Maths Skills (MTCS) G5071

Lecture 3

Kingsley Sage
Room 5C16, Pevensey III
khs20@sussex.ac.uk

Lecture 3

- Introduction to differential calculus …

(L) Isaac Newton, Copyright Royal Society, and (R) Gottfried Leibniz

A mathematical function

- Roughly speaking, a function can be thought of as taking as "input" some value(s) and producing as "output" another value
- How does that compare to a computer program function?
- The general notation is $y=f(x)$ where x is the name of the input variable, or argument, f is the name of the function, and y is the name of the output variable
- Usually, x is called the independent variable and y is called the dependent variable
- $y=x+2$ can also be written $f(x)=x+2$, you will commonly see both forms.

Examples of simple functions

- Many functions take a real number (written as a decimal value like 1.95) as input:

  $y = sin(x)$
  $y = log(x)$
  $y = exp(x)$ or $y=e^x$
  $y = 3x + 2$ this is a linear function
  $y = 3x^2 + x – 5$ this is a polynomial function

- NB: You need order of composition I.e. brackets, then multiplication and division, and then addition and subtraction. Remember BODMAS ..

Simple functions

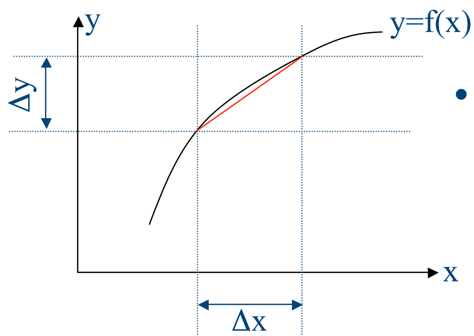
- For example, the last function $y=3x^2+x-5$ could have been written $f(x)=3x^2+x-5$.
- Even "if x is greater than 1 then $y=0$, else $y=x$" is a function. How might you write it more formally?
- As we shall see, MATLAB can *evaluate* a function (sometimes only approximately) and *visualise* it by drawing a graph.

Differentiation

- The basic idea of the differential calculus is that of a rate of change.
- For example, a function whose graph is a straight line like $y=3x+2$ means that any change in x produces a change 3 times as big in y.
- The slope of the line is 3 in this case for every value of x (it's a straight line).
- When we have a curve instead of a straight line (like many useful functions) the idea of the change in y produced by a small change in x turns out to be consistent and valuable …

Differentiation

- The change in *y* is divided by the change in *x*, as we use smaller and smaller changes, settles down to a steady value called the derivative of *y* with respect to *x*. This is usually written *dy/dx* or *f'(x)*



- As the change get infinitely small, it is characterised by the slope of the tangent to the curve at a point *x*. This slope is the numerical derivative.

Differentiation

- We could choose to do all differentiation numerically, using approximation techniques.
- But the differential calculus provides a set of convenient rules that allows us to compute the derivative value directly for many interesting and useful classes of functions,
- and for other specific functions there are well established rules.

Rules for differentiation

- Some rules for specific functions:
  - If $y=sin(x)$, then $dy/dx=cos(x)$
  - It is possible to work these out from first principles, but usually we look them up in a table or a textbook, or use a *symbolic* computing package (such as MATLAB)
- Rules for specific classes of functions:
  - if $y=x^n$, then $dy/dx = nx^{n-1}$
  - This applies to a class of functions – you just substitute the value of *n* for your application
  - So is $y=x^4$, then $dy/dx=4x^3$
  - So if $y=nx$, then $dy/dx = n$ (simple linear function)

Rules for differentiation

- Rules for products:
  - If a function can be written down as two functions multiplied together, and you can differentiate each of the two functions, then you differentiate the function itself using the rule:
  - If $y = f(x).g(x)$ then $dy/dx = f(x).dg(x)/dx + g(x).df(x)/dx$
  - $df(x)/dx$ means dy/dx *for* $y=f(x)$
  - For example, if $y=3x.sin(x)$, then $dy/dx=3x.cos(x) + 3sin(x)$
- The chain rule:
  - If a function can be written as one function applied to the result of another function, then the derivative of the whole thing uses the rule:
  - If $y=f(g(x))$, then $dy/dx = df(z)/dz.dg(x)/dx$ evaluated for $z=g(x)$
  - For example, if $y=sin(x^2)$ then $dy/dx=2x.cos(x^2)$ where $z=x^2$

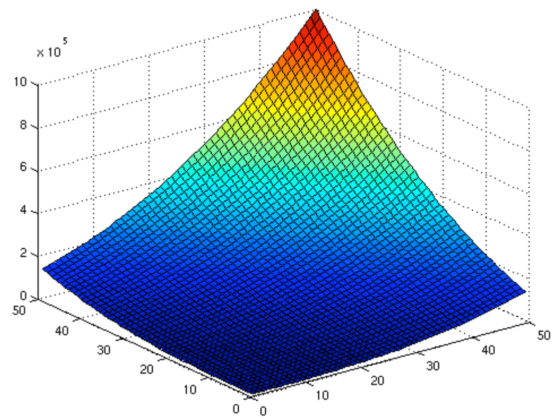
Rules for differentiation

- Arrrrrrrggggggghhh …
- Quoting the Hitchhikers' Guide to the Galaxy "Don't Panic".
- If you can't make sense of the rules, then the problem might well lie in the notation for functions, and in remembering what each symbol stands for. You need to write out every step until you can do this without thinking.

More complex functions

- For many applications, the idea of a function needs to be generalised to functions of more than one real variable (think of computer program functions here).
- A function of two variables might be written as $z=x+y$ or $f(x,y) = x+y$. You can think of x and y as inputs and z as the output.
- For $z=x+y$, it is useful to visualise the function as a surface or landscape, x and y represent position on a 2-D plane, and z represents the height above that plane (or below it if it is negative).
- MATLAB is very good at displaying these surfaces …

More complex functions



More complex functions

- For functions of more than two variables, there is no simple way to visualise the whole function, but such functions are often discussed.
- If a function has many arguments, you might see something like $y=f(x_1,x_2,\ldots,x_N)$ meaning that *f* is a function of *N* variables, which are distinguished by subscripts rather than having completely different names.
- This kind of thing is very common in neural network analysis. In MATLAB, these arguments can be conveniently represented as a vector.

Partial differentiation

- It is often necessary to know something about how the value of a function with several inputs is changed by small changes to one of its arguments – I.e. we need to differentiate it (somehow).
- Can you think of any examples?

Partial differentiation

- The basic idea is quite simple. Consider the function $z=xy$ and suppose that instead of being a variable, y simply stood for a fixed value – let's say 5.
- Then the function would be $z=5x$ and so it would follow that $dz/dx=5$ (the straight line equation again).
- If we didn't know the particular value of y, but we did know that it was fixed, we could still write $dz/dx=y$, with the understanding that y was being treated as a fixed quantity rather than a variable.
- This derivative, found by pretending that y is a fixed quantity, is called the *partial derivative* with respect to x.

Partial differentiation

- In order to distinguish this from an ordinary derivative, some special notation is used; a curly *d* instead of a normal d, looking like $\partial y/\partial x$.
- The partial derivative tells us how a function is affected by a small perturbation in one of its inputs – this can be very useful.
- What's the drawback with the partial derivative?

Summation

- Finally, you should be able to read the notation for forming sums – I.e. adding a set of things together. This uses the capital Greek letter Σ:

$$\sum_{k=1}^{k=5} kx = x + 2x + 3x + 4x + 5x$$

- In general, there is some variable (in this case *k*) that takes a set of values (in this case 1,2,3,4, and 5)

Summation

- For each of these values, an expression involving the variable is evaluated, and the results added together.
$\sum_{j=0}^{2} x_j^2 = x_0^2 + x_1^2 + x_2^2$
- In the form in which it is being used here, the variable takes integer values, starting from the one specified below the Σ, and going up to the value specified above.
- It is extremely common for the summation variable to form a subscript in the expression:

$$\sum_{j=0}^{2} x_j^2 = x_0^2 + x_1^2 + x_2^2$$

Next time …

- Introduction to applications
  - Artificial Neural Networks ANNs
  - Learning