2.19.20

1

Module 4

Multivariate Taylor.

So far in the module, we've been introduced to concepts of Power Series, and been showed how to use it to approximate function

We've also seen how Power Series can be used to give you some indication of the size of the error that results from using these approximations, which is very useful when applying numerical methods

The last idea here that we going to very briefly look at is, is upgrading the Power Series from one dimension to its more general multivariate form.

Just to recap the notational options from the previous session

We saw we can re-express the Taylor series from a form that emphasises building the approximation of a function at point p to a totally equivalent form, that emphasises using that function to evaluate other point, that are small distance $\Delta x$ away.

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(p)}{n!}(x-p)^n \qquad f(x+\Delta x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(x)}{n!}\Delta x^n$$

Just to make sure you can see that they are the same we can write out the first 4 terms one above the other to make the comparison clear.

$$f(x) = f(p) + f'(p)(x-p) + \tfrac{1}{2} f''(p)(x-p)^2 + \tfrac{1}{6} f^{(3)}(p)(x-p)^3 + \ldots$$

$$f(x+\Delta x) = f(x) + f'(x)\Delta x + \tfrac{1}{2} f''(x)\Delta x^2 + \tfrac{1}{6} f^{(3)}(x)\Delta x^3 + \ldots$$

In this session we going to continue using the $\Delta x$ notation as it is more compact, which will come in handy when we are in higher dimensions

$$f(x+\Delta x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(x)}{n!} \Delta x^n$$

We won't be working through the multivariate Taylor series derivation in any detail, as all we really need to take away from this session, is what a multidimensional Power Series will look like, both as an equation and in graph.

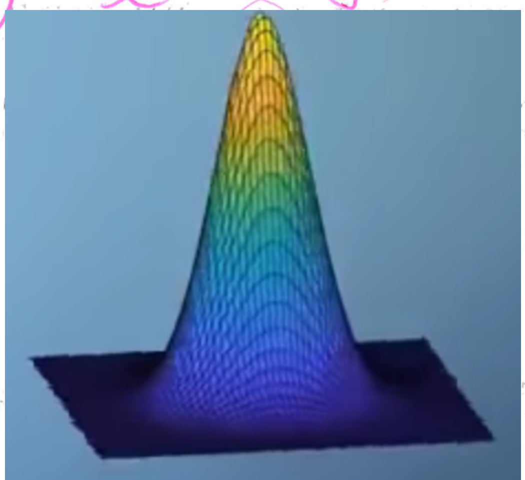
So keeping in mind ~~our~~ our 1 dimensional expression,
Lets start by looking at 2Dim case, where
f is now a function of the 2 variables $x, y$

$$f(x+\Delta x, y+\Delta y) = ???$$

So our truncated Taylor Series expression will enable us to approximate the function at some nearby point $x+\Delta x, y+\Delta y$.

Lets look at the 2Dim (simple) gaussian function

$$f(x,y) = e^{-(x^2+y^2)}$$



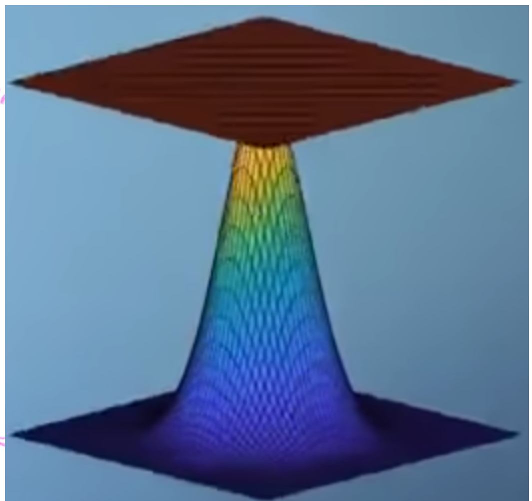
Its a nice, well behaved function, with a simple maximum, at point 0,0

In a 1 Dim analysis, our Power Series will give us an expression for a 1 Dim function.

However, in a 2 Dim Case, our Power Series will of Course now give us a 2 Dim ~~(form)~~ function which we will more intuitively refer to as a surface

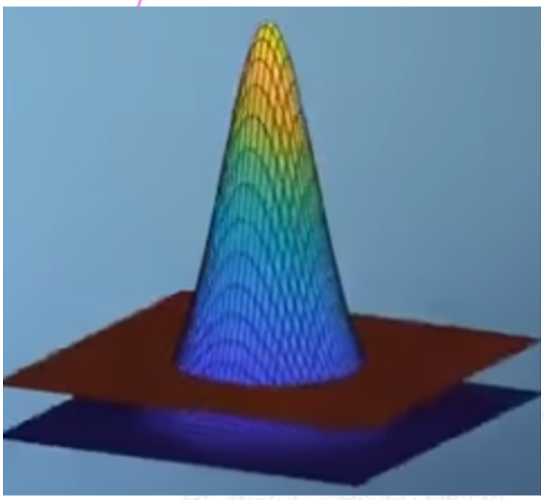
Just as with 1 Dim Case, our zeroth order of approximation are simply a value of function at that point applied everywhere. Which means in 2Dim/D, this will just be a flat surface

So a zeroth order approximation at the peak will look like this,



$g_0(x+\Delta x, y+\Delta y)$

And a zeroth order approximation somewhere on the side of the "bell", and will look like the:



This is fairly straightforward, But now lets think about the 1st order approximation.

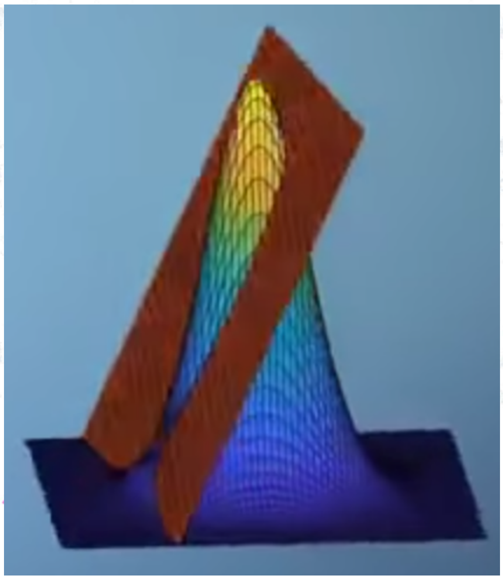
Drawing the analogy from the 1D Case again, the 1st order approximation, should have a height and gradient

So we still expecting a straight surface, but this time it can be at an angle

If we are calculating it ~~at~~ the peak, it will not be a very exciting case, as its a turning point, so the gradient ~~the~~ is 0

However, lets look again at the point at side of Slope

Taking a first order approximation again gives us a surface with same height and gradient at point



Finally, lets look at the 2nd order approximation, for these 2 points

We now expecting, some kind of parabolic Surface

However, if I calculate it at peak, nothing seems to appear

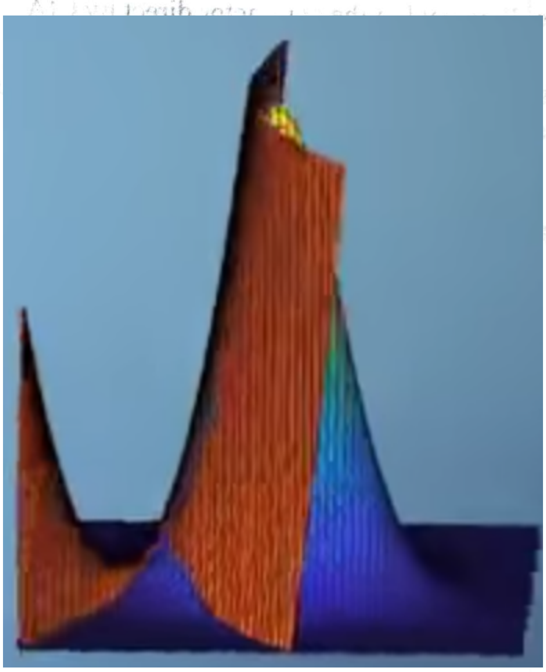
But, thats just because we've calculated

Created a parabola inside our bell Curve.

Do we need to look from underneath, to even see it.

Finally, if we plot the 2nd order approximation at point on side of Bell Curve, we end up with some interesting Saddle function.

But it is fairly clear, by eye, that the gradient of the curvature are matching up nicely at that point



although this approximation is evidently not useful outside a fairly small region around the point

We now going to take a look at how we would write expressions for the functions

So, If we want to build a Taylor Series Expansion, of the 2D m function f, at point x, y and ~~then~~ then use ~~it~~ it to evaluate the function at point $x+\Delta x$, $y+\Delta y$, our zeroth order approximation is just a flat surface with same height as the function at our expansion point

The first order approximation, incorporate the gradient information in two directions

$f(x,y)$

~~Ordance~~

Once again we thinking about how ~~Rise~~ ~~run~~

rise = gradient × run

Notice here the compactness I am using the partial symbol, with subscript to signify a derivative w.r.t a certain variable

$$+\left(\partial_x f(x,y)\Delta x + \partial_y f(x,y)\Delta y\right)$$

If we look now at the second order approximation we can see that the coefficient of $\frac{1}{2}$ still remains as per the 1Dim case

But now we have 3 terms, all of which are second derivatives

$$+\frac{1}{2}\left(\partial_{xx} f(x,y)\Delta x^2 + 2\partial_{xy} f(x,y)\Delta x\Delta y + \partial_{yy} f(x,y)\Delta y^2\right)$$

The 1st has been differentiated w.r.t x twice.

The last w.r.t y twice

and middle term has been differentiated w.r.t each of x and y

Now, there are off course higher order terms but we have already got all more than enough to play with here

So lets look at again the 1st order term.

Its the sum of the products of the 1st derivatives with step size

But hopefully this will remind you about the discussion of the Jacobian.

So we can actually express this as just the Jacobian multiplied by vector containing $\Delta x, \Delta y$

$$\therefore \; + \left(\partial_x f(x,y)\Delta x + \partial_y f(x,y)\Delta y\right)$$

$$\Rightarrow \left[\partial_x f(x,y), \partial_y f(x,y)\right]\begin{bmatrix}\Delta x \\ \Delta y\end{bmatrix} \Rightarrow J_f \Delta \mathbf{x}$$

Where we can write as $\Delta$ bold $x$, where bold $x$ signifies a vector, containing these 2 terms

Finally, if we look at the second order term in same way, and notice that these second derivatives can be collected into a matrix which we previously defined as our Hessian

Then to make the sum we need, we now need to multiply our $\Delta x$ vector by the Hessian

$$+\frac{1}{2}\left(\partial_{xx} f(x,y)\Delta x^2 + 2\partial_{xy} f(x,y)\Delta x\Delta y + \partial_{yy} f(x,y)\Delta y^2\right)$$

$$\Rightarrow \frac{1}{2}\begin{bmatrix}\Delta x, \Delta y\end{bmatrix}\begin{bmatrix}\partial_{xx} f(x,y) & \partial_{xy} f(x,y)\\ \partial_{yx} f(x,y) & \partial_{yy} f(x,y)\end{bmatrix}\begin{bmatrix}\Delta x\\ \Delta y\end{bmatrix}$$

$$\Rightarrow \frac{1}{2}\Delta x^t H_f \Delta x$$

And then again, by the transpose of $\Delta x$ vector

And that's it

So now we have a more compact expression for the 2nd order multivariate Taylor series expansion which brings together so much of our Calculus and LA skills and makes use of the Jacobian and Hessian concepts

$$\boxed{f(x+\Delta x) = f(x) + J_f\Delta x + \frac{1}{2}\Delta x^t H_f \Delta x + \ldots}$$

Although we have been talking about the 2 Dim Case in this session so far, we can actually have any number of Dimensions contained in our $\Psi$, $H$ or $\Delta x$ terms

So this immediately generalizes from 2D to multidimensional hypersurfaces