More Derivative Examples

(SPEECH)
In

(DESCRIPTION)
Text, Basics of Neural Network Programming. More derivatives examples. Website, deep learning, dot, A.I.

(SPEECH)
this video, I'll show you a slightly more complex example where the slope of the function can be different to different points in the function.

Let's start with an example.

(DESCRIPTION)
New slide, Intuition about derivatives.

(SPEECH)
You have plotted the function $f(a) = a^2$.

Let's take a look at the point $a=2$.

So a^2 or $f(a) = 4$.

Let's nudge a slightly to the right, so now $a=2.001$.

$f(a)$ which is a^2 is going to be approximately 4.004.

It turns out that the exact value, you call the calculator and figured this out is actually 4.004001.

I'm just going to say 4.004 is close enough.

So what this means is that when $a=2$, let's draw this on the plot.

So what we're saying is that if $a=2$, then $f(a) = 4$ and here is the x and y axis are not drawn to scale.

Technically, does vertical height should be much higher than this horizontal height so the x and y axis are not on the same scale.

But if I now nudge a to 2.001 then $f(a)$ becomes roughly 4.004.

So if we draw this little triangle again, what this means is that if I nudge a to the right by 0.001, $f(a)$ goes up four times as much by 0.004.

So in the language of calculus, we say that a slope that is the derivative of $f(a)$ at $a=2$ is 4 or to write this out of our calculus notation, we say that d/da of $f(a) = 4$ when $a=2$.

Now one thing about this function $f(a) = a^2$ is that the slope is different for different values of a.

This is different than the example we saw on the previous slide.

So let's look at a different point.

If $a=5$, so instead of $a=2$, and now $a=5$ then $a^2=25$, so that's $f(a)$.

If I nudge a to the right again, it's tiny little nudge to a, so now $a=5.001$ then $f(a)$ will be approximately 25.010.

So what we see is that by nudging a up by .001, $f(a)$ goes up ten times as much.

So we have that $d/da\ f(a) = 10$ when $a=5$ because $f(a)$ goes up ten times as much as a does when I make a tiny little nudge to a.

So one way to see why did derivatives is different at different points is that if you draw that little triangle right at different locations on this, you'll see that the ratio of the height of the triangle over the width of the triangle is very different at different points on the curve.

So here, the slope=4 when a=2, a=10, when a=5.

Now if you pull up a calculus textbook, a calculus textbook will tell you that d/da of f(a), so $f(a) = a^2$, so that's d/da of a^2.

One of the formulas you find are the calculus textbook is that this thing, the slope of the function a^2, is equal to 2a.

Not going to prove this, but the way you find this out is that you open up a calculus textbook to the table formulas and they'll tell you that derivative of 2 of a^2 is 2a.

And indeed, this is consistent with what we've worked out.

Namely, when a=2, the slope of function to a is 2x2=4.

And when a=5 then the slope of the function 2xa is 2x5=10.

So, if you ever pull up a calculus textbook and you see this formula, that the derivative of $a^2=2a$, all that means is that for any given value of a, if you nudge upward by 0.001 already your tiny little value, you will expect f(a) to go up by 2a.

That is the slope or the derivative times other much you had nudged to the right the value of a.

Now one tiny little detail, I use these approximate symbols here and this wasn't exactly 4.004, there's an extra .001 hanging out there.

It turns out that this extra .001, this little thing here is because we were nudging a to the right by 0.001, if we're instead nudging it to the right by this infinitesimally small value then this extra every term will go away and you find that the amount that f(a) goes out is exactly equal to the derivative times the amount that you nudge a to the right.

And the reason why is not 4.004 exactly is because derivatives are defined using this infinitesimally small nudges to a rather than 0.001 which is not.

And while 0.001 is small, it's not infinitesimally small.

So that's why the amount that f(a) went up isn't exactly given by the formula but it's only a kind of approximately given by the derivative.

To

(DESCRIPTION)
New slide, More derivative examples.

(SPEECH)
wrap up this video, let's just go through a few more quick examples.

The example you've already seen is that if $f(a) = a^2$ then the calculus textbooks formula table will tell you that the derivative is equal to 2a.

And so the example we went through was it if (a) = 2, f(a) = 4, and we nudge a, since it's a little bit bigger than f(a) is about 4.004 and so f(a) went up four times as much and indeed when a=2, the derivatives is equal to 4.

Let's look at some other examples.

Let's say, instead the $f(a) = a^3$.

If you go to a calculus textbook and look up the table of formulas, you see that the slope of this function, again, the derivative of this function is equal to $3a^2$.

So you can get this formula out of the calculus textbook.

So what this means?

So the way to interpret this is as follows.

Let's take a=2 as an example again.

So f(a) or $a^3=8$, that's two to the power of three.

So we give a a tiny little nudge, you find that f(a) is about 8.012 and feel free to check this.

Take 2.001 to the power of three, you find this is very close to 8.012.

And indeed, when a=2 that's $3x2^2$ does equal to 3x4, you see that's 12.

So the derivative formula predicts that if you nudge a to the right by tiny little bit, f(a) should go up 12 times as much.

And indeed, this is true when a went up by .001, f(a) went up 12 times as much by .012.

Just one last example and then we'll wrap up.

Let's say that f(a) is equal to the log function.

So on the right log of a, I'm going to use this as the base e logarithm.

So some people write that as log(a).

So if you go to calculus textbook, you find that when you take the derivative of log(a).

So this is a function that just looks like that, the slope of this function is given by 1/a.

So the way to interpret this is that if a has any value then let's just keep using a=2 as an example and you nudge a to the right by .001, you would expect f(a) to go up by 1/a that is by the derivative times the amount that you increase a.

So in fact, if you pull up a calculator, you find that if a=2, f(a) is about 0.69315 and if you increase f and if you increase a to 2.001 then f(a) is about 0.69365, this has gone up by 0.0005.

And indeed, if you look at the formula for the derivative when a=2, d/da f(a) = 1/2.

So this derivative formula predicts that if you pump up a by .001, you would expect f(a) to go up by only 1/2 as much and 1/2 of .001 is 0.0005 which is exactly what we got.

Then when a goes up by .001, going from a=2 to a=2.001, f(a) goes up by half as much.

So, the answers are going up by approximately .0005.

So if we draw that little triangle if you will is that if on the horizontal axis just goes up by .001 on the vertical axis, log(a) goes up by half of that so .0005.

And so that 1/a or 1/2 in this case, 1a=2 that's just the slope of this line when a=2.

So that's it for derivatives.

There are just two take home messages from this video.

First is that the derivative of the function just means the slope of a function and the slope of a function can be different at different points on the function.

In our first example where f(a) = 3a those a straight line.

The derivative was the same everywhere, it was three everywhere.

For other functions like f(a) = a^2 or f(a) = log(a), the slope of the line varies.

So, the slope or the derivative can be different at different points on the curve.

So that's a first take away.

Derivative just means slope of a line.

Second takeaway is that if you want to look up the derivative of a function, you can flip open your calculus textbook or look up Wikipedia and often get a formula for the slope of these functions at different points.

So that, I hope you have an intuitive understanding of derivatives or slopes of lines.

Let's go into the next video.

We'll start to talk about the computation graph and how to use that to compute derivatives of more complex functions.