MATH7019 — Applied Technological Maths 311

J.P. McCarthy

September 3, 2020

Contents

0.1 Introduction

Lecturer

J.P. McCarthy

Office

Meetings before class by appointment via email only.

Email & Web:

jpmccarthymaths@gmail.com and *http://jpmccarthymaths.com*

This page will comprise the webpage for this module and as such shall be the venue for course announcements including definitive dates for the tests. This page shall also house such resources as links (such as to exam papers), as well supplementary material. Please note that not all items here are relevant to MATH7019; only those in the category 'MATH7019'. Feel free to use the comment function therein as a point of contact.

Module Objective

This module covers: curve fitting; static beam differential equations; Taylor series; normal distributions, statistical inference.

Module Content

Curve Fitting & Mathematical Modelling

Applications of curve fitting. Lagrange Interpolation. Determination of curves of best-fit in the least squares sense. Non-linear laws. Correlation. Use of models.

Static Beam Differential Equations

Step functions. Static beam differential equations including simply supported, fixed ends, and cantilevers.

Taylor Series

Review of single variable calculus. Taylor Series of one variable. Euler Method and Three Term Taylor. Review of functions of several variables and partial differentiation. Differentials with applications to analysis of rounding error.

Probability & Statistics

Random variables. Laws of probability. Normal distribution. Applications to engineering problems. Sampling, and Hypothesis Testing, for population means and population proportions.

Assessment

Either four 25% assessments or two 30% and two 20% or two 30% and one 40%. Unless in truly exceptional circumstances, late assignments will be assigned a mark of ZERO. You may hand in partially completed reports or email reports.

Exercises

There are many ways to learn maths. Two methods which aren't going to work are

1. reading your notes and hoping it will all sink in
2. learning off a few key examples, solutions, etc.

By far and away the best way to learn maths is by doing exercises, and there are two main reasons for this. The best way to learn a mathematical fact/ theorem/ etc. is by using it in an exercise. Also the doing of maths is a skill as much as anything and requires practise.
There are exercises in the notes for your consumption. The webpage may contain a link to a set of additional exercises. Past exam papers are fair game. Also during lectures there will be some things that will be *left as an exercise*. How much time you can or should devote to doing exercises is a matter of personal taste but be certain that effort is rewarded in maths.

Reading

Your primary study material shall be the material presented in the lectures; i.e. the lecture notes. Exercises done in tutorials may comprise further worked examples. While the lectures will present everything you need to know about MATH7019, they will not detail all there is to know. Further references are to be found in the library. Good references include:

- K A Stroud, Dexter J Booth 2013, Engineering Mathematics, 7th Ed., Palgrave Macmillan Hampshire [ISBN: 978113703120]
- John Bird 2017, Higher Engineering Mathematics, 7th Ed., Routledge/Taylor & Francis London, New York [ISBN: 9781138673571]
- Douglas C. Montgomery, George C. Runger 2014, Applied Statistics and Probability for Engineers, 6th Ed., John Wiley & Sons Inc New York [ISBN: 9781118744123]

The webpage may contain supplementary material, and contains links and pieces about topics that are at or beyond the scope of the course. Finally the internet provides yet another resource. Even Wikipedia isn't too bad for this area of mathematics! You are encouraged to exploit these resources; they will also be useful for further maths modules.

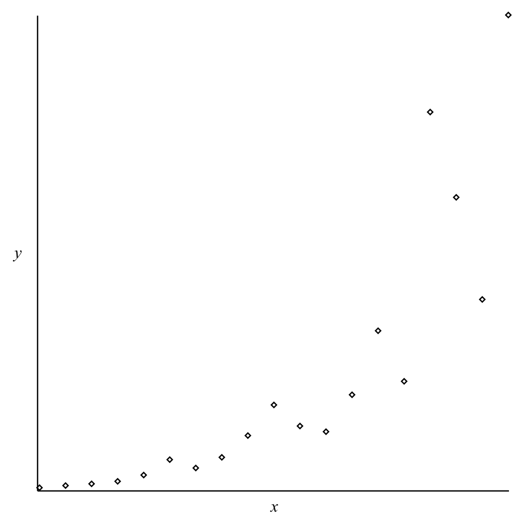
0.2 Learning Outcomes

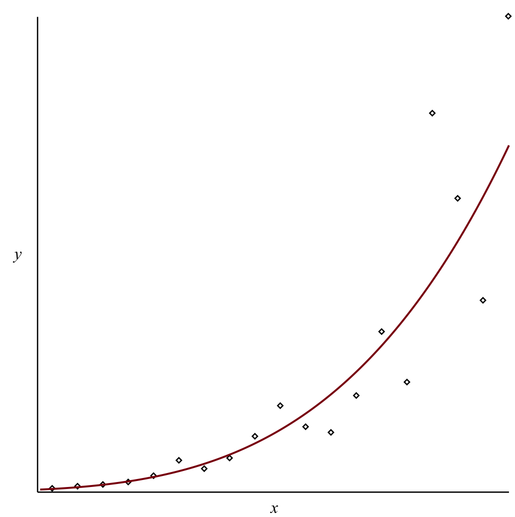
This module has four/five learning outcomes and a student who has achieved these learning outcomes should be equipped to complete the following representative tasks. In this module these tasks will all be put in an engineering context.

Find and use the Curve of Best Fit to a data set

Given some (x, y) data:
Find the curve in a given family (say $y = ax^b$) that is the curve of best fit[1]:

[1]this will be what is known as 'in the least squares sense'; or rather with this particular family of curves, in the 'log-linear least squares sense'





Understand the Concept of a Differential Equation

At the very least given a differential equation (boundary value problem), e.g.

$$\frac{d^2M}{dx^2} = -18\delta(x-1); \qquad M(0) = M(6) = 0,$$

it should be understood that it's solution $y = M(x)$ is a function which can be plotted:

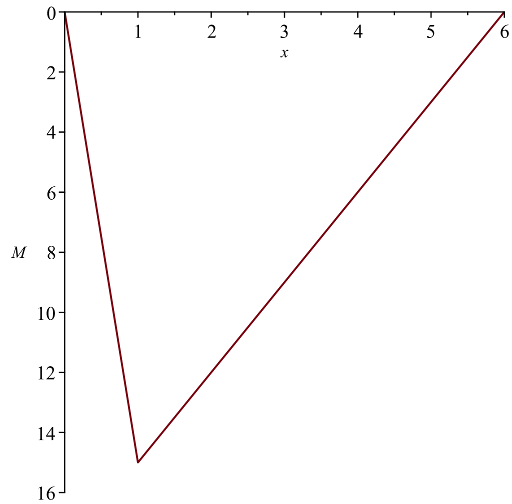
Solve Static Beam Equations

Given the load per unit length at a distance x from the left of the beam:

$$w(x) = f(x),$$

and the type of beam (simply supported, fixed ends, or cantilevered), solve the differential equations

$$\frac{d^2M}{dx^2} = -w(x)$$
$$EI\frac{d^4y}{dx^4} = -w(x)$$



for the bending moment $M = M(x)$ as a function of x and the deflection $y = y(x, E, I)$ as a function of x, E, and I.

Use Taylor Series to approximate solutions of Differential Equations

When solving an ordinary differential equation, e.g.

$$\frac{dV}{dx} = -w(x),$$

and $w(x)$ is given by a non-integrable formula, or given by data, then the differential equation cannot be solved exactly for $V = V(x)$. Instead we can make do with approximations to V at discrete points, e.g. $x = 0, 0.1, 0.2, \ldots, 5.9, 6.0$.

Use the Normal Distribution to solve Engineering Problems

If you want to find the average strength of a component used in a structure you take a random sample and estimate from there. How confident can we be that this approximation is accurate? We answer this question.

Suppose someone claims that the component's quality has disimproved over the past year. How can we test this claim? We answer this question.

Chapter 1

Curve Fitting & Mathematical Modelling

I never guess. It is a capital mistake to theorise before one has data. Insensibly one begins to twist facts to suit theories, instead of theories to suit facts.

Sir Arthur Conan Doyle, author of Sherlock Holmes stories

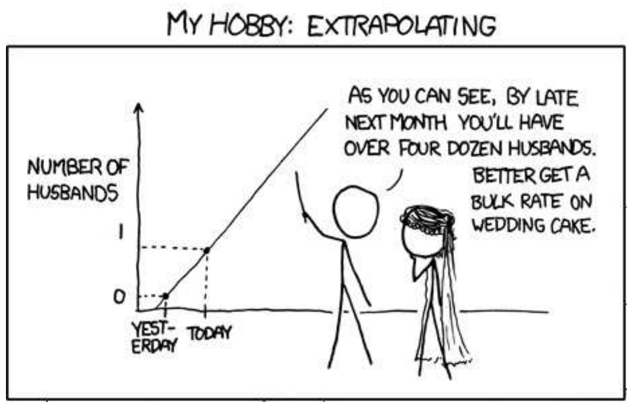


Figure 1.1: Careful now.

Calculus Review I: Functions

In engineering, the way to think about it is to imagine a function as a relationship between variables x and y such that for each value of x there is a corresponding value of y. We say that *y is a function of x* and we may even say that *y depends* on x, and we write $y = f(x)$ or $y = y(x)$. The number $f(x)$ is the *value of f at x*. For example, the area of a square, A, is a function of its sidelength s, and we write $A = A(s)$. We visually represent functions using their *graphs*. The graph of f is all of the points of the form $(x, f(x)) = (\text{input}, \text{output})$.
Some of the functions that we will encounter will be of the form:

What this means is that whatever number we feed into the function, x, the expression on the right gives the recipe for calculating $f(x)$. We will also see functions that depend on more than a single variable in a later chapter.

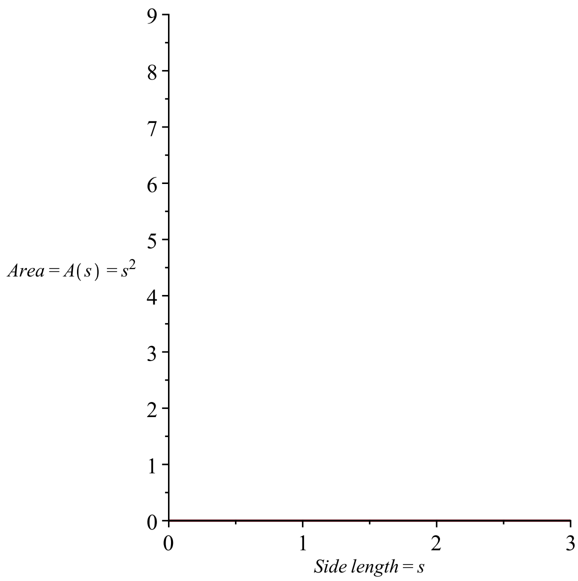


Figure 1.2: Usually you control the x-variable, the *independent* variable, while you perhaps measure and record the y variable, the *dependent* variable. The dependent variable is *a function of* the independent variable.

Examples that occur in MATH7019

1. The unit cost of training a new trainee, C, depends on the number of trainees, T; and we write $C = C(T)$.

2. The temperature of a component, θ, depends on the time it has been used, t; and we write $\theta = \theta(t)$.

3. The depth of a pile driving, D, depends on the number of blows, N; and we write $D = D(N)$.

4. * The yearly sales of a product, S, depends on the number of years the product has been on the market, t; and we write $S = S(t)$.

5. * For beams, we have that the load per unit length, w; the shearing force, V; the bending moment, M; the slope, y'; and the deflection, y; ALL depend on the distance from one end of the beam, x; and we write $w = w(x)$, $V = V(x)$, $M = M(x)$, $y' = y'(x)$ and $y = y(x)$.

Introduction

Very often data is collected and the formulaic/function relationship between the sets of data is not known. For example the temperature of an object at various times could be measured but there might not be any known formula or causal relationship between temperature and time.

Consider the following abstract example:

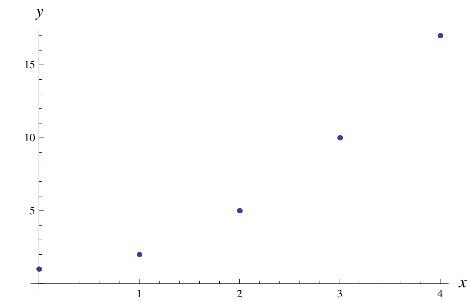


Figure 1.3: For the *independent variable* $x = 0, 1, 2, 3, 4$, the variable y was measured and recorded. What might the value of y at $x = 2.5$ (interpolation)? At $x = 5$ (extrapolation)?

Reasonable questions at this point are what would be a good estimate for the value of y at $x = 2.5$ and what about $x = 5$? An estimation of the value of y at $x = 2.5$ is known as an *interpolation* while an estimate of the value of y at $x = 5$ is known as an *extrapolation*. Could we suggest a functional/formulaic relationship between x and y? This is another use of curve fitting — we take measurements of a variable x and a variable $y = y(x)$ that depends on x. We plot y vs x and come up with a possible functional relationship between the variables $y = f(x)$ by looking at the graph... in other words we can try and figure out a formula for $y = y(x)$ from the data. Later we will learn how to test how well data fits a formula.

On the other hand we might know that x and y are related by a law of the form $y = ax^2 + b$ where a and b might be some parameters. In order to calculate a and b we can take measurements of x and y and fit a curve to the data and hence find out the values of a and b. This kind of thing can be done to measure the Young's Modulus of an alloy.

1.1 Lagrange Polynomial Interpolation

How many lines pass through a given point $P(x_1, y_1)$ on the plane? How about through P and some other point $Q(x_2, y_2)$?

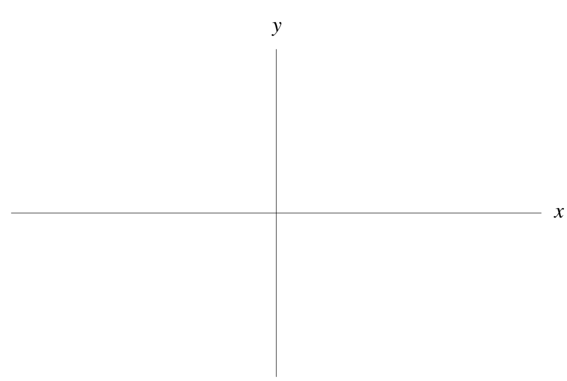


Figure 1.4: An infinite number of lines pass through P; only one of these, however, also passes through Q.

How about quadratics/parabolas (degree-two polynomials, $y = ax^2 + bx + c$)? Given points P, Q and $R(x_3, y_3)$ is there more than one quadratic through P and Q but only one through P, Q *and* R?

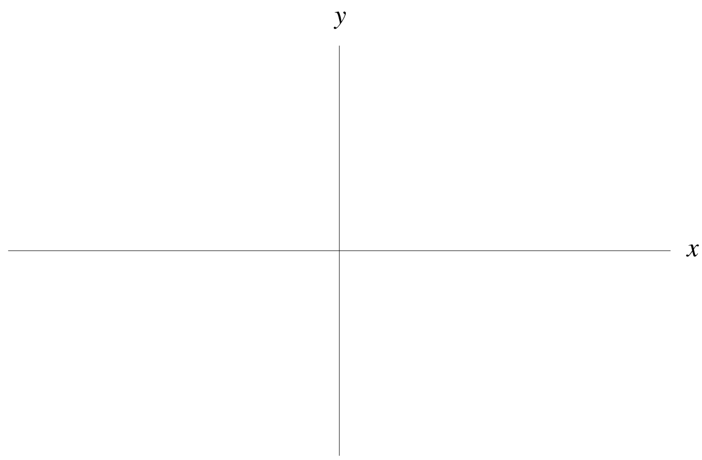


Figure 1.5: $2 + 1 = 3$ points defines a unique degree 2 polynomial. In other words given three points there is a unique quadratic/parabola $q(x)$ whose graph goes through all three points.

A bit of heavy algebra shows that in fact we have the following:

The Polynomial Unisolvence Theorem

Given a set of $d+1$ points on the plane *there is a unique polynomial $P = a_d x^d + \cdots$ of degree at most d that passes through all of these points.*

The rest of the analysis that we do can be generalised easily, but for MATH7019 we will just use it for fitting a quadratic, degree-2 polynomial to three points.

The Polynomial Unisolvence Theorem: MATH7019 Version

Given a set of 3 points on the plane *there is a unique quadratic/ parabola $q(x) = ax^2 + bx + c$ that passes through all of these points.*

Suppose that $(x_0, f(x_0))$, $(x_1, f(x_1))$, $(x_2, f(x_2))$ are three points in the plane. Suppose that you can find a quadratic $\ell(x) = ax^2 + bx + c$ such that

$$\ell(x_0) = f(x_0)\,,\ \ell(x_1) = f(x_1)\,,\ \ell(x_2) = f(x_2),$$

presented for/by us as:

x	x_0	x_1	x_2
y	$f(x_0)$	$f(x_1)$	$f(x_2)$

then $\ell(x)$ has to be the unique interpolating quadratic:

$$\ell(x) = \frac{(x-x_1)(x-x_2)}{(x_0-x_1)(x_0-x_2)}f(x_0) + \frac{(x-x_0)(x-x_2)}{(x_1-x_0)(x_1-x_2)}f(x_1) + \frac{(x-x_0)(x-x_1)}{(x_2-x_0)(x_2-x_1)}f(x_2) \qquad (1.1)$$

We might call this the *Lagrange Interpolating Polynomial* of degree 2. The more general Lagrange Interpolating Polynomial of degree $n-1$ for n points is given by:

$$\ell(x)=\sum_{i=1}^{n}\left(\prod_{\substack{j=1\\ i\neq j}}^{n}\frac{x-x_j}{x_i-x_j}f(x_j)\right). \tag{1.2}$$

Like $\sum$ means add up; $\prod$ means multiply.

Examples

1. **Autumn 2018** In order to predict the number of blows N required to pile drive to a depth of$D$ m in a given soil, the following recordings were made

depth, D/ m	0	1	2	3	4
blows, N	10	25	65	155	390

Using three suitable data points, use *Lagrange interpolation* to provide a rough estimate for the number of blows required to reach a depth of 2.5 m.

Solution: Perhaps use the points:

depth, D/ m	1	2	3
blows, N	25	65	155

Note that $N=f(D)$ — $N\sim y$ depends on $D\sim x$:

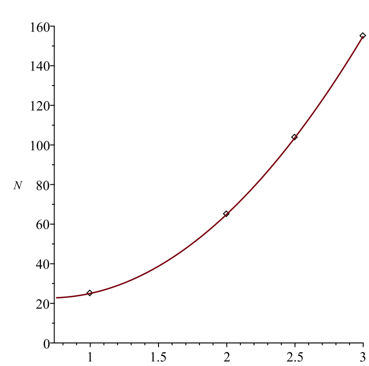


Figure 1.6: For Autumn 2018, we see that there is only one quadratic going through the three points. We use this quadratic to estimate/interpolate that $N\approx 104$ for $D=2.5$.

2. **Autumn 2017** The shear stresses of five specimens taken at various depths in a clay stratum are listed below.

depth, D, in metres	1.9	4.2	5.8	8.1	10.0
shear stress, τ, in kilopascals	14.4	19.2	33.5	71.8	76.6

Use Lagrange Interpolation to estimate the shear stress at a depth of $D = 6$.

Solution: Perhaps use the points:

depth, D, in metres	4.2	5.8	8.1
blows, shear stress, τ, in kilopascals	19.2	33.5	71.8

Note that $\tau = f(D)$ — $\tau \sim y$ depends on $D \sim x$:

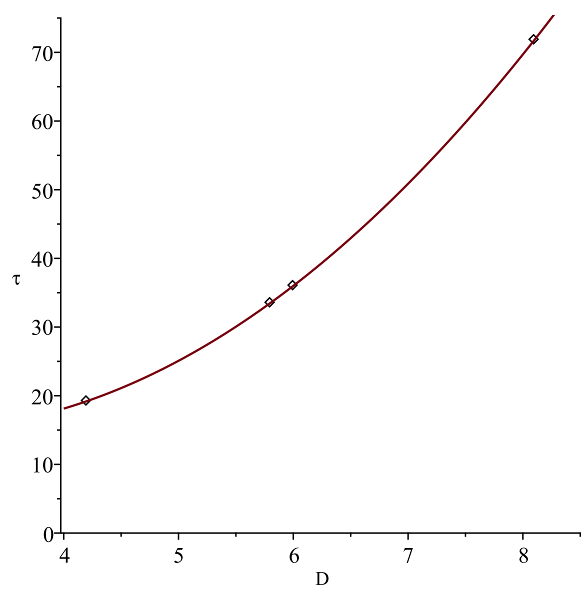


Figure 1.7: For Autumn 2017, we see that there is only one quadratic going through the three points. We use this quadratic to estimate/interpolate that $\tau \approx 36$ for $D = 6$.

Exercises

1. **Autumn 2015** The owner of a new manufacturing engineering firm with rapid growth is keeping track of the monthly sales, S, in hundreds, in the months after launch:

months after start-up, m	2	4	6	8	10	12
monthly sales, $S/100$	1	2	5	11	20	35

Use Lagrange Interpolation to estimate $S(9)$, the sales in month nine. **Ans:** ≈ 15.

2. **MATH7021 Summer 2014** In order to predict the number of blows N required to pile drive to a depth of D m in a given soil, the following recordings were made

depth, $D/$ m	0	1	2	3	4
blows, N	10	25	65	155	390

Using three suitable data points, use *Lagrange interpolation* to provide a rough estimate for the number of blows required to reach a depth of 2.5 m. **Ans:** $90 < \ell(2.5) < 105$.

3. **Autumn 2016** For a beam of fixed material, cross-section and total load, the span L and the maximum deflection δ were measured and recorded.

L/m	4.0	5.0	6.0	7.0	8.0
δ/mm	1.92	3.75	6.48	10.29	15.36

Having access to this data, an engineer wanted to interpolate a value for δ for a length $L = 6.5$ m beam. Use Lagrange Interpolation to estimate δ for $L = 6.5$ m. **Ans:** 8.25 mm.

---

4. By using Lagrange Interpolation estimate the value of y at $x = 1.5$ where

$$f(2) = 4\,,\ f(3) = 9\,,\ f(4) = 16.$$

Can you do this question in your head using the Polynomial Unisolvence Theorem? **Ans:** 2.25.

1.2 Least Squares Curve Fitting

Given a larger data set there are a number of situations where interpolating — fitting through — the data is not the correct thing to do. There is one quick but technical reason why not all data can be interpolated using a polynomial. If the data consists of $\{(1,1),(1,2),(2,1)\}$, with two data points with the same x-value

Calculus Review II: Lines, Differentiation and Max/Min Problems

What do we need to define a line:

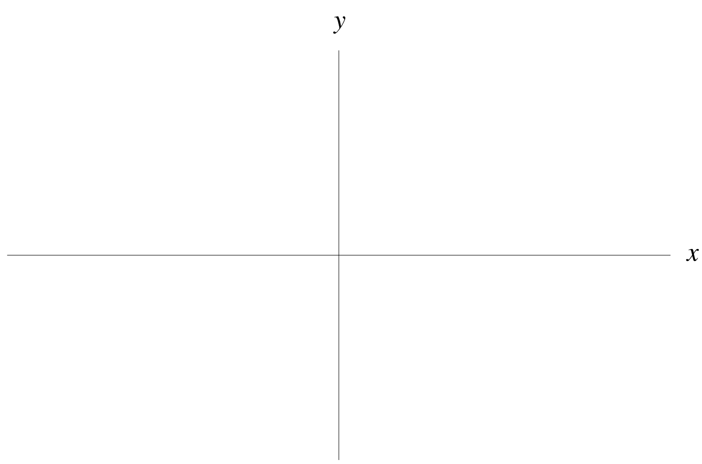


Figure 1.8: Two points defines a line. What does it take for a point (x, y) to be on this line?

A point (x, y) is on the line if and only if the slope from (x_1, y_1) to (x, y) is equal to the slope of the line. What is slope?

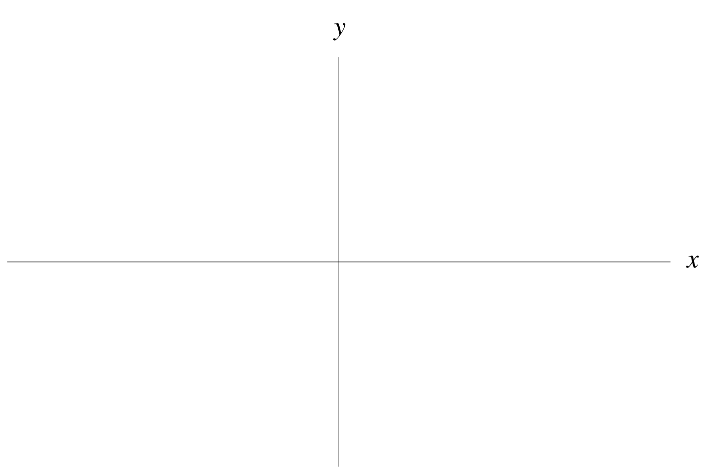


Figure 1.9: The slope of a line is the ratio of how much you go up, as you across: slope $= |\uparrow|/|\rightarrow|$

Consider again the line defined by the point $(2, 3)$ and slope 4:

$$y - 3 = 4(x - 2)$$

This shows that any equation of the form:

for m, c constants is another way of writing the equation of the line. m is the slope. What about c?

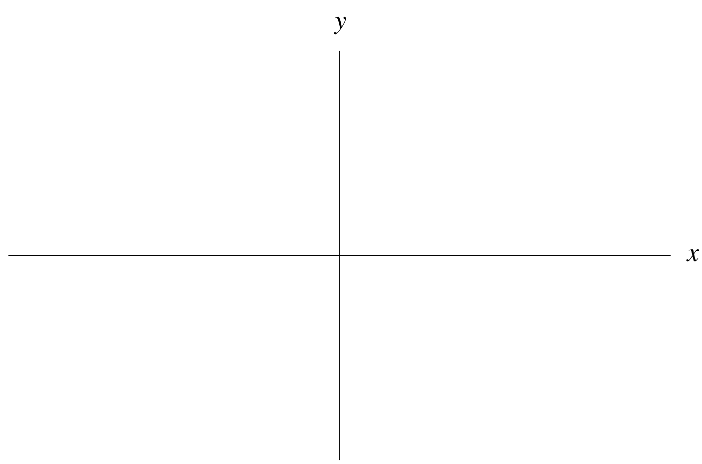


Figure 1.10: When the x-coordinate is zero, the graph of the line cuts the y-axis and so c is the y-cut.

So c is where the line cuts the y-axis. I would recommend that inasmuch as possible write your lines in this form.

The Derivative of a Function

In first year, ye developed the idea of the *derivative* of a function that would allow ye to study the following problems:

1. **Tangents** Most of these elementary functions are *smooth.* This means that *locally* (near a point), they are well-approximated by lines:

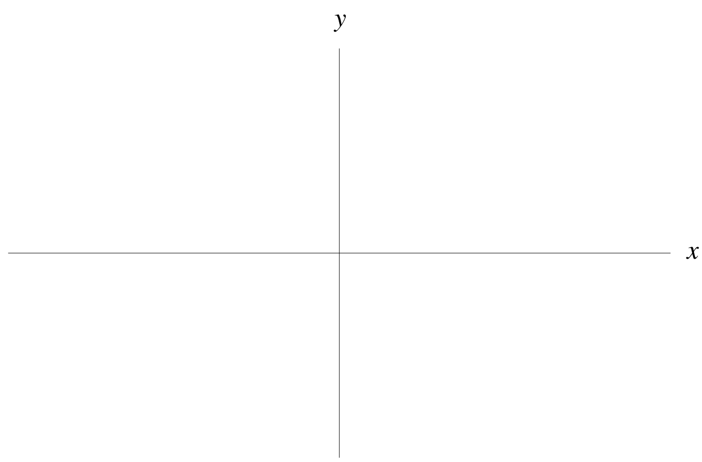


Figure 1.11: Near the point $x = 1$, the curve $y = \ln x$ is well-approximated by the line $y = x - 1$.

To find the equation of this *tangent* line, which has equation

$$y - y_1 = m(x - x_1)$$

it is necessary to find the slope of the tangent. This led us to the following:

$$\text{slope (of tangent to } y = f(x) \text{ at } x = a) \;:= f'(a)$$

where

$$f'(x) \equiv \frac{dy}{dx} = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}.$$

This allows us to approximate functions *locally*[1] by lines; i.e. using their tangents. Hence the derivative of a line is the slope:

$$\frac{d}{dx}(mx + c) = m.$$

[1] near a specific point

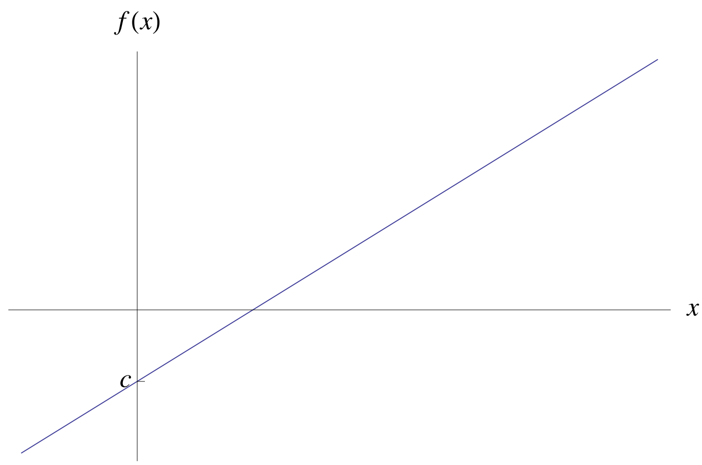


Figure 1.12: The graph of a line. Note that the slope morryah the derivative is constant. Note also that $f(0) = m(0) + c = c$.

2. **Local Maxima/Minima** Suppose we have a function of the form $y = f(x)$. Can we find its (local) maxima and minima? The derivative allows us to do so:

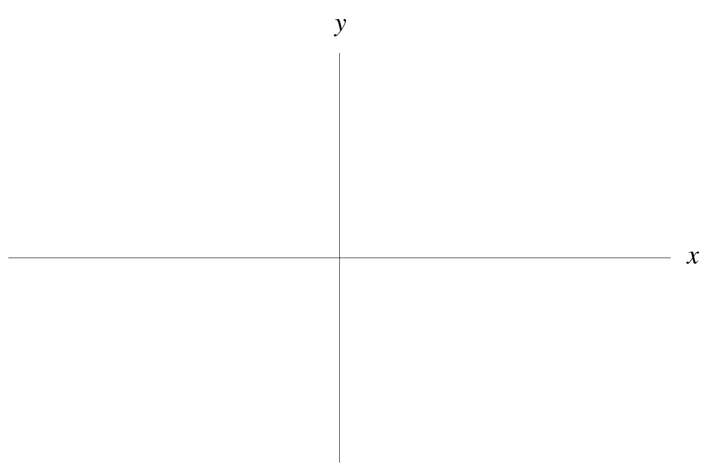


Figure 1.13: At a turning point, the slope of the tangent to the curve is zero: $f'(x) = 0$. We can use the second derivative to determine whether the turning point is a (local) maximum or minimum.

1.2.1 Theory

Suppose that you have a set of data $\{(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)\}$ and you plot your data:

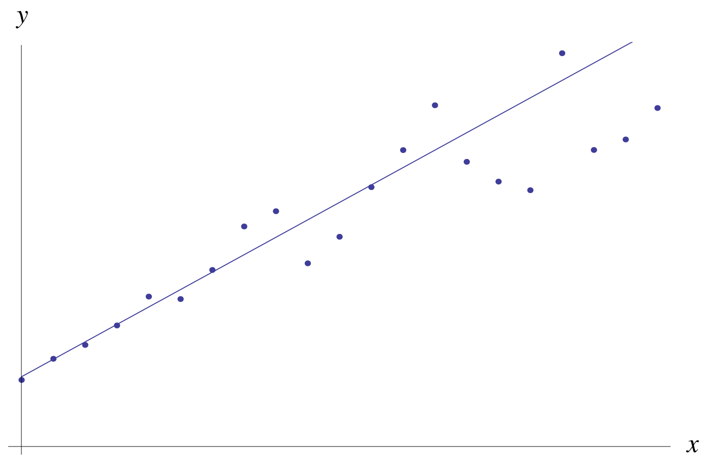


Figure 1.14: The data appears to have a linear relationship $Y = aX + b$ (cf. $y = mx + c$). Perhaps the measuring process was not precise or accurate. Is there a line of best fit for this data as shown? Is there a theoretical *true* behaviour as represented by the straight line?

