Module 2 59:27. ①

Jacobian

Previously we saw we can differentiate functions of multiple variable and ~~it~~ it isn't much harder than the unit variable calculus at the start

Here, we going to introduce the $\boxed{\text{Jacobian}}$ which brings in some of the ideas of LA. to build these partial derivatives into something particularly useful

The concept of the Jacobian can be applied to variety of different problems

But in the context of getting started with optimization of ML, there is a particular scenario that comes up a lot, which is the Jacobian of a single function of many variables.

In short, if you have a function with many variables: $\Rightarrow$ f of $x_1, x_2 \ldots$

$$f(x_1, x_2, x_3 \ldots)$$

Then the Jacobian is simply a vector, where each entry is the partial derivative of f w.r.t each one of those variables in turn.

$$J = \left[\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \frac{\partial f}{\partial x_3}, \ldots\right]$$

By convention, we write this as a row vector, rather than a column vector for reason that will become clear later.

Lets start by looking at simple functions, to see how straightforward building a Jacobian can be.

Consider the function:

$$f(x,y,z) = x^2y + 3z$$

To build a Jacobian, we just find each of the partial derivative of the function one by one

w.r.t x

$$\frac{\partial f}{\partial x} = 2xy$$ [everything else is constants]

w.r.t y

$$\frac{\partial f}{\partial y} = x^2$$

w.r.t z

$$\frac{\partial f}{\partial z} = 3.$$

$$J = (2xy, x^2, 3) \quad A$$

Bringing all of it together A, But what does this tell us?

$\Rightarrow$ we have an algebraic expression for a vector, which when we give it a specific x, y, z coordinate, will

(9)

return a vector pointing in direction of steepest slope of this function.

The vector for this particular function has a constant contribution in the z direction, which does not depend on the location selected.

eg at point 0,0,0, we can see that our Jacobian will just be

$$J(0,0,0) = [0, 0, 3]$$

∴ Our Jacobian is vector of length 3, pointing directly in the z direction.

Some of the numerical method that will be discussed later in course, will require us to calculate the Jacobian in 100's of dimensions.

However even for the 3D example we just solved, graphical representation is quite difficult.

So we will revert to 2 Dimensions, to see what's going on.

But to keep things interesting, we will look at a particularly complicated, but attractive function:

Here's the equation we going to plot, which's shown briefly, just to so we can see, even though it looks huge we'll agree that just with tools we covered, we really could calculate the partial derivative, But we would not really learn anything new from grinding through it.

Instead we will be shown the results graphically:

$$z(x,y) = 3(1-x)^2 e^{-x^2-(y+1)^2} - 10\left(\frac{x}{5} - x^3 - y^5\right)e^{-x^2-y^2} - \frac{1}{3}e^{-(x+1)^2-y^2}$$

[See ......] picture

- - - - - - -

Because of the nature of this particular function, we know that nothing interesting happens outside of the region shown [so we can forget about it for now]

Although the plot is fairly clear, our intuition about the gradient is a bit lacking in this format, so lets briefly make things 3D, where the value of z is also represented by the height

So, as indicated by start of ~~[illegible]~~ scene, the Jacobian is simply a vector that we can calculate for each location of this plot, that points in direction of steepest uphill slope

Further more, the steeper the slope the greater the magnitude of the Jacobian at that point. 7

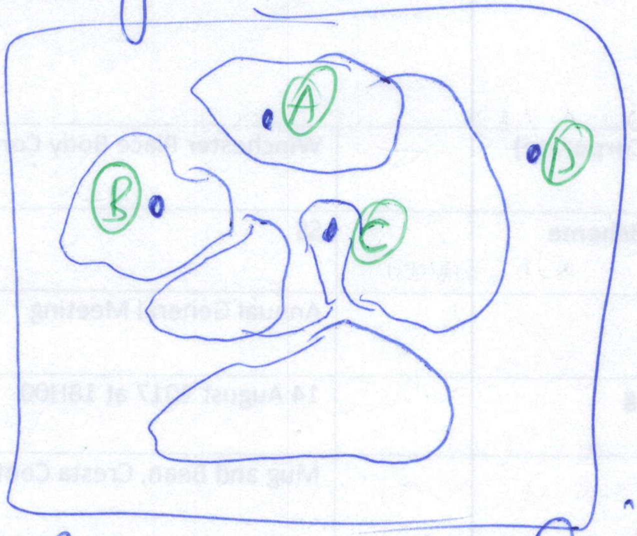
Hold the image of the 3D space in your mind, as we now go back to 2D, and rather than showing all of the grid points, we use to plot the graph, lets instead convert to contour plot, where just like a map of actual ~~act~~ mountains, we will draw a line along the regions of the same height, which for us means the same value of z

This removes all the clutter from plot, which is useful for the final step which will be adding lots of Jacobian vectors on top of our contour plot.

However, before doing it, lets look at four points an see ~~if~~ if my intuition is intune, by guessing

8

which one will have the Jacobian with the largest magnitude



Overlaying the Jacobian vector field, we can see that they are clearly all pointing uphill; away from the ~~low~~ dark regions, and towards the high bright regions.

Also we see that where the contour lines are tightly packed this is where we find our largest Jacobian vectors such as at point A.

Whereas at the peak of mountain and bottom of valleys or even out on the wide flat planes our gradient's and therefore our Jacobians are small.

Hope, that these clear 2D example will give us confidence to trust maths, when we come up against much higher dimensions and problems later in case. (9)