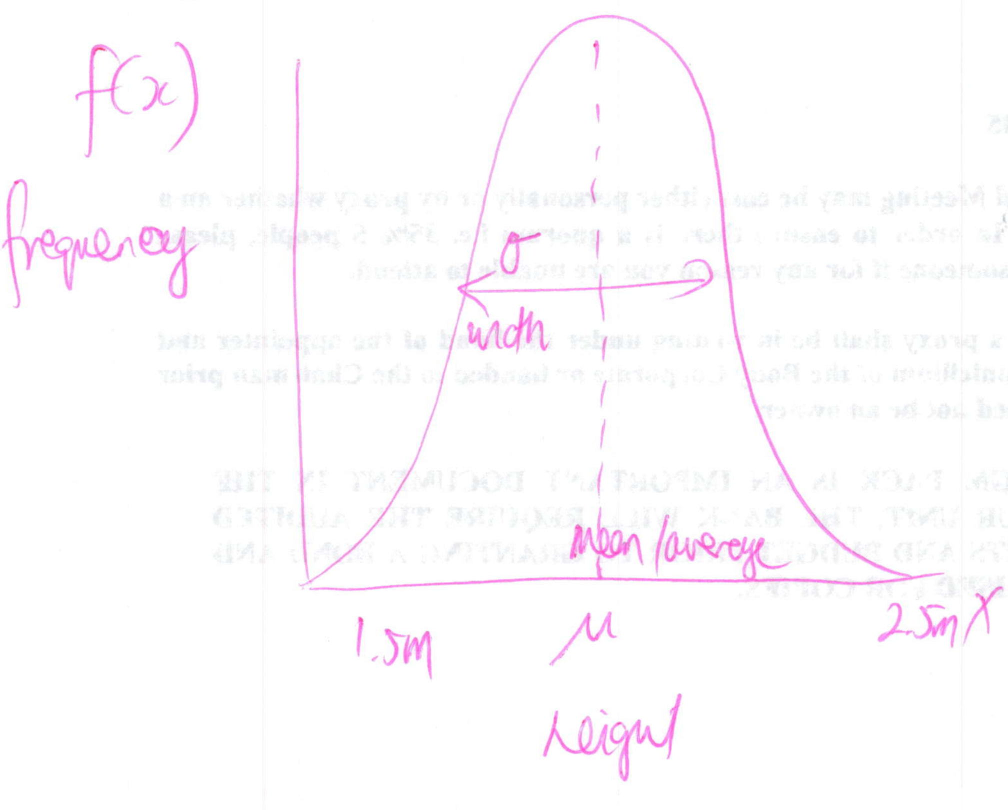
1

Module 5
Intro

David Dye here again!

Here, in this module, we'll start to use the Calculus ~~which~~ we have done, and put it together with vectors, in order to solve equations

In this first session, we'll look at a simple case, where we need to find the gradient, ~~the~~ the derivative, in order to solve equations, using, what's called, the Newton-Raphson method



Say we have, we have the distribution of lengths again with mean, average → mu (μ) and width (sigma σ) and we want to fit an equation to ~~fit~~ that distribution

So we don't have to ~~fit~~ ...... bother about carrying round all the data points

we just have a model with 2 (~~data points~~) parameters, a mean and width

And we can just do everything using just the model and that will be loads faster and simpler, and would let us make predictions and so on

and would be much, much nicer,

But, how do we find the right parameters for the model

How do we find the best μ and σ we can?

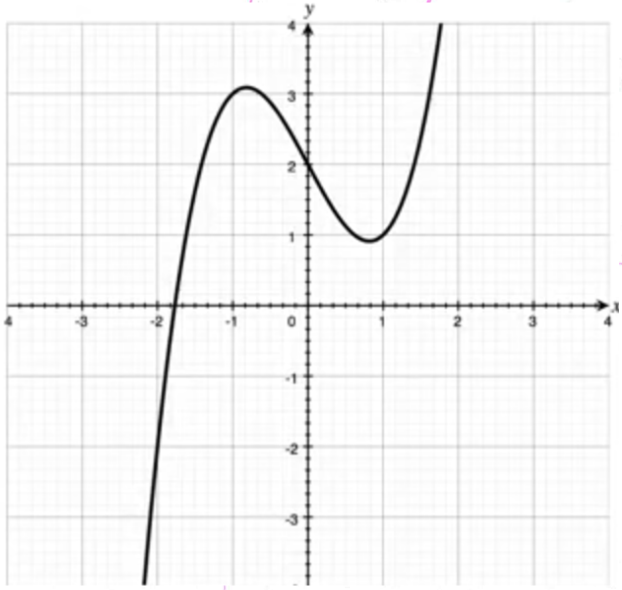
What we going to do is find some ~~expes~~ expression for how well the model fits the data and then look at how that goodness of fit ~~varying~~ varies as the fitting parameters μ and σ varying.

So, ~~try~~ try to solving an equation where the fitting parameters variable in the equation

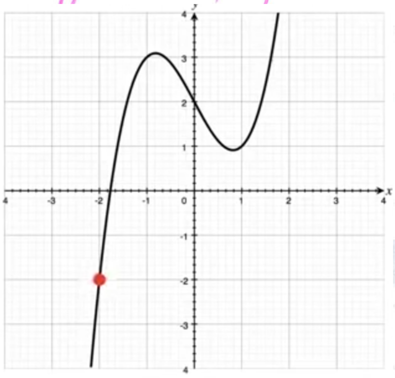
But in order to get there, in next module, we first need to do a bit more calculus

So let look at the equation of line.

$y = x^3 \cancel{+} -2x + 2$



If we differentiate this equation, we get quadratic equation: $3x^2-2$



$$\frac{dy}{dx} = 3x^2 - 2$$

And that that quadratic will have 2 solutions

∴ two turning points will exist

one maximum, one minimum

Say I do not know what the equation will look like?

and I do not have enough computer resources

to graph out the values at every point

.... and exist in so many dim, I cannot visualise it at all!

But say I only need to find the solution to equation where $y=0$

So where $x^3 - 2x + 2 = 0$? $\quad [x^3 - 2x + 2 = 0]$

we can see that there's actually only 1 solution here on this graph, but there could be more depending on the exact equation we want to solve.

Now, say that I have an idea that I want to hunt for a solution somewhere near some initial guess (red dot •)

For instance, the content is ~~very~~ pretty small and positive, so my guess that I may need a slightly negative value of x, say -2, maybe somewhere near a solution

Now, if I can evaluate the value of the function at my guess of $x = -2$, I find the function has a value of -2

and if I ask what the gradient is at that value $x = -2$, I find that the gradient is positive and is 10

Now, I can extrapolate the gradient, to the intercept with y axis, ~~which is~~ will be my first guess of the solution of equation, ... where it find intercept with y-axis.

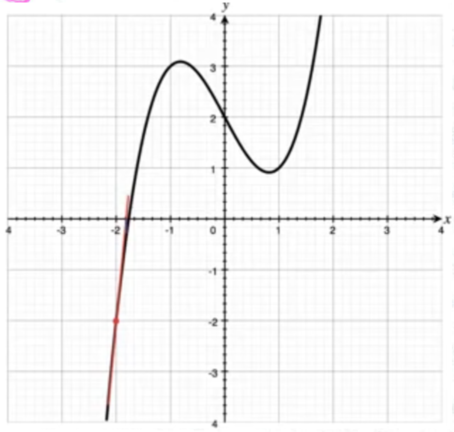
So, I can use that value of x at the intercept as a new estimate for what the solution to the equation is.

Effectively I am guessing that the function is a straight line
and I am using the gradient to extrapolate to find the solution.

It isn't really a straight line, of course, but 1st order approximation would be, that is a straight line, and we use that to update our guess, then go again and evaluate

So I can write down an equation for my guess: x_{i+1}, based on my previous guess, at term x_i by term as been my original guess, minus the value of function, divided by it's gradient

$$f(x_{i+1}) = f(x_i) - \frac{f(x_i)}{f'(x_i)}$$



So lets see how it plays out..

We can make a table starting with our initial guess $x=0$, then we can find the gradient ~~at~~ and intercept.

i	x_i	$y(x_i)$	$\frac{dy(x_i)}{dx}$
0	-2	-2	10
1	1.8	-0.23	7.7
2	-1.77	-0.005	7.4
3	-1.769	-2.3E-6	

and use that to generate a new guess
~~for this example~~, -
In this case $-2 - (-2)$ divided by 10, gives us 1.8
then we can evaluate the result for that guess,
and find that it's just a little less than 0, -0.23
and gradient is 7.7

So we've gone from been out by 2 on y, by
just been out by 0.23. So we got 10 times better
in our estimate, just in our first go.
If we carry on, then we get the next guess for x_2
~~and here~~ is -1.77 and that's just 0.005 away
from axis, which is really close
And if we have another go, after this
3 iterations we get an answer for x
-1.769, which is just 2.3×10^{-6} from axis
So in just 3 iterations we pretty much solve problem

This method is called Newton-Raphson method

$$\boxed{f(x_{i+1}) = f(x_i) - \frac{f(x_i)}{f'(x_i)}}$$

and it's really pretty neat.

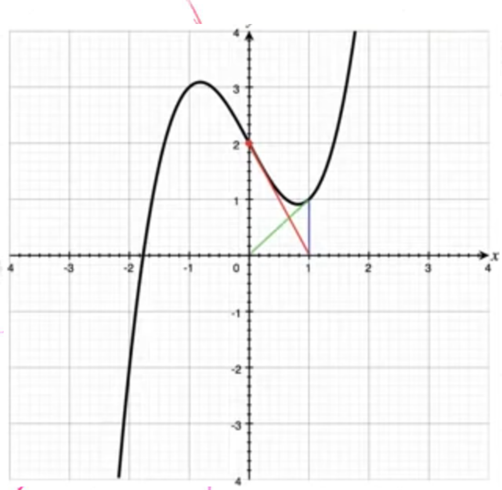
To solve an equation, all we need to be able to do, is evaluate it and differentiate it
we don't need to graph and visualize it everywhere, Calculate it ~~tott~~ lots and lots times
and don't need to solve it algebraically ~~ott~~ either

Which if we have lots of Dim to dataset and big multilevel function we want to fit to that data, it will be much too expensive to solve it analytically, and even plot it out for all possible values for variables

This sort of method, where we try to ~~solve~~ a solution and evaluate it, and try it with new guess, and evaluate that, and again, again, called iteration.

This is a very fundamental computational approach

There are some things that can sometimes go wrong with this method, so lets briefly look at ~~it~~ these



$$y = x^3 - 2x + 2$$

$$f(x_{i+1}) = f(x_i) - \frac{f(x_i)}{f'(x_i)}$$

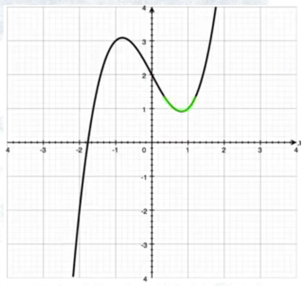
$$\frac{dy}{dx} = 3x^2 - 2$$

i	x_i	$y(x_i)$	$\frac{dy(x_i)}{dx}$
0	0	2	-2
1	1	1	1
2	0	2	

Say I started off with guess of $x = 0$, which evaluates to $y = 2$

Then find gradient for that, and extrapolate it.

it takes me away from solution to the other side of turning point, and gives me a new guess at $x=1$



When we evaluate that one, then I get value for y at $x=1$, of 1, when I find the gradient and extrapolate that, then new estimate lands me back at $x=0$

So I have a problem, so I have magically landed in a closed loop, where my estimates just cycle back between $x=0$ and $x=1$

and I never get close, and I never even go anywhere near the solution of $x=-1.769$.

Theres another problem, which that if I am close to a turning point here,  to a minimum and maxima, then because my gradient will be very small, when I divide by gradient in NR equation my Next estimate will take me zapping off to some ~~?~~ crazy value, there of it won't converge and dive off somewhere

So that the Newton Raphson method.,
we iterate to a solution to the equation, by
each making an a new estimate from the
solution, using the gradient to extrapolate
towards the solution, and going again, again, again etc

Most ~~of~~ of time, it works well, as a means
to step towards the solution

Recap:

→ look at method for using just the gradient,
to step our way towards solving a problem
— Called Newton Raphson method
— Really powerful way to solve an equation
— just by evaluating it and gradient a few times
— Its as if you standing on hill, in fog, you knew
your height and knew locally what's going on
around you, but you can't see the whole
landscape around you, i.e. ~~Cant~~ don't
know how to get down mountain

So what you do, you guess based on how steep the hill is, locally around you, which way to go, and you take step blindfolded, then when you get there, you ask again what height you at, and keep making more step down hill

Points, you don't need to know what the landscape look like, (the function) ⇒ just need to altimeter (altitude meter), the value of function, to feel with your toe, what the gradient is locally

Next session, look at how to apply this, where we have variable's, not just x, and that will involve finding the gradient vector

i.e. how to go down a hill on a contour plot.