Now what does *line of best fit* mean? One possible thing we could is choose the line such that the sum of the errors $\sum \delta_i$ is a minimum. What are these errors given by? That is we want to find the values of a and b that minimise the function

$$s(a, b) = \sum_{i=1}^{n} |Y_i - (aX_i + b)| = \sum \delta_i.$$

This function does not yield easily to analysis and it does not give any weighting to particularly large errors. A way to resolve both of these issues is to instead look at the sum of the *squares* of the errors, $\sum \delta_i^2$. It turns out that this function is much easier to handle and also magnifies the effect of large errors:

$$S(a, b) = \sum_{i=1}^{n} (Y_i - aX_i - b)^2 = \sum \delta_i^2.$$

Now how do we find the maxima and minima of functions of several variables? Calculus of a single variable generalises neatly to that of functions of several variables and was explored briefly in MATH6040 ($z = f(x, y)$).

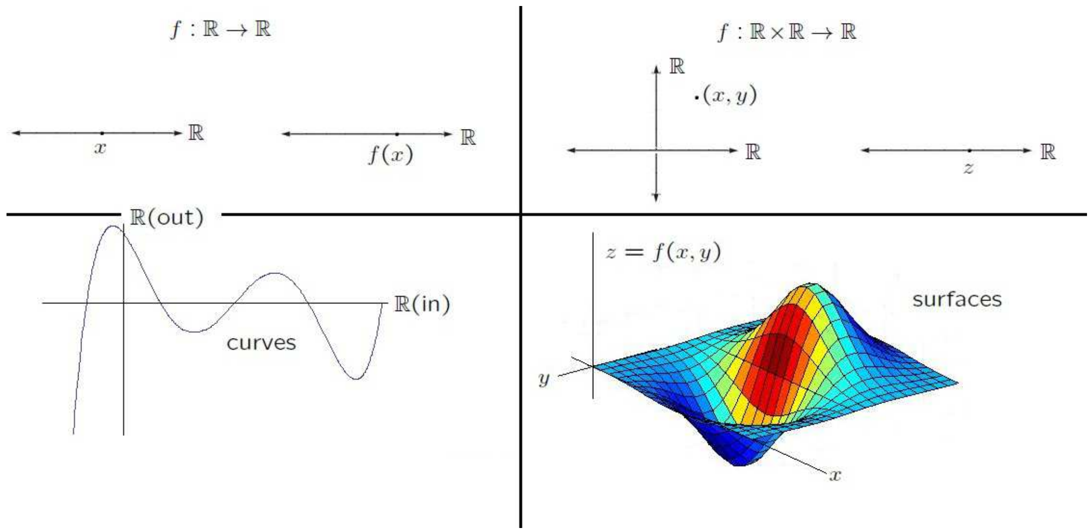


Figure 1.15: Functions of a single variable are represented geometrically by curves while functions of two variables are represented geometrically by surfaces. Here we have a function $S = S(a, b)$ that we want to minimise; i.e. we want to find a slope, a, and y-intercept, b, such that the sum of the squared deviations in a minimum.

For example, consider the data

$$(X, Y) = \{(0, 1), (1, 3), (2, 5), (3, 8), (4, 11)\}.$$

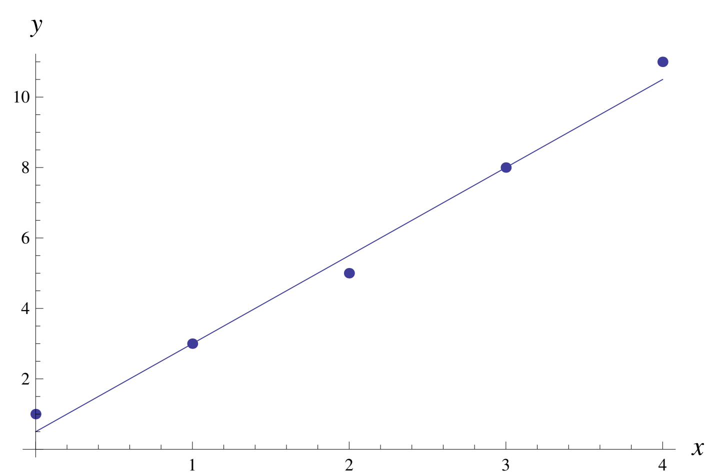


Figure 1.16: This data can be plotted and we can try and fit a line of best fit. Note that at e.g. $x = 2$, the error is given by $\delta_3 = 2a + b - 5$.

What we can do to try and find this line of best fit is to write down for this data the function $S(a, b)$:

$$\begin{aligned}
S(a,b) &= \sum_{i=1}^{n}(Y_i - aX_i - b)^2 \\
&= (1 - a(0) - b)^2 + (3 - a - b)^2 + (5 - 2a - b)^2 + (8 - 3a - b)^2 + (11 - 4a - b)^2 \\
&= 220 - 162a + 30a^2 - 56b + 20ab + 5b^2.
\end{aligned}$$

Now this function can be plotted. At each coordinate (a, b), we have a height of $S(a, b)$:

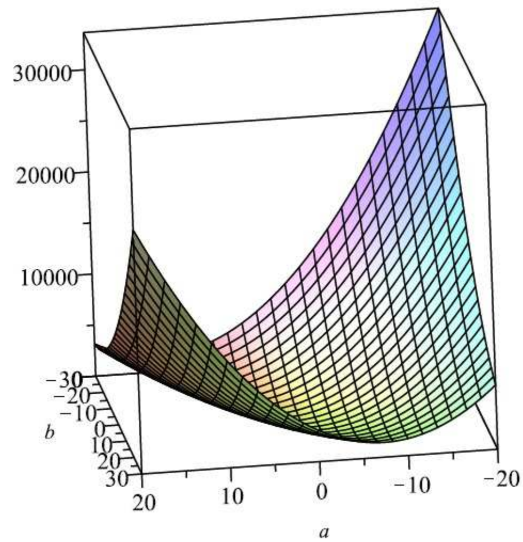


Figure 1.17: This function $S(a, b)$ has a point where $S(a, b)$ — the sum of the squared deviations — is minimised. How do we find the minimum?

Recall that $\dfrac{\partial S}{\partial a}$, the *partial derivative of S with respect to a*, is the derivative of $S(a, b)$ with respect to a, **keeping** b constant. The partial derivative of $S(a, b)$ with respect to b is defined similarly.

One picture suggests to us that we can locate local extrema of surfaces by finding when we have extrema with respect to a and b coinciding:

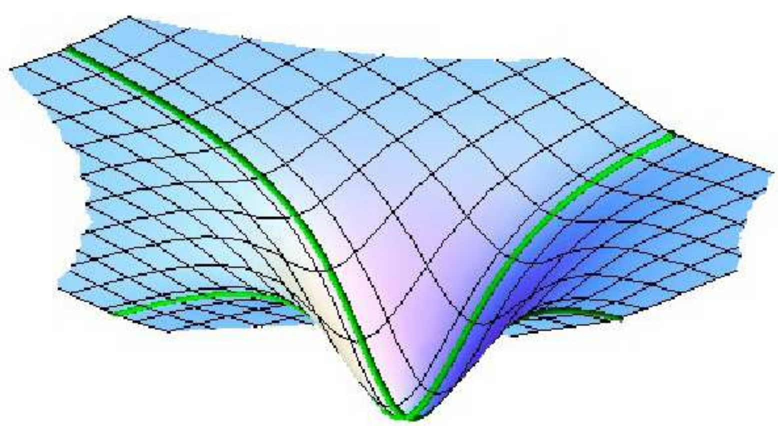


Figure 1.18: Here we can see that at a local minimum of a surface both the cross-sections describe curves that are also 'experiencing' local minima.

In other words, we have that minima occur when

$$\frac{\partial S}{\partial a} = \frac{\partial S}{\partial b} = 0, \text{ and}$$

$$\frac{\partial^2 S}{\partial a^2} > 0, \quad \text{and} \quad \frac{\partial^2 S}{\partial b^2} > 0.$$

Compare this with first year where we learnt that we have a min when $\dfrac{dy}{dx} = 0$ and $\dfrac{d^2y}{dx^2} > 0$:

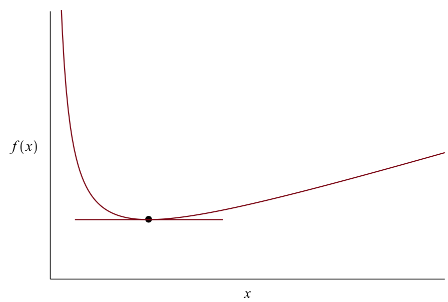


Figure 1.19: The minimum of a function of a single variable $y = f(x)$.

For the data above it turns out that for $a = 5/2$ and $b = 3/5$ both $\dfrac{\partial S}{\partial a}$ and $\dfrac{\partial S}{\partial b}$ are equal to zero and the sum of the squared deviations is a minimum:

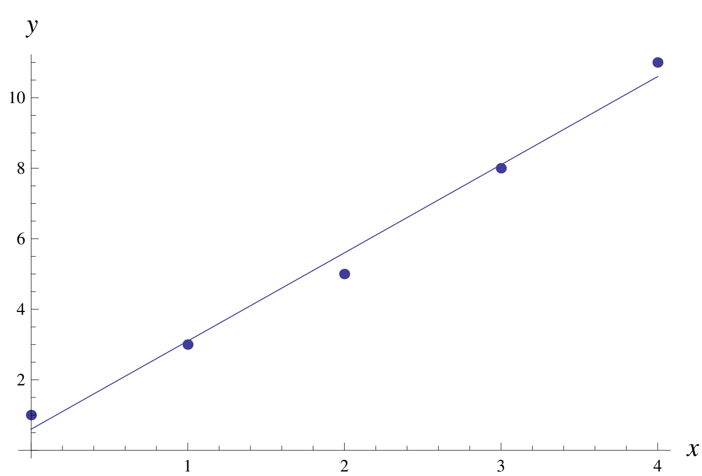


Figure 1.20: The curve $Y = \frac{5}{2}x + \frac{3}{5}$ is the unique line that minimises $\sum_i \delta_i^2$.

Now we won't explicitly write down $S(a, b)$ but instead take advantage of the following. So consider some data $\{(X_i, Y_i)\}$. We want to find the values of a and b such that partial derivatives of $S(a, b) = \sum(Y_i - aX_i - b)^2$ vanish *simultaneously*. First differentiate partially with respect to b...

Well by the Sum Rule for differentiation we can differentiate term-by-term. For each term we can use the Chain Rule. Recall that with respect to b, a is a constant. All of the X_i and Y_i are constants:

$$\frac{\partial S}{\partial b} = \sum 2(Y_i - aX_i - b)(-1)$$

We are interested in when this is equal to zero. Firstly we can split the sum:

$$\begin{aligned} \sum 2(Y_i - aX_i - b)(-1) &\stackrel{!}{=} 0 \\ \Rightarrow -2\sum Y_i + 2\sum(aX_i + b) &= 0 \\ \Rightarrow 2\sum Y_i &= 2a\sum X_i + 2\sum b \\ \Rightarrow \sum Y_i &= a\sum X_i + \sum b. \end{aligned}$$

Now differentiate $S(a,b) = \sum(Y_i - aX_i - b)^2$ with respect to a and solve $\frac{\partial S}{\partial a}$ equal to zero:

$$\begin{aligned} \frac{\partial S}{\partial a} = \sum 2(Y_i - aX_i - b)(-X_i) &\stackrel{!}{=} 0 \\ \Rightarrow -\sum 2X_iY_i + 2\sum(aX_i^2 + bX_i) &= 0 \\ \Rightarrow 2\sum X_iY_i &= 2a\sum X_i^2 + 2b\sum X_i \\ \Rightarrow \sum X_iY_i &= a\sum X_i^2 + b\sum X_i. \end{aligned}$$

Now we have *simultaneous equations* in a and b:

$$\begin{aligned} \sum Y &= \sum aX + \sum b \\ \sum XY &= \sum aX^2 + \sum bX \end{aligned}$$

Note that all of the X_i and Y_i are actual numbers: we are solving these equations for a and b. We call these the *normal equations*. How is the second equation related to the first?
Incidentally we haven't shown that this process minimises $S(a,b)$ but we can ensure that it *always* does by showing that

$$\frac{\partial^2 S}{\partial a^2} \text{ and } \frac{\partial^2 S}{\partial b^2}$$

are both positive (an exercise).

Further Remark: Maximum Likelihood

There is an alternative way of looking at this. Suppose that we assume there is a linear relationship between Y and X but with some *noise*, ε:

$$Y = aX + b + \varepsilon.$$

We can, given a fixed model $Y = aX + b$, write down the probability of finding a certain set of data:

$$\mathbb{P}[\text{data} \mid Y = aX + b].$$

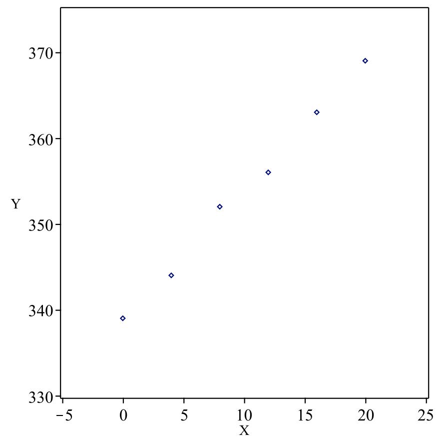
This is called the *likelihood function* of the model. Given that in curve fitting we have data, how about finding the curve $Y = aX + b$ that maximises the likelihood for a given data set? It turns out under reasonable assumptions that the least squares method gives the same answer as this likelihood method.

This means the least squares method gives the model that is most likely to produce the observed data.

Winter 2019

The below table, illustrated in the plot, lists the average carbon dioxide level, Y, in the atmosphere, measured in parts per million, X years after 1980.

years after 1980, X	0	4	8	12	16	20
Y in ppm	339	344	352	356	363	369



i. Considering the plot above, is it necessary to use a method as complex as Lagrange Interpolation to accurately estimate the average CO_2 level in 1998? [ANSWER PART A OR PART B AND NOT BOTH PARTS]

   A. If yes, use Lagrange Interpolation, with three suitable points, to estimate the average CO_2 level for $X = 18$.

   B. If no, accurately estimate the average CO_2 level for $X = 18$ using a method simpler than that of Lagrange Interpolation.

   Solution: No. Use straight-line interpolation:

(ii.) Looking at the plot, it is believed that Y and X have a relationship of the form:

$$Y = a \cdot X + b.$$

Find the best values of the constants a and b in the *Least Squares sense*. Use four significant figures for all calculations.

Solution: The normal equations are given by

Set up a table with Y, X, XY and X^2

							$\sum$
X	0	4	8	12	16	20	60
Y	339	344	352	356	363	369	2123
XY	0	1376	2816	4272	5808	7380	21652
X^2	0	16	64	144	256	400	880

This gives normal equations

$$60a + 6b = 2123$$
$$880a + 60b = 21652$$

which gives

iii. According to this model, when will the CO_2 level exceed 415 parts per million?

Solution: We solve $Y > 415$:

iii. Produce a *rough* sketch of your model $Y = a \cdot X + b$. Your plot should include labelled axes, a line, and a y-intercept.

Solution:

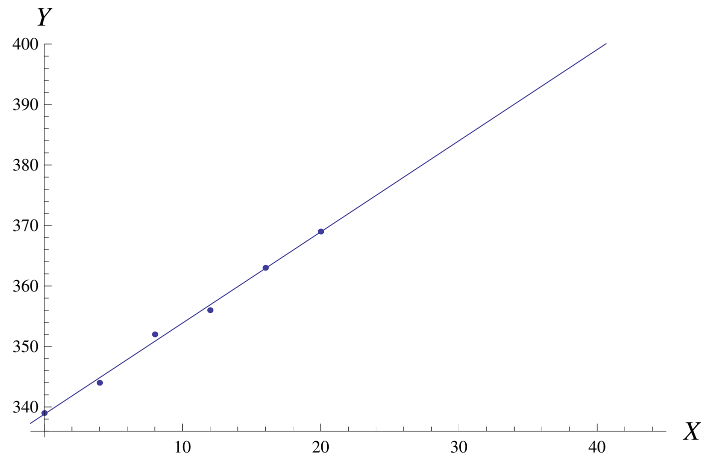


Figure 1.21: Here we see the data together with the fitted line.

However there is an easier way to remember the 'normal equations' that generalises neatly.

1.2.2 In Practise

In the following, X and Y are the data we know about and we want to find constants a and b. The above theory generalises to fitting curves of the form:

$$Y = a \cdot m_a + b \cdot m_b.$$

Call the functions m_a and m_b — and they are functions of X — the *multipliers* of a and b. The normal equations are:

$$\sum m_a Y = a \sum m_a^2 + b \sum m_a m_b$$
$$\sum m_b Y = a \sum m_a m_b + b \sum m_b^2$$

So you multiply $Y = am_a + bm_b$ by m_a — and add up; and by m_b — and add up; and these are the normal equations.

Normal Equations

Suppose that $\{(X_i, Y_i) : i = 1, 2, \ldots, n\}$ is a set of points on the plane. The following is a list of possible curves that might fit this data and the relevant normal equations that must be solved to find a and b.

1. $Y = aX^2 + b$

$$\sum Y = \sum aX^2 + \sum b$$
$$\sum X^2 Y = \sum aX^4 + \sum bX^2$$

2. $Y = aX^2 + bX$

$$\sum XY = \sum aX^3 + \sum bX^2$$
$$\sum X^2 Y = \sum aX^4 + \sum bX^3$$

3. $Y = \dfrac{a}{X} + b$

$$\sum Y = \sum \frac{a}{X} + \sum b$$

$$\sum \frac{Y}{X} = a\sum\left(\frac{1}{X^2}\right) + b\sum\left(\frac{1}{X}\right) = \sum\frac{a}{X^2} + \sum\frac{b}{X}$$

4. $Y = \dfrac{a}{X} + bX$

$$\sum \frac{Y}{X} = \sum \frac{a}{X^2} + \sum b$$

$$\sum XY = \sum a + \sum bX^2$$

Suppose that $\{(X_i, Y_i, Z_i) : i = 1, 2, \ldots, n\}$ is a set of points in *space*. If you want to fit a curve of the form

$$Z = aX + bY + c$$

to this data to find a, b and c you must solve the normal equations

$$\begin{aligned} \sum Z &= \sum aX + \sum bY + \sum c \\ \sum XZ &= \sum aX^2 + \sum bXY + \sum cX \\ \sum YZ &= \sum aXY + \sum bY^2 + \sum cY \end{aligned}$$

Proof. Left as an exercise •

Revision: Cramer's Rule

Suppose that a linear system/set of simultaneous equations has the form

$$\begin{aligned} a_{11}x + a_{12}y &= b_1 \\ a_{21}x + a_{22}y &= b_2 \end{aligned}$$

This can be written as $A\mathbf{v} = \mathbf{b}$,where A is a 2×2 matrix, $\mathbf{b}$ is the vector of constants and $\mathbf{v}$ is the vector of variables:

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

$$\Rightarrow x = \frac{\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}} \qquad y = \frac{\begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}$$

where the determinant of a 2×2 matrix is 'right-down-diagonal minus left-down diagonal':

These are written briefly as:

$$x = \frac{D_x}{D}, \; y = \frac{D_y}{D}. \tag{1.3}$$

If you only want to find x *or* y Cramer's Rule is useful. If the matrix of coefficients and/or the constants are decimals it is also a good idea to use Cramer's Rule. Cramer's Rule does generalise to simultaneous equations with more than two variables (although we won't require it).

Example

Use Cramer's Rule to solve the linear system

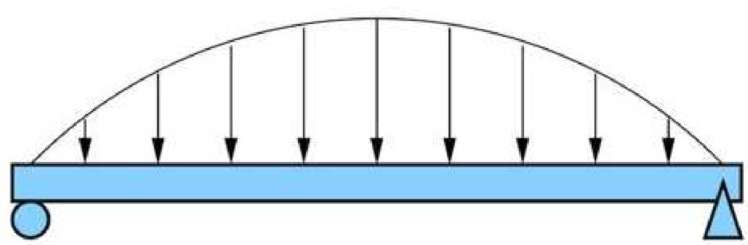
$$0.431x - 1.21y = 8.08$$
$$10.9x + 9.89y = 0.0128$$

Solution: First we find $x = \dfrac{D_x}{D}$:

Now we find $y = \dfrac{D_y}{D}$:

Examples

1. **Winter 2018** An engineer was tasked with finding the maximum deflection due to a symmetric load on a beam of span 5 m.



She did not have a formula for the load per unit length at each point, so instead decided to make some measurements of distance from one end of the beam, x, and of the load at that point, w, and fit a curve to the measurements. She decided that the curve resembled a '$\bigcap$' parabola and so decided to fit a curve of the form

$$w = Ax^2 + Bx,$$

to the data:

x/m	0	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
w/kN m^{-1}	0.0	8.6	15.3	20.0	22.9	23.8	22.8	20.0	15.2	8.6	0.0

(a) By fitting $w = Ax^2 + Bx$ to this data, find the best values of the constants A and B in the *Least Squares sense*. Use at least three significant figures for all calculations. [HINT: $\sum wx = 392.8$, $\sum x^3 \approx 378.1$, $\sum x^2 = 96.25$, $\sum wx^2 \approx 1170$, $\sum x^4 \approx 1583$]

Solution: We write down the normal equations and solve for A:

and B:

(b) Hence find the *location* of maximum load according your model $w = Ax^2 + Bx$.

Solution: Where is the maximum of a function found?

(c) Given that the load is symmetric, comment briefly on your answer to ii.

Solution: As expected, it is at the...

(d) Using the values of A and B found in part i., estimate the total load by evaluating

$$w_T = \int_0^5 (Ax^2 + Bx)\, dx.$$

Solution: We evaluate

2. **Winter 2015** The below table lists the marginal cost, C, for taking N workers to an offshore job.

number of workers, N	5	10	15	20	25
marginal cost, C, in thousands of euro	14.05	10.10	10.15	11.20	12.65

It is believed that C and N have a relationship of the form:

$$C = a \cdot N + \frac{b}{N}.$$

(a) Find the best values of the constants a and b in the *Least Squares sense*. Use four significant figures for all calculations.

[HINT: $\sum CN = 863.75$, $\sum N^2 = 1375$, $\sum(C/N) = 5.563$, $\sum(1/N^2) = 0.059$]

Solution: The normal equations are given by:

Using the hint this gives normal equations:

Therefore the model is

(b) Using differentiation **only**, find the number of workers N such that C is minimised.

Solution: We solve

We could calculate $C(11)$ and $C(12)$ to see which is the minimum... Also note that

$$\frac{d^2C}{dN^2} > 0,$$

so we do have a minimum.

Marking Scheme: Summer 2014

Tymar Engineering are a civil engineering firm in the midst of expansion. Due to economies of scale, the cost in thousands C of training in a new employee reduces when the number of trainees T increases. In order to get an idea of how much a sector expansion will cost, the HR manager was asked to produce data on previous sector expansions:

number of trainees, T	5	10	15	20
cost of training per trainee, C	3.5	2.8	2.5	2.4

It is believed that C and T have a relationship of the form:

$$C = a + \frac{b}{T}.$$

i. Find the best values of the constants a and b in the *least squares sense*. Use three places of decimals for all calculations (today I ask instead for four significant figures).

[5 Marks]

Solution: The normal equations are given by

$$\sum C = \sum a + b \sum \frac{1}{T}$$
$$\sum \frac{C}{T} = a \sum \frac{1}{T} + b \sum \frac{1}{T^2}. \qquad \mathbf{[1]}$$

Therefore we tabulate **[2]**

					$\sum$
T	5	10	15	20	-
C	3.5	2.8	2.5	2.4	11.2
$\frac{1}{T}$	0.20	0.10	0.067	0.05	0.417
$\frac{C}{T}$	0.7	0.28	0.167	0.12	1.267
$\frac{1}{T^2}$	0.04	0.01	0.004	0.003	0.057

This yields

$$\begin{aligned}
4a + 0.417b &= 11.2 \\
0.417a + 0.057b &= 1.267 \qquad [\mathbf{1}] \\
\Rightarrow 4a &= 11.2 - 0.417b \\
\Rightarrow a &= \frac{11.2 - 0.417b}{4} \\
\Rightarrow 0.417\left(\frac{11.2 - 0.417b}{4}\right) + 0.057b &= 1.267 \\
\Rightarrow 0.417(11.2 - 0.417b) + 0.228b &= 5.068 \\
\Rightarrow 4.67 - 0.174b + 0.228b &= 5.068 \\
\Rightarrow 0.054b &= 0.398 \\
\Rightarrow b &= 7.37 \\
\Rightarrow a &= \frac{11.2 - 0.417(7.37)}{4} = 2.032 \qquad [\mathbf{1}]
\end{aligned}$$

This gives an answer of

$$C(T) = 2.032 + \frac{7.37}{T}.$$

ii. Hence find the number of trainees T required to get the cost of training an employee down to €2,200.

[1 Mark]

Solution: We want to find the value of T such that $C(T) = 2.2$:

$$\begin{aligned}
2.032 + \frac{7.37}{T} &= 2.2 \qquad [\mathbf{1}] \\
\Rightarrow 2.032T + 7.37 &= 2.2T \\
\Rightarrow 0.168T &= 7.37 \\
\Rightarrow T &= 43.869
\end{aligned}$$

So we need $T = 44$.

Marking Scheme

When a force of 100 kN is applied to an object of length L_0, with a cross sectional area of 1 cm^2, made out of an alloy with a Young's Modulus of E GPa, the length is transformed to L_1. In order to measure the Young's Modulus, E, of a particular alloy, the following recordings were made

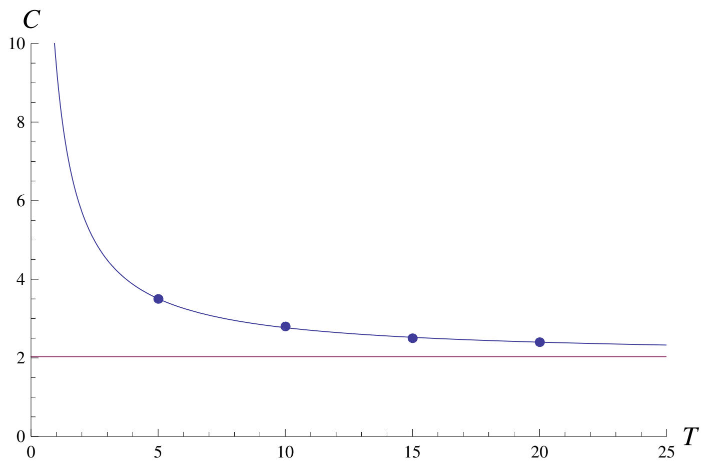


Figure 1.22: Here we see the data and the fitted curve. Note that as T gets big, the cost of training a trainee reaches a basement figure of €2,032 each.

L_0/m	2	4	6	8	10
L_1/m	2.011	4.019	6.032	8.037	10.054

The formula governing the relationship between L_1 and L_0 can be written as

$$L_1 = \left(1 + \frac{1}{E}\right) L_0 + \varepsilon,$$

where ε is a small corrective constant to account for measurement error and non-linear behavior.

i. Write the above relationship in the form $Y = aX + b$, clearly identifying the constants a and b and the variables Y and X.

[2 Marks]

Solution:

$$\underbrace{L_1}_{Y} = \underbrace{\left(1 + \frac{1}{E}\right)}_{a} \underbrace{L_0}_{x} + \underbrace{\varepsilon}_{b} \qquad \textbf{[2]}$$

ii. Find the best values of the constants a and b in the *least squares sense*. Make all calculations correct to four significant figures.

[6 Marks]

Solution: The normal equations are given by

$$\sum XY = a\sum X^2 + b\sum X$$
$$\sum Y = a\sum X + \sum b \qquad \textbf{[1]}$$

Hence we tabulate **[2]**

						$\sum$
X	2	4	6	8	10	30
Y	2.011	4.019	6.032	8.037	10.05	30.15
XY	4.022	16.08	36.19	64.30	100.5	221.1
X^2	4	16	36	64	100	220

Hence we have

$$221.1 = 220a + 30b$$
$$30.15 = 30a + 5b \qquad [\mathbf{1}]$$
$$\Rightarrow -180.9 = -180a - 30b$$
$$\Rightarrow 40.2 = 40a$$
$$\Rightarrow a := 1.005$$
$$\Rightarrow 221.1 = 221.1 + 30b$$
$$\Rightarrow 30b = 0 \Rightarrow b = 0 \qquad [\mathbf{1}]$$
$$\Rightarrow Y = 1.005X \qquad [\mathbf{1}]$$

iii. Hence calculate the Young's Modulus of the alloy.

[2 Marks]

Solution: We have

$$1 + \frac{1}{E} = 1.005 \qquad [\mathbf{1}]$$
$$\Rightarrow E + 1 = 1.005E$$
$$\Rightarrow 0.005E = 1$$
$$\Rightarrow E = 200 \qquad [\mathbf{1}]$$

Exercises:

1. **Autumn 2020** The size of a reservoir should depend on the water flow in the river that is being impounded. Often, records of such flow data are not available while meteorological data on precipitation is. Therefore, it is useful to determine a relationship between flow and precipitation. This relationship can then be used to estimate flows for a given precipitation. The following data is available for a river that is to be dammed.

precipitation, P, in cm	88.9	108.5	104.1	139.7	127
flow, F, in cubic metres per second	14.6	16.7	15.3	23.2	19.5

It is believed that F and P have a relationship of the form:

$$F = a \cdot P + b.$$

(a) Find the best values of the constants a and b in the *Least Squares sense*. Use four significant figures for all calculations. [HINT: $\sum P = 568.2$, $\sum F = 89.3$, $\sum P^2 \approx 66160$, $\sum FP \approx 10420$] **Ans:** $F = 0.1711P - 1.580$.

(b) Estimate the water flow F when precipitation is $P = 120$. **Ans:** 18.95 m^3/s.

(c) What precipitation level P predicts a flow of $F = 25$? **Ans:** 155.3 cm.

2. **Winter 2017** For a given spring, the spring constant, k, depends on the number of coils, the shear modulus of the alloy the spring is made from, the radius of the wire and the radius of the coils in metres. A number of these quantities are difficult to measure accurately and so it is sometimes easier to measure the spring constant using data from an experiment. Hence the following measurements were made of an oscillating spring with a 'weight' attached:

mass, M/kg	1	2	3	4	5	6
Period-Squared of Oscillation, T^2/s^2	0.4	1.1	1.6	2.0	2.4	3.0

The formula governing the relationship between T^2 and M can be written as

$$T^2 = \frac{4\pi^2}{k}M + \varepsilon,$$

where ε is a small corrective constant to account for non-linear behavior. This is of the form

$$y = mx + c,$$

with $y = T^2$, $m = \dfrac{4\pi^2}{k}$, $x = M$, and $c = \varepsilon$:

x	1	2	3	4	5	6
y	0.4	1.1	1.6	2.0	2.4	3

(a) By fitting $y = mx + c$ to this data, find the best values of the constants m and c in the *Least Squares sense*. Use four significant figures for all calculations. [HINT: $\sum y = 10.5$, $\sum x = 21$, $\sum x^2 = 91$, $\sum xy = 45.4$] **Ans:** $T^2 = 0.4943x + 0.02$

(b) Hence calculate the spring constant (the units are N/m). **Ans:** 79.87 N/m

3. **Autumn 2017** The shear stresses of five specimens taken at various depths in a clay stratum are listed below.

depth, D, in metres	1.9	4.2	5.8	8.1	10.0
shear stress, τ, in kilopascals	14.4	19.2	33.5	71.8	76.6

It is believed that τ and D have a relationship of the form:

$$\tau = a + bD^2.$$

(a) Using all of the data and not just three data points, find the best values of the constants a and b in the *Least Squares sense*. Use four significant figures for all calculations. [HINT: $\sum \tau = 215.5$, $\sum D^2 = 220.5$, $\sum \tau D^2 \approx 13890$, $\sum D^4 \approx 15760$] **Ans:** $\tau = 11.06 + 0.7266D^2$.

(b) Use this model to predict the shear stress, τ, for $D = 6$. **Ans:** 37.22 kPa.

4. **Autumn 2016** After establishing market share with initially low prices, a company selling bio-tanks to private homeowners has been steadily increasing prices. Unfortunately the rising prices have been met with decreasing sales

price, P, in thousands	5.0	5.5	6.0	6.5	7.0
sales, S, in hundreds of units	5.50	4.77	4.17	3.65	3.21

It is believed that S and P have a relationship of the form:

$$S = a + \frac{b}{P}.$$

(a) Find the best values of the constants a and b in the *Least Squares sense*. Use four significant figures for all calculations. [HINT: $\sum S = 21.3$, $\sum(1/P) = 0.846$, $\sum(S/P) = 3.685$, $\sum(1/P^2) = 0.145$] **Ans:** $S = -3.125 + 43.64/P$

(b) According to this model, what price P should the bio-tanks be in order to achieve sales of $S = 3$? **Ans:** €7,125.

(c) Check the validity of the model for very large prices.

1.2.3 Pearson's Correlation Coefficient r

Consider the two sets of data below:

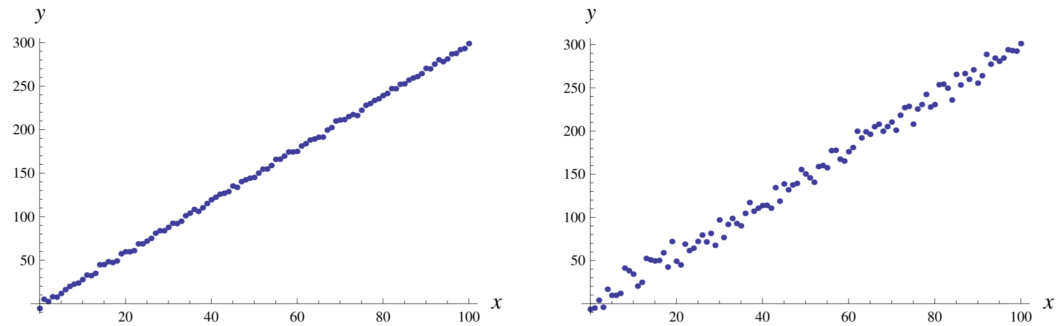


Figure 1.23: Although both sets of data are well-fitted by the same line, it is clear that the first set of data will have a high-quality fit while the second will be a less-quality fit.

To measure the goodness of fit here we will use what we will call simply call the *correlation coefficient*, r^2, and we consider it a measurement of the goodness of fit. We shall apply it when fitting linear data $Y = aX + b$ to n data points $\{(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)\}$. The formula is given by:

$$r = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}}.$$

You will not be asked to calculate a correlation coefficient (well maybe using Excel — in your first assignment) but you do need to understand that the closer to plus or minus one the better the fit.

Examples

1. Tymar Engineering are a civil engineering firm in the midst of expansion. Due to economies of scale, the cost in thousands C of training in a new employee reduces when the number of trainees T increases. In order to get an idea of how much a sector expansion will cost, the HR manager was asked to produce data on previous sector expansions:

number of trainees, T	5	10	15	20
cost of training per trainee, C	3.5	2.8	2.5	2.4

   We can examine the correlation of $C = a \cdot \frac{1}{T} + b$. Here we have $X \sim \dfrac{1}{T}$ and $Y \sim C$ so we must first convert the data to $(1/T, C)$ rather than (T, C).

If we scatter-plot C vs $\dfrac{1}{T}$ we get:

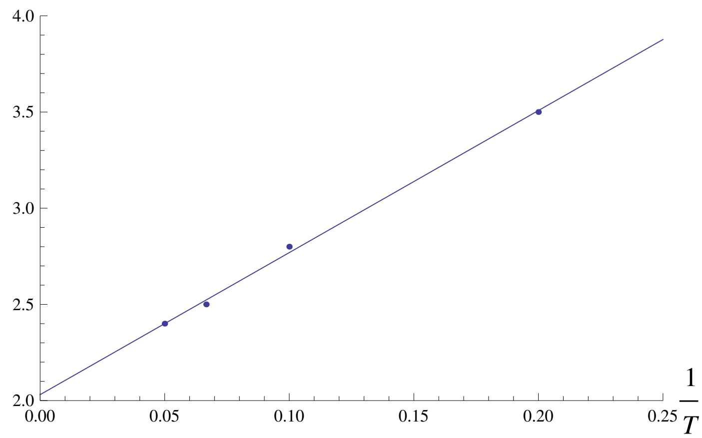


Figure 1.24: How strong is this straight-line-relationship? Note that when doing a least squares fit we actually fit $C = \frac{a}{T} + b$ to C vs T rather see Figure 1.22.

If we apply the formula to get an r-goodness-of-fit/correlation coefficient we find that $r \approx 0.99896$ indicating a very good fit.

2. When a force of 100 kN is applied to an object of length L_0, with a cross sectional area of 1 cm^2, made out of an alloy with a Young's Modulus of E GPa, the length is transformed to L_1. In order to measure the Young's Modulus, E, of a particular alloy, the following recordings were made

L_0/m	2	4	6	8	10
L_1/m	2.011	4.019	6.032	8.037	10.054

We can consider the correlation for

$$L_1 = \left(1 + \frac{1}{E}\right) \cdot L_0 + \varepsilon,$$

where ε is a small corrective constant to account for measurement error and non-linear behavior.

Exercise: **Autumn 2015** The owner of a new manufacturing engineering firm with rapid growth is keeping track of the monthly sales, S, in hundreds, in the months after launch:

months after start-up, m	2	4	6	8	10	12
monthly sales, $S/100$	1	2	5	11	20	35

ii. The producer wants to fit a curve to the data to predict future sales figures. She is considering the model $S_1 = am^2 + b$ but also the model $S_2 = am^3 + b$. Suppose the Pearson Correlation Coefficient for the first model is 0.9871 and for the second model is 0.9998. Which model should the producer use?

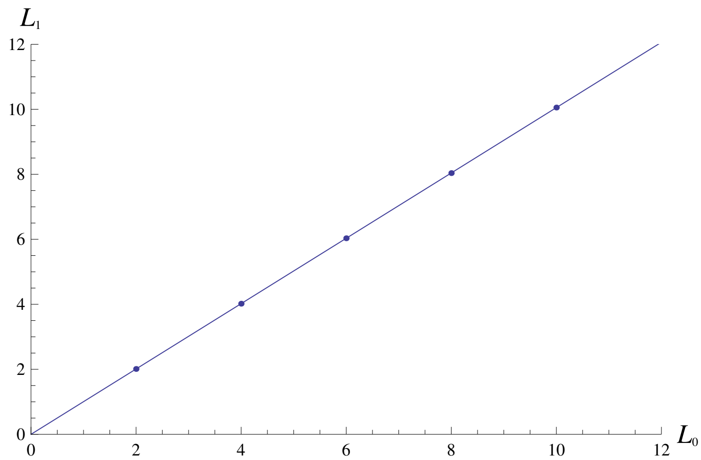


Figure 1.25: How strong is this straight-line-relationship in Example 2? In this case *very* strong! In fact, $r = 1$ to five decimal places.

1.2.4 Non-Linear Laws

Many quantities share a linear relationship:

In particular, if we know that some process is a linear one, but we don't know the *parameters* a and b, then we can make measurements of Y and X, and apply the method of least squares to find the best possible approximations to a and b. However not all mathematical relationships take such a tractable form. For example, it can be shown that the decay of a radioactive material is modeled by:

In this case N is the number of particles at a time t, a is the initial number of particles and b is a constant. If we were to take measurements of N at various times t, and plotted them, we would get something like:

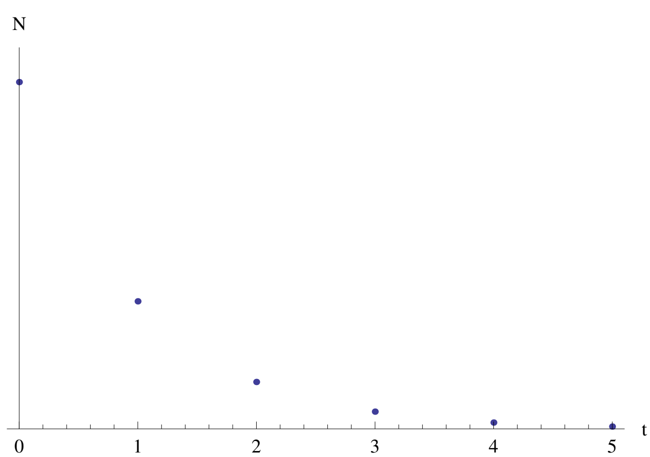


Figure 1.26: In Radioactive Decay, the number of particles reduced exponentially.

However we can transform the multiplication into addition and the powers into multiplication in the decay equation and write it in a linear form by taking logarithms.

Review of Logs

We define the *natural logarithm to base e* as the inverse function of $f(x) = e^x$. Another way of looking at logarithms is to consider them as a way of converting multiplication to addition; division to subtraction; and powers to multiplication — by way of the *laws of logs*:

Examples

Write each of the following in linear form, $y = mx + c$. Note that by definition $\ln(e^n) = n$.

1. $Y = a \cdot b^X$

   Solution:

2. $Y = aX^b$

   Solution:

3. $Y = ae^{bX}$.

   Solution:

The whole time we want

$$\underbrace{\text{variable}_1}_{y} = \underbrace{\text{constant}_1}_{m} \cdot \underbrace{\text{variable}_2}_{x} + \underbrace{\text{constant}_2}_{c}. \tag{1.4}$$

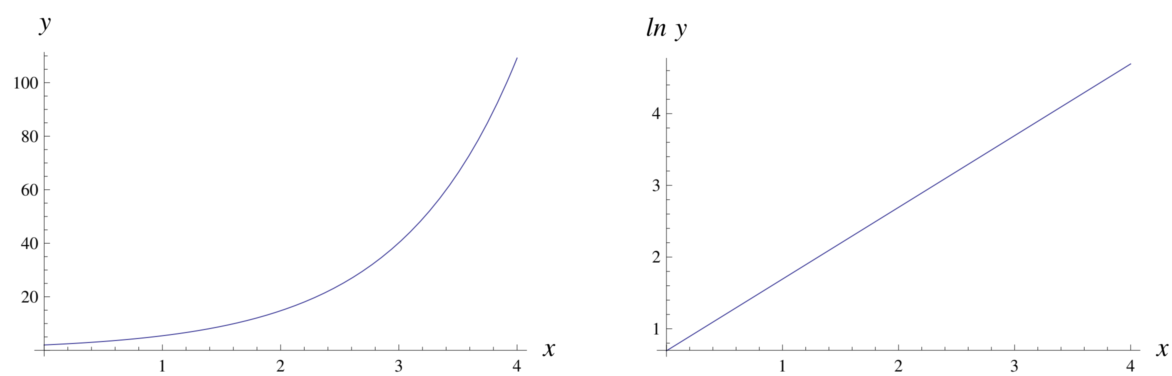


Figure 1.27: Linearisation converts data — such as the $y = 2e^x$ here on the left — into linear data. The data on the right is $\ln y$ vs x, specifically $\ln y = 1x + \ln 2$.

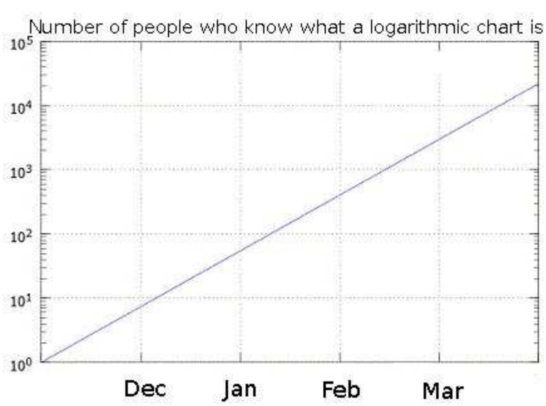


Figure 1.28: Do you know what this is referring to?

Exercises
For the following, a and b are the constants. Please write in linear form $y = mx + c$; identifying y, m, x, c:

$$R = aV^b.$$
$$M = ae^N.$$
$$R = aS^x.$$

To fit a non-linear curve to data we do the following:

1. Write the non-linear relationship in linear form $y = mx + c$ identifying y, m, x and c.
2. Tabulate the 'new' x and y values.
3. Apply the Linear Least Squares Method; i.e.

$$\sum y = m \sum x + \sum c$$
$$\sum xy = m \sum x^2 + c \sum x$$

4. Convert back to the original constants if needs be.

Examples

1. **Autumn 2015** Consider a light beam of span L with a load per unit length of $w(x)$. The deflection of the beam at any point $y(x)$ is given as the solution the fourth order differential equation

$$EI \cdot \frac{d^4y}{dx^4} = -w(x),$$

and from this a formula for the maximum deflection

$$\delta = f(w_i, L, E, I),$$

can be found in terms of the loads w_i, the span L, the Young's Modulus of the alloy, E, and the second moment of area of the beam, I.

It is not always possible to solve the fourth order differential equation or find where the deflection is at a maximum. In this case, an *empirical law* based on data is one way of generating a formula for the maximum deflection. Suppose an engineer measured and recorded the maximum deflections, δ of cantilevered beams of the same Young's Modulus and second moment of area, under a load $w(x)$ that depends on x and L only, for different values of L:

L/m	4	5	6	7
δ/mm	5.12	12.50	25.92	48.02

After plotting the data, the engineer believes the data has a relationship of the form

$$\delta = a \cdot L^N.$$

(a) Write the relationship in linear form, $y = m \cdot x + c$.

Solution:

(b) Hence find the best value of N in the (log-linear) *Least Squares* sense. Use four significant figures for all calculations.

Solution: The normal equations are given by

If we transform the data according to $x = \ln L$ and $y = \ln \delta$ and tabulate x^2 and xy:

					$\sum$
x	1.386	1.609	1.792	1.946	6.733
y					
x^2					
xy					

This gives normal equations

Solving these, we get $m = N = 3.940$.

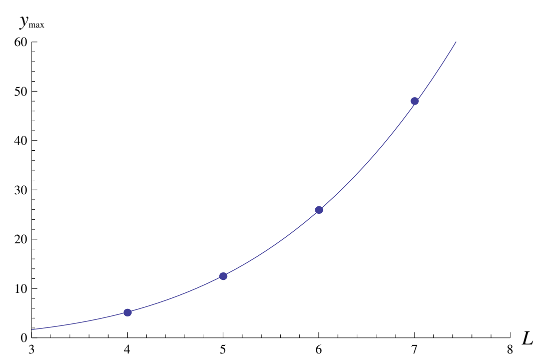


Figure 1.29: Just by fitting a curve to the data, we can find an *empirical rule* for the deflection: in this case $\delta = 0.02216L^{3.94} \approx \frac{1}{50}L^4$. Note that if we double the span, $L_0 \to L_1 = 2L_0$, the deflection is multiplied by *16*: $\delta_1 = a(2L_0)^4 = 16aL_0 = 16\delta_0$

2. **MATH7021 Summer 2014** In order to predict the number of blows N required to pile drive to a depth of D m in a given soil, the following recordings were made

depth, D/ m	0	1	2	3	4
blows, N	10	25	65	155	390

ii. A more sophisticated[2] method of interpolating $N(2.5)$ is to model the relationship between N and D using a function such as

$$N = a \cdot e^{bD}.$$

Write the relationship in linear form, $y = m \cdot x + c$ and find the best values of the constants a and b in the *log-linear least squares sense*, and hence estimate the number of blows required to reach a depth of 2.5 m. Use three significant figures for all calculations.

[2] than Lagrange Interpolation

Solution: First we write the relationship in the form $y = mx + c$:

The normal equations are given by

$$\sum y = m \sum x + \sum c$$
$$\sum xy = m \sum x^2 + c \sum x$$

Hence we tabulate

						$\sum$
$D = x$	0	1	2	3	4	10
$\ln N = y$	2.30	3.22	4.17	5.04	5.97	20.7
x^2	0	1	4	9	16	30
xy	0	3.22	8.34	15.1	23.9	50.6

Hence we have

So we have $N(2.5)$ approximated by:

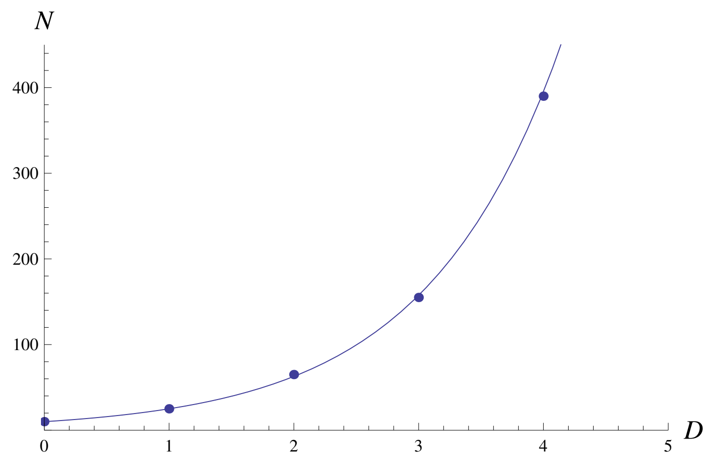


Figure 1.30: Perhaps as you go deeper, the ground compacts and you need more and more blows to go one metre further.

Marking Scheme

In order to be able to match production with demand, the manufacturers of a new concrete mix examined the sales of their product, in tonnes, over the last four years.

years after launch, t/years	0	1	2	3	4
sales, S/tonnes	100	180	420	860	1210

Using a model such as $S = a \cdot b^t$, or otherwise, extrapolate the sales for $t = 5$ years after launch. Start by writing the relationship in linear form, $y = m \cdot x + c$. (Today I ask for all calculations to be correct to three/four significant figures).

[7 Marks]

Solution: First we linearise the relationship:

$$\begin{aligned} &\ln(S) = \ln(ab^t) \\ \Rightarrow &\ln S = \ln a + \ln b^t \\ \Rightarrow &\underbrace{\ln S}_{y} = \underbrace{\ln b}_{m} \cdot \underbrace{t}_{x} + \underbrace{\ln a}_{c} \qquad \textbf{[1]} \end{aligned}$$

The normal equations are given by

$$\begin{aligned} \sum xy &= m \sum x^2 + c \sum x \\ \sum y &= m \sum x + \sum c. \qquad \textbf{[1]} \end{aligned}$$

Hence we tabulate: **[1]**

						$\sum$
$t = x$	0	1	2	3	4	10
$\ln S = y$	4.61	5.19	6.04	6.76	7.10	29.7
xy	0	5.19	12.08	20.28	28.40	65.95
x^2	0	1	4	9	16	30

Therefore the normal equations are given by:

$$
\begin{aligned}
65.95 &= 30m + 10c \\
29.7 &= 10m + 5c \qquad [\mathbf{1}] \\
\Rightarrow -59.4 &= -20m - 10c \\
\Rightarrow 6.55 &= 10m \\
\Rightarrow m &= 0.66 \\
\Rightarrow 65.95 &= 19.8 + 10c \\
\Rightarrow 10c &= 46.15 \\
\Rightarrow c &= 4.62 \qquad [\mathbf{1}]
\end{aligned}
$$

Now note that $m = \ln b = 0.66 \Rightarrow b = 1.93$ and $c = \ln a = 4.62 \Rightarrow a = 101.49$ [**1**] so we have

$$
\begin{aligned}
S &= 101.49(1.93)^t \\
\Rightarrow S(t) &= 101.49(1.93)^5 = 2717.75. \qquad [\mathbf{1}]
\end{aligned}
$$

Worked Example: MATH7021 Summer 2013

World concrete production has been increasing over the five decades from 1971 to 2011:

decades after 1971, t	0	1	2	3	4
concrete production in Gt, C	3.8	5.9	7	10.7	18.8

It is believed that t and C have a relationship of the form:

$$C = k\,R^t.$$

i. Find the best values of the constants k and R in the *least squares sense*. Start by writing the relationship in linear form, $y = m \cdot x + c$. Use two places of decimals for all calculations (today I ask instead for four significant figures).

ii. Hence make a prediction for world concrete production in 2021.

Solution: Firstly linearise the relationship:

$$
\begin{aligned}
\ln C &= \ln(k \cdot R^t) \\
&= \ln k + \ln(R^t) \\
\Rightarrow \underbrace{\ln C}_{=:y} &= \underbrace{\ln(R)}_{=:m} \cdot \underbrace{t}_{=:x} + \underbrace{\ln k}_{=:c}
\end{aligned}
$$

The Normal Equations are given by

$$
\begin{aligned}
\sum y &= m \sum x + \sum c \\
\sum xy &= m \sum x^2 + c \sum x.
\end{aligned}
$$

Hence we tabulate xy and x^2:

						$\sum$
$t = x$	0	1	2	3	4	10
$\ln C = y$	1.34	1.78	1.95	2.37	2.93	10.37
xy	0	1.78	3.90	7.11	11.72	24.51
x^2	0	1	4	9	16	30

Hence we have the normal equations

$$\begin{aligned} 10.37 &= 10m + 5c \\ 24.51 &= 30m + 10c \\ \Rightarrow -20.74 &= -20m - 10c \\ \Rightarrow 3.77 &= 10m \Rightarrow m = 0.38 \\ \Rightarrow 5c &= 10.37 - 10(0.38) \\ \Rightarrow c &= \frac{10.37 - 3.8}{5} = 1.31. \end{aligned}$$

Recall that $m = \ln R$ and $c = \ln k$ so

$$\begin{aligned} \ln k = 1.31 &\Rightarrow k = e^{1.31} \approx 3.71 \\ \ln R = 0.38 &\Rightarrow R = e^{0.38} \approx 1.46 \\ &\Rightarrow C(t) = 3.71 \cdot 1.46^t \end{aligned}$$

Now 2021 is five decades after 1971 so we want to find $C(5)$:

$$C(5) = 3.71 \cdot (1.46)^5 \approx 24.61 \text{ Gt}.$$

Marking Scheme: Autumn 2016

For a beam of fixed material, cross-section and total load, the span L and the maximum deflection δ were measured and recorded.

L/m	4.0	5.0	6.0	7.0	8.0
δ/mm	1.92	3.75	6.48	10.29	15.36

(b) Looking at the data in its entirety, it is reasonable to postulate a relationship between δ and L of the form:

$$\delta = A \cdot L^N.$$

Write the relationship in the form $y = mx + c$.

[2 Marks]

Solution:

$$\begin{aligned} \ln(\delta) &= \ln(A \cdot L^N) \\ &= \ln(A) + \ln(L^N) = \ln(A) + N \cdot \ln(L) \\ \Rightarrow \underbrace{\ln(\delta)}_{y} &= \underbrace{N}_{m} \cdot \underbrace{\ln(L)}_{x} + \underbrace{\ln(A)}_{c} \qquad \textbf{[2]} \end{aligned}$$

(c) Hence use the (log-linear) Least Squares Method to find the best values of the constants A and N. Use four significant figures for all calculations.

[7 Marks]

Solution: We transform the data to $(x, y) = (\ln L, \ln \delta)$ [**1**] and consider the normal equations for $y = mx + c$:

$$\sum y = m \sum x + \sum c$$
$$\sum xy = m \sum x^2 + c \sum x, \qquad [\mathbf{1}]$$

and so we tabulate [**1**]:

						$\sum$
$\ln L = x$	1.386	1.609	1.792	1.946	2.079	8.813
$\ln \delta = y$	0.652	1.322	1.869	2.331	2.732	8.906
x^2	1.921	2.589	3.211	3.787	4.322	15.83
xy	0.904	2.127	3.349	4.356	5.680	16.42

Therefore the normal equations are given by:

$$8.906 = 8.813m + 5c$$
$$16.42 = 15.83m + 8.813c \qquad [\mathbf{1}]$$
$$\Rightarrow 78.49 = 77.67m + 44.07c$$
$$\Rightarrow -82.1 = -79.15m - 44.07c$$
$$\Rightarrow -3.61 = -1.481m$$
$$\Rightarrow m = \frac{3.61}{1.481} = 2.439. \qquad [\mathbf{1}]$$
$$\Rightarrow 8.906 = 8.813(2.439) + 5c$$
$$\Rightarrow 8.906 = 21.5 + 5c$$
$$\Rightarrow 5c = -12.59 \Rightarrow c = -2.518 \qquad [\mathbf{1}]$$
$$\Rightarrow \ln A = -2.518 \Rightarrow A = e^{-2.518} = 0.081 \qquad [\mathbf{1}]$$
$$\Rightarrow \delta = 0.081 \cdot L^{2.438}.$$

(d) Use this *empirical law* to estimate the maximum deflection for such a beam of length $L = 6.5$ m.

[2 Marks]

Solution: We calculate

$$\delta(6.5) = 0.081(\underbrace{6.5}_{[\mathbf{1}]})^{2.438} = 7.784 \text{ mm} \qquad [\mathbf{1}]$$

Exercises:

1. For various beams, the span L and the maximum deflection δ were measured and recorded.

L/m	2.0	4.0	5.0	8.0	10.0
δ/mm	1.15	1.20	1.26	1.32	1.38

The values of L and y are related by a formula of the type $\delta = AL^N$ where A and N are constants.

(a) Write the relationship in linear form, $y = m \cdot x + c$.

(b) Hence use the log-linear Least Squares Method find the best values of the constants A and N. Make all corrections correct to four significant figures. **Ans:** $\delta \approx 1.05\, L^{0.11}$.

2. **Winter 2019** The Richter Scale is defined such that a magnitude $M + 1$ earthquake is ten times stronger than a magnitude M earthquake (in a certain sense). What people don't know is that there is an empirical law that says that magnitude $M + 1$ earthquakes are, for some constant b, b times *rarer* than magnitude M earthquakes. This is the *Gutenberg–Richter Law*, which says that, approximately, the number of earthquakes in a region, N, of magnitude M is given by:

$$N = a \cdot b^M.$$

Where an interval, e.g. $5 < M \leq 6$, is represented by a point, e.g. $M = 5.5$, global data from the 20th century shows:

magnitude, M	4.5	5.5	6.5	7.5	8.5
average annual global frequency, N	6200	800	120	18	1

(a) Write the relationship $N = a \cdot b^M$ in the form $y = m \cdot x + c$; where y and x are variables, and m and c are constants.

(b) Hence use the (log-linear) Least Squares Method to find the best values of the constants a and b. Use at least three significant figures for all calculations. [HINT: a is large] **Ans:** $N \approx 36,000,000 \cdot (0.139)^M$.

(c) Use your model to estimate how much rarer a magnitude $M = 8$ earthquake is than a magnitude $M = 5$ earthquake. **Ans:** ≈ 372 times rarer.

3. **Winter 2014** An engineer analysed a gradation test of a sample of Anthracite and produced the following table of data:

grain size, g/mm	0.6	1.2	1.4	1.7	2.0	2.5
passing percentage, P	1	3	17	47	77	97

The engineer plotted the data and it had an S shape. However the four middle data points appeared to have the shape of an exponential curve:

$$P = Ae^{kg}.$$

The engineer needed to find the grain size $g_{\frac{1}{2}}$ such that $P = 50$ and so considered the relevant data only:

g/mm	1.2	1.4	1.7	2.0
passing percentage, P	3	17	47	77

(a) Write the relationship $P = Ae^{kg}$ in linear form, $y = mx + c$.

(b) Hence find the best values of A and k in the (log-linear) *Least Squares* sense. Use four significant figures for all calculations. **Ans:** $A \approx 0.04817$ and $k \approx 3.851$.

(c) Solve $P = 50$ to find the value of $g_{\frac{1}{2}}$. **Ans:** ≈ 1.803 mm.

4. **Autumn 2017** The following model is frequently used in environmental engineering to model the effect of temperature in °C, T, on biochemical reaction rates, k:

$$k = a \cdot b^T.$$

Suppose the following measurements were made:

temperature, T, in ° C	6	12	18	24	30
rate, k, per day	0.14	0.20	0.31	0.46	0.69

(a) Write the relationship in the form $y = mx + c$, identifying y, m, x and c.

(b) Hence use the (log-linear) Least Squares Method to find the best values of the constants a and b. Use four significant figures for all calculations. **Ans:** $k = 0.092 * (1.069)^T$.

(c) Use the model to estimate the reaction rate at $T = 17$. **Ans:** 0.286 per day.

(d) Prove that a is the predicted rate at 0 °C.

5. **Winter 2017** As a member of *Engineers Without Borders*, you are working in a community that has contaminated water. At time, $t = 0$, you add a disinfectant to a cistern with a very high concentration, c, of contaminating bacteria. You make the following measurements at several times thereafter:

time, t/hours	2	4	6	8	10
concentration, c/ (units/100 mL)	430	190	80	35	16

The water is safe to drink when the concentration falls below 5 units per 100 ml.

(a) It is reasonable to postulate a relationship between c and t of the form:

$$c = a \cdot e^{bt}.$$

Write the relationship in the form $y = mx + c$, identifying y, m, x and c.

(b) Hence use the (log-linear) Least Squares Method to find the best values of the constants a and b. Use four significant figures for all calculations. **Ans:** $c = 972.6\, e^{-0.4125t}$

(c) Hence estimate the time at which the drinking water will be safe. **Ans:** $12.78 \approx 13$ hours later.

6. A lake has been contaminated with *E. coli* bacteria and the growth is being monitored by a team of engineers:

number of days after first detection, t	0	1	2	3	4
colony-forming units per ml, C	1.92	3.75	6.48	10.29	15.36

(a) Use Lagrange Interpolation to estimate the *E. coli* level for $t = 5$. **Ans:** 21.69 using $t = 2, 3, 4$.

(b) Looking at the data in its entirety, it is reasonable to postulate a relationship between C and t of the form:

$$C = a \cdot e^{bt}.$$

Write the relationship in the form $y = mx + c$.

(c) Hence use the (log-linear) Least Squares Method to find the best values of the constants a and b. Use four significant figures for all calculations. **Ans:** $C(t) = 2.111e^{0.517t}$.

(d) Use the model to estimate the *E. coli* level for $t = 5$. **Ans:** 28.

(e) Complete the following in your answer book:

____ growth is slower than ____ growth therefore the answer for part ____ is less than the answer for part ____.

[HINT: Lagrange Interpolation uses a t^2 and so is *quadratic*. The fit in part (b) is e^{bt} and so exponential.

7. **Autumn 2020** *Benford's Law* is an empirical law that says that for many types of numerical data, the first digit is not uniformly distributed among $\{1, 2, \ldots, 9\}$ but rather there is a much larger proportion of ones than twos and twos than threes, etc. The first digits of 11, 69.5 and 987.2 are one, six and nine, respectively. In 1972, Hal Varian suggested that the law could be used to detect possible fraud in lists of socio-economic data submitted in support of public planning decisions. The following (incomplete) set of data was extracted from a database of invoices.

digit, D	1	2	3	4	5
Proportion, P, of bills starting with D	0.301	0.176	0.125	0.097	0.079

(a) Use Lagrange Interpolation to estimate the proportion of bills starting with $D = 6$. **Ans:** ≈ 0.071.

(b) It is reasonable to postulate a relationship between P and D of the form:

$$P = a \cdot D^b.$$

Write the relationship in the form $y = mx + c$, identifying y, m, x and c.

(c) Hence use the (log-linear) Least Squares Method to find the best values of the constants a and b. Use four significant figures for all calculations. **Ans:** $P = 0.3073 \cdot D^{-0.8327}$.

(d) Use the model to estimate the proportion of bills with first digit $D = 6$. **Ans:** 0.06844.

1.3 Chapter Summary

Normal Equations

For $y = m \cdot x + c$ are

$$\sum y = m \sum x + \sum c$$
$$\sum xy = m \sum x^2 + c \sum x$$

Logarithms

$$\ln(x \cdot y) = \ln x + \ln y$$

$$\ln(x^n) = n \cdot \ln x$$

$$\ln(e^x) = x$$

Lagrange Interpolation

$$\ell(x) = \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)} \cdot f(x_0) + \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)} \cdot f(x_1) + \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x_2 - x_1)} f(x_2)$$

Need to Knows

1. *Lagrange Interpolation* is the fitting of a quadratic curve through three points $\{(x_0, f(x_0), (x_1, f(x_1)), (x_2, f(x_2)))\}$. The formula allows one to find for any point x the corresponding y value on the curve $y = \ell(x)$.

2. *Least Squares* fits a curve

$$Y = a\,\theta_1(x) + b\,\theta_2(x) = a\,\theta_1 + b\,\theta_2$$

to a data set $\{(X_i, Y_i) : i = 1, \ldots, N\}$ such that the sum of the squared deviations:

$$\sum_i \delta_i^2 = \sum_i |Y_i - (a\,\theta_1(X_i) + b\,\theta_2(X_i))|^2$$

is minimised. These constants a & b are found by solving the *normal equations*:

$$\sum \theta_1 Y = a \sum \theta_1^2 + b \sum \theta_1\theta_2$$
$$\sum \theta_2 Y = a \sum \theta_1\theta_2 + b \sum \theta_2^2.$$

3. *Pearson's Correlation Coefficient, r* gives a measure of the goodness of fit of data to a curve (in the least squares sense). The closer the value to ± 1 the better the fit.

4. *Non-Linear Least Squares* is fitting a not-necessarily-linear curve $Y = f(X)$ to data $\{(X_i, Y_i) : i = 1, \ldots, N\}$. To fit a non-linear curve to data we do the following:

   (a) Write the non-linear relationship in linear form $y = mx + c$ identifying y, m, x and c.

   (b) Tabulate the 'new' x and y values.

   (c) Apply the Linear Least Squares Method; i.e.

$$\sum y = m \sum x + \sum c$$
$$\sum xy = m \sum x^2 + c \sum x$$

   (d) Convert back to the original constants if needs be.

Chapter 2

Static Beam Differential Equations

What is the origin of the urge, the fascination that drives physicists, mathematicians, and presumably other scientists as well? Psychoanalysis suggests that it is sexual curiosity. You start by asking where little babies come from, one thing leads to another, and you find yourself preparing nitroglycerine or solving differential equations. This explanation is somewhat irritating, and therefore probably basically correct.

David Ruelle

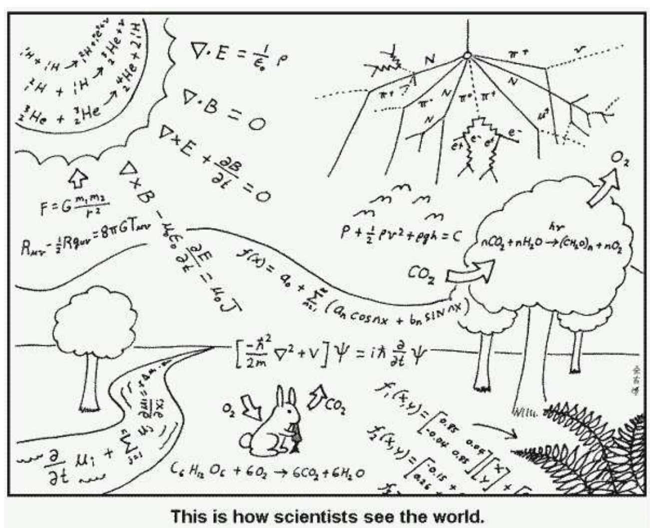


Figure 2.1: Pretty much all of your formulae come from solving differential equations.

2.0.1 Motivation

Here is a screenshot from a YouTuber *Engineer4Free*:

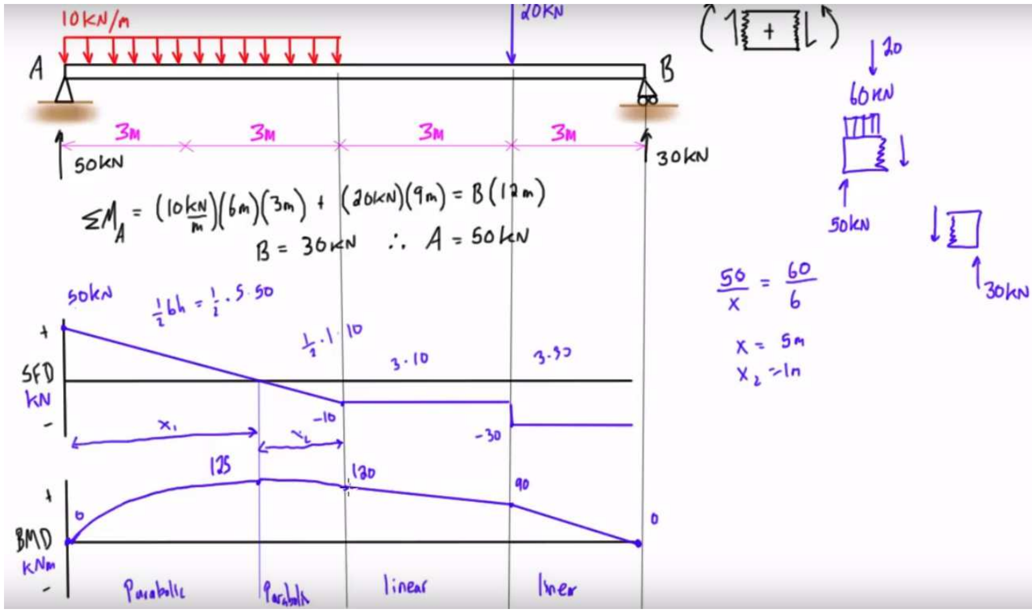


Figure 2.2: This approach gives you a *qualitative* understanding of the bending moment but not enough *quantitative* information.

In this chapter we will learn how to, given a load, find *formulas* for the shear and bending moment in terms of the distance from the support A. We will learn the theory of how to do this for a great many different types of loads — not just udls, point loads, and linear loads. In the final section of this chapter, which really belongs to the final chapter, we learn how to generate approximate solutions when the methods of this chapter fail.

Calculus Review III: Differentiation and Antidifferentiation

At the moment all we need to know about differentiation is

1. **Definition** Let $a_n, a_{n-1}, \dots, a_2, a_1, a_0 \in \mathbb{R}$. A *polynomial of degree* n is a function of the form

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_2 x^2 + a_1 x + a_0.$$

   As example of a polynomial is $f(x) = x^5 - 3x^3 + 2x^2 + 1$.

2. **Main Idea/Properties** A polynomial of degree n has n roots (solutions of $f(x) = 0$) — some of which may be complex, some of which may be repeated. In general, a degree n polynomial has n roots — it is equal to 0 for n different values of x. For example, the example of the polynomial above is equal to zero five times.

   However if all the roots are real and distinct then the polynomial cuts the x-axis n times. The derivative of a polynomial of degree n is a polynomial of degree $n-1$;

$$\text{e.g.}\quad \frac{d}{dx}(x^5 - 3x^3 + 2x^2 + 1) = 5x^4 - 3(3x^2) + 2(2x) = 5x^4 - 9x^2 + 4x,$$

   which has potentially $n-1$ real roots and hence $n-1$ points where $f'(x) = 0$ — potentially $n-1$ turning points.

As an example note that quadratics are degree two polynomials and have one turning point.

3. **Derivative** We differentiate a polynomial using the Sum, Scalar & Power Rules:

$$\frac{d}{dx}(ax^n) = a\frac{d}{dx}x^n = a(nx^{n-1}) = anx^{n-1}.$$

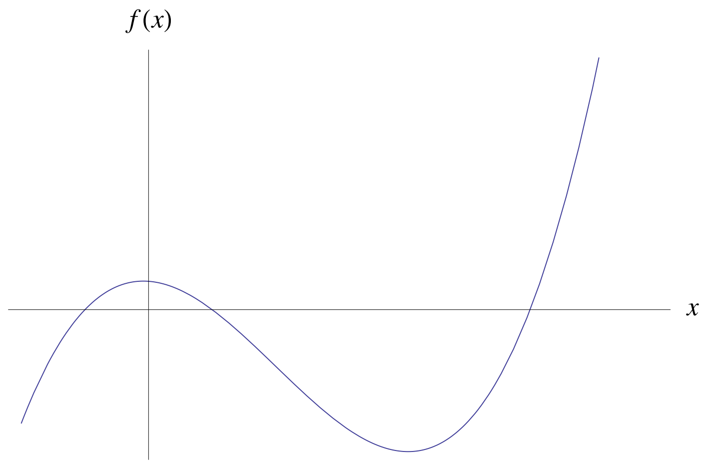


Figure 2.3: This is an example of a cubic: $ax^3 + bx^2 + cx + d$. Note that it has *three* real roots and *two* turning points because $f'(x) = \text{slope} =\sim x^2$. In some sense this is typical behaviour of polynomials.

Given $F(x)$, we know how to compute its derivative $F'(x)$. Now consider the inverse problem: given the derivative of a function, find the function itself:

Examples

1. Suppose that $f(x) = \cos x$. Find an antiderivative of $f(x)$.

   Solution:

   This example is typical. Once we have found one particular antiderivative $F(x)$ of a function $f(x)$, then all antiderivatives of $f(x)$ are given by the formula $F(x) + C$, where C is an arbitrary constant.

2. Suppose that $f(x) = x^5$. Final *all* antiderivatives of f.

   Solution:

   If we are given some extra numerical information, this will pin down the value of C.

3. A curve satisfies $\dfrac{dy}{dx} = 3x^2$ and passes through the point $(2,\ 5)$[1]. Find the equation of the curve.

   Solution: A particular antiderivative of $3x^2$ is x^3. Hence all functions of the form $x^3 + C$ are anti-derivatives.

Proposition

Suppose that u and v have anti-derivatives and that $k \in \mathbb{R}$. *Then we have*

1. $$\int ku(x)\,dx = k\int u(x)\,dx.$$

2. $$\int [u(x) \pm v(x)]\,dx = \int u(x)\,dx \pm \int v(x)\,dx.$$

Proof. Follows easily from the scalar and sum rules for differentiation •

Proposition

Suppose that $n \in \mathbb{Q}$, $n \neq -1$. *Then*

$$\int x^n\,dx = \frac{x^{n+1}}{n+1} + C.$$

Proof. Simply differentiate the right-hand side •

Examples

1. Find
   $$\int (x^4 + 4x^3)\,dx.$$

   Solution:

2. Find
   $$\int (x^2 - 4x + 2)\,dx.$$

   Solution:

[1] recall that we write the x-coordinate first and then the y coordinate. Therefore this $(2,5)$ means that $y = 5$ when $x = 2$. This might also we written as $y(2) = 5$ from the $y = y(x)$ notation — the y-value at $x = 2$ is 5.

Theorem

Let $F(x)$ be an antiderivative of $f(x)$. *Then every antiderivative of f has the form $F(x) + C$ for some constant C.*

Remark

This is a geometrically obvious fact:

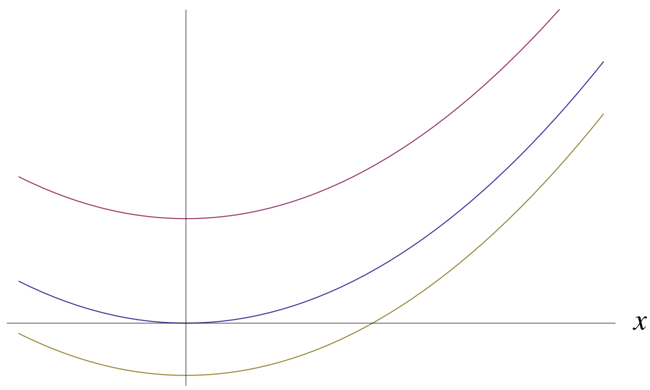


Figure 2.4: Shifting a graph up or down does not change its slope (here we have x^2, x^2+2, x^2-1). Shifting a graph up or down is equivalent to adding a (positive or negative) constant.

Notation: the *antiderivative* $\int f(x)\,dx$ means the antiderivatives[2] of $f(x)$.

Separable Differential Equations

A differential equation is an equation containing a function $y = f(x)$ and one or more derivatives, e.g.,

$$y' = x^2,$$
$$\frac{d^2y}{dx^2} + x\frac{dy}{dx} = \sin x.$$

Note that the solutions are *functions* — not numbers. Note the notation $y' = \dfrac{dy}{dx}$. Most laws in physics and engineering are differential equations.

The *order* of a differential equation is the order of the highest derivative that appears;

$$y' = x^2 \quad \text{is first order}$$
$$\frac{d^2y}{dx^2} + \left(\frac{dy}{dx}\right)^3 = 1 \quad \text{is second order}$$
$$(\cos x)\left(\frac{d^2y}{dx^2}\right)^3 + \frac{dy}{dx} = y \quad \text{is second order.}$$

As an example, the beam equation is given by

$$EI \cdot \frac{d^4y}{dx^4} = -w(x),$$

[2]written as $F(x) + C$ elsewhere

and so is 4th order. We will also see the following differential equations:

$$\begin{aligned} \frac{dV}{dx} &= -w(x) \\ \frac{d^2M}{dx^2} &= -w(x) \\ EI \cdot \frac{d^2y}{dx^2} &= M(x) \end{aligned}$$

A *separable* first-order differential equation is one that can be written in the form

$$\frac{dy}{dx} = \frac{f(x)}{g(y)}$$

In this situation we can *separate the variables*:

$$g(y)\,dy = f(x)\,dx.$$

Each side can now be anti-differentiated... well hopefully; if not numerical approximations to values of $y(x_i)$ will have to do:

$$\int g(y)\,dy = \int f(x)\,dx.$$

In this module, all of the differential equations are going to be of the form

$$\begin{aligned} \frac{dy}{dx} &= F(x) \\ \Rightarrow dy &= F(x)\,dx \\ \Rightarrow \int dy &= \int F(x)\,dx \\ \Rightarrow y(x) &= \int F(x)\,dx \end{aligned}$$

i.e. so all we actually have to do is anti-differentiate both sides from the start and not worry about separating.

Further Remarks: Picard's Existence Theorem

There is a theorem in the analysis of differential equations which states that if a differential equation is suitably *nice* in an interval about the boundary condition then not only does a solution exist but it is unique. This allows us to define functions as solutions to differential equations. For example, an alternate definition of the exponential function, e^x, is the unique solution to the differential equation:

$$\frac{dy}{dx} = y(x)\,,\ \ y(0) = 1.$$

Exercises

1. Solve $y'(x) = 3x^2 + 2x - 7$ **Ans**: $y(x) = x^3 + x^2 - 7x + C$

2. The point $(3, 2)$ is on a curve, and at any point (x, y) on the curve the tangent line has slope $2x - 3$. Find the equation of the curve. **Ans**: $y(x) = x^2 - 3x + 2$

Figure 2.5: When you differentiate/antidifferentiate $y = e^x$ you get e^x again.

2.1 Impulse & Step Functions

2.1.1 Introduction

In this section we examine *impulse* ($\sim$ point loads) and *step functions* ($\sim$ u.d.l.s). They arise naturally in the theory of beams. They model various *discontinuous* phenomena.

2.1.2 Step Functions

For example, consider a simple switch that has two states: on (1) and off (0). Suppose the switch is turned on at time $x = 0$. We use the *Heaviside Function* to model this:

$$u(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } x \geq 0. \end{cases} \tag{2.1}$$

The u stands for $u$nit step; e.g. $u(2) = 1$ and $u(-2) = 0$.

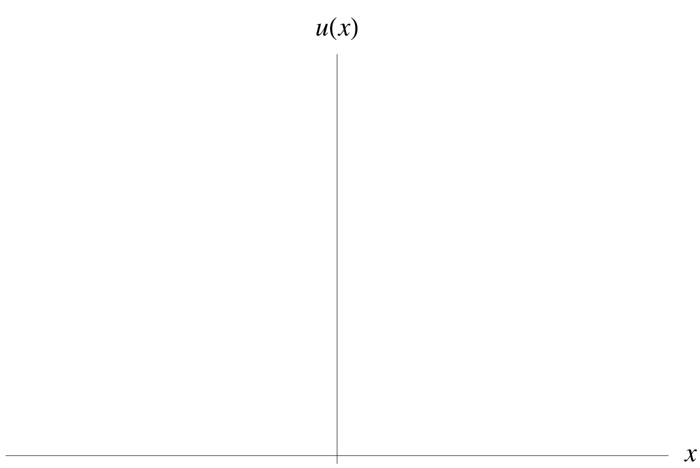


Figure 2.6: The graph of the Heaviside Function.

It is no problem writing down a switch which starts at a time $x = 2$:

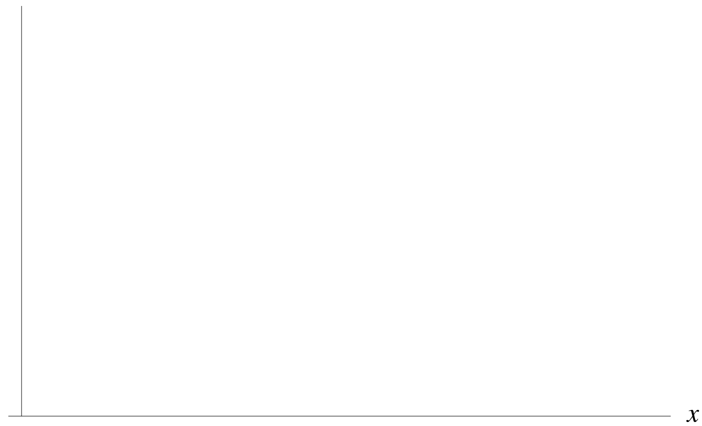


Figure 2.7: The graph of the Heaviside Function $u(x - 2)$.

We can also use a combination of Heaviside functions to write an expression for a switch that is on between times $x = 2$ and $x = 4$:

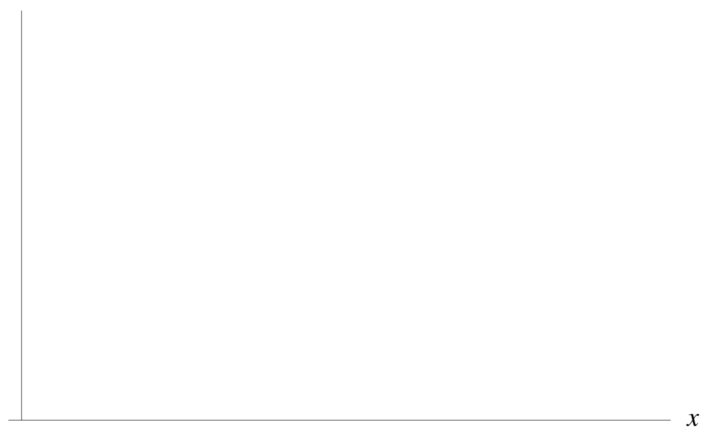


Figure 2.8: This function is equal to $u(x-2)-u(x-4)$. It is an exercise to show that this function is equal to one for $2 \leq x < 4$ and zero elsewhere.

Of course the output value doesn't have to be one, it could be 3 say; or we could have a switch equal to 4 between $x = 2$ and $x = 5$:

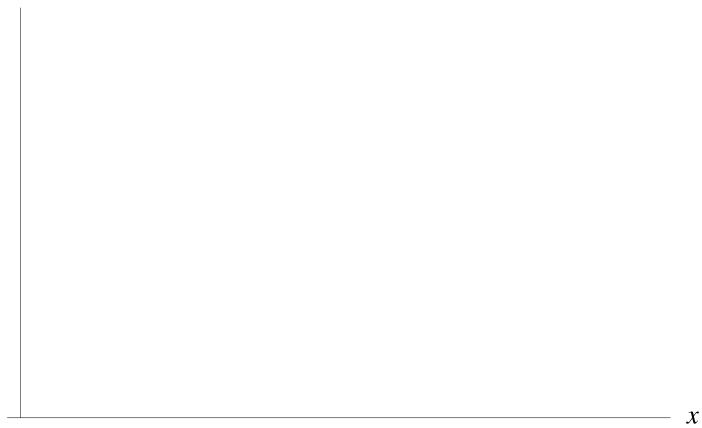


Figure 2.9: The graphs of $3u(x)$ and $4u(x-2) - 4u(x-5)$.

What is the derivative/slope of $u(x)$? The slope is 0 except at $x = 0$ where it is infinite. This is the *Dirac Delta Function* or the *Impulse Function*:

$$\delta(x) = \begin{cases} \text{“}\infty\text{”} & \text{if } x = 0 \\ 0 & \text{otherwise} \end{cases}, \tag{2.2}$$

like $\delta(0) = \text{“}\infty''$ and $\delta(x) = 0$ for $x \neq 0$, so we have

$$\frac{d}{dx}u(x) = \delta(x),$$

and recalling that $\int \cdot \, dx$ is anti-differentiation:

$$\int \delta(x-a)\, dx = u(x-a) + C.$$

There is a more precise definition that can be made such that the antiderivative of $\delta(x)$ is $u(x)$. Using this function we can model a 'jolt' or a kick acting at the time $t = a$ or a point force acting at position $x = a$ of magnitude F is given by $F \cdot \delta(x-a)$.

We can also anti-differentiate the Heaviside function (this is what we will need to do). First some notation:

2.1.3 Notation: Macauley Bracket

$$[x] := x \cdot u(x) = \begin{cases} x & x \geq 0 \\ 0 & x < 0 \end{cases} \tag{2.3}$$

Note now that $[f(x)]$ is a function whose value if $f(x)$ when $f(x) \geq 0$ and zero otherwise. So, for example $[+2] = 2$ and $[-2] = 0$. In essence, $[x]$ is the function which 'chops off' the negative part of the graph of $f(x)$.

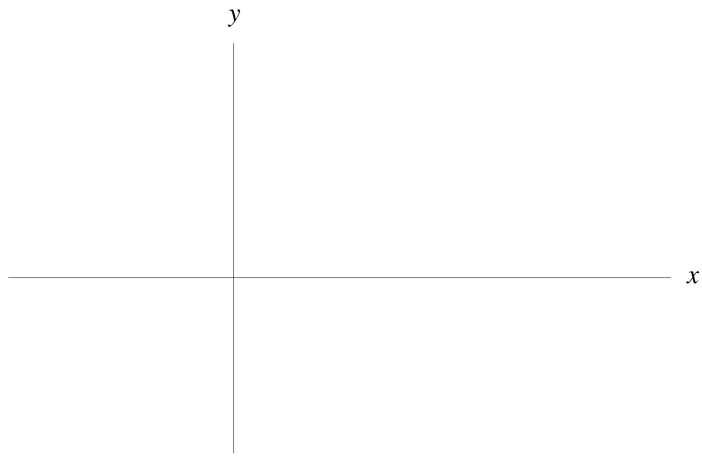


Figure 2.10: The graph of a function $f(x)$ and $[f(x)]$.

2.1.4 Anti-Differentiation of $u(x)$

The antiderivative of $u(x)$ is equal to $[x] = x \cdot u(x)$, that is

$$\int u(x-a)\,dx = [x-a] + C.$$

Proof. We have that

$$[x-a] = \begin{cases} x-a & \text{if } x \geq a \\ 0 \text{ if } x < a \end{cases}$$
$$\Rightarrow \frac{d}{dx}[x-a] = \begin{cases} 1 & \text{if } x \geq a \\ 0 \text{ if } x < a \end{cases}$$
$$= u(x-a) \qquad \bullet$$

2.1.5 Antiderivative of $[x-a]$

The antiderivative of $[x-a]$ is equal to $\dfrac{[x-a]^2}{2}$.

Proof. We have that

$$\frac{[x-a]^2}{2} = \begin{cases} \dfrac{x^2-2ax+a^2}{2} & \text{if } x \geq a \\ 0 \text{ if } x < 0 \end{cases}$$
$$\Rightarrow \frac{d}{dx}\frac{[x-a]^2}{2} = \begin{cases} x-a & \text{if } x \geq a \\ 0 \text{ if } x < 0 \end{cases}$$
$$= [x-a] \qquad \bullet$$

Similarly we can show that

$$\int [x-a]^n\,dx = \frac{[x-a]^{n+1}}{n+1} + C, \tag{2.4}$$

which is very similar to

$$\int x^n \, dx = \frac{x^{n+1}}{n+1} + C.$$

Examples

Solve the following differential equations. This means find $y = y(x)$.

1. $\dfrac{dy}{dx} = 2u(x) - 2u(x-1)$.
2. $\dfrac{d^2y}{dx^2} = 3u(x-2) + \delta(x-1)$.
3. $\dfrac{d^2y}{dx^2} = [x-2]^2 - [x-4]^2$

Solution: We simply anti-differentiate the appropriate number of times in each case:

1. $\Rightarrow$

2. $\Rightarrow$

3. $\Rightarrow$

Exercises:

1. Find the general solution of the following second order differential equations

   (a) $\dfrac{d^2y}{dx^2} = 6x + 2$ **Ans:** $y(x) = x^3 + x^2 + C_1x + C_2$.

   (b) $\dfrac{d^2M}{dx^2} = -6$ **Ans:** $M(x) = -3x^2 + C_1x + C_2$.

   (c) $EI\dfrac{d^2y}{dx^2} = -60[x - 3]$ **Ans:** $y(x) = \dfrac{1}{EI}\left(-10[x - 3]^3 + C_1x + C_2\right)$.

2. Find the general solution of the following second order separable differential equations:

   (a) $\dfrac{d^2M}{dx^2} = -\delta(x - 1) - \delta(x - 5)$ **Ans:** $M(x) = -[x - 1] - [x - 5] + C_1x + C_2$.

   (b) $\dfrac{d^2M}{dx^2} = -18$ **Ans:** $M(x) = -9x^2 + C_1x + C_2$.

   (c) $\dfrac{d^2M}{dx^2} = -36u(x - 1) + 36u(x - 5) - 5\delta(x - 4)$
   Ans: $M(x) = -18[x - 1]^2 + 18[x - 5]^2 - 5[x - 4] + C_1x + C_2$.

   (d) $\dfrac{d^2M}{dx^2} = -\delta(x - 1) - \delta(x - 3) - 144u(x - 5) + 144u(x - 8) - 3\delta(x - 7)$
   Ans: $M(x) = -[x - 1] - [x - 3] - 72[x - 5]^2 + 72[x - 8]^2 - 3[x - 7] + C_1x + C_2$.

   (e) $\dfrac{d^2M}{dx^2} = -x - 2 - 10\delta(x - 2)$ **Ans:** $M(x) = -\dfrac{x^3}{6} - x^2 - 10[x - 2] + C_1x + C_2$.

   (f) $\dfrac{d^2M}{dx^2} = -3x - 3 - 72u(x - 2)$ **Ans:** $M(x) = -\dfrac{x^3}{2} - \dfrac{3}{2}x^2 - 36[x - 2]^2 + C_1x + C_2$.

3. Now find the particular solutions to the first two differential equations above under the following initial conditions.

   (a) $M'(0) = 1$ and $M(0) = 0$. **Ans:** $M(x) = -[x - 1] - [x - 5] + x$.

   (b) $M'(0) = 54$ and $M(0) = 0$. **Ans:** $M(x) = -9x^2 + 54x$.

   Note after we do the next section we can look back and see that these are the beam equations for:

   (a) A simply supported beam of length 6 m with point loads of magnitude 1 kN at $x = 1$ and $x = 5$.

   (b) A simply supported beam of length 6 m with a uniform load of 18 kN m^{-1}.

2.2 Applications to Beams

2.2.1 Six Variables

1. distance along the beam, x:

2. load per unit length, w, at a point:

3. shear force, V, at a point x_0 is the sum of the loads from $x = 0$ to $x = x_0$:

4. the bending moment, M, at a point x_0 is the sum of the moments from $x = 0$ to $x = x_0$.

5. the slope, y', at a point:

6. the deflection, y, at a point x_0.

In this section we learn how to formulate and solve beam equations so that we may calculate/ estimate the deflection of a beam due to the loads on it. This theory known variously as the *Euler-Bernoulli Beam Theory*, *Engineer's Beam Theory* or *Classical Beam Theory*. It is a simplification of the theory of elasticity which, after it was used in the design of the Eiffel Tower and Ferris wheels, became a cornerstone of engineering.

The theory makes a number of underlying assumptions. First we must show a beam in a deformed state and an undeformed state:

The beams for which we apply the model are assumed to be:

1. slender: their length is much greater than their width
2. isotropic and homogeneous: the material behaves the same at all points in the beam and in all directions
3. constant cross-section

The assumptions of the model (Kirchhoff's Assumptions) are:

1. Normals remain straight: they do not bend
2. Normals remain unstretched: they do not change length
3. Normals remain normal: they stay perpendicular to the neutral plane

Now using these assumptions we can derive an equation relating the *deflection* $y(x)$ at a point x and the *load per unit length* $w(x)$. This equation (whose derivation is outside the remit of our course. If you really want to see how they are derived Google 'Spivak physics pdf' and have a look at Lecture 8) is given by:

where E is Young's Modulus and I is the second moment of area of the beam. Note that in all cases we use kilo-Newtons (kN) rather than Newtons (N). Solving this equation requires *four* anti-differentiations.

The bending moment, $M(x)$, and the shearing force, $V(x)$, at a distance x from the left-hand side of a beam are related to the load per unit length $w(x)$ and the shearing force $V(x)$ by the differential equations

$$\frac{dM}{dx} = V(x), \quad \frac{dV}{dx} = -w(x).$$

Definition

The *shearing force* at the point x is the sum of the forces at or to the left of the point x.

Ye probably know more about shearing than me — either now or after doing structural stuff next semester — and we just need this working definition. These can be combined to form the second order differential equation

$$\frac{d^2M}{dx^2} = -w(x), \tag{2.5}$$

which implies that

$$EI\frac{d^2y}{dx^2} = +M(x).$$

Hence if we know the load we can calculate the bending moment, $M(x)$. Then we can appeal to $EI\,y''(x) = M(x)$ to solve for the deflection.

Remark: Sign Conventions

Later on, in accordance with J.J. Murphy's work, whenever I plot the bending moment, $M(x)$, I will plot it as negative (below the x-axis). J.J. tells me that "*the bending moments is plotted below the x-axis, on the tension face of the beam. This is to coincide with the placing of main reinforcement in reinforced concrete beams (as concrete is weak in tension, reinforcement is principally required on the 'tension face').*"

We have four main types of loads:

1. *Uniformly distributed loads, U.D.L. over the entire beam*

   Mathematically we have

   $$w_{\text{UDL}}(x) = w_0 \tag{2.6}$$

   where w_0 is the constant value of the UDL; e.g. $w(x) = 18$ kilonewtons per metre $= 18$ kN m^{-1}

   Example

   Find the bending moment due to a U.D.L. of 36 kN m^{-1} across a beam of length 7 m where $M(0) = 0$ and $M'(0) = 126$.

   Solution: With a picture we see $w(x) = 36$ a constant. We anti-differentiate twice:

   Now we apply the boundary conditions. First $M(0) = 0$:

   $$M(0) = 0 = C_2.$$

   Now $M'(0) = 126$:

   $$M'(0) = C_1 = 126.$$

   So the answer is

   $$M(x) = -18x^2 + 126x.$$

2. *U.D.L. over a segment of the beam.*

   Where w_0 is the magnitude of the UDL across $a \leq x \leq b$:

   $$w_{\text{UDL}\,:x=a\to b}(x) = w_0[u(x-a) - u(x-b)] = w_0 u(x-a) - w_0 u(x-b). \tag{2.7}$$

Example

i. Write down the load, $w(x)$, due to a UDL of 18 kN m^{-1} from $x = 2$ to $x = 3$ on a beam of length 5 m.

ii. Now find the bending moment due to this load. Write your answer in terms of M_A, the bending moment at $x = 0$, and R_A the reaction/shearing force at $x = 0$.

Solution: We are going to solve the differential equation $\frac{d^2M}{dx^2} = -w(x)$ for the bending moment $M(x)$, so first draw a picture so that we can see what the loading, $w(x)$, is:

Now write down the bending moment equation, $M''(x) = -w(x)$, and anti-differentiate twice:

Now apply the boundary conditions — that is $M(0) = M_A$, the bending moment at A, and $M'(0) = R_A$, the shearing/reaction force at A — recall that $M'(x)$ is the shearing force:

Therefore the answer is

$$M(x) = -9[x - 2]^2 + 9[x - 3]^2 + R_A x + M_A.$$

Remark

Note that when a UDL continues up to $x = L$ m, there is absolutely no need to 'turn off' the UDL. For example, consider a beam of length 6 m subject to a UDL of 32 kN m^{-1} between $x = 4$ m and $x = 6$ m:

Whether you take

$$w(x) = 32u(x-4) \text{ or } w(x) = 32u(x-4) - 32u(x-6),$$

you will get the same answer. This is because $u(x-6) = 0$ for $x < 6$; i.e. points on the beam. Therefore you are better off not 'turning these off'. If you turn it off your answer will still be correct but you will have an extra, unnecessary term in your answer.

3. *Point Loads*

Mathematically we have, for a point load of weight w_0 at a point $x = a$:

$$w(x) = w_0 \cdot \delta(x-a) \tag{2.8}$$

Examples

(a) i. Write down the load per unit length due to a point mass of 13 kN at $x = 3$ on a beam of length 6 m.

ii. Now solve $M''(x) = -w(x)$ to find the bending moment.

Solution: We simply have $w(x) = 13\,\delta(x-3)$ so we have the bending moment equation:

When solving $M''(x) = -w(x)$, we will always have the first constant of integration here equal to R_A and the second equal to M_A. Although this example doesn't ask for it, we can nearly always give our answer in terms of R_A and M_A:

$$M(x) = -13[x-3] + R_A x + M_A$$

(b) *Linear loads over the entire beam*

Here we have $w(0) = w_A$ and $w(L) = w_B$. It is not hard to show that a line has equation

$$\underbrace{\text{OUTPUT}}_{y} = \underbrace{(\text{SLOPE} \times \text{INPUT})}_{mx} + \underbrace{Y\text{-INTERCEPT}}_{c}.$$

The slope is the ratio of how much you go up as you go across (rise/run):

$$\text{SLOPE} = \frac{w_B - w_A}{L} \tag{2.9}$$

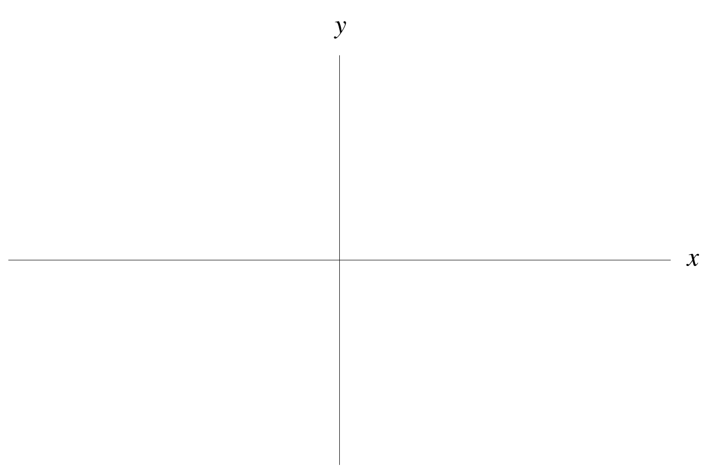


Figure 2.11: The curve $y = mx + c$ is a line of slope m and y-intercept c.

Therefore we have

$$w(x) = \frac{w_B - w_A}{L}x + w_A \tag{2.10}$$

(c) i. Write down the load on a 8 m beam due to a linear load that varies from 10 kN m^{-1} at $x = 0$ to 26 kN m^{-1} at $x = 8$.

ii. Hence solve $M''(x) = -w(x)$ to find the bending moment.

Solution: First we draw:

From this we see that the slope is two and the 'y-intercept' is 10 so we have that $w(x) = 2x + 10$ so we have the bending moment equation:

Remark

In fact, in MATH7019 where there will never be a point load at A, when solving the differential equation

$$\frac{d^2M}{dx^2} = -w(x),$$

the first constant of integration is *always*[3] R_A and the second constant of integration is *always* M_A.

Further Remark

We prove this here. Starting with $M''(x) = -w(x)$, integrate with respect to t from $t = 0$ to $t = x$:

$$\begin{aligned} \int_0^x \frac{d^2M}{dx^2}(t)\,dt &= -\int_0^x w(t)\,dt \\ \Rightarrow \frac{dM}{dx} &= -\int_0^x w(t)\,dt + C_1. \end{aligned}$$

However $M'(x) = V(x)$ and letting $x = 0$ on both sides gives $C_1 = V(0)$ which is nothing but the shear at A i.e. $C_1 = R_A$. A similar calculation shows that $C_2 = M(0) = M_A$.
Exercises For each of the following:

- Draw a picture.
- Write down $w(x)$ (summary on p. 109).
- Write down $M''(x) = -w(x)$.
- Anti-differentiate twice.

[3]this won't happen in Chapter 4 questions either, but if the 'primitive' antiderivative of $w(x)$ is non-zero at zero then the first constant is not equal to R_A, e.g. with $w(x) = e^x$ or $\cos x$.

What form does the bending moment take for a beam of span 5 m where there is

1. a U.D.L. of 12 kN m^{-1} between $x = 2$ and $x = 5$. **Ans:** $M(x) = -6[x-2]^2 + R_A x + M_A$.
2. there is a point load of 10 kN at $x = 2$. **Ans:** $M(x) = -10[x-2] + R_A x + M_A$.
3. a linear load varying from 8 kN m^{-1} at $x = 0$ to 18 kN m^{-1} at $x = 5$. **Ans:** $M(x) = -\frac{x^3}{3} - 4x^2 + R_A x + M_A$.
4. there are point loads of 10 kN and 12 kN at $x = 2$ and $x = 3$ respectively. **Ans:** $M(x) = -10[x-2] - 12[x-3] + R_A x + M_A$.
5. there is a point load of 12 kN at $x = 3$ and a U.D.L. of 10 kN m^{-1} between $x = 2$ and $x = 5$. **Ans:** $M(x) = -12[x-3] - 5[x-2]^2 + R_A x + M_A$.

2.2.2 Further Remark: Macauley's Method

Note that if you are sure what you are doing you can write down the bending moment straight away. Personally I would prefer to solve the $M''(x) = -w(x)$ equation but this is an option for you. A very valid question is: why am I showing you how to solve these beam problems from first principles when ye can just look up a table? I have a friend who did Structural Engineering who works for Liebherr. He told me that if he were to solve problems from first principles in his office that he would be told where to go. However, he said that the people who write the computer programmes that they use would know the theory behind all the tables and that these were the people on the big bucks.

If you are going to use Macauley's Method I think we should show the following to be sure about what you are doing.

Macauley's Method

When solving the differential equation

$$\frac{d^2M}{dx^2} = -w(x), \tag{2.11}$$

the following hold:

1. **If the bending moment due to a load $w_1(x)$ is $M_1(x)$, and the bending moment due to a load $w_2(x)$ is $M_2(x)$,** *then the bending moment due to the load $w_1(x) + w_2(x)$ is given by $M_1(x) + M_2(x)$.*
2. *The bending moment due to a U.D.L. of w_0 kN m^{-1} is given by*

$$M(x) = -\frac{w_0}{2}x^2 + R_A x + M_A, \tag{2.12}$$

where R_A is the reaction or shearing force at $x = 0$ and M_A is the bending moment at $x = 0$.

3. *The bending moment due to a U.D.L. of w_0 kN m^{-1} applied between points $x = a$ and $x = b$ (with $a < b$) is given by*

$$-\frac{w_0}{2}[x-a]^2 + \frac{w_0}{2}[x-b]^2 + R_A x + M_A \tag{2.13}$$

4. *The bending moment due to a point load of magnitude w_0 kN at $x = a$ is given by*

$$-w_0[x-a] + R_A x + M_A \tag{2.14}$$

5. *The bending moment due to a linear load varying from w_A at $x = 0$ to w_B at $x = L$ is given by*

$$-\frac{(w_B - w_A)x^3}{6L} - \frac{w_A x^2}{2} + R_A x + M_A \tag{2.15}$$

Proof:

1. This follows from the fact that anti-differentiation is additive; e.g. $\int (u(x) + v(x))\, dx = \int u(x)\, dx + \int v(x)\, dx$.

2. The load is given by $w(x) = w_0$, a constant and so we solve:

$$\begin{aligned} \frac{d^2M}{dx^2} &= -w_0 \\ \Rightarrow \frac{dM}{dx} &= -w_0 x + R_A \\ \Rightarrow M(x) &= -\frac{w_0}{2}x^2 + R_A x + M_A. \end{aligned}$$

3. The load is of the form $w(x) = w_0 u(x-a) - w_0 u(x-b)$. Therefore we anti-differentiate twice to find the bending moment:

$$\begin{aligned} \frac{d^2M}{dx^2} &= -w_0 \cdot u(x-a) + w_0 \cdot u(x-b) \\ \Rightarrow \frac{dM}{dx} &= -w_0[x-a] + w_0[x-b] + R_A \\ \Rightarrow M(x) &= -\frac{w_0}{2}[x-a]^2 + \frac{w_0}{2}[x-b]^2 + R_A x + M_A. \end{aligned}$$

4. The load is equal to $w(x) = w_0\delta(x-a)$. Anti-differentiate twice to find the bending moment:

$$\begin{aligned} \frac{d^2M}{dx^2} &= -w_0 \cdot \delta(x-a) \\ \Rightarrow \frac{dM}{dx} &= -w_0 \cdot u(x-a) + R_A \\ \Rightarrow M(x) &= -w_0 \cdot [x-a] + R_A x + M_A. \end{aligned}$$

5. Left as an exercise.

The three different types of beams that we look at are simply supported beams, fixed end beams and cantilever beams. They differ mathematically only in terms of the boundary conditions. The master equation involves four anti-differentiations so we must have four boundary conditions for each.

Calculus Review IV: Quadratics/Parabolas

This section is necessary because if we have a linear load, $w(x) = mx + c$, then $V(x) = M'(x)$ will be a parabola, $ax^2 + bx + c$.

1. **Definition** Let $a, b, c \in \mathbb{R}$. A quadratic is a function of the form
$$f(x) = ax^2 + bx + c.$$
An example of a quadratic is $f(x) = x^2 + 1$.

2. **Main Idea/Properties** A quadratic either has a $\bigcup$ shape (when $a > 0$) or a $\bigcap$ shape (when $a < 0$). It has two *roots* given by the $\dfrac{-b \pm \sqrt{\cdots}}{2a}$ formula. If they are both *real* (when $b^2 - 4ac > 0$), then the graph cuts the x-axis at two points. The graph is symmetric about the max/min. Hence the max/min can be found by looking at $f'(x) = 0$ or else be found at the midpoint of the roots. If $b^2 - 4ac < 0$ then the roots contain a $\sqrt{(-)}$ — *complex roots.*

3. **Derivative** The derivative of a quadratic is a line!
$$\frac{d}{dx}(ax^2 + bx + c) = a(2x) + b(1) + 0 = 2ax + b.$$

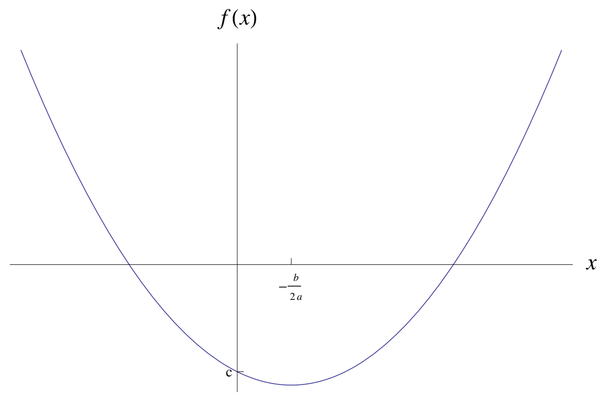


Figure 2.12: The graph of a $+x^2$ quadratic with $a > 0$. Note that the slope goes from negative to zero to positive — like a line. This quadratic has two real roots — where the graph cuts the x-axis — and the minimum occurs at $-\dfrac{b}{2a}$. At this point the tangent is horizontal. This point can be found by differentiating $ax^2 + bx + c$, e.g. getting the slope, and setting it equal to zero.

We need to be able to find the roots of a quadratic $f(x) = ax^2 + bx + c$. In MATH7019 the quadratic will usually not have simple roots and then the usual technique for finding the roots — factorising — will not suffice. There is a formula for these cases. Of course the formula may always be used to extract the two roots but as you might see in MATH7021 if at all possible you should factorise.

Only Two Roots???

In each case we have seen that there are two roots — is this always going to be the case?

Proposition

Suppose *$f(x) = ax^2 + bx + c$. Then $f(x)$ may be re-written as:*
$$f(x) = a\left(x + \frac{b}{2a}\right)^2 + \left(c - \frac{b^2}{4a}\right).$$

Proof. Multiply out •

Remark

It can be shown that because of this, the graph of $f(x)$ is similar to the graph of $f(x) = x^2$ (it is got by translating and stretching the axes.):

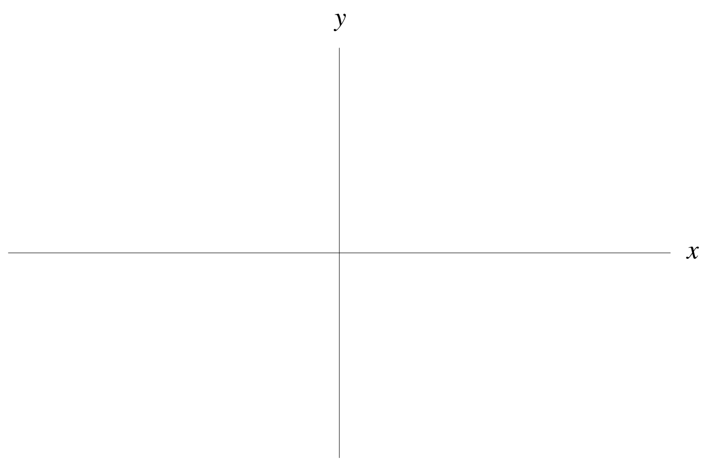


Figure 2.13: Quadratic functions look like x^2, $\bigcup$ — hence they can only cut the x-axis at most twice.

In fact we can go further and solve this for x.

Proposition

Suppose $f(x) = ax^2 + bx + c$**.** *Then the roots of* $f(x)$ *are given by:*

$$x_\pm = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}. \tag{2.16}$$

Proof. Solve

$$a\left(x + \frac{b}{2a}\right)^2 + \left(c - \frac{b^2}{4a}\right) = 0.$$

Remark

One of the most annoying things that I see is people not writing down the formula properly. Note the following are, for example, *NOT* the correct formula:

$$x_\pm = -b \pm \frac{\sqrt{b^2 - 4ac}}{2a}$$

$$x_\pm = \frac{-b}{2a} \pm \sqrt{b^2 - 4ac}$$

$$x_\pm = -b \pm \sqrt{\frac{b^2 - 4ac}{2a}}$$

$$x_\pm = \frac{-b \pm \sqrt{(b^2 - 4)ac}}{2a}.$$

If you are using a calculator, for say $x^2 + 7x - 11 = 0$, to take your time like this:

$$\begin{aligned} x_{\pm} &= \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \\ &= \frac{-(7) \pm \sqrt{(7)^2 - 4(1)(-11)}}{2(a)} = \frac{-7 \pm \sqrt{93}}{2} \\ &= \frac{-7}{2} \pm \frac{\sqrt{93}}{2} \approx -8.322,\ 1.322. \end{aligned}$$

In other words remember BEDMAS.

The other problem I see is people doing stuff like

$$\sqrt{-81} = -\sqrt{81} = -9,$$

this is more rubbish: whenever you have $\sqrt{\text{a negative number}}$ you have no (real) roots.

2.2.3 Simply Supported Beams

A simply supported beam looks as follows:

Figure 2.14: For simply supported beams, there is nothing keeping the beam 'straight' at the supports... the bending moment M is zero at the ends. Of course, the deflection y is zero at the ends.

We have the following *four* boundary conditions:

1. the bending moment at each of the ends is zero: i.e. $M(0) = 0$ and $M(L) = 0$.
2. the deflection at both ends are zero: i.e. $y(0) = 0$ and $y(L) = 0$.

In addition, if the load is symmetric about the centre $x = L/2$ then we also have $R_A = R_B = W_T/2$, where W_T is the total load.

Examples

1. A light beam of span 5 m is *simply supported* at its end points and carries a uniformly distributed load (U.D.L.) of 18 kN m^{-1} along the beam. By solving the differential equation

$$\frac{d^2M}{dx^2} = -w(x).$$

find the Bending Moment at any point along the beam.

Solution: First a sketch:

We have that $w(x) = 18$, and so we anti-differentiate twice to find the general solution:

Now we have to apply some boundary conditions. The bending moment about $x = 0$:

$$M(0) = -9(0)^2 + R_A(0) + M_A = 0 \Rightarrow M_A = 0$$

This will always be the case for Simply Supported Beams. We also know that $M(5) = 0$. Hence we apply this boundary condition:

Note that we could have also seen that the load is symmetric so that the reaction forces at A and B are half the load each. The total load $5 \times 18 = 90$ so $R_A = 45 = R_B$. Either way, this yields the solution:

$$M(x) = -9x^2 + 45x \tag{2.17}$$

Here I plot the load, the shearing, and the bending moment on the one graph:

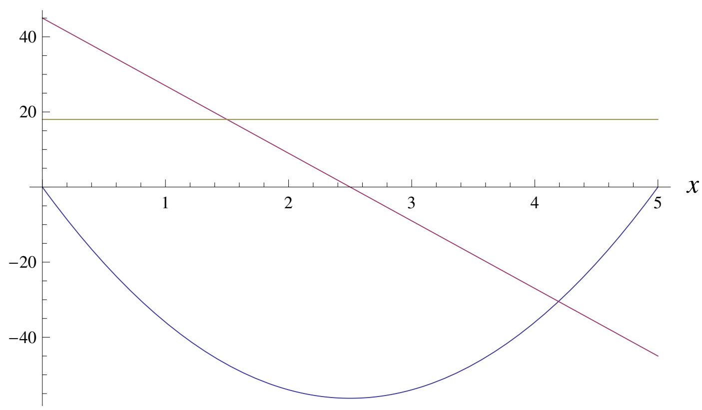


Figure 2.15: A plot of the load, shearing and the bending moment. Recall that the bending moment has been plotted below the x-axis even though it is positive.

2. **Winter 2012** A light beam of span 4 m is simply supported at its endpoints. At the point $x = 1$ m there is a load of 36 kN. Between the points $x = 2$ m and $x = 4$ m there is a U.D.L. of 30 kN m^{-1}.

   (i) Solve $M''(x) = -w(x)$ to find the Bending Moment $M(x)$ in terms of step functions ($u(x-a)$ and $[x-a]$).

   (ii) The deflection y at any point on the beam is found by solving the differential equation

$$EI\frac{d^2y}{dx^2} = M(x).$$

   Solve the differential equation.

   (iii) What is the deflection of the beam at $x = 3$ m?

[1 Mark]

Solution: First we draw a picture:

The load per unit length is:

As the UDL goes to the end of the beam there is no need to include $-30u(x-4)$ as this term is zero for any point on the beam (i.e. any point of interest).

(i) To find the bending moment we solve the differential equation

To solve this we anti-differentiate twice

Now because we are simply supported we know that the bending moment at $x = 0$ and $x = 4$ m is zero. The first boundary condition implies that $M_A = 0$. Now we apply $M(4) = 0$:

so we have $R_A = 42$ so

$$M(x) = -36[x-1] - 15[x-2]^2 + 42x.$$

(ii) We have

$$EI\frac{d^2y}{dx^2} = -36[x-1] - 15[x-2]^2 + 42x.$$

To solve this we anti-differentiate twice:

To find C_1 and C_2 we must apply the boundary conditions that $y = 0$ at $x = 0$ and $x = 4$. $y(0) = 0$ implies that $C_2 = 0$. Now we apply $y(4) = 0$:

(iii) This is just $y(3)$:

3. **Winter 2011** A light beam of span 5 m is simply supported[4] at its end points and carries a load that varies uniformly with x the distance from one end of the beam. The load varies from 18 kN m^{-1} at $x = 0$ to 12 kN m^{-1} at $x = 5$.

   i. Find a formula for the load per unit length.

   ii. By solving the differential equation

$$\frac{d^2M}{dx^2} = -w(x), \tag{2.18}$$

   find the bending moment M at any point along the beam.

   iii. Also find the maximum value of the bending moment.

   Solution: First a picture:

   Now $-w(x) = \frac{6}{5}x - 18$. We anti-differentiate twice to find the bending moment:

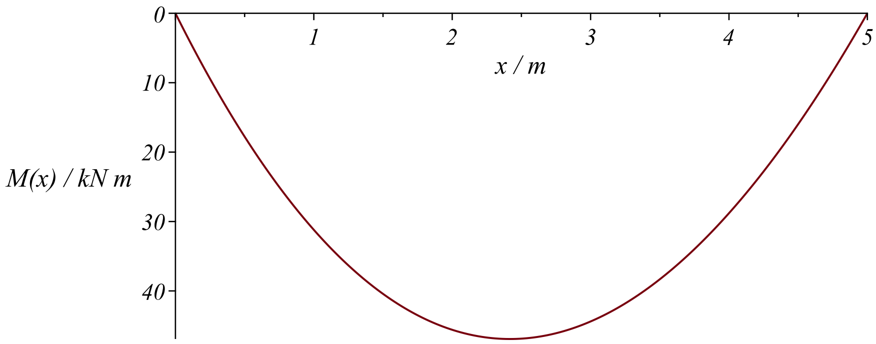
   Now we apply the boundary conditions that $M(0) =: M_A = 0 = M(5)$. We may *not* use R_A equals half the total load as the load is not symmetric:

   This yields a bending moment function

$$M(x) = \frac{1}{5}x^3 - 9x^2 + 40x \tag{2.19}$$

[4] so that $M_A = 0$

The bending moment looks like



The maximum occurs when the slope is zero...

Obviously the + here is too big and the maximum is found at $15 - 5\sqrt{57}/3 \approx 2.417$. Throw this into $M(x)$ to find the maximum bending moment:

$$M(2.42) = \frac{1}{5}(2.42)^3 - 9(2.42)^2 + 40(2.42) \approx 42.9269 \text{ kN m}.$$

Remark: Normally I would throw the surd into the bending moment function. The use of decimals introduces *rounding errors* into our calculations and pure mathematicians consider unrestricted decimal approximation very gravely indeed. However, at the maximum the bending moment function is quite flat so moving a little bit away from the true value doesn't change $M(x)$ that much:

$$M(x_{\text{max}}) \approx M(x_{\text{max}} + \varepsilon) \tag{2.20}$$

In fact in this example $M(x_{\text{max}})$ and $M(2.42)$ agree to four places of decimals. However doing the same thing around $x = 1$ m is *significant*. In particular $M(1) \approx 31.2$ kN m but $M(1 + 1") \approx 31.789$ kN m — a not insignificant difference — over 500 N m in fact about the torque needed to hold 50 kg stable at an arm's length.

Indeed we will see in Chapter 4, that this error is approximated by:

$$\begin{aligned} \Delta M \approx |dM|_{\text{max}} &= \left|\frac{\partial M}{\partial x}\right| \Delta x \\ &= 0 \times \Delta x = 0 \qquad \text{at max} \\ &= \underbrace{|0.6x^2 - 18x + 40|}_{\text{at } x=1} \times 0.025 \approx 0.6 \text{ kN m} \end{aligned}$$

4. **Winter 2011** A light beam of span 6 m is simply supported of its endpoints. Between the points $x = 2$ m and $x = 5$ m there is a U.D.L. of 72 kN m^{-1}.

   i. Express the bending moment $M(x)$ in terms of step functions by solving $M''(x) = -w(x)$ or otherwise.

   ii. The deflection $y(x)$ at any point on the beam is found by solving the differential equation

$$EI\frac{d^2y}{dx^2} = M(x) \tag{2.21}$$

   Solve this differential equation.

Solution: First as always, a picture, and the loading:

Now we write down and solve $M''(x) = -w(x)$:

Here the loading is not symmetric so the reaction isn't shared equally among the two points. But, as the load is simply supported the bending moment at both of the ends is zero:

$$M(0) =: M_A = 0 = M(6)$$

We can use these two equations to find R_A and M_A:

Which yields

$$M(x) = -36[x-2]^2 + 36[x-5]^2 + 90x \tag{2.22}$$

Remark: Again, as the picture suggests, most of the reaction force is concentrated at B due to the asymmetry. In fact $R_B = 3(72) - 90 = 126$ kN.

Now anti-differentiating twice:

That is we have

$$EIy(x) = -3[x-2]^4 + 3[x-5]^4 + 15x^3 + C_1 x + C_2 \tag{2.23}$$

Now we apply the boundary condition that $y(0) = 0 = y(6)$:

Now we can write our final answer:

$$y(x) = \frac{1}{EI}\left(-3[x-2]^4 + 3[x-5]^4 + 15x^3 - \frac{825}{2}x\right) \tag{2.24}$$

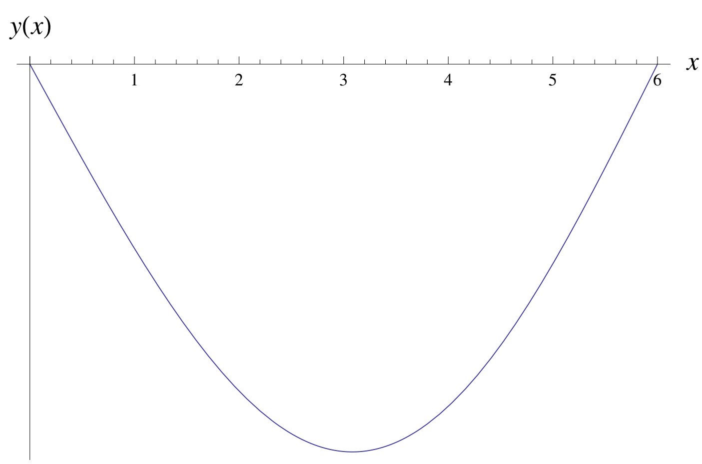


Figure 2.16: A plot of the deflection... due to the asymmetry the deflection is greatest to the right of the midpoint.

Marking Scheme: Winter 2013

A light beam of span 6 m is simply supported at its endpoints. At the point $x = 1$ m there is a point load of 36 kN. Between the points $x = 2$ m and $x = 4$ m there is a U.D.L. of 72 kN m^{-1}.

i. Either by solving the differential equation $M''(x) = -w(x)$, where $w(x)$ is the loading, or otherwise, express the bending moment $M(x)$ in terms of step functions.

[3 Marks]

Solution: We have that the loading is given by $w(x) = 36\delta(x-1) + 72u(x-2) - 72u(x-4)$ **[1]**. Now we solve the differential equation via two anti-differentiations:

$$\begin{aligned} M''(x) &= -36\delta(x-1) - 72u(x-2) + 72u(x-4) \\ \Rightarrow M'(x) &= -36u(x-1) - 72[x-2] + 72[x-4] + R_A \\ \Rightarrow M(x) &= -36[x-1] - 36[x-2]^2 + 36[x-4]^2 + R_A x + M_A \qquad \textbf{[1]} \end{aligned}$$

Now we use the boundary conditions that come from the beam being simply supported: $M(0) = M(6) = 0$. That $M(0) = M_A = 0$ and now we look at $M(6) = 0$:

$$\begin{aligned} M(6) &\stackrel{!}{=} 0 = -36[5] - 36[4]^2 + 36[2]^2 + 6R_A \\ \Rightarrow 6R_A &= 180 + 36(16) - 36(4) \\ \Rightarrow 6R_A &= 612 \\ \Rightarrow R_A &= 102. \qquad \textbf{[1]} \\ \Rightarrow M(x) &= -36[x-1] - 36[x-2]^2 + 36[x-4]^2 + 102x. \end{aligned}$$

Alternate Solution: Using Macauley's Method we have

$$M(x) = -36[x-1] - 36[x-2]^2 + 36[x-4]^2 + R_A x + M_A. \qquad \textbf{[2]}$$

The beam is simply supported so $M_A = 0$. Now we use the boundary condition $M(6) = 0$ to get $R_A = 102$ **[1]**.

ii. The deflection, $y(x)$, at any point on the beam is found by solving the differential equation

$$EI\frac{d^2y}{dx^2} = M(x).$$

Solve the differential equation for $y(x)$.

[8 Marks]

Solution: We anti-differentiate the differential equation twice to find $EI \cdot y(x)$:

$$\begin{aligned} EI\frac{d^2y}{dx^2} &= -36[x-1] - 36[x-2]^2 + 36[x-4]^2 + 102x \qquad \textbf{[1]} \\ \Rightarrow EI\frac{dy}{dx} &= -18[x-1]^2 - 12[x-2]^3 + 12[x-4]^3 + 51x^2 + C_1 \qquad \textbf{[1]} \\ \Rightarrow EI \cdot y(x) &= -6[x-1]^3 - 3[x-2]^4 + 3[x-4]^4 + 17x^3 + C_1 x + C_2 \qquad \textbf{[1]} \end{aligned}$$

We use the boundary conditions $y(0) = y(6) = 0$ **[1]** to find C_1 and C_2. The first, $y(0) = 0$ yields $C_2 = 0$ **[1]**. Now we apply $y(6) = 0$:

$$\begin{aligned} EI \cdot y(6) &\stackrel{!}{=} 0 = -6[5]^3 - 3[4]^4 + 3[2]^4 + 17(6)^3 + 6C_1 \qquad \textbf{[1]} \\ \Rightarrow 6C_1 &= 6(125) + 3(256) - 3(16) - 17(216) \\ \Rightarrow 6C_1 &= -2,202 \qquad \textbf{[1]} \\ \Rightarrow C_1 &= -367 \\ \Rightarrow y(x) &= \frac{1}{EI}\left(-6[x-1]^3 - 3[x-2]^4 + 3[x-4]^4 + 17x^3 - 367x\right). \qquad \textbf{[1]} \end{aligned}$$

iii. At $x_1 \approx 2.959$ m we have $y'(x_1) = 0$ and $y''(x_1) > 0$. What can we conclude about the deflection at x_1?

[1 Mark]

Solution: The deflection is greatest at x_1.

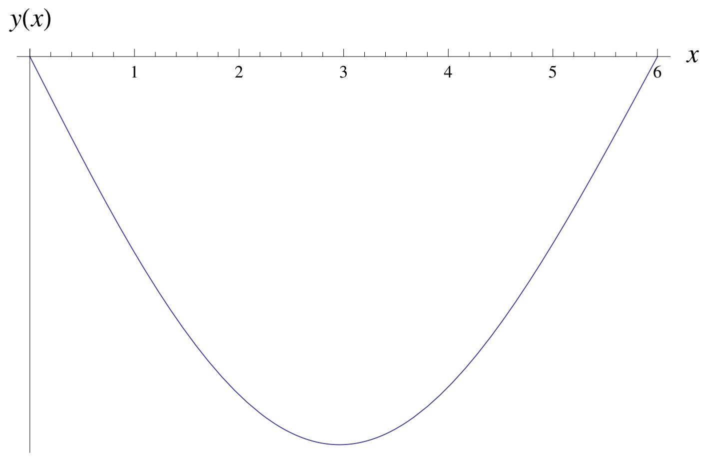


Figure 2.17: The maximum deflection is to the left of the midpoint (why?).

Marking Scheme: Winter 2014

A light beam of span 6 m is simply supported at its endpoints. At the points $x = 1$ and $x = 5$ m there are point loads of 18 kN. Between the points $x = 2$ and $x = 4$ m there is a U.D.L. of 24 kN m^{-1}.

i Draw a sketch of the beam and hence explain why $R_A = 42$.

[2 Marks]

Solution: The nice sketch **[1]** will exhibit the symmetry in the problem and hence R_A and R_B share the total load of $18 + 18 + 2 \times 24 = 84$ **[1]** so that $R_A = 42$.

ii. By solving the differential equation

$$\frac{d^2M}{dx^2} = -w(x),$$

where $w(x)$ is the load per unit length, or otherwise, find the bending moment $M(x)$.

[3 Marks]

Solution: We have

$$\begin{aligned} w(x) &= 18\delta(x-1) + 24u(x-2) - 24u(x-4) + 18\delta(x-5) \\ \Rightarrow \frac{d^2M}{dx^2} &= -18\delta(x-1) - 24u(x-2) + 24u(x-4) - 18\delta(x-5) \\ \Rightarrow \frac{dM}{dx} &= -18u(x-1) - 24[x-2] + 24[x-4] - 18u(x-5) + R_A \qquad \textbf{[1]} \end{aligned}$$

However we know that $R_A = 42$ from previous part.

$$M(x) = -18[x-1] - 12[x-2]^2 + 12[x-4] - 18[x-5] + 42x + M_A \qquad \textbf{[1]}$$

however as the beam is simply supported we have $M_A = 0$ **[1]** to give

$$M(x) = -18[x-1] - 12[x-2]^2 + 12[x-4] - 18[x-5] + 42x.$$

iii. Calculate the bending moment at $x = 50$ cm.

[2 Marks]

Solution: We calculate

$$\underbrace{M(0.5)}_{[1]} = -18[-0.5] - 12[-1.5]^2 + 12[-3.5]^2 - 18[-4.5] + 42(0.5) = 21 \text{ kN m. } \mathbf{[1]}$$

iv. The deflection, $y(x)$, at any point on the beam is found by solving the differential equation

$$EI \cdot \frac{d^2y}{dx^2} = M(x).$$

Solve the differential equation for $y(x)$.

[5 Marks]

Solution: Continuing from above we have

$$\begin{aligned} EI \cdot y''(x) &= -18[x-1] - 12[x-2]^2 + 12[x-4]^2 - 18[x-5] + 42x \qquad \mathbf{[1]} \\ \Rightarrow EI \cdot y'(x) &= -9[x-1]^2 - 4[x-2]^3 + 4[x-4]^3 - 9[x-5]^2 + 21x^2 + C_1 \qquad \mathbf{[1]} \\ \Rightarrow EI \cdot y(x) &= -3[x-1]^3 - [x-2]^4 + [x-4]^4 - 3[x-5]^3 + 7x^3 + C_1 x + C_2 \qquad \mathbf{[1]} \end{aligned}$$

As the beam is simply supported $y(0) = 0 \Rightarrow C_2 = 0$ **[1]**. Now we look at $y(6) = 0$:

$$\begin{aligned} EI \cdot y(6) &= -3(5)^3 - 4^4 + 2^4 - 3(1) + 7(6)^3 + 6C_1 \stackrel{!}{=} 0 \qquad \mathbf{[1]} \\ \Rightarrow C_1 &= -149, \end{aligned}$$

to give an answer of

$$y(x) = \frac{1}{EI}\left(-3[x-1]^3 - [x-2]^4 + [x-4]^4 - 3[x-5]^3 + 7x^3 - 149x\right).$$

v. From symmetry, we expect to find the maximum deflection at $x = 3$ m. Use the slope, $y'(x)$, to verify this.

[2 Marks]

Solution: We calculate

$$\begin{aligned} \underbrace{y'(3)}_{[1]} &= \frac{1}{EI}\left(-9(4) - 4(1) + 4(0) + 0 + 21(9) - 149\right) \\ &= \frac{1}{EI}(0) = 0 \checkmark \qquad \mathbf{[1]} \end{aligned}$$

Marking Scheme: Autumn 2014

A light beam of span 5 m is simply supported at its end points and carries a load that varies uniformly with x the distance from one end of the beam. The load varies from 18 kN m^{-1} at $x = 0$ to 12 kN m^{-1} at $x = 5$.

i. Find a formula for the load per unit length, $w(x)$.

ii. By solving the differential equation

$$\frac{d^2M}{dx^2} = -w(x),$$

find the bending moment M at any point along the beam.

iii. Find the maximum value of the bending moment.

[9 Marks]

Solution: We have

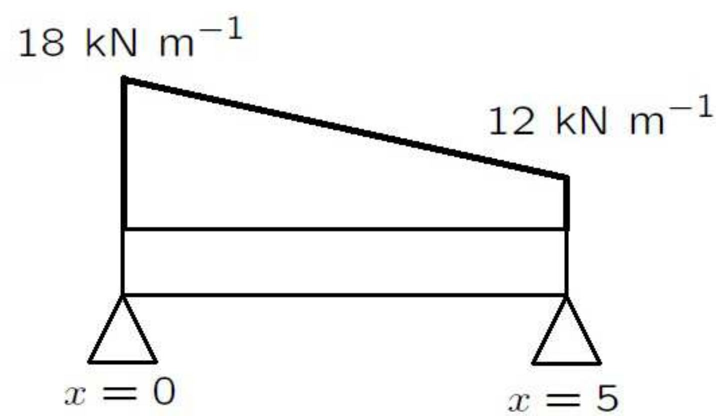


Figure 2.18: Using $y = mx + c$ we have $w(x) = -\frac{6}{5}x + 18$.**[2]**

We anti-differentiate this differential equation twice to find $M(x)$:

$$\frac{d^2M}{dx^2} = \frac{6}{5}x - 18$$
$$\Rightarrow \frac{dM}{dx} = \frac{6x^2}{10} - 18x + R_A$$
$$\Rightarrow M(x) = \frac{1}{5}x^3 - 9x^2 + R_Ax + M_A \qquad \mathbf{[1]}$$

We apply the boundary conditions $M(0) = 0$ and $M(5) = 0$ which gives us $M_A = 0$ **[1]** and

$$M(5) \stackrel{!}{=} 0 = \frac{1}{5}(5)^3 - 9(5)^2 + R_A(5)$$
$$\Rightarrow 5R_A = 225 - 25 = 200$$
$$\Rightarrow R_A = 40 \qquad \mathbf{[1]}$$
$$\Rightarrow M(x) = \frac{1}{5}x^3 - 9x^2 + 40x.$$

The maximum of the bending moment will occur where $M'(x) = 0$ **[1]**

$$\begin{aligned} M'(x) &= \frac{3}{5}x^2 - 18x + 40 \overset{!}{=} 0 \\ \Rightarrow 3x^2 - 90x + 200 &= 0 \qquad \textbf{[1]} \end{aligned}$$

This does not factorise so we look at the $-b \pm \cdots$ formula:

$$\begin{aligned} x &= \frac{90 \pm \sqrt{(-90)^2 - 4(3)(200)}}{2(3)} = \frac{90 \pm \sqrt{5700}}{6} \\ &= \frac{90 \pm 10\sqrt{57}}{6} = 15 \pm \frac{5}{3}\sqrt{57} \\ &\approx 2.417 \text{ or } 27.58 \qquad \textbf{[1]} \end{aligned}$$

Clearly $x = 27.58$ is not physical. We should ensure that $x = 2.417$ is a maximum either by noting the '$\sim$' geometry of the $+x^3$ or by showing that $M''(2.417) = \frac{6}{5}(2.417) - 18 < 0$ which is clear. The maximum bending moment is

$$M(2.417) = \frac{1}{5}(2.417)^3 - 9(2.417)^2 + 40(2.417) = 46.927 \text{ kN m } \textbf{[1]}$$

Exercises

1. A light beam of span 5 m is simply supported at its end points and carries a uniformly distributed load of 18 kN m^{-1} along the beam. By solving the differential equation

$$\frac{d^2M}{dx^2} = -w(x),$$

find the bending moment at any point along the beam. **Ans:** $M(x) = -9x^2 + 45x$.

2. A light beam of span 5 m is simply supported at its endpoints. Between the points $x = 2$ m and $x = 5$ m, there is a UDL of 30 kN m^{-1}.

   i. Express the bending moment in terms of a step function by solving $M''(x) = -w(x)$, or otherwise.

   ii. By solving the differential equation $EIy''(x) = M(x)$, find the deflection at any point on the beam. **Ans:** $y(x) = \dfrac{1}{EI}\left(-\dfrac{5}{4}[x-2]^4 + \dfrac{9}{2}x^3 - \dfrac{369}{4}x\right)$.

3. **Autumn 2020** A light beam of span 6 m is simply supported at its endpoints. At the points $x = 1$ and $x = 5$ m there are point loads of 18 kN.

   i Draw a sketch of the beam and hence explain why $R_A = 18$.

   ii. By solving the differential equation

$$\frac{d^2M}{dx^2} = -w(x),$$

   where $w(x)$ is the load per unit length, or otherwise, find the bending moment $M(x)$. **Ans:** $M(x) = -18[x-1] - 18[x-5] + 18x$.

   iii. Calculate the bending moment at $x = 50$ cm. **Ans:** 9 kN m.

iv. The deflection, $y(x)$, at any point on the beam is found by solving the differential equation

$$EI \cdot \frac{d^2y}{dx^2} = M(x).$$

Solve the differential equation for $y(x)$.
Ans: $y(x) = \frac{1}{EI}\left(-3[x-1]^3 - 3[x-5]^3 + 3x^2 - 45x\right)$

v. From symmetry, we expect to find the maximum deflection at $x = 3$ m. Use the slope, $y'(x)$, to verify this.

4. A light beam of span 6 m is simply supported at its endpoints. At the point $x = 3$ m there is a load of 72 kN. Between the points $x = 4$ m and $x = 6$ m there is a UDL of 72 kN m^{-1}.

(a) By solving the differential equation

$$\frac{d^2M}{dx^2} = -w(x),$$

where $w(x)$ is the load per unit length, or otherwise, find the bending moment $M(x)$.
Ans: $M(x) = -72[x-3] - 36[x-4]^2 + 60x$.

(b) Explain the true statement:

By considering the physics/geometry/engineering of the problem, we know that the reaction at $x = 0$ m, R_A, is less than 108 KN.

(c) The deflection, $y(x)$, at any point on the beam is found by solving the differential equation

$$EI \cdot \frac{d^2y}{dx^2} = M(x).$$

Solve the differential equation for the deflection, $y(x)$, in terms of EI. **Ans:** $y(x) = \frac{1}{EI}\left(-12[x-3]^3 - 3[x-4]^4 + 10x^3 - 298x\right)$.

(d) Find, in terms of EI, the slope of the beam at $x = 0$ m. **Ans:** $-298/EI$

5. **Winter 2016** A light beam of span 7 m, simply supported at its endpoints, carries a load varying uniformly between 7 kN m^{-1} at $x = 0$ m and 35 kN m^{-1} at $x = 7$ m.

(a) Write down an expression for the load per unit length at a distance x along the beam, $w(x)$. **Ans:** $w(x) = 4x + 7$.

(b) By solving the differential equation

$$\frac{d^2M}{dx^2} = -w(x),$$

where $w(x)$ is the load per unit length, find the bending moment $M(x)$. **Ans:** $M(x) = -\frac{2}{3}x^3 - \frac{7}{2}x^2 + \frac{343}{6}x$.

(c) Explain using the physics/geometry/engineering of the situation why $R_A < \frac{1}{2}(147)$.
Hint: Find the total load.

(d) Find the *location* of the maximum bending moment. **Ans:** ≈ 3.875 m.

(e) The deflection, $y(x)$, at any point on the beam is found by solving the differential equation

$$EI \cdot \frac{d^2y}{dx^2} = M(x).$$

Solve the differential equation for $y(x)$.

Ans: $y(x) = \frac{1}{EI}\left(-\frac{1}{30}x^5 - \frac{7}{24}x^4 + \frac{343}{36}x^3 - \frac{2008}{7}x\right).$

2.2.4 Fixed Ends

A fixed end beam of length L looks as follows:

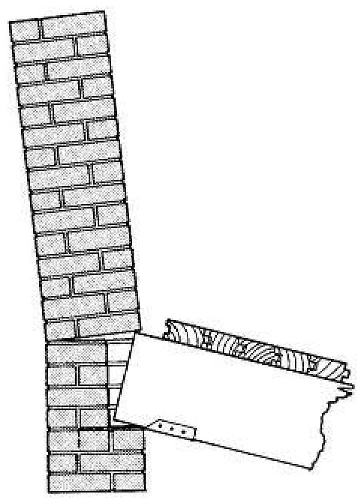


Figure 2.19: For fixed end beams, there *is* a non-zero moment keeping the beam 'straight' at the supports... this means that the bending moment M is *not* zero at the ends. However this moment causes the slope y' to be zero at the ends. Of course, the deflection y is zero at the ends.

We have the following *four* boundary conditions:

1. the deflection at both ends are zero: i.e. $y(0) = 0$ and $y(L) = 0$.
2. the slope at both ends is zero: i.e. $y'(0) = 0$ and $y'(L) = 0$.

Note that $M_A \neq 0 \neq M_B$ necessarily as the wall exerts a bending moment. We will always have that $y(0) = 0 \Rightarrow C_2 = 0$ and $y'(0) = 0 \Rightarrow C_1 = 0$. This is because we will always have — in MATH7019 at least

$$y(x) = \frac{1}{EI}\left(\text{ stuff that equals zero at zero } + C_2\right),$$

and so

$$0 = y(0) = \frac{C_2}{EI} \Rightarrow C_2 = EI \cdot y(0) = 0, \tag{2.25}$$

always in MATH7019. Similarly we will always have — in MATH7019 at least

$$y'(x) = \frac{1}{EI}\left(\text{ stuff that equals zero at zero } + C_1\right),$$

which implies

$$0 = y'(0) = \frac{C_1}{EI} \Rightarrow C_1 = EI \cdot y'(0) = 0, \qquad (2.26)$$

for fixed ends and cantilevers.

Examples

1. **Winter 2013** A light beam of 6 m has both ends embedded in walls. Draw a picture of an unloaded beam with fixed ends which explains why the boundary conditions are given by $y(0) = y(6) = 0$ and $y'(0) = y'(6) = 0$.

   Solution:

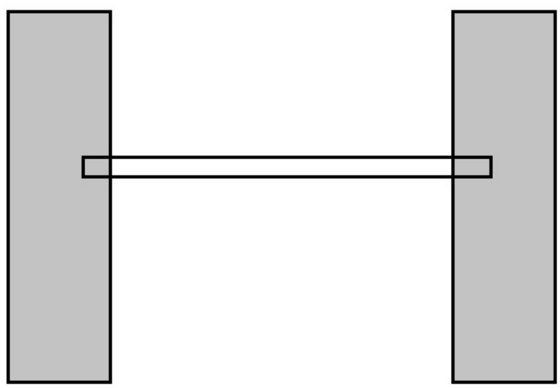


Figure 2.20: The deflections at the walls are zero as are the slopes.

2. **Autumn 2014** A light beam of span 6 m is fixed at its endpoints. At the points $x = 1$ m and $x = 5$ m there are point loads of 36 kN and 36 kN respectively.

   i. y solving the differential equation $M''(x) = -w(x)$, where $w(x)$ is the loading, or otherwise, express the bending moment $M(x)$ in terms of step functions, the reaction force $M'(0) = R_A$ and the bending moment $M(0) = M_A$.

   Solution: First a little sketch:

   We have that the loading is given by $w(x) = 36\delta(x-1) + 36\delta(x-5)$.

Now we solve the differential equation via two anti-differentiations:

Alternate Solution: Using Macauley's Method we have

$$M(x) = -36[x-1] - 36[x-5] + R_A x + M_A.$$

ii. The deflection, $y(x)$, at any point on the beam is found by solving the differential equation

$$EI\frac{d^2y}{dx^2} = M(x).$$

Solve the differential equation for $y(x)$.

Solution: We anti-differentiate the differential equation twice to find $EI \cdot y(x)$. Note that we have two constants now (R_A and M_A) — we will have two more (C_1 and C_2):

We use the boundary conditions $y(0) = y'(0) = 0$ to find C_1 and C_2. The first, $y(0) = 0$ yields $C_2 = 0$ and the second, $y'(0) = 0$ yields $C_1 = 0$ so we have

$$EI \cdot y(x) = -6[x-1]^3 - 6[x-5]^3 + \frac{R_A}{6}x^3 + \frac{M_A}{2}x^2.$$

Now we apply $y(6) = 0$ and $y'(6) = 0$ to generate two equations in R_A and M_A which we can solve:

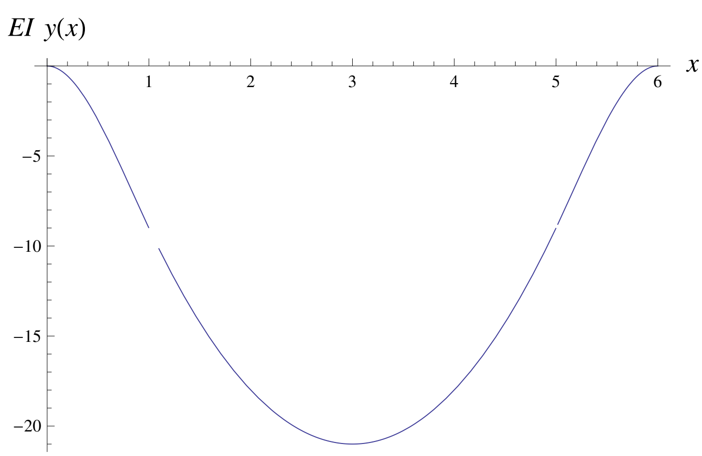


Figure 2.21: Here we can see the symmetry in the deflection and the fact that both the deflection and slope is zero at the ends of the beam. The larger the value of EI, the less the deflection. Note that $y'(x) = 0$ at the point of maximum deflection.

iii. Draw the beam, including the fixed ends and the point loads. Where do you expect the maximum deflection to be? Verify your guess by looking at $y'(x)$ at the point in question.

Solution:

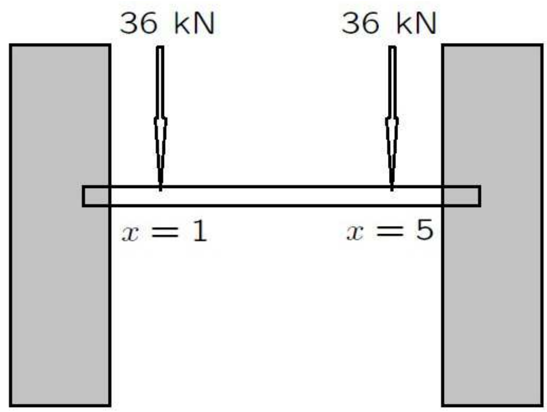


Figure 2.22: By symmetry we expect to find the maximum deflection at $x = 3$ m.

We verify this by showing that $y'(3) = 0$:

3. **Winter 2011** A light beam of span 6 m has both ends embedded in walls. At the point $x = 2$ m there is a load of 36 kN. Between the points $x = 4$ m and $x = 6$ m there is a U.D.L. of 72 kN m^{-1}.

   i. Either by solving $M''(x) = -w(x)$, or otherwise, express the bending moment $M(x)$ in terms of step functions, the reaction force $M'(0) = R_A$ and the bending moment $M(0) = M_A$.

   ii. Solve the differential equation
   $$EI\frac{d^2y}{dx^2} = +M(x), \tag{2.27}$$
   to find the deflection at any point on the beam.

Solution: First draw a picture:

We will write down the loading function:
$$w(x) = 36\delta(x-2) + 72u(x-4),$$
and anti-differentiate $-w(x)$ twice to find $M(x)$:

In the case of a simply supported beam we have the boundary condition $M(0) = 0 = M(6)$ **but this is not the case when the ends are fixed**. You must carry around an M_A and R_A for now. Now we need to solve
$$\begin{aligned} EI\frac{d^2y}{dx^2} &= M(x) \\ &= -36[x-2] - 36[x-4]^2 + R_A x + M_A \end{aligned}$$

Thus
$$EI \cdot y(x) = -6[x-2]^3 - 3[x-4]^4 + \frac{R_A}{6}x^3 + \frac{M_A}{2}x^2 + C_1 x + C_2$$
Four unknowns is tough but we have four boundary conditions which will generate four equations in R_A, M_A, C_1, C_2 which we should then be able to solve. First up let's use $y(0) = 0 = y'(0)$: this gives $C_2 = 0$ and $C_1 = 0$.

Now we apply $y(6) = 0$ and $y'(6) = 0$ to generate simultaneous equations for R_A and M_A:

Marking Scheme: Autumn 2015

Consider a light beam if span 5 m carrying a load varying linearly from
12 kN m^{-1} at $x = 0$ m to 17 kN m^{-1} at $x = 5$ m. By solving the second order differential equation

$$\frac{d^2M}{dx^2} = -w(x),$$

where $w(x)$ is the load per unit length, the bending moment is given by

$$M(x) = -\frac{x^3}{6} - 6x^2 + R_A x + M_A.$$

(a) Suppose that the beam is simply supported. Using the boundary condition $M(5) = 0$, it can be shown that $R_A = 205/6$. Hence, find the *location* of the maximum bending moment. [5 Marks]

Solution: To find the location of the maximum bending moment we find where $M'(x) = V(x) = 0$ [**1**]:

$$M'(x) = -\frac{x^2}{2} - 12x + \frac{205}{6} \stackrel{!}{=} 0 \qquad [\mathbf{1}]$$

$$\Rightarrow x^2 + 24x - \frac{205}{3} = 0 \qquad [\mathbf{1}]$$

$$\Rightarrow x_{\pm} = -\frac{24 \pm \sqrt{24^2 - 4(-205/3)}}{2} = -26.57\,,\ 2.57, \qquad [\mathbf{1}]$$

but of course only $x = 2.57$ m is on the beam. [**1**]

(b) Suppose that the beam is fixed at both ends.

By solving the second order differential equation

$$EI \cdot \frac{d^2y}{dx^2} = -\frac{x^3}{6} - 6x^2 + R_A x + M_A,$$

find the deflection at any point on the beam.

[10 Marks]

Solution: In this case the beam is fixed so we have $y(0) = 0 \Rightarrow C_2 = 0$ [**1**] and $y'(0) = 0 \Rightarrow C_1 = 0$ [**1**]. Anti-differentiating once we have

$$\begin{aligned} EI \cdot y'(x) &= -\frac{x^4}{24} - 2x^3 + \frac{R_A}{2}x^2 + M_A x + C_1 \qquad [\mathbf{1}] \\ \Rightarrow EI \cdot y(x) &= -\frac{x^5}{120} - \frac{x^4}{2} + \frac{R_A}{6}x^3 + \frac{M_A}{2}x^2 + C_1 x + C_2 \qquad [\mathbf{1}] \\ \Rightarrow EI \cdot y(x) &= -\frac{x^5}{120} - \frac{x^4}{2} + \frac{R_A}{6}x^3 + \frac{M_A}{2}x^2. \end{aligned}$$

No we apply the boundary conditions $y(5) = 0$ [**1**] and $y'(5) = 0$ [**1**].

$$\begin{aligned} EI \cdot y(5) = -\frac{5^5}{120} - \frac{5^4}{2} + \frac{R_A}{6}5^3 + \frac{M_A}{2}5^2 &\stackrel{!}{=} 0 \\ \Rightarrow \frac{125}{6}R_A + \frac{25}{2}M_A &= \frac{8125}{24} \\ \Rightarrow 500R_A + 300M_A &= 8125 \\ \Rightarrow 20R_A + 12M_A &= 325. \qquad [\mathbf{1}] \end{aligned}$$

and to get a second equation in R_A and M_A we look at:

$$\begin{aligned} EI \cdot y'(5) = -\frac{5^4}{24} - 2(5)^3 + \frac{R_A}{2}(5)^2 + M_A(5) &\stackrel{!}{=} 0 \\ \Rightarrow -5^4 - 48(5)^3 + 12(25)R_A + 120M_A &= 0 \\ \Rightarrow 60R_A + 24M_A &= 1325. \qquad [\mathbf{1}] \end{aligned}$$

Decimal approximations are frowned upon but acceptable!

We have

$$\begin{aligned} -60R_A - 36M_A &= -97560R_A + 24M_A \qquad = 1325 \\ \Rightarrow -12M_A &= 350 \\ \Rightarrow M_A &= -\frac{350}{12} \approx 29.17 \text{ kN m} \qquad [\mathbf{1}] \end{aligned}$$

$$\begin{aligned} \Rightarrow 20R_A + 12\left(-\frac{350}{12}\right) &= 325 \\ \Rightarrow R_A &= \frac{135}{4} \approx 33.75. \qquad [\mathbf{1}] \end{aligned}$$

so we have

$$y(x) = \frac{1}{EI}\left(-\frac{x^5}{120} - \frac{x^4}{2} + \frac{135}{24}x^3 - \frac{175}{12}x^2\right)$$

INTERLEAVED Exercises

1. **Winter 2019** A light beam of span 6 m is fixed at its endpoints under a constant load of 18 kN m^{-1}

   (a) By solving the differential equation

   $$\frac{d^2M}{dx^2} = -w(x),$$

   where $w(x)$ is the load per unit length, or otherwise, find the bending moment $M(x)$ in terms of R_A and M_A. **Ans:** $M(x) = -9x^2 + R_Ax + M_A$

   (b) The deflection, $y(x)$, at any point on the beam is found by solving the differential equation

   $$EI \cdot \frac{d^2y}{dx^2} = M(x).$$

   Solve the differential equation for $y(x)$. **Ans:** $y(x) = \frac{1}{EI}\left(-\frac{3}{4}x^4 + 9x^3 - 27x^2\right)$

   (c) From symmetry, we expect to find the maximum deflection at $x = 3$ m. Use the slope, $y'(x)$, to verify this.

   (d) Calculate the bending moment at $x = 50$ cm. **Ans:** -29.4 kN m

   (e) Also find the maximum bending moment. **Ans:** 27 kN m

2. A simply supported light beam of span 5 m carries a UDL of 24 kN m^{-1} between $x = 1$ and $x = 4$ m.

   i. Solve $M''(x) = -w(x)$.

   ii. Hence solve $EIy''(x) = M(x)$. **Ans:** $y(x) = \dfrac{1}{EI}\left(-[x-1]^4 + [x-4]^4 + 6x^3 - 99x\right)$.

   iii. Find the slope of y at the midpoint of the beam.

3. **Autumn 2011** A light beam of span 6 m has both ends embedded in walls. Between the points $x = 3$ m and $x = 6$ m there is a U.D.L. of 24 kN m^{-1}.

   i. Express the bending moment $M(x)$ in terms of step functions, the reaction force $M'(0) = R_A$ and the bending moment $M(0) = M_A$.

   ii. Solve the differential equation $EI\dfrac{d^2y}{dx^2} = M(x)$ to find the deflection y at any point on the beam. **Ans:** $y(x) = \dfrac{1}{EI}\left(-[x-3]^4 + \dfrac{9}{4}x^3 - \dfrac{45}{4}x^2\right)$.

4. **Autumn 2009** A light beam of span 6 m is simply supported at its endpoints. At the points $x = 2$ m and $x = 4$ m, there are point loads: each equal to 36 kN.

   i. Express the Bending Moment $M(x)$ in terms of step functions by solving $M''(x) = -w(x)$, or otherwise.

   ii. Solve the differential equation

   $$EI\frac{d^2y}{dx^2} = M(x)$$

   to find the deflection y at any point.

   Ans: $y(x) = \dfrac{1}{EI}\left(-6[x-2]^3 - 6[x-4]^3 + 6x^3 - 144x\right)$.

5. **Summer 2008** At the points $x = 2$ m and $x = 4$ m on a beam of span 6 m there are loads of 18 kN and 36 kN, respectively. Both ends of the beam are embedded in walls.

   i. Express the bending moment $M(x)$ in terms of step functions, the reaction force $M'(0) = R_A$ and the bending moment $M(0) = M_A$.

   ii. Find the deflection y at any point on the beam by solving the differential equation $EIy''(x) = M(x)$.
   Ans: $y(x) = \frac{1}{EI}\left(-3[x-2]^3 - 6[x-4]^3 + \frac{34}{9}x^3 - 16x^2\right)$.

6. **Winter 2010** A light beam of span 6 m is simply supported at its endpoints. At the point $x = 2$ m there is a load of 72 kN. Between the points $x = 4$ m and $x = 6$ m there is a UDL of 72 kN m^{-1}.

   i. By solving $M''(x) = -w(x)$, or otherwise, express the Bending Moment $M(x)$ in terms of step functions.

   ii. Solve the differential equation

   $$EI\frac{d^2y}{dx^2} = M(x)$$

   to find the deflection $y(x)$ at any point x on the beam.
   Ans: $y(x) = \frac{1}{EI}\left(-12[x-2]^3 - 3[x-4]^4 + 12x^3 - 296x\right)$.

7. **Autumn 2013** A light beam of 4 m has both ends embedded in walls. At the point $x = 1$ m there is a load of 36 kN. Between the points $x = 2$ m and $x = 4$ m there is a U.D.L. of 30 kN m^{-1}.

   i. Express the Bending Moment $M(x)$ in terms of step functions, the reaction force $M'(0) = R_A$ and the bending moment $M(0) = M_A$.. **Ans:** $M(x) = -36[x-1] - 15[x-2]^2 + R_Ax + M_A$

   ii. The deflection $y(x)$ at any point on the beam is found by solving the differential equation

   $$EI\frac{d^2y}{dx^2} = M(x).$$

   Solve the differential equation.
   Ans: $y(x) = \frac{1}{EI}\left(-6[x-1]^3 - \frac{5}{4}[x-2]^4 + \frac{111}{16}x^3 - \frac{131}{8}x^2\right)$.

   iii. What is the deflection of the beam at $x = 1$ m. **Ans:** $-\frac{151}{16EI}$.

8. **Autumn 2011** A light beam of span 4 m is simply supported at its end points and carries a load that varies in a uniform way along the beam. The load varies from 8 kN m^{-1} at $x = 0$ m to 20 kN m^{-1} at $x = 4$ m.

   i. Find a formula for the load per unit length $w(x)$.

   ii. By solving the differential equation

   $$\frac{d^2M}{dx^2} = -w(x),$$

   find the bending moment $M(x)$ at any point along the beam. **Ans:** $M(x) = -x^3/2 - 4x^2 + 24x$.

9. **Summer 2007** Between the points $x = 0$ m and $x = 2$ m on a beam of span 6 m there is a U.D.L. of 36 kN m^{-1}. Both ends of the beam are embedded in walls.

   i. By solving $M''(x) = -w(x)$, or otherwise, express the bending moment $M(x)$ in terms of step functions, the reaction force $M'(0) = R_A$ and the bending moment $M(0) = M_A$.

   ii. By solving the differential equation $EIy''(x) = M(x)$ find the deflection y at any point on the beam. **Ans:** $y(x) = \frac{1}{EI}\left(-\frac{3}{2}x^4 + \frac{3}{2}[x-2]^4 + \frac{98}{9}x^3 - 22x^2\right)$.

2.2.5 Cantilevers

A cantilever looks like:

and as long as there is no *point load* on the end of the beam the following boundary conditions can be seen:

1. There is no deflection at the fixed end so $y(0) = 0 \Rightarrow C_2 = 0$.
2. There is no slope at the fixed end so $y'(0) = 0 \Rightarrow C_1 = 0$.
3. There is no bending moment at the free end so $M(L) = 0 \Rightarrow EI \cdot y''(L) = 0 \Rightarrow M_A < 0$. Alternatively, one can note that for equilibrium the nett moment must be zero so that $M_A = -\sum M_i$.
4. There is no reaction force at the free end so $R_A = \sum(\text{loads})$. This is derived by demanding that the total force must be zero.

This last boundary condition is replaced by $|V(L)| = mg$ if there is a point load of mass m at the end of the beam (this won't happen in MATH7019).

It is clear from the physics/geometry/engineering of the situation that the maximum deflection occurs at $x = L$ and the question that will be given will be find the deflection at the end of the beam in terms of E & I. It might prove useful here to solve directly

$$EI\frac{d^4y}{dx^4} = EIy^{(iv)}(x) = -w(x),$$

rather than embracing Macauley's Method. Then the first two boundary conditions means that the last two constants of integration are zero. Then look at the last boundary condition to find that $R_A = \sum(\text{loads})$ and all that is left is to find M_A!

Examples

1. Find the maximum deflection of a light cantilever beam of span 6 m with a uniformly distributed load of 36 kN m^{-1} between $x = 2$ m and $x = 4$ m.

   Solution: We begin with a sketch and an expression for the loading

   This gives a differential equation that we antidifferentiate four times:

$$EIy^{(iv)}(x) = -36u(x-2) + 36u(x-4)$$

   Now we apply the boundary conditions. Note that R_A is just equal to the total load. You may assume this in future. Now we look at $M(6) = EI \cdot y''(6) = 0$:

   This means that we have

$$y(x) = \frac{1}{EI}\left(-\frac{3}{2}[x-2]^4 + \frac{3}{2}[x-4]^4 + 12x^3 - 108x^2\right).$$

   Now the maximum deflection may be found at $x = 6$ m (note $EI \sim 100$):

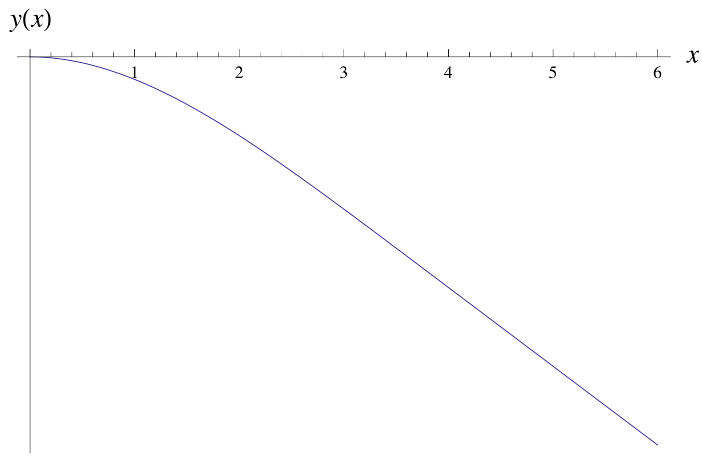


Figure 2.23: Textbook Cantilever Deflection. Note that $y'(0) = 0$ and $y(0) = 0$. $y_{\max}$ occurs at $x = 6$.

2. Find the maximum deflection of a light cantilever beam of span 6 m with a uniformly distributed load of 48 kN m^{-1} between $x = 1$ m and $x = 3$ m and a point load of 18 kN at $x = 5$ m.

   Solution: We begin with a sketch and an expression for the loading

   to give a differential equation that we antidifferentiate four times:

$$EIy^{(iv)}(x) = -48u(x-1) + 48u(x-3) - 18\delta(x-5)$$

   Now we apply the boundary conditions. The first two boundary conditions, because the beam is fixed at $x = 0$ m yield $C_1 = C_2 = 0$. Now R_A is equal to the total load $2 \times 48 + 18 = 114$.

Now we look at $M(6) = EI \cdot y''(6) = 0$:

This means that we have

$$y(x) = \frac{1}{EI}\left(-2[x-1]^4 + 2[x-3]^4 - 3[x-5]^3 + 19x^3 - 141x^2\right).$$

Now the maximum deflection may be found at $x = 6$ m:

Remark

Note that in all cases we have

$$\begin{aligned} EI \cdot y(x) &= F(x) + C_2 \qquad \text{and} \\ EI \cdot y'(x) &= G(x) + C_1, \end{aligned}$$

where F and G are just functions of x such that $F(0) = 0 = G(0)$. Therefore if we input $x = 0$ into the any of these formulae we get:

$$\begin{aligned} EI \cdot y(0) &= 0 + C_2 \qquad &&\Rightarrow C_2 = EI \cdot y(0) \Rightarrow y(0) = \frac{C_2}{EI} \qquad \text{and} \\ EI \cdot y'(0) &= 0 + C_1 \qquad &&\Rightarrow C_1 = EI \cdot y'(0) \Rightarrow y'(0) = \frac{C_1}{EI}. \end{aligned}$$

In particular, because $EI > 0$ and never zero, this explains why

- for simply supported, $C_1 < 0$ and $C_2 = 0$. Let us draw a simply supported beam under some load:

  Now, *geometrically* — i.e. with respect to this diagram — what does $y(0)$ represent:

  Now the deflection at $x = 0$ m has to be zero and so

  $$C_2 = EI \cdot y(0) = EI(0) = 0 \text{ also.}$$

Now, *geometrically*, what does $y'(0)$ represent:

Now is this slope positive or negative? Hence

$$C_1 = EI \cdot y'(0) \sim (+)(-) = (-),$$

and so C_1 is negative. Note that

$$y'(0) = \frac{C_1}{EI}$$

is derived independently of the values of E and I and so increasing E and/or I decreases the slope. What does a larger E mean? What does a larger I mean?

- for fixed ends and cantilevers, $C_1 = C_2 = 0$. Let us draw a beam fixed at $x = 0$ m (it may or may not be fixed at $x = L$ also):

As before the deflection at $x = 0$ m is zero hence $C_2 = 0$ also. Now what is $y'(0)$ equal to here? Hence

$$C_1 = EI \cdot y'(0) = EI \cdot (0) = 0.$$

Marking Scheme: Autumn 2015

Find, in terms of EI, the maximum deflection of a light cantilever beam of span 6 m with a constant U.D.L. of 12 KN m^{-1} across the entire beam.

The deflection, $y(x)$, is the solution the fourth order differential equation

$$EI \cdot \frac{d^4y}{dx^4} = -w(x).$$

Equivalently, the deflection can be found by first finding the bending moment, $M(x)$, using Macauley's method, or by solving the second order differential equation:

$$\frac{d^2M}{dx^2} = -w(x).$$

The deflection, $y(x)$, at any point on the beam is then found by solving the differential equation

$$EI \cdot \frac{d^2y}{dx^2} = M(x).$$

[10 Marks]

Solution: We have $y(0) = 0 \Rightarrow C_2 = 0$ [**1**], $y'(0) = 0 \Rightarrow C_1 = 0$ [**1**], $V(6) = 0 \Rightarrow R_A = 12 \times 6 = 72$ kN [**1**]. If they want, students can take $M_A = -(72(3)) = -216$. We have

$$w(x) = 12,$$

so we look at

$$\begin{aligned}
EI \cdot \frac{d^4y}{dx^4} &= -12 \\
\Rightarrow EI \cdot \frac{d^3y}{dx^3} &= -12x + 72 \qquad [\mathbf{1}] \\
\Rightarrow EI \cdot \frac{d^2y}{dx^2} &= -6x^2 + 72x + M_A, \qquad [\mathbf{1}]
\end{aligned}$$

where we used that $R_A = \sum(\text{loads})$. Now we have $M(6) = 0$ [**1**] and $M(x) = EI \cdot y''(x)$:

$$\begin{aligned}
M(6) = EIy''(6) &= -6(36) + 72(6) + M_A \overset{!}{=} 0 \\
\Rightarrow M_A &= -216, \qquad [\mathbf{1}]
\end{aligned}$$

so that, and we will use $C_1 = C_2 = 0$:

$$\begin{aligned}
EI \cdot y''(x) &= -6x^2 + 72x - 216 \\
\Rightarrow EI \cdot y'(x) &= -2x^3 + 36x^2 - 216x + C_1 \qquad [\mathbf{1}] \\
\Rightarrow EI \cdot y(x) &= -\frac{x^4}{2} + 12x^3 - 108x^2 + C_1x + C_2 \qquad [\mathbf{1}] \\
\Rightarrow EI \cdot y(6) &= -\frac{6^4}{2} + 12(6)^3 - 108(6)^2 + 0 + 0 = -1944 \qquad [\mathbf{1}] \\
\Rightarrow y_{\max} &= -\frac{1944}{4EI}.
\end{aligned}$$

Marking Scheme: Winter 2014

Find, in terms of EI, the maximum deflection of a light cantilever beam of span 4 m with a U.D.L. of 18 kN m^{-1} between the points $x = 1$ and $x = 2$ m and a point load of 24 kN at $x = 3$ m.

The deflection, $y(x)$, is the solution the fourth order differential equation

$$EI \cdot \frac{d^4y}{dx^4} = -w(x).$$

Equivalently, the deflection can be found by first finding the bending moment, $M(x)$, using Macauley's method, or by solving the second order differential equation:

$$\frac{d^2M}{dx^2} = -w(x).$$

The deflection, $y(x)$, at any point on the beam is then found by solving the differential equation

$$EI \cdot \frac{d^2y}{dx^2} = M(x).$$

[11 Marks]

Solution: We have $y(0) = 0 \Rightarrow C_2 = 0$ [**1**], $y'(0) = 0 \Rightarrow C_1 = 0$ [**1**], $V(4) = 0 \Rightarrow R_A = 18 \times 1 + 24 = 42$ kN [**1**]. If they want, students can take $M_A = -(18(1.5) + 24(3)) = -99$. We have

$$w(x) = 18u(x-1) - 18u(x-2) + 24\delta(x-3),$$

so we look at

$$\begin{aligned}
EI\cdot\frac{d^4y}{dx^4} &= -18u(x-1)+18u(x-2)-24\delta(x-3) \qquad \mathbf{[1]}\\
\Rightarrow EI\cdot\frac{d^3y}{dx^3} &= -18[x-1]+18[x-2]-24u(x-3)+R_A \qquad \mathbf{[1]}\\
\Rightarrow EI\cdot\frac{d^2y}{dx^2} &= -9[x-1]^2+9[x-2]^2-24[x-3]+42x+M_A, \qquad \mathbf{[1]}
\end{aligned}$$

where we used that $R_A=\sum(\text{loads})$. Now we have $M(4)=0$ **[1]** and $M(x)=EI\cdot y''(x)$:

$$\begin{aligned}
M(4)=EIy''(4)&=-9(3)^2+9(2)^2-24(1)+42(4)+M_A\overset{!}{=}0\\
&\Rightarrow M_A=-99, \qquad \mathbf{[1]}
\end{aligned}$$

so that, and we will use $C_1=C_2=0$:

$$\begin{aligned}
EI\cdot y''(x) &= -9[x-1]^2+9[x-2]^2-24[x-3]+42x-99\\
\Rightarrow EI\cdot y'(x) &= -3[x-1]^3+3[x-2]^3-12[x-3]^2+21x^2-99x+C_1 \qquad \mathbf{[1]}\\
\Rightarrow EI\cdot y(x) &= -\frac{3}{4}[x-1]^4+\frac{3}{4}[x-2]^4-4[x-3]^3+7x^3-\frac{99}{2}x+C_1+C_2 \qquad \mathbf{[1]}\\
\Rightarrow EI\cdot y(4) &= -\frac{3}{4}(3)^4+\frac{3}{4}(2)^4-4+7(4)^3-\frac{99}{2}(16)=-1587 \qquad \mathbf{[1]}\\
\Rightarrow y_{\max} &= -\frac{1587}{EI}.
\end{aligned}$$

INTERLEAVED Exercises

1. Find the maximum deflection of a light cantilever beam of span 6 m with a constant load of 60 kN m^{-1}. **Ans:** $-\dfrac{9720}{EI}$
2. **Autumn 2010** A light beam of span 4 m is simply supported at its endpoints. At the points $x=1$ m and $x=3$ m, there are point loads of 36 kN and 72 kN, respectively.
   i. By solving $M''(x)=-w(x)$, or otherwise, express the Bending Moment $M(x)$ in terms of step functions.
   ii. Solve the differential equation
   $$EI\frac{d^2y}{dx^2}=M(x)$$
   to find the deflection y at any point.
   Ans: $y(x)=\dfrac{1}{EI}\left(-6[x-1]^3-12[x-3]^3+7.55x^3-76.5x\right)$.
3. Find the maximum deflection of a light cantilever beam of span 6 m with a uniformly distributed load of 48 kN m^{-1} between $x=3$ m and $x=6$ m. **Ans:** $-\dfrac{6642}{EI}$.
4. **Winter 2019** Find, in terms of EI, the maximum deflection of a light cantilever beam of span 6 m with two point loads: one of magnitude 18 kN, at $x=1$ m, and another of magnitude 36 kN, at $x=4$ m. **Ans:** $-\dfrac{1395}{EI}$

5. **Winter 2017** A light beam of span 6 m, fixed at both ends, carries a uniformly distributed load of 18 kN m^{-1} between $x = 1$ and $x = 4$ m.

   (a) Using the fact that the ends of the beam are fixed, write down the values of A, B, and C:

   The bending moment at the ends is not necessarily equal to $\underline{\ A\ }$. *In particular,* M_A *is not equal to* $\underline{\ A\ }$. *The slope at* $x = 0$ *and* $x = 6$ *m is equal to* $\underline{\ B}$, *and the deflection at* $x = 0$ *and* $x = 6$ *m is equal to* $\underline{C}$.

   (b) Explain using the physics/geometry/engineering of the situation why $R_A > R_B$.

   (c) By solving the differential equation

   $$\frac{d^2M}{dx^2} = -w(x),$$

   where $w(x)$ is the load per unit length, find the bending moment $M(x)$ in terms of R_A and M_A. $M(x) = -9[x-1]^2 + 9[x-4]^2 + R_Ax + M_A$

   (d) The deflection, $y(x)$, at any point on the beam is found by solving the differential equation

   $$EI\frac{d^2y}{dx^2} = M(x).$$

   Solve the differential equation for $y(x)$.
   Ans: $y(x) = \dfrac{1}{EI}\left(-\dfrac{3}{4}[x-1]^4 + \dfrac{3}{4}[x-4]^4 + \dfrac{265}{48}x^3 - \dfrac{327}{16}x^2\right)$

   (e) At $x_1 \approx 2.892$, $y'(x_1) = 0$. What can you conclude?

6. **Winter 2015** Find, in terms of EI, the maximum deflection of a light cantilever beam of span 6 m with a U.D.L. of 24 kN m^{-1} between the points $x = 4$ and $x = 5$ m and a point load of 18 kN at $x = 2$ m. **Ans:** $-\dfrac{1287}{EI}$

7. **Autumn 2020** Find, in terms of EI, the maximum deflection of a light cantilever beam of span 4 m with a linearly varying load, varying from 18 kN m^{-1} at $x = 0$, to 34 kN m^{-1} at $x = 4$. **Ans:** $-\dfrac{1484.8}{EI}$

8. **Summer 2008** A light beam of span 6 m has both ends embedded in walls. At the point $x = 2$ m there is a load of 36 kN. Between the points $x = 4$ m and $x = 6$ m there is a U.D.L. of 18 kN m^{-1}.

   i. By solving $M''(x) = -w(x)$, or otherwise, express the bending moment $M(x)$ in terms of step functions, the reaction force $M'(0) = R_A$ and the bending moment $M(0) = M_A$.

   ii. Solve the differential equation $EI\dfrac{d^2y}{dx^2} = M(x)$ to find the deflection y at any point on the beam. **Ans:** $y(x) = \dfrac{1}{EI}\left(-6[x-2]^3 - \dfrac{3}{4}[x-4]^4 + 5x^3 - 19x^2\right)$

2.3 Beam Equations Summary

The deflection of a beam, $y(x)$, is the solution the fourth order differential equation

$$EI \cdot \frac{d^4y}{dx^4} = -w(x),$$

where $w(x)$ is the load per unit length.

Equivalently, the deflection can be found by first finding the bending moment, $M(x)$, using Macauley's method, or by solving the second order differential equation:

$$\frac{d^2M}{dx^2} = -w(x).$$

The deflection, $y(x)$, at any point on the beam is then found by solving the differential equation

$$EI \cdot \frac{d^2y}{dx^2} = M(x).$$

Note that

1. *Point Loads* are of the form $w(x) \sim \delta(x-a)$,
2. *Uniformly Distributed Loads* are of the form $w(x) \sim u(x-a)$,
3. *Linear(ly varying) Loads* are of the form $w(x) \sim mx + c$.

Note that

$$\int \delta(x-a)\,dx = u(x-a) + C,$$

$$\int u(x-a)\,dx = [x-a] + C,$$

$$\int [x-a]^n\,dx = \frac{1}{n+1}[x-a]^{n+1} + C.$$

Note that there are four anti-differentiations, hence four constants of integration and hence four boundary conditions are required. First we detail at what point what constants of integration are picked up:

$$\begin{array}{llll}
 & EI \cdot \dfrac{d^4y}{dx^4} = \dfrac{d^2M}{dx^2} & = -w(x) & \text{pick up the constant} \\
\underset{\int \cdot\, dx}{\longrightarrow} & EI \cdot \dfrac{d^3y}{dx^3} = \dfrac{dM}{dx} & = V(x) & R_A \\
\underset{\int \cdot\, dx}{\longrightarrow} & EI \cdot \dfrac{d^2y}{dx^2} = M(x) & & M_A \\
\underset{\int \cdot\, dx}{\longrightarrow} & EI \cdot \dfrac{dy}{dx} & & C_1 \\
\underset{\int \cdot\, dx}{\longrightarrow} & EI \cdot y(x) & & C_2
\end{array}$$

Now we present the boundary conditions that allow you to calculate these constants:

	$y(0)$	$y(L)$	$y'(0)$	$y'(L)$	$y''(0)/M(0)$	$y''(L)/M(L)$
SS	0	0	<0	>0	0	0
$\Rightarrow$	$C_2=0$	C_1	$C_1<0$	-	$M_A=0$	R_A
FE	0	0	0	0	<0	-
$\Rightarrow$	$C_2=0$	R_A/M_A	$C_1=0$	R_A/M_A	$M_A<0$	-
C'ver	0	<0	0	<0	<0	0
$\Rightarrow$	$C_2=0$	-	$C_1=0$	-	$M_A<0$	M_A

- $C_2 = 0$ for ALL.
- $C_1 = 0$ for FE and C — but NOT SS.
- $M_A = 0$ for SS ONLY.

2.4 Numerical Solution of First Order Differential Equations

Motivation

Consider the following load on a fixed-end beam:

Figure 2.24: Yes this is a solid version of the McDonald's Logo.

Suppose you want a formula for the maximum deflection, δ, of such a beam under such a load. The formula should depend on the length of the beam L, and the properties of the beam used: these are given by the Young's Modulus E and the second moment of area I. The material comprising the logo can be captured by a load per unit length factor w. This means we are looking for a formula of the form:

$$\delta = \delta(w, L, E, I),$$

Theoretically, this isn't that difficult a problem. Find the load per unit length at each point along the beam, $w(x)$, and solve the fourth order **differential equation**

$$EI \cdot \frac{d^4y}{dx^4} = -w(x).$$

We have a problem... we don't have a formula for $w(x)$...

One thing we could do is to use the methods of Chapter 1 to fit a curve to the data and antidifferentiate that. Assuming the material is homogenous (same density throughout), the load per unit length will be proportional to the height.

So measure the height of the logo at perhaps 10 cm intervals:

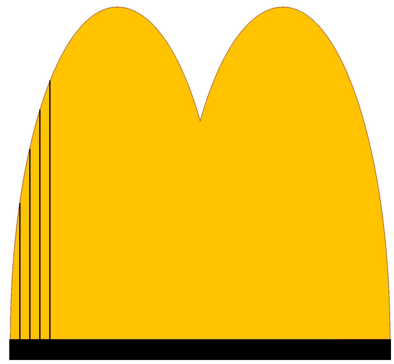


Figure 2.25: We measure some (x, w) data.

Now what we can do is fit a curve to this data!

However, for a great many 'bespoke' loads this isn't a very natural approach. For example, a load might not have much symmetry nor pattern — unlike the McDonald's load. In this section we will see how to approximate the shear at *discrete* points, say $x = 0, 0.1, 0.2, 0.3, \ldots$ Suppose we are interested in a McDonald's load on a beam of span $L = 6$ m with a total load of $w_T = 190$ kN. As before, (x, w) data must be measured and recorded: an engineer goes off and makes a data set, say:

x/m	0.0	0.1	0.2	0.3	0.4	0.5	$\cdots$
$w(x)$/kN m^{-1}	0.0	3.9	7.6	11.1	14.4	17.5	$\cdots$

Now we want to find the shear force. Using the methods of this section we find that it might look something like:

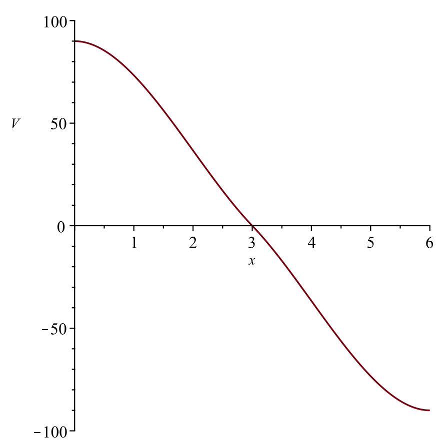


Figure 2.26: Here we see that $V(0) = 90 = R_A = w_T/2$ and that the shear is equal to zero at the midpoint $x = 3$. This implies that the maximum bending moment occurs there.

There is another possibility that poses a problem — that we have a load given by a function that we cannot anti-differentiate. This gives two problematic scenarios with the differential equation $V'(x) = -w(x)$:

- $w(x)$ is given by function which has no antiderivative.
- $w(x)$ is given by data rather than a formula.

As we said, it's impossible to solve $V'(x) = -w(x)$ in the sense of obtaining an explicit formula for the solution. In this section, we show that, despite the absence of an explicit solution, we can still learn a lot about the solution through a graphical approach (direction fields) or a numerical approach (Euler's Method and the Three-Term-Taylor Method). The rest of this section is focussed on $V'(x) = -w(x)$ but the theory applies to any first order differential equation.

2.4.1 Direction Fields

Suppose we are asked to sketch the graph of the shear given:

$$\frac{dV}{dx} = -w(x) = -5\cos(x^2/6) - 5\ ,\ \ V(0) = 20.$$

Note this means $V = 20$ when $x = 0$. We don't know a formula for the solution, so how can we possibly sketch its graph? Let's think about what the differential equation means. The equation $V' = -w(x)$ tells us that the slope at any point (x, V) on the graph of $V(x)$ is equal minus the load per unit length. In particular, because the curve passes through the point $(0, 20)$, its slope there must be

$$-5\cos(0) - 5 = -10.$$

So a small portion of the solution curve near the point $(0, 20)$ looks like a short line segment through $(0, 20)$ with slope -10:

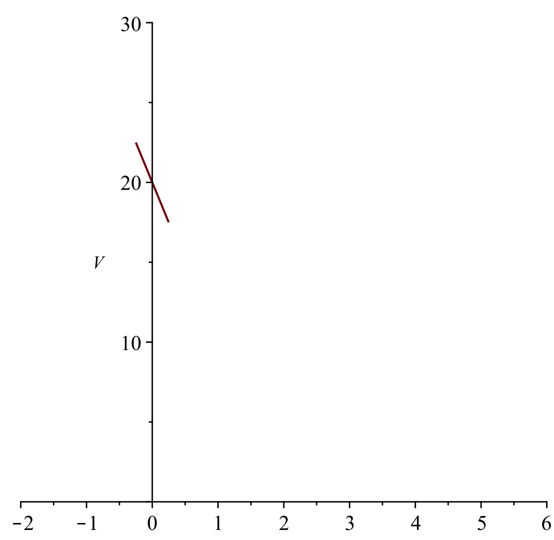


Figure 2.27: Near the point $(0, 20)$ — got from $V(0) = 20$ — the slope of the solution curve is -10.

As a guide to sketching the rest of the curve, let's draw short line segments at a number of points (x, V) with slope $-w(x)$. The result is called a *direction field*.

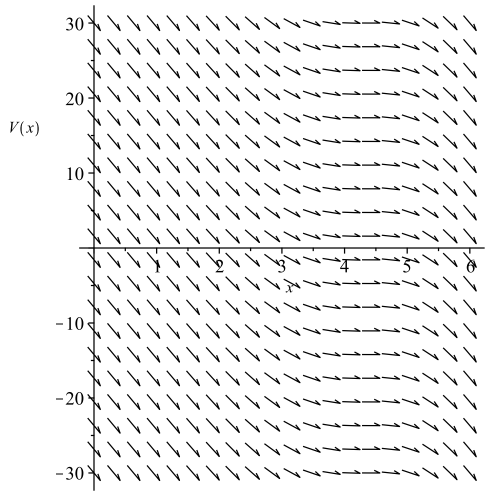


Figure 2.28: For example, the line segment at the point $(2, 5)$ has slope ≈ -8.929. The direction field allows us to visualise the general shape of the solution by indicating the direction in which the curve proceeds at each point.

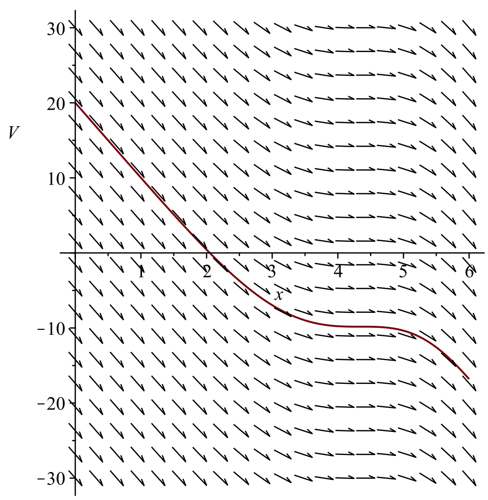


Figure 2.29: We can sketch the solution curve through the point $(0, 20)$ by following the direction field. Notice that we have drawn the curve so that it is parallel to nearby line segments.

2.4.2 Euler's Method

The basic idea behind direction fields can be used to find numerical approximations to solutions of differential equations. We illustrate the methods on the initial-value problem that we used to introduce direction fields:

$$\frac{dy}{dx} = -5\cos(x^2/6) - 5\ , \ \ V(0) = 20.$$

The differential equation tells us that $V'(0) = -10$, so the solution curve has slope -10 at the point $(0, 20)$. As a first approximation to the solution we could use the linear approximation $L(x) = -10x + 20$. In other words we could use the tangent line at $(0, 20)$ as a rough approximation to the solution curve.

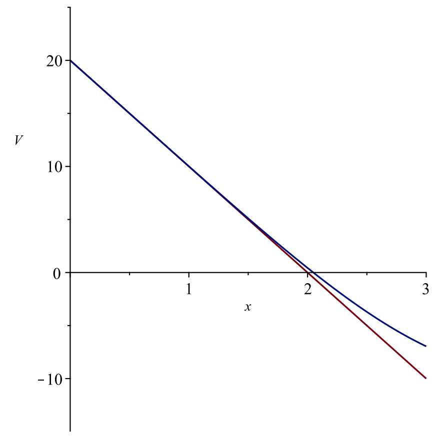


Figure 2.30: The tangent at $(0, 20)$ approximates the solution curve for values near $x = 0$.

Euler's idea was to improve on this approximation by proceeding only a short distance along this tangent line and then making a correction by changing direction according to the direction field:

Euler's method says to start at the point given by the initial value and proceed in the direction indicated by the direction field. Stop after a short time, look at the slope at the new location, and proceed in that direction. Keep stopping and changing direction according to the direction field. Euler's method does not produce an exact solution to the initial-value problem — it gives approximations. But by decreasing the step size (and therefore increasing the amount of corrections), we obtain successively better approximations to the correct solution.

For the general first-order initial-value problem $m = y'(x) = F(x, y)$, $y(x_0) = y_0$, our aim is to find approximate values for the solution at equally spaced numbers x_0, $x_1 = x_0 + h$,$x_2 = x_0 + 2h = x_1 + h$..., where h is the step size. The differential equation tells us that the slope at (x_0, y_0) is $y' = F(x_0, y_0)$:

This shows us that the approximate value of the solution when $x = x_1 = x_0 + h$ is

$$\begin{aligned} y_1 &= y_0 + h \cdot F(x_0, y_0) \\ &= y_0 + h \cdot y_0'. \end{aligned}$$

Similarly,

$$y_2 = y_1 + hF(x_1, y_1)$$

We will use the notation

$$y_k' := F(x_k, y_k).$$

Euler's Method

If

$$\frac{dy}{dx} = y'(x) = y' = F(x, y) \ , \ \ y(x_0) = y_0$$

is an initial value problem then the Euler Method using step size h gives approximations to $y(x_0 + kh) = y_k$:

$$\underbrace{y(x_{n+1})}_{\text{exact}} \approx \underbrace{y_{n+1}}_{\text{next}} = \underbrace{y_n}_{\text{previous}} + \overbrace{h}^{\text{step}} \cdot \underbrace{y'_n}_{\text{prev slope}} \tag{2.28}$$

for $n \geq 0$.

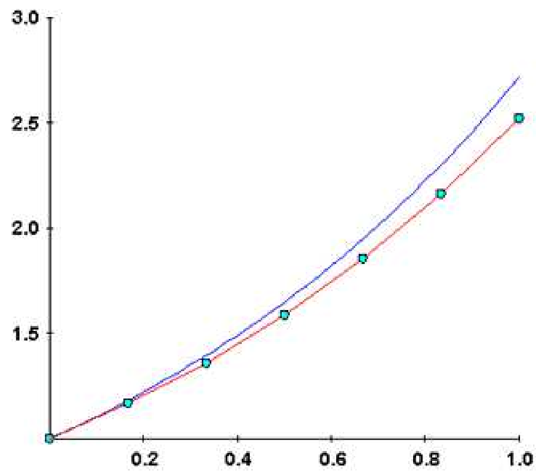


Figure 2.31: Euler's Method starts at some initial point (here $(x_0, y_0) = (0, 1)$), and proceeds for a distance h (in this plot $h = 1/6$.) at a slope that is equal to the slope at that point $y' = x_0 + y_0$. At the point $(x_1, y_1) = (x_0 + h, y_1)$, the slope is changed to what it is at (x_1, y_1), namely $x_1 + y + 1$, and proceeds for another distance h until it changes direction again. It turns out the error is $\propto h$: so that if you halve the step-size you halve the error. The exact value $y(1)$ is approximated by y_6, and we write $y(1) \approx y_6$.

Examples

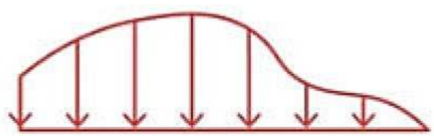
1. **Winter 2014** Suppose that if the loading on a beam of span L is given by $w(x) = 20e^{-x^2}$, then the shear, $V(x)$, is given as the solution of the initial value problem

$$\frac{dV}{dx} = -20e^{-x^2} \ , \ \ R_A = V(0) = 35.$$

Use a numerical method with a step size of $h = 0.5$ to estimate the value of $V(1)$, the shear at $x = 1$ m.

Solution: Where the rôle of y is played by V, and so $V'(x) = -20e^{-x^2}$, using the Euler Method we have

2. **Autumn 2020** Suppose that the loading on a beam of span 6 m looks like:



There is no formula for $w(x)$, and so a numerical method must be used to estimate e.g. the shear. The shear, $V(x)$, is given as the solution of the initial value problem

$$\frac{dV}{dx} = -w(x)\,, \quad V(0) = R_A.$$

From measurements know that $w(0) = 10$, $w(0.1) = 18$ and $R_A = 54$.

Use a numerical method with a step size of $h = 0.1$ to numerically estimate the value of $V(0.2)$, the shear at $x = 20$ cm.

Solution:

Further Remark: Implementing the Euler Method in Microsoft Excel

It is possible to show, although I don't do it here, that the solution above is not very accurate. To make it more accurate the step size, h, should be decreased dramatically. If, however, the step size is reduced from $h = 50$ cm to $h = 1$ cm, then the number of iterations required goes from two to 100. Therefore, in the real world, a software package is used to implement the Euler (and similar) Method(s).

Here we see how to solve the above problem more accurately using *Microsoft Excel.* Just so I can fit everything in on a screen shot, I use $h = 4$ cm. First of all you set up the table of x values: Next enter the initial value of $V(x)$, namely $V(0) = 35$. Also input (being ultra-careful with signs — Excel doesn't follow BEMDAS)

```
=-20*exp(-(A2^2))
```

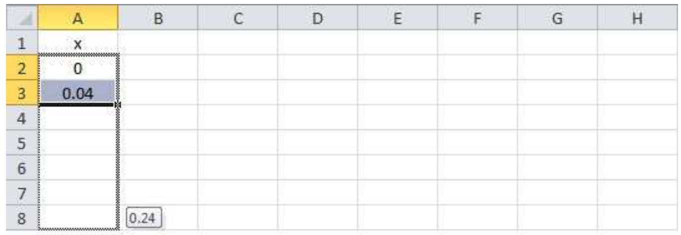


Figure 2.32: As per the question — which wants $V(1)$ — this should be continued down to $x = 1$ m.

into cell C2. Now, as shown, drag the formula the whole way down:

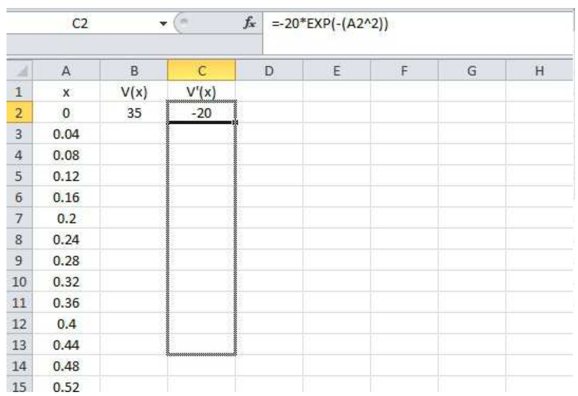


Figure 2.33: Now we need to implement the formula $V_{k+1} = V_k + h \cdot V'_k$.

Now in cell B3 input, following $y_{k+1} = y_k + h \cdot y'_k$:

```
=B2+0.04*C2
```

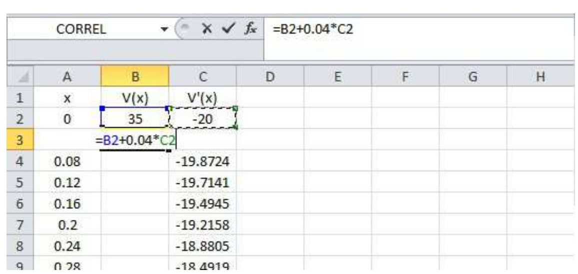


Figure 2.34: Now just drag down the B column to find $V(1) \approx 19.81$ kN. This is far more accurate than the 'two-stepper' done above.

Calculus Review VI: Implicit Differentiation

Suppose there is a derivative in terms of x, y given by:

$$\frac{dy}{dx} = F(x, y).$$

Can we find the second derivative in terms of x and y? There is a(n MATH6040) technique called *implicit differentiation*, which allows us to do this. Briefly, we implicitly say that $y = y(x)$, so when we differentiate, for example y^2, we must see it as $[y(x)]^2$ so it would have derivative:

$$\frac{d}{dx}[y(x)]^2 = 2(y) \cdot \frac{dy}{dx},$$

where we used the *Chain Rule*:

$$\frac{d}{dx}o(v(x)) = o'(v(x)) \cdot v'(x)$$

then if both sides are differentiated with respect to x then the equation reads:

$$\frac{d}{dx}\left(\frac{dy}{dx}\right) = \frac{d}{dx}(F(x,y)) \Rightarrow \frac{d^2y}{dx^2} = G\left(x, y, \frac{dy}{dx}\right)$$

Just remember:

- The derivative of y with respect to x is just dy/dx.
- The derivative of a function of y must be differentiated as a chain rule.

$$\frac{d}{dx}f(y(x)) = f'(y) \cdot \frac{dy}{dx}$$

- The product rule for a terms that are multiplied together.

2.4.3 Three Term Taylor Method — Parabolas rather than Lines

The Three-Term-Taylor Method is a more accurate method of approximating the solutions to ordinary differential equations. It uses parabolas rather than lines. Given a differential equation, we will be given the initial/boundary condition that $y(x_0) = y_0$. Suppose that we assume that our differential equation has an *analytic solution*[5] at x_0. Then we know that for points *close* to x_0 — so that $|x - x_0|$ small — we have

$$y(x) = \overbrace{\underbrace{y(x_0) + y'(x_0)(x - x_0)}_{\text{tangent}} + \frac{y''(x_0)}{2}(x - x_0)^2}^{\text{tangent parabola}} + \frac{y'''(x_0)}{3!}(x - x_0)^3 + \cdots$$

An example:

$$e^x = 1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3 + \underbrace{\frac{1}{24}x^4 + \cdots}_{\approx 0 \text{ for } x \approx 0}.$$

When we are particularly close to x_0, in particular $|x - x_0| < 1$, then the higher powers of $(x - x_0)$ in this *Taylor Series expansion* will be getting smaller and smaller and if we want we just take the first three terms as an approximation:

$$y(x) \approx y(x_0) + y'(x_0)(x - x_0) + \frac{y''(x_0)}{2}(x - x_0)^2 \sim ax^2 + bx + c.$$

[5] more on this in Chapter 4

Usually what we want to do here is find the solution at multiples of a step size h here so what we usually do is write $x_1 - x_0 := h \Rightarrow x_1 = x_0 + h$ and so

We can find the value of $y'(x_0)$ because we know that $y'(x_0) = F(x_0, y_0)$. Do we know what $y''(x_0)$ is? Well $y''(x)$ is the derivative with respect to x of $y'(x)$ — morryah the derivative of y' with respect to x. We assume that y depends on x so we write $y'(x) = y'$ and differentiate implicitly. Therefore we may write

$$y_1 := y(x_0 + h) \approx y(x_0) + y'(x_0)h + \frac{y''(x_0)}{2}h^2.$$

Like Euler's method we can compute an approximate solution from $y(x_0) = y_0$ using step size h using:

$$\underbrace{y(x_0 + (k+1)\cdot h)}_{\text{exact}} \approx y_{k+1} = \overbrace{y_k + y_k' h}^{\text{Euler}} + y_k'' \frac{h^2}{2} \tag{2.29}$$

where the notation is the same to that of the Euler Method. We can see that the Euler Method is a *Two* Term Taylor Method, and we must understand that the Three Term Taylor Method is more accurate.

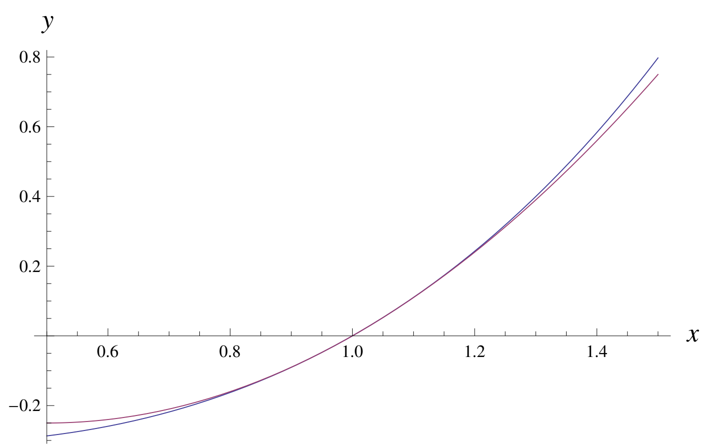


Figure 2.35: The Three Term Taylor Method takes advantage of the fact that smooth functions are locally very-well approximated by parabolas. For example, here we see that $2e^{x-1} - x - 1$ is well-approximated by a parabola near $(1, 0)$.

Example: Winter 2019

Suppose that the loading on a beam of span 6 m is given by $w(x) = 20e^{-0.1x}$ so that the shear, $V(x)$, is given by the solution of the initial value problem

$$\frac{dV}{dx} = -20e^{-0.1x}, \quad V(0) = 49.6.$$

i. Use a numerical method with a step size of $h = 0.1$ to find V_2, an estimate, correct to *five* significant figures, of the value of $V(0.2)$, the shear at $x = 20$ cm.

ii. The exact solution of the initial value problem is $V(x) = 200e^{-0.1x} - 150.4$. Hence find, correct to *five* significant figures, the exact value of $V(0.2)$.

iii. Find the percentage error in the approximation $V_2 \approx V(0.2)$.

Solution:

i. Using Euler's Method:

ii.

1. We calculate:

Figure 2.36: The following are $h = 0.5$ Euler and Three-Term-Taylor Method Approximations to the exact solution. The Three-Term-Taylor Method is almost indiscernable from the exact solution for this problem.

Exercises

1. **Winter 2016** Suppose that a light beam of span 6 m carries a symmetric load of total weight 100 kN. Then the shear, $V(x)$, is given as the solution of the initial value problem

$$\frac{dV}{dx} = -w(x)\,, \ \ V(0) = 50.$$

Suppose that an analytic formula for $w(x)$ is unknown but it is known that $w(0) = 5$ kN m^{-1} and $w(0.5) = 10$ kN m^{-1}. Use a numerical method with a step size of $h = 0.5$ to estimate the value of $V(1)$, the shear at $x = 1$ m. Use four significant figures for intermediate calculations. **Ans:** 42.5 kN.

2. **Winter 2015** Suppose that the loading on a beam of span 6 m is given by

$$w(x) = 10\cos\left(\frac{(x-3)^2}{2\pi}\right).$$

Then the shear, $V(x)$, is given as the solution of the initial value problem

$$\frac{dV}{dx} = -10\cos\left(\frac{(x-3)^2}{2\pi}\right)\,, \ \ V(0) = 24.4.$$

Use a numerical method with a step size of $h = 0.1$ to estimate the value of $V(0.2)$, the shear at $x = 20$ cm. For those using the Three Term Taylor Method, you may use

$$\frac{d}{dx}\left(-10\cos\left(\frac{(x-3)^2}{2\pi}\right)\right) = +\frac{(x-3)}{\pi}\sin\left(\frac{(x-3)^2}{2\pi}\right).$$

[HINT: Use RADIANS not degrees. Use four significant figures for intermediate calculations.]
Ans: Using Euler's Method. 24.03 kN

3. **Autumn 2017** Consider the following load on a light, simply supported beam of span 6 m:

$$w(x) = 10 \cdot e^{\sin(\pi x/6)}.$$

This function doesn't have an elementary anti-derivative. Therefore, the shear force at x, $V(x)$, the solution to

$$\frac{dV}{dx} = -10 \cdot e^{\sin(\pi x/6)}\,; \ \ V(0) = 59.29,$$

has to be calculated using a numerical method. Use Euler's Method with a step size of $h = 0.1$ to numerically estimate the value of $V(0.2)$, the shear at $x = 20$ cm. Use RADIANS and not degrees. **Ans:** 57.24 kN.

4. **Autumn 2016** The shear, $V(x)$, is given as the solution of the initial value problem

$$\frac{dV}{dx} = -w(x)\,, \ \ V(0) = R_A.$$

Suppose that we don't have an analytic formula for $w(x)$ but know that $w(0) = 10$, $w(0.1) = 20$ and $R_A = 54$. Use Euler's Method with a step size of $h = 0.1$ to numerically estimate the value of $V(0.2)$, the shear at $x = 20$ cm. **Ans:** 51 kN.

Chapter 3

Probability & Statistics

The probable is what usually happens.

Aristotle.

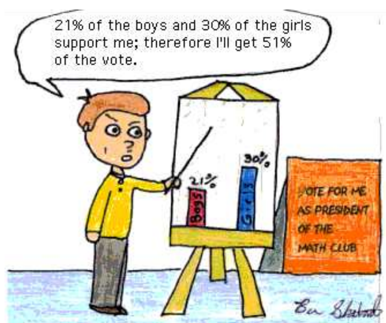


Figure 3.1: More than anything else in the world, people are crap at statistics.

3.0.1 Motivation

In the last chapter, and in later in Chapter 4, we never measured *how* accurate our approximate solutions were. However *statistics*, particularly sampling theory, can often tell us when our approximations are good enough. For example, suppose we have a business which constructs a machine component. Suppose the company ordering the component wishes to know what 'stress-level' the component can take. Due to natural variations some samples will have a larger tolerance than others — so how can we approach the business and say that our components can take a stress level of S? In practise we can't, but we can make statements along the line of:

On average, our components can withstand a stress-level of S.

However we can't go around testing every single one of the components produced. So what we do is we take a sample of 100 or 1,000 of these components away and have them tested. In this module we will see that we can be 'quite' confident that the average ability to withstand stress of all the components we produce is very well estimated by the sample average. *Sampling Theory* makes precise this idea.

3.1 Random Variables

The central concept of probability theory is that of a *random variable*. Examples of random variables include:

1. the outcome of a coin flip or dice roll
2. the outcome of a random selection of a card from a deck
3. the number of plants which grow to maturity in a glass house
4. the time spent on hold when calling the tax office
5. the height of an Irish male chosen at random (*normal distribution*)

Associated to a random variable there is a *probability distribution*. A good way to introduce a random variable is to write $X := \cdots$ (X is defined equal to $\cdots$) and say $X \sim \cdots$ (X is distributed as $\cdots$). The technical definition of a probability distribution is actually quite difficult and for us the following definition will have to do:

Definition

The probability distribution $\mathbb{P}$ of a random variable X is a function

$$\mathbb{P} : \{\text{'events' concerning } X\} \to [0,1]. \tag{3.1}$$

For example, if X is a coin toss, then $\{\text{all possible outcomes concerning } X\} = \{H, T, \text{none}, \text{either}\}$, and we have $\mathbb{P}[\text{neither}] = 0$, $\mathbb{P}[H] = \mathbb{P}[T] = \frac{1}{2}$, $\mathbb{P}[\text{either}] = 1$.

In a 'finite' system we usually define the probability of A occurring

$$\mathbb{P}[A] = \mathbb{P}[X \in A] = \frac{\#\text{ ways that } A \text{ can happen}}{\text{total } \#\text{ of outcomes}} \tag{3.2}$$

if each outcome is equally likely.

Remarks

1. in a finite system[1], outcomes that *never* happen are assigned a probability of 0 and outcomes that *always* happen are assigned a probability of 1.

   Examples:

   (a) Let X = a dice roll:

   (b) Let Y = a five card poker hand:

2. suppose that A and B are mutually exclusive events. This means that A and B can't happen at the same time. Then

$$\mathbb{P}[A \text{ or } B] = \mathbb{P}[A] + \mathbb{P}[B] \tag{3.3}$$

(OR $\rightsquigarrow$ +)

[1] with an infinite number of possible outcomes things are a little trickier

Example: Let X = a random selection from a deck of cards. Now

3. let A be some outcome. Now either A occurs or not-A *always* happens — and also these events are mutually exclusive:

   Example: Let X = three coin flips. What is the probability of getting at least one tail? Getting at least one tail is the same as not-(HHH). Therefore

4. when A and B are *independent* outcomes (the occurrence of one outcome makes it neither more nor less probable that the other occurs) we have

$$\mathbb{P}[A \text{ and } B] = \mathbb{P}[A]\mathbb{P}[B]. \tag{3.4}$$

   Example: Draw a card X_1 from the deck and replace it at random. Now draw another card X_2. What is the probability of two aces; $X_1 = X_2 = A$? What is the probability of getting two aces if select without replacement?

   Solution: If we select an ace the first time the chances of selecting an ace again (with replacement) are equal:

   However if we select *without replacement* the probability of a second ace is diminished and instead we must look at the probability:

$$\mathbb{P}[X_1 = A \text{ and } X_2 = A] = \mathbb{P}[X_1 = A] \cdot \mathbb{P}[X_2 = A | X_1 = A] \tag{3.5}$$

   where $\mathbb{P}[A|B]$ is read *the probability of A given that B is true.* Then

Another example of *non*-independents would be the event of Dublin winning the All-Ireland and Ciarán Kilkenny winning the Man of the Match in the final.

Example

A deck of 52 cards contains 13 cards in each of the spade (♠), heart (♡), diamond (♢) and club (♣) suits. What is the probability that two cards dealt randomly from the deck will both be spades?

Solution: First off it is not true that $P[\spadesuit\&\spadesuit] = P[\spadesuit]P[\spadesuit]$ as choosing a spade the second time around is less likely when one has already been taken out of the deck.:

3.1.1 Outcomes, Events, Probabilities, Random Variables

The usual way mathematicians think about probability is something like the following. First of all we have a set of *outcomes*, Ω. Say for a dice:

$$\Omega = \{1, 2, 3, 4, 5, 6\}.$$

Then we have a set of *events*, $\mathcal{F}$, which can basically thought of as combinations of outcomes. So for a coin, where $\Omega = \{H, T\}$, the events are, roughly:

$$\mathcal{F} = \{\text{nothing happens}, H, T, H \text{ or } T\}.$$

We have a *probability*, $\mathbb{P} : \mathcal{F} \to [0, 1]$, which assigns to each event a probability between zero and one. So for example, for a single coin toss:

$$\mathbb{P}[\{H\}] = \frac{1}{2}.$$

We can consider *random variables*, which are functions $X : \Omega \to E$, where E is said to be a *measurable* space. For example, the set $[36] = \{1, 2, \dots, 36\}$ is a measurable space and where $\Omega = \{(1, 1), (1, 2), \dots, (6, 6)\}$ is the set of outcomes when two dice are rolled

$$X(\omega) = \text{multiply together the two dice scores}$$

is a random variable. The probability that $X \in S$ is given by:

$$\mathbb{P}[X \in S] = \mathbb{P}\left[\underbrace{\{\omega \in \Omega \colon X(\omega) \in S\}}_{\in \mathcal{F}}\right].$$

For example, where

$$X : \{(1, 1), (1, 2), \dots, (6, 6)\} \to [36],$$

is the product, e.g. $X((3, 4)) = 3 \times 4 = 12$,

$$\mathbb{P}[X \in \{30\}] = \mathbb{P}\left[\{(5, 6), (6, 5)\}\right] = \frac{2}{36}.$$

Examples

1. **Winter 2019** Consider two factories, A, and B, run by two different companies, using different raw materials, different production methods, in different parts of the world. Consider the following *random variables*: the lengths, ℓ_1, ℓ_2, of two randomly selected rods from the same batch produced by Factory A; and the length, ℓ_3, of a randomly selected rod from Factory B. Consider the events E_1, E_2, E_3, that, respectively, $\ell_1 < 8$ m, $\ell_2 < 8$ m, and $\ell_3 < 8$ m.

   i. Are events E_1 and E_2 *independent*? Justify your answer.

   ii. Are events E_1 and E_3 *independent*? Justify your answer.

   Solution: E_1 and E_2 are *not* independent because:

   E_1 and E_3 are independent because:

2. **Autumn 2020** Consider the *random variable*: the length, ℓ, of a randomly selected rod produced by a factory. Consider the events E_1, E_2, E_3, that, respectively, $\ell < 8$ m, $\ell \neq 8$ m, and $\ell > 8$ m.

   (a) Are events E_1 and E_2 *mutually exclusive*? Justify your answer.

   (b) Are events E_1 and E_3 *mutually exclusive*? Justify your answer.

   Solution: E_1 and E_2 are *not* mutually exclusive because:

   E_1 and E_3 are mutually exclusive because:

3.2 Normal Distribution

A machine is filling a very fine powder into containers. Let X = the fill put into the next one. Note that X is usually *NOT* a whole number. Due to the variability of the machine, the value of X is not certain. A *continuous random variable* is a random variable that can take on any value in a continuous range. For a continuous random variable there are just too 'many' possible outcomes and all the probabilities $\mathbb{P}[X = k]$ are zero. Hence we must change tack.

Suppose that in a particular population of males the following distribution of heights are found:

Height/cm	<150	150-160	160-170	170-180	180-190	190-200	>200
Relative Frequency	0.03	0.1	0.2	0.33	0.2	0.1	0.03

This data could be summarised in a relative frequency graph:

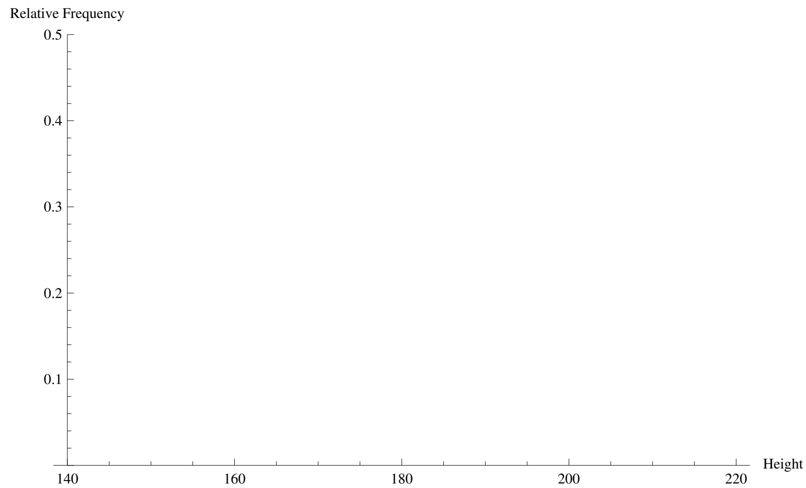


Figure 3.2: The heights of males in a population.

What if we make the bins narrower?

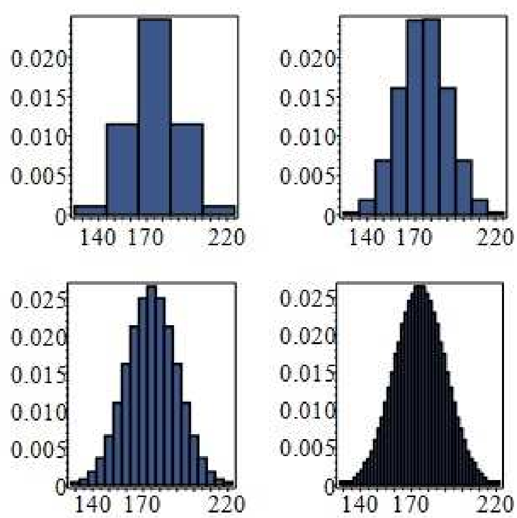


Figure 3.3: Bin widths of 20, 10, 5, and 2.

What we speculate is that there is a probability density function that can tell us not only how what proportion of people have a height between 170 and 180 cm but between 170 and 172 cm or any other interval:

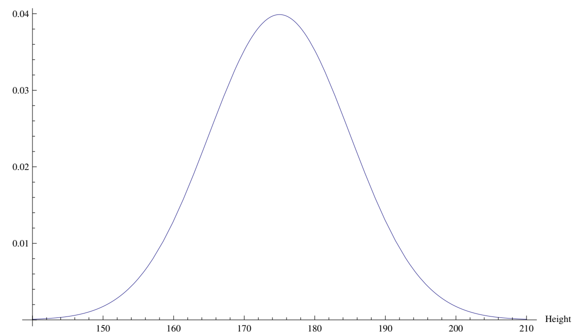


Figure 3.4: The *probability density function* of the heights. From above, we have 0.33 of heights $\approx$ 33% of area. The total area is one. The relative frequency is $f(h) \propto e^{-(h-175)^2} = \dfrac{1}{e^{(h-175)^2}}$.

Many examples of continuous data are unimodal (peaked, bell-shaped) and symmetric (about the peak) and can be shown to be of a certain form. Such data is called *normal* and we say it has a *normal distribution*. By and large, data with a dominant average with deviations from the mean just as likely to be positive or negative tend to have this shape:

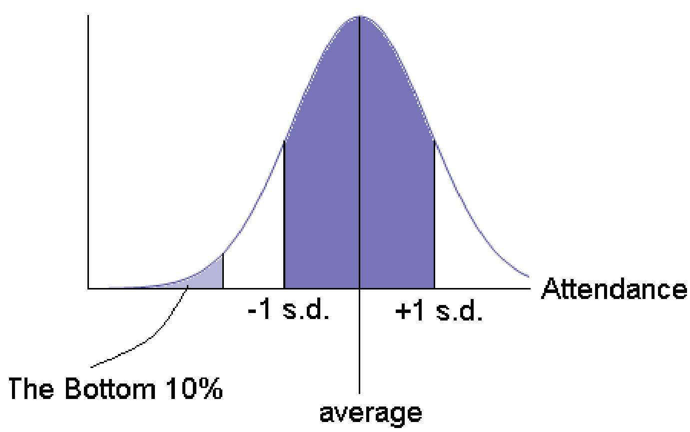


Figure 3.5: Note that for a normal distribution approximately two thirds of the data is found within a distance of one standard deviation from the average. There are two bits of information which summarise the data: the mean-average, μ (said *me-euh*); and the standard deviation, σ (said *sigma*) — a measure of the spread of the data. We write $N[\mu, \sigma]$ for the distribution.

Examples of random variables likely to conform to the normal distribution are:

1. A particular experimental measurement subject to several random errors.
2. Snowfall in a Norwegian town.
3. The heights of woman belonging to a certain population.
4. The length/mass/etc. of a manufactured product.

But be careful! Maybe snowfall isn't normally distributed. Wave height certainly isn't: read about *rogue waves*.

3.2.1 The z-Distribution

In practise, the normal distribution is very useful in that real-life calculations are very easy to handle because all normal distributions are related to each other in that all of the are rescaling of a particular normal distribution — the Universal Distribution if you will. This is the normal distribution with mean $\mu = 0$ and standard deviation $\sigma = 1$:

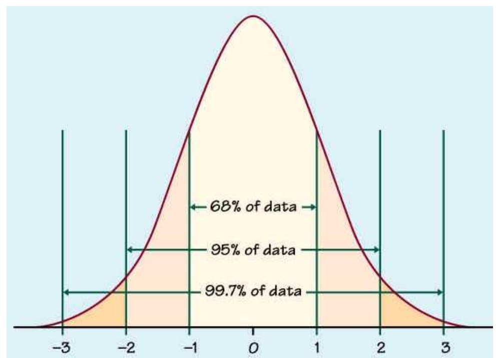


Figure 3.6: All normal distributions are just transforms of $z = N[0, 1]$. More specifically they all have *probability density function* $f(x) = \dfrac{1}{\sigma\sqrt{2\pi}}\exp(-(x-\mu)^2/2\sigma^2)$ — $N[0, 1]$ has this pdf with $\mu = 0$ and $\sigma = 1$.

This ease comes from the fact that we can transform from $X \sim N[\mu, \sigma]$ to $z \sim N[0, 1]$ by the following *z-transform*

$$z = \frac{X - \mu}{\sigma} \tag{3.6}$$

This is a transform that converts the X-distribution to a z-distribution (Note that $\mu \to 0$ by (3.6) as you would hope.):

We can also go from z- back to X- via

$$X = \mu + z\sigma \tag{3.7}$$

z is the number of standard deviations σ that X is away from μ: $z < 0$ for $X < \mu$ and vice versa. It is not clear why we are doing this but the following fact makes it all clear:

Fundamental Calculation of Normal Distributions

Suppose that $X \sim N[\mu, \sigma]$ and we want to calculate the probability

$$\mathbb{P}[x_1 \leq X \leq x_2].$$

Then we can transform $x_1 \to z_1$ and $x_2 \to z_2$ using (3.6); and in this case

$$\mathbb{P}[x_1 \leq X \leq x_2] = \mathbb{P}[z_1 \leq z \leq z_2] \tag{3.8}$$

Proof. Write the density as $f_{\mu,\sigma}$. We have, by definition,

$$\mathbb{P}[x_1 \leq X \leq x_2] = \int_{x_1}^{x_2} \frac{1}{\sigma\sqrt{2}} e^{-(x-\mu)^2/2\sigma^2}\, dx.$$

Do the 'u'-substitution $z = (x-\mu)/\sigma$. Note that

$$\frac{dz}{dx} = \frac{1}{\sigma} \Rightarrow \frac{dx}{\sigma} = dz,$$

and also $-(x-\mu)^2/2\sigma^2 = -z^2/2$. Also the limit $x_1 \to (x_1 - \mu)/\sigma = z_1$, and similarly $x_2 \to z_2$. Now implement to u-substitution to get

$$\begin{aligned}\mathbb{P}[x_1 \leq X \leq x_2] &= \int_{z_1}^{z_2} \frac{1}{1\sqrt{2\pi}} e^{-(z-0)^2)/2(1)^2}\, dz \\ &= \int_{z_1}^{z_2} f_{0,1}(z)\, dz = \mathbb{P}[z_1 \leq z \leq z_2] \quad \bullet\end{aligned}$$

Now how do we calculate $\mathbb{P}[z_1 \leq z \leq z_2]$? Well we can't right away but what we can do is integrate the frequency distribution of $N[0,1]$ from $-\infty$ to z_1 to calculate $\mathbb{P}[z \leq z_1]$... too much work, too difficult? Yes absolutely: that is why we use a table of values.

This z is the number of standard deviations from the mean.

Examples

1. What z value has a 20% probability of being exceeded?

   Solution: We are looking for the z_1 such that $\mathbb{P}[z \geq z_1] = 0.2$:

   Alternatively we can look at $\mathbb{P}[z \leq z_1] = 0.8$ and read off the table and get $z = 0.84$.

2. **Winter 2016** Due to natural variance, when a river engineer measures the ammonia level in a river sample, the results are not uniform. Historic records suggest that, when tested, the ammonia level in one river is normally distributed with a mean of 0.200 mg/L and a standard deviation of 0.040 mg/L. Assuming that the statistics have not changed, calculate the probability that a randomly selected sample from this river has an ammonia level

   i. less that 0.140 mg/L.
   ii. between 0.140 mg/L and 0.260 mg/L.

Solution: Let X = the ammonia level in the sample. $X \sim \mathrm{N}[0.2, 0.04]$.

i. We need to find $\mathbb{P}[X \leq 0.14]$. We find the corresponding z-value:

   Therefore we want $\mathbb{P}[z \leq -1.5]$. This is a short left tail. So by symmetry the same as the short right tail $\mathbb{P}[z \geq 1.5]$:

   Which is $1 - \mathbb{P}[z \leq 1.5] = 1 - 0.9332 = 0.0668$.

ii. From the previous part, and symmetry ($0.14/0.26 = 0.2 \pm 0.06$), we know:

$$\mathbb{P}[0.14 \leq X \leq 0.26] = \mathbb{P}[-1.5 \leq z \leq 1.5].$$

   We can find this by doing a long left tail minus a short left tail:

3. **Winter 2010** The lengths of pipes are assumed to be normally distributed with a mean value of 5 m and with a standard deviation of 0.003 m. What percentage of pipes have a length (i) less than 5.002 m and (ii) between 4.99 m and 5.01 m.

   Solution: Let $X \sim N[5, 0.003]$ be the length of a pipe.

   i. We wish to find

Now this is the same as $\mathbb{P}[X \leq 5.002] = \mathbb{P}[z \leq z_1]$ where

This can be read exactly from the tables:

ii. We wish to find

Now this is the same as $\mathbb{P}[4.99 \leq X \leq 5.01] = \mathbb{P}[z_1 \leq z \leq z_2]$ where

So we want $\mathbb{P}[-3.33 \leq z \leq 3.33]$. Let us draw a picture of this

Applied Example: Allowing for Exceptional Snowfall in Norway

Assuming a normal distribution, suppose that the mean annual snowfall in a Norwegian town is 31 inches with a standard deviation of 11 inches (inch = 2.52 cm). How many inches of snow should a structure be able to withstand using the following rules:

1. 'one in two hundred years' snowfall?

2. 'one in one hundred years' of snowfall *times* 1.5?

Solution: First of all call the snowfall distribution by $X \sim N[31, 11]$.

The snowfall distribution looks like

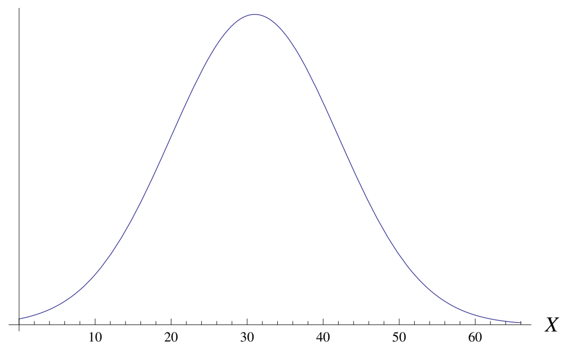


Figure 3.7: In approximately two-thirds of years, the annual snowfall is between 20 and 42 inches.

1. In this case we want to find an A such that $\mathbb{P}[X \geq A] = \frac{1}{200} = 0.005$.

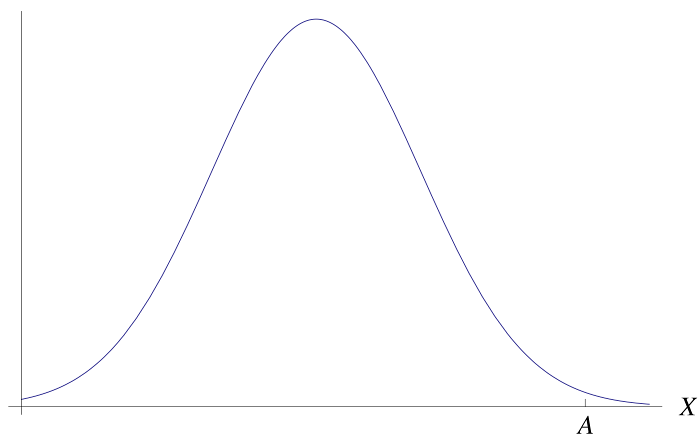


Figure 3.8: For part 1., we want to find an A such that there will be more than A inches of snowfall only every two hundred years.

   We of course transform this to a problem about $z \sim N[0,1]$ and we find from the tables $\mathbb{P}[z \geq 2.58] = 0.0049$ so our z value is 2.58 and so our 'one in two hundred year' snow event is

$$X = \mu + z\sigma = 31 + 2.58(11) = 59.38''.$$

   The engineer should makes the structure capable of withstanding 60 inches of snowfall.

2. This time we want to find an A such that $\mathbb{P}[X \geq A] = \frac{1}{100} = 0.01$. We transform this to a problem about $z \sim N[0,1]$ and we find from the tables $\mathbb{P}[z \geq 2.33] = 0.0099$ so our z value is 2.33 and so our 'one in one hundred year' snow event is

$$X = \mu + z\sigma = 31 + 2.33(11) = 56.63''.$$

   Under this rule, the engineer should makes the structure capable of withstanding $1.5 \times 56.63 \approx 85$ inches of snowfall. This is in fact a 'one in 200,000 years' event!! Perhaps the 1.5 rule is in place because the assumption that the snowfall is normally distributed is only that: an assumption. For example, the model gives a positive probability for *negative* snowfall — although this can only help our engineer.

Exercises: These answers have a health warning.

1. z is a random variable with standard normal distribution. Find the value of z_1 for which $\mathbb{P}[z > z_1] = 0.0808$. **Ans**: 1.4

2. z is a random variable with standard normal distribution. Find $\mathbb{P}[1 < z < 2]$. **Ans**: 0.136

3. z is a random variable with standard normal distribution. Calculate $\mathbb{P}[-2.13 < z \leq 1.46]$. **Ans**: 0.9113

4. z is a random variable with standard normal distribution. Find $\mathbb{P}[z < -0.46]$. **Ans**: 0.3228

5. **Winter 2012** The number of defective items in a given day might be normally distributed with a mean of 8.2 and a standard deviation of 2.12. If this is the case, calculate the probability of defective being

   i. greater than 10 a day **Ans**: 0.1977

   ii. between 7 and 9 a day **Ans**: 0.3614

6. **Autumn 2010** The weights have bricks are assumed to be normally distributed with a mean value of 1 kg and with a standard deviation of 0.006 kg. Calculate the percentage of bricks that weigh (i) less than 1.02 kg and (ii) between 0.99 kg and 1.01 kg. **Ans**: 0.9996, 0.9044,

7. **Autumn 2017** Assume that every third class road in the country gets resurfaced with a certain frequency. For example, say the L3504 in Cavan gets resurfaced every nine years and the L8282 in Wicklow gets resurfaced every five years. Assume further that the NRA claims that the period between resurfacing is normally distributed with a mean of 6 years and a standard deviation of 1.5 years. Calculate the probability that for a random third class road the period between resurfacing is

   i. less that four years. **Ans**: 0.0918

   ii. between four and eight years. **Ans**: 0.8164

---

8. **Autumn 2009** The masses of bricks are assumed to be normally distributed with a mean value of 1.02 kg and with a standard deviation of 0.006 kg. Calculate the percentage of bricks that have a mass greater than 1.033 kg. **Ans**: 0.0151

9. **Winter 2008** Cement is packed into bags with a mean weight of 50 kg and a standard deviation of 0.03 kg. Calculate the probability that a bag will weigh between 49.98 kg and 50.02 kg. **Ans**: 0.4971

3.3 Sampling Theory

Consider the problem of finding the mean-average length, μ, of all the metallic rods produced in a factory. Plainly this is impossible. However we could *approximate* this population mean-average by taking a random sample of say n rods from the population. This general process is called *sampling* and is used to make inferences of a whole *population* just by looking at a *random sample* of the data. Suppose we measure these rods to have lengths

$$l_1, l_2, \ldots, l_n.$$

We could then find the mean-average of the sample:

$$\bar{l} = \frac{l_1 + \cdots + l_n}{n} = \frac{1}{n}\sum_{i=1}^{n} l_i,$$

Now we could take the sample mean $\bar{l}$ as an estimate of the population mean μ; $\bar{l} \approx \mu$. How accurate is this? Consider all the possible samples we could have taken from the population:

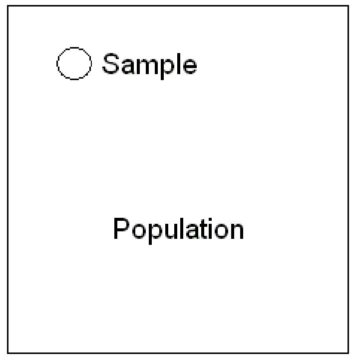


Figure 3.9: There are many, many ways of choosing n rods from a population of perhaps millions. If we look at the sample mean-average as a random variable, then the mean-average of the sample mean-averages is equal to the population mean-average.

Each random sample has a sample mean and so the sample mean (of a random sample) is a random variable. It turns out that *the average of the sample means is equal to the population mean.* That is for the example of the rods, $\bar{\bar{l}} = \mu$. We can show that the standard deviation — a measure of spread/deviation — of the sample means from the population mean is given by $\sigma/\sqrt{n}$, where σ is the standard deviation of the population and n is the sample size. That is two factors influence the accuracy of $\mu \approx \bar{x}$

- the sample size, n,
- the spread/deviation of the data, summarised by σ.

For the standard deviation we use ($\delta_i := |x_i - \mu|$ so $\delta_i^2 = (x_i - \mu)$):

$$\sqrt{\frac{\sum(x_i - \mu)^2}{N}} = \sqrt{\overline{\delta^2}} = \sigma \approx s = \sqrt{\frac{\sum(x - \bar{x})^2}{n-1}}. \tag{3.9}$$

The sample must however be *random*. For example, suppose that we want to estimate the number of smokers in CIT/Ireland by sampling from just this class. This is not a random sample (why?). Various sources cite the proