

Module 4

Multivariate Taylor.

So far in the module, we've been introduced

to Concept of Power Series, and Seen Shown + Definition

how to use it to approximate general

We've also seen how Page Sizes can

be used to give you some indicators of the size of the error that results

from very low approximations, with
1.100 ordinary numerical

is very helpful when applying numerical methods

The last idea here that we going to try

Brendy lost cat is upgrading to Power

Sew from one dimension to another

More general multivariate form.

Just to recap the notational option from
the previous session

We saw we can re-express the Taylor series
from a form that emphasizes building up
approximation of a function at point P
to a totally equivalent form, that emphasizes
using that function to evaluate other
points, that are small distance Δx away.

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

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

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

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

In this lesson we going to continue using the Δx notation
as it's 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
lesson, is what a multidimensional function
will look like, both as an equation and
in graph.

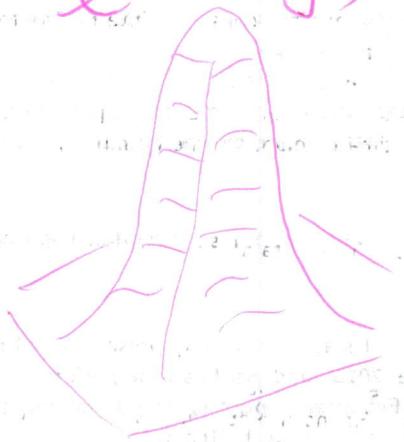
So keeping in mind our our 1 dimensional expressi,
lets start by looking at 2D in case, where
 f is now a function of the 2 variables x, y

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

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

Let's look at the 2D in (single) gaussian function

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



It's a nice, well behaved function, with a single
maximum at point $0,0$

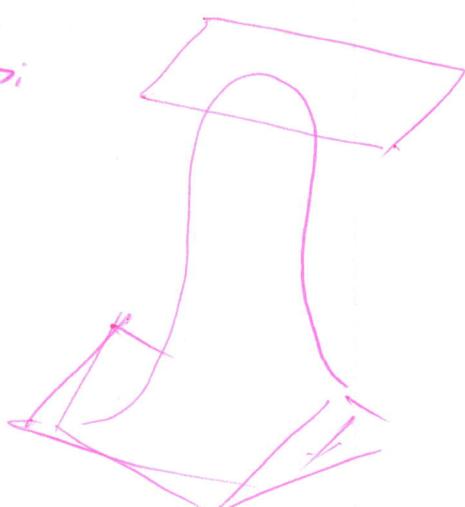
In a 1D analysis, our Power Series will give us an expression for a 1D function.

However, in a 2D case, our Power Series will now give us a 2D (or $f(x,y)$) function, which we will more intuitively refer to as a surface.

Just as with 1D case, our zeroth order of approximation are simply a value of function at that point applied everywhere.

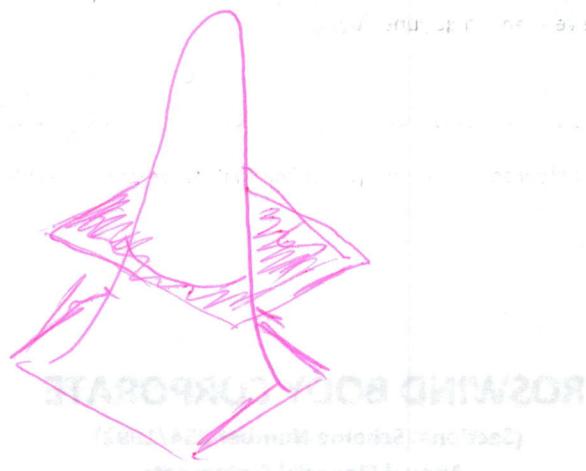
Which means in 2D, this will just be a flat surface

So a zeroth order approximation at the point 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 this:



This is fairly straightforward, But now let's think
about the 1st order approximation.

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

So we still expect 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 it
is a turning point, so the gradient ~~is~~ is 0

However, let's 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, let's look at the second order approximation, for the 2 points

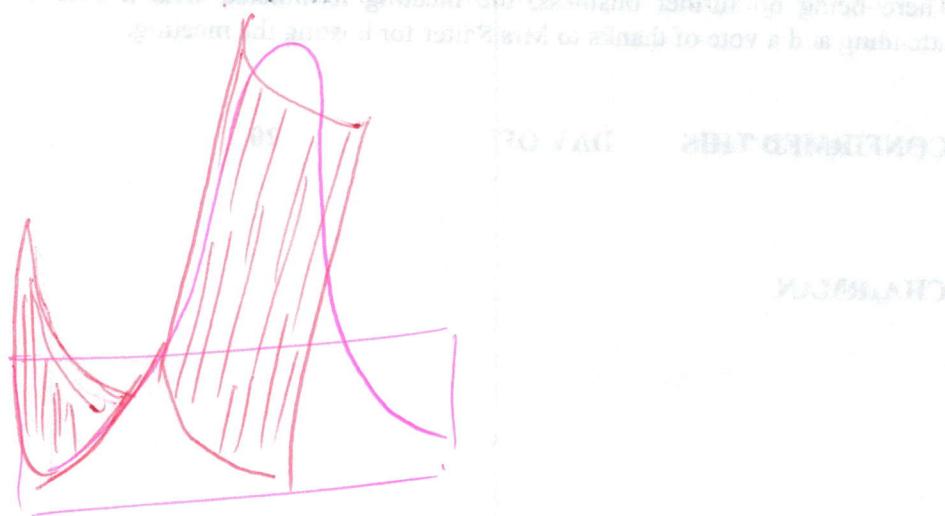
We know expecting some kind of parabolic surface however, if I calculate it at peak, nothing seems to appear but that just because we've calculated created a parabola inside our bell curve.

⑧

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 direction.

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

If we want to build a Taylor Series Expansion,
of the 2Dm function f , at point (x, y)
and then use it to evaluate the function
at point $x + \Delta x, y + \Delta y$, or z^{th} 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

$$f(x, y)$$

Once again we thinking about how
use = gradient \times n

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

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

If we look now at the second order approximation we can see that the coefficient of $\frac{1}{2}$ still remains as per the 1st order case. But now we have 3 terms, all of which are second derivatives.

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

The 1st has been differentiated wrt x twice. The last wrt y twice. And middle term has been differentiated wrt each of x and y.

Now, here 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.

It's the sum of the products of the 1st derivatives
with step size

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

So we can actually express this as just the
Jacobian multiplied by vector containing Δx , Δy

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

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

When we write a Δ bnd x , where

bnd x , signifies a vector, containing those
2 terms

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

(12)

Now to make the sum we need, we now need to multiply our Δx vector by the Hessian

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

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

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

And then again, by the transpose of Δx vector

And that's it

Now we have a nice compact expression for the second order multivariate Taylor series expansion which brings together so much of our Calculus and LA skills and makes just the Jacobian and Hessian concepts

$$f(x + \Delta x) = f(x) + J_f \Delta x + \frac{1}{2} \Delta x^T H_f \Delta x.$$

Although we have been talking about the 2D vs
3D case in the previous slides, we can actually
have any number of dimensions contained
in our f , H or \mathbf{x} terms.

Do we immediately generalize from 2D to
multidimensional hyper surfaces?