Implicit Differentiation

[section 2.7 of book]

Compare the two equations that describe familiar curves

\[ y = x^2 + 3 \quad \text{(parabola)} \]
\[ y^2 + x^2 = 3 \quad \text{(circle)} \]

The first defines \( y \) as an *explicit* function of \( x \), because every value of \( x \) gives rise to a single value of \( y \). The second however defines \( y \) as an *implicit* function of \( x \) because each value of \( x \) does not give a unique value for \( y \).

How do we differentiate such implicit functions? For some we can re-arrange to find an implicit function, but this is not always possible and it is easier to use the chain rule.

For example, for the circle above

\[ \frac{d}{dx}(y^2 + x^2) = \frac{d}{dx} 3 \]

and so

\[ \frac{d}{dx} y^2 + 2x = 0 \]

to find \( d/dx(y^2) \) use the chain rule, so that

\[ \frac{d}{dx} y^2 = 2y \frac{dy}{dx} \]

putting this back into the original equation we obtain

\[ 2y \frac{dy}{dx} + 2x = 0 \]
and we can rearrange to give
\[
\frac{dy}{dx} = -\frac{x}{y}
\]
which gives us an expression for \(dy/dx\) as we require.

**Infinitessimals/differentials and linear approximations**

[section 2.9 of book]
Sometimes we need a quick and simple estimate of the change in a function \(f(x)\) that results when we change \(x\) slightly.

If we change \(x\) by \(x \rightarrow x + \Delta x\) then the resulting change in \(y = f(x)\) is just \(\Delta y = f(x + \Delta x) - f(x)\). As shown in the figure.

![Diagram showing linear approximation](image)

However if the change is very small then we can use the tangent line as an approximation to the curve and instead find \(dy\) as an approximation to \(\Delta y\). Therefore, if the slope of the tangent line is \(m = f'(x)\), and using the equation \((y = mx + c)\) for a straight line, we can write
\[
\Delta y \approx f'(x) \Delta x
\]
and in terms of infinitessimals,
\[
dy = f'(x)dx
\]
Optimization: finding maxima, minima and critical points

We’ve introduced the idea of the differential of a curve at a point $x_0$ as being the slope of the tangent line at that point. We now use this idea to examine properties of curves.

We note first that if a function, $f(x)$ has a maxima or a minima at some point, then the slope of the curve is zero AT THAT POINT because the tangent is flat. We call such a point an extrema of the function. However if there is an asymptote, the slope of the curve is infinite because it’s pointing vertically up (or down) AT THAT POINT. So, to examine the properties of the curve over some interval $I$, we compute it’s differential, $f'(x)$ and determine whether or not the differential vanishes or goes to infinity at any point in the interval.

However once we have found a point where the function has a tangent with a slope of zero, how do we tell if it’s a maximum or a minimum? Two possible ways are:

- **The first derivative test**
  Compute the slope of the function on either side of the extrema. If the function is increasing on the left side ($dy/dx > 0$), and decreasing on the right ($dy/dx < 0$) then we have a maxima, if it is decreasing on the left and increasing on the right then we have a minima.

- **The second derivative test**
  If the second derivative of the function $f''(x) < 0$, then we have a local maxima. If $f''(x) > 0$ then we have a local minima.

Finding roots of equations numerically

Often we need to find the roots, or zeros, of a function, by which I mean that we are trying to find where the function crosses the $x$-axis.

For some functions, such as quadratic and cubics, there exist explicit formulae for computing their roots, but for most functions there is no general way of finding them. Instead we turn to approximate, or numerical, techniques.

We will talk about two similar techniques, one that always converges but can be very slow, and one that is fast but doesn’t always converge.
The bisection method

Consider trying to find the roots of

\[ y = f(x) \]

The point of this technique is to find two points \( a \) and \( b \), one that is above and one that is below the root\(^1\), and then progressively move each point, \( a \) and \( b \), so that we narrow down the interval between them while keeping them on either side of the root.

We choose our first values for \( a \) and \( b \), call them \( a_1 \) and \( b_1 \) by guessing at some numbers and then testing to see whether \( f(a_1) \) has the opposite sign to \( f(b_1) \) which means that they are on either side of the root.

Note that if, say, \( f(a_1) \) is positive, and \( f(b_1) \) is negative then their product is negative too. And so to test whether they both lie on either side of the root all we need to do is calculate \( f(a_1) \times f(b_1) \) and show that it is negative.

Once we have found a suitable initial choice for \( a_1 \) and \( b_1 \) we can start to zoom in on the root.

We first find the midpoint between \( a_1 \) and \( b_1 \)

\[ m_1 = \frac{a_1 + b_1}{2} \]

and we test whether \( f(a_1) \times f(m_1) \) or whether \( f(m_1) \times f(b_1) \) is negative, to determine whether the root lies between \( a_1 \) and \( m_1 \) or between \( m_1 \) and \( b_1 \).

---

\(^1\)Use the mean value theorem to show that the root lies between them
If \( f(a_1) \times f(m_1) < 0 \) then the root lies between \( a_1 \) and \( m_1 \) and so we make a new interval by setting \( a_2 = a_1 \) and \( b_2 = m_1 \) and find a new midpoint

\[
m_2 = \frac{a_2 + b_2}{2}
\]

Similarly if \( f(m_1) \times f(b_1) \) is negative then we set \( a_2 = m_1 \) and \( b_2 = b_1 \) and again find the new midpoint.

We then test \( m_2 \) to find which of \( f(a_2) \times f(m_2) \) or \( f(m_2) \times f(b_2) \) is negative. We then continue by setting \( a_3 \) and \( b_3 \) equal to the resulting boundaries, and so on until we have reached a desired degree of accuracy.

So, the next question is how do we decide how accurate our answer is at each iteration? This is actually pretty straightforward: at the \( n^{th} \) iteration we’ve shown that the root lies between \( a_n \) and \( b_n \). This means the midpoint \( m_n \) is within half of the distance between \( a_n \) and \( b_n \) of the true root. Therefore, the quantity

\[
h_n = \frac{b_n - a_n}{2}
\]

tells us how close \( m_n \) is to the correct answer.

This \emph{bisection} method always converges to the correct answer, although it can be extremely slow.

**Newton’s method**

Newton’s method rests upon the idea that when you are close to a curve, the tangent line is a good approximation to the curve.

This time we start with a single guess at the root, call it \( a_1 \).

We find the next approximation, \( a_2 \) to the root by finding the intersection with the \( x \) axis of the tangent line to \( f(x) \) at the point \( a_1 \), as shown in the figure.

Therefore if the slope of the tangent line is \( f'(x) \) then the new guess is

\[
a_2 = a_1 - \frac{f(a_1)}{f'(a_1)}
\]

We can then continue the iteration by finding \( a_3 \) from the tangent line at \( a_2 \) until we have reached the desired degree of accuracy.

Newton’s method does not always converge to the correct answer, but it can be significantly faster than the bisection method.
Differential Equations

We frequently encounter problems of the form

\[ \frac{dy}{dx} = f(x) \]

where we have a differentiated quantity on the LHS and some function of \( x \) on the right. Such problems are called differential equations. The problem then becomes: “how do we invert the differentiation to find \( y \) as a function of \( x \)?”

Simple, first order, ordinary differential equations (ODE’s) can be solved by integration, eg:

\[ \frac{ds}{dt} = 10 \]

is an ODE for the speed of a vehicle travelling at 10 mph, where \( s \) is the distance.

Integrate to find the total distance gone from time \( t = 0 \) to \( t = 1 \) hour

\[
\int_0^1 ds = \int_0^1 10 \, dt = 10t \bigg|_0^1 = 10 \text{ miles}
\]
Exponential growth and decay

ODE’s that have the form
\[
\frac{dN}{dt} = kN
\]
describe exponential growth (if \( k > 0 \)) or decay (\( k < 0 \)).

Again solve by simple integration
\[
\frac{1}{N} \, dN = k \, dt
\]
\[
\int \frac{1}{N} \, dN = \int k \, dt
\]
\[
\ln(N) = kt + C
\]
\[
N(t) = \exp(kt + C) = \exp(kt) \exp(C) = K \exp(kt)
\]
where, since \( C \) is a constant, \( K = \exp(C) \) is also a constant.

What actually is \( K \)?

Think about setting \( t = 0 \), ie when the system first starts off. Then we have:
\[
N(t = 0) \equiv N(0) = K \exp(k \times 0) = K
\]
and so \( K = N(0) \) is the initial value for the problem. [In fact we can use the initial value of any ODE to find the constants of integration.]

Separable differential equations

Any ODE that can be manipulated into the form where all \( x \)’s and \( dx \)'s are one one side of the equation, while all \( y \)’s and \( dy \)'s are on the other is called a separable differential equation, and can be integrated directly.

Examples
\[
\frac{dy}{dx} = x^2(y + 1) \quad \Rightarrow \quad \frac{1}{y + 1} \, dy = x^2 \, dx
\]
\[
\frac{dy}{dx} = (y + 1) \quad \Rightarrow \quad \frac{1}{y + 1} \, dy = dx
\]
\[
\frac{dy}{dx} = y^2 \sin(x) \quad \Rightarrow \quad \frac{1}{y^2} \, dy = \sin(x) \, dx
\]
and in the same way as for exponential growth, use the initial values (if given) to find numerical values for the constant of integration.
Integration

For the midterm I’m going to assume that you understand that integration is the ‘inverse’ of differentiation, and I’m more concerned with your being able to practically solve integrals.

With this in mind, some common errors... for which you would be penalized for... are:

Forget infinitessimals

If one side of an equation is multiplied by an infinitessimal, then the other side must also be multiplied by an infinitessimal too, for example:

\[ dy = 7 \sin(x) \, dx \]

and so the integral becomes

\[ \int dy = \int 7 \sin(x) \, dx \]

and so

\[ y = -7 \cos(x) + C \]

Constants of integration

When indefinite integrals (i.e. those that do not have limits) are evaluated, the solution must include a constant of integration:

\[ \int f(x) \, dx = F(x) + C \]

but definite integrals (i.e. those that have limits) do not have a constant of integration

\[ \int_a^b f(x) \, dx = F(b) - F(a) \]

Integrating \( f(3x), \ f(4x), \text{ etc} \)

This is a special case of integration by substitution. To integrate (e.g.) \( \sin(5x) \) make the substitution \( u = 5x \), so:

\[ du = 5 \, dx \quad \text{or} \quad dx = \frac{1}{5} du \]
and therefore

\[
\int \sin(5x) \, dx = \int \sin(u) \frac{1}{5} \, du = \\
= -\frac{1}{5} \cos(u) \\
= -\frac{1}{5} \cos(5x)
\]

Integration by substitution

Some integrals may be transformed to a simpler integral by substitution. When transforming an integral

- **Step 1** transform the variable, e.g. \( u = \sin(x) \)
- **Step 2** transform the infinitessimal, e.g. \( du = \cos(x) \, dx \)
- **Step 3** (for definite integrals) transform the limits, e.g. \( u_1 = \sin(x_1), \, u_2 = \sin(x_2) \)