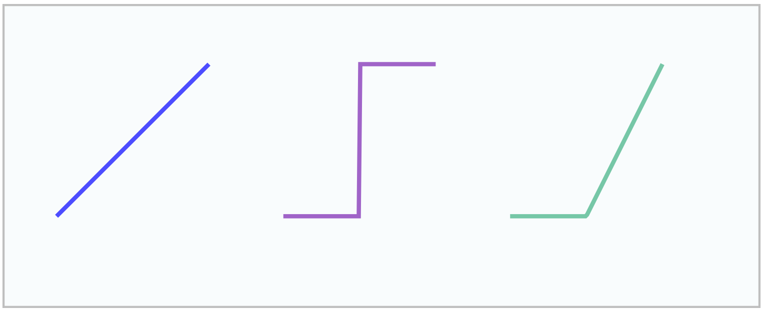


Figure 2.1: Line $f_{/}(x)$, Step $f_{\lrcorner\ulcorner}(x)$, Hockey $f_{\_/}(x)$

2.1 Line

In the real world, we can observe many things that follow a straight line. We can describe these as **line**ar. The mathematical counterpart of this is a function known as a `linear function`. We will simply call it *Line*. In the same spirit, we will call the other functions *Step* and *Hockey*. The goal of this informal naming is not to be humorous, but to make the names easy to remember throughout the rest of the book through their visual association.

2.1.1 Single Variable

We will start with a scenario where the function takes a single numerical **input** and produces a single numerical **output**. The single numerical value is also known as a `scalar`. The simplest formulation for a linear mapping is:

$$f_{/}(x) = x$$

This version of Line is probably the most boring and laziest function in the world. It simply returns the same value it was given. Let's get back to shopping with another example: you see peaches priced at 1 per piece. If you buy 5, the total cost is also 5. This relationship can be described by Line, it takes the quantity as input and returns it as output, where the output now represents the price. Let us make Line more interesting by adding a term b, which gives the new equation:

$$y = f(x + b) = x + b$$

We can express the same relationship in a notation that highlights the presence of the term b:

$$f_{/}(x; b) = x + b$$

Here, the function still takes a single input x, and the term b is written after the semicolon to emphasize that it is a separate, fixed parameter of the function. Now, the function returns the input number x plus a constant b. Every input value is adjusted by a fixed amount b, which is why this term is called `bias`.

Returning to our shopping example, imagine receiving a voucher from the supermarket with a value of 3. If you buy 5 peaches, your checkout price will be 5 **minus** the voucher amount of 3, totaling 2. Here, the voucher acts as the bias, and it enters the equation with a negative sign $b = -3$. Examples of plots with a bias of 3 and -3 are shown in Figure 2.2. The effect of the bias can be described in two equivalent ways: it can be seen as either an **up or down** shift or a **left or right** shift.

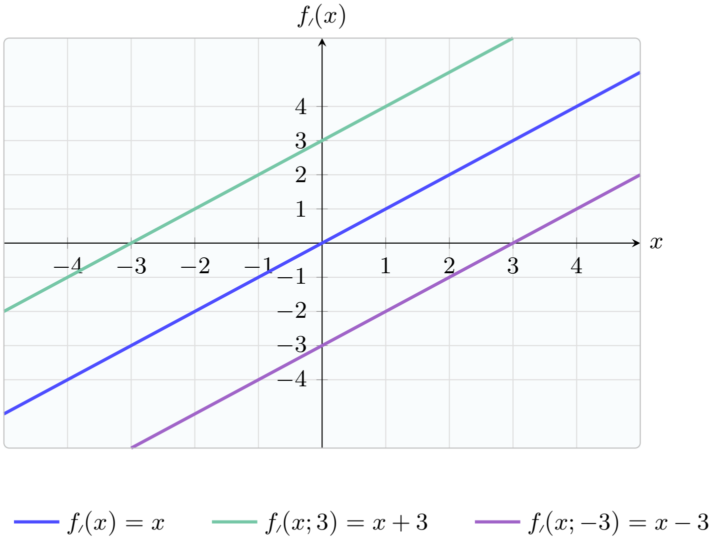


Figure 2.2: Linear function with different bias terms.

We'd now like to rotate Line (without the bias) to make it steeper or flatter. This can be done by multiplying the input x by a factor w, giving us:

$$y = wx$$

As before, we can express this using function notation in two ways:

$$y = f_{/}(wx) = wx \text{ or } f_{/}(x; w) = wx,$$

where the second form shows w explicitly as a separate, fixed parameter of the function. The parameter w is called `weight`. Back to the peaches, bad news for your wallet. The price has now jumped to 2 per peach. This means that buying five peaches will cost $2 \cdot 5 = 10$. The 2 in front of the quantity represents the weight w. A visualization of Line for different weights is shown in Figure 2.3. As indicated, the weight controls the rotation of Line. Note that you might be familiar with the letter m from your school days instead of w, but it's the same thing: w is

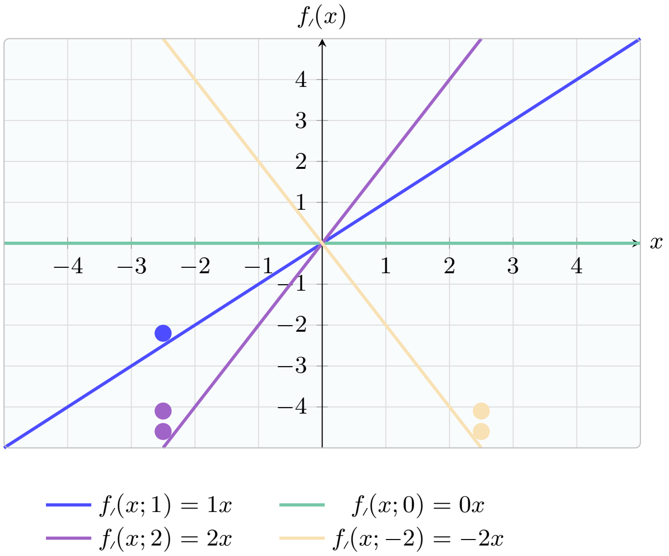


Figure 2.3: Linear function with different weights.

really just an m lying on its back.

To build an intuition for the term *weight*, imagine that Figure 2.3 represents a playground seesaw centered at point $(0, 0)$. A weight of 0 corresponds to a perfectly balanced seesaw, keeping it horizontal, as illustrated with the green line in the figure. Adding a weight of 1 on the **left** hand side tilts the seesaw, causing it to sink lower where the weight is and rise higher on the other side. Figure 2.3 illustrates this with the blue line and a single blue weight in the shape of a circle. If we add two weights, the seesaw will tilt even more, as shown by the pink line with the two weights. Negative weights can also be represented by placing the weight to the **right** of the y-axis, as shown by the orange line.

If we now combine both the weight and the bias, the full expression becomes:

$$y = f_{/}(wx + b) = wx + b \text{ or } f_{/}(x; w, b) = wx + b, \tag{2.1}$$

where, as before, the weight w and bias b are listed after the semicolon, with the weight always appearing first. This notation emphasizes that w and b are fixed parameters of the function, while x remains the variable input. As you can see, these parameters adjust Line in various ways, rotating and shifting it. Some examples are shown in Figure 2.4. Think of the weight as a knob that rotates the line, while the bias acts like a slider that moves it up or down in parallel. I have illustrated this in Figure 2.5.

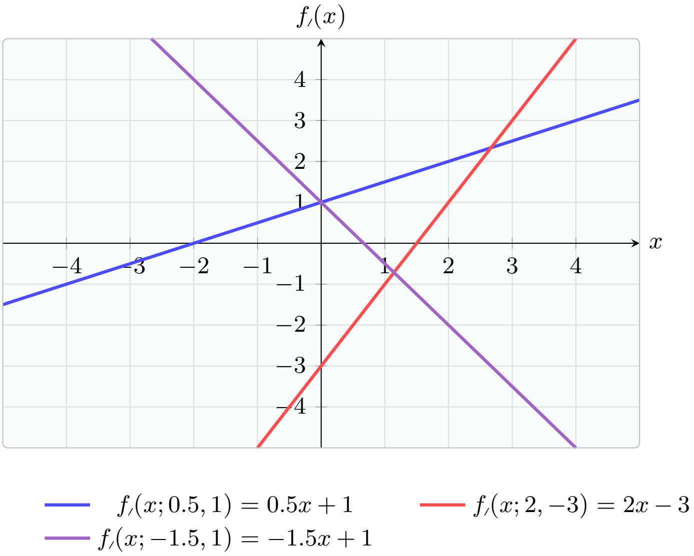


Figure 2.4: Linear function with different weights and biases.

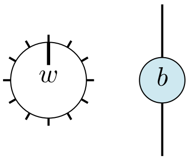


Figure 2.5: Weight knob for rotating, bias slider for parallel shifting of a linear function.

2.1.2 Diagrams

There are different ways to understand a function. First, there's the formula, the "recipe" for how it works. Most people find formulas a bit dry (okay, very dry). Then, there's a chart that shows the function's output for a range of inputs, a kind of "photo" of the formula. While helpful, it doesn't reveal much about what's actually happening inside the function.

Think of it like baking bread. The formula is the written recipe:

> *Take 3 glasses of water and 900 g of flour. Mix them in a bowl and bake at 200 degrees for 50 minutes. Out comes the bread.*

The recipe tells you exactly what steps to follow. The chart, on the other hand, is more like a picture of the ingredients and the final loaf. It shows what goes in and what comes out, but it doesn't explain the process of turning raw ingredients into the finished product.

That's where a third approach comes in: a diagram, a visual breakdown of the process that illustrates key steps and ingredient quantities. Our bread recipe is now shown in Diagram 2.6. No words are needed.

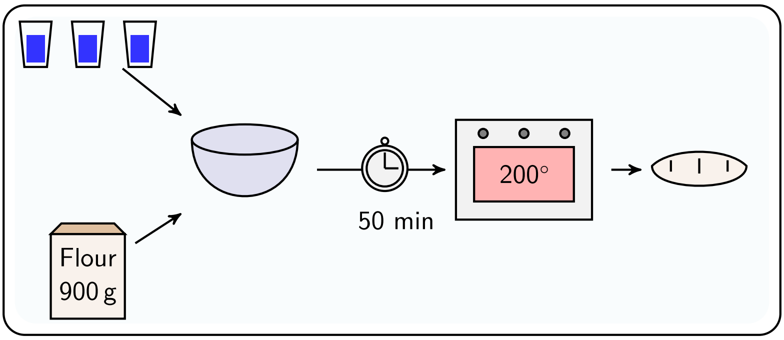


Diagram 2.6: Bread making process.

Back to functions. Moving forward, we'll use all three methods to understand functions, making the journey more visual and engaging rather than just relying on formulas. A diagram is essentially a visual blueprint that illustrates the key parameters and flow of a function. Humans naturally process visuals and patterns more easily than abstract concepts, making diagrams a powerful tool for understanding. They serve as a foundation for the concepts introduced in the upcoming chapters, and you'll see a lot of them. Take the time to grasp them now, it will pay off later. Keep in mind that a diagram is just another way of representing the raw formula. It doesn't add any extra computation or functionality; it simply makes the structure easier to understand.

Let us continue with the previous example and create a diagram to represent our linear function $f_{/}(2x - 3) = 2x - 3$. This function takes a linear, that is weighted and biased input, and returns it unchanged. However, in the sections ahead, we will look at situations where this linear input is passed into other functions, such as $f_{\_/}$, to introduce nonlinearity, meaning the output will no longer follow a straight line.

We will now visualize the function in a diagram. To begin, we rewrite the equation by using $x2$ instead of $2x$ and use colors to highlight the individual components, making the structure and flow of the computation easier to follow. The equation now becomes

$$f_{/}(\ x \cdot 2\ -3)$$

We can represent this equation as shown in Diagram 2.7, where the colors from the formula are clearly reflected. The diagram should be read from left to right and can be seen as a pipeline that progressively feeds results forward. The circles are called `nodes`, a term we will use very often. The arrows in between connect the

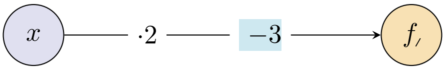


Diagram 2.7: $f_{/}(x \cdot 2 - 3)$

nodes. The diagram begins with the indigo circle, representing the **input** node x, the starting point of the pipeline. From there, an arrow indicates the direction to follow. The first step along the path is multiplying x by the weight 2. At this stage, the intermediate result in the pipeline is:

$$x \cdot 2$$

The next step in the pipeline is the bias -3, which is added to the current result. At this stage, the total is

$$x \cdot 2 - 3$$

since adding a negative number is the same as subtracting. Finally, we feed this result into the function $f_{/}(x) = x$ which is illustrated visually in the orange **output** node, completing the representation of the formula.

While the current diagram captures all the key elements, it appears a bit verbose and cluttered. To make it cleaner, we will simplify the notation by removing the multiplication $(\cdot)$ sign. It will be understood that each input x is multiplied by the number shown first, then the bias is added, and the result is passed to the function shown in the output node. To streamline things further, we will use the visual symbol / instead of writing out $f_{/}$. The final, simplified version of the diagram is shown in Diagram 2.8.

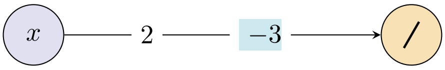


Diagram 2.8: $f_{/}(x \cdot 2 - 3)$

The background colors were chosen for a reason:

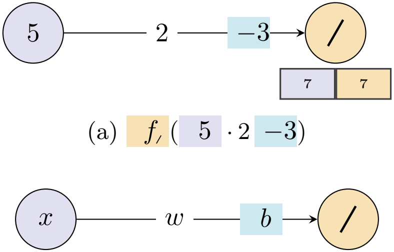
- the **in**put x is **in**digo,
- the **w**eight w is **w**hite,
- the **b**ias b is **b**lue,
- the **o**utput is **o**range.

This makes it easy to remember the different variables going forward throughout the rest of the book.

Let us now visualize the diagram for a concrete input, such as $x = 5$, which results in

$$5 \cdot 2 - 3 = 7$$

as shown in Figure a) of Diagram 2.9. Under the output node, you'll see a small helper box. On the left side, it shows the input to the function (in indigo), and on the right, the output after the function is applied (in orange). In this case, the input and output are the same given that we are looking at $f_{/}(x) = x$. This process of taking specific inputs and passing them through the function to compute the result is called a `forward pass` and is a concept we will revisit later in more complex settings. A representation with generic parameters is shown in Figure b) of Diagram 2.9, where the helper box is omitted since we are not considering concrete inputs.



(a) $f_{/}(5 \cdot 2 - 3)$

(b) $f_{/}(x \cdot w + b)$

Diagram 2.9: Linear function with scalar input and output.

And with that, we've completed the introduction to our linear function diagrams. As you can hopefully see, these visual diagrams are very helpful for understanding how a function works internally. They offer a perspective that formulas alone often can't provide. Plus, they're colorful and undeniably beautiful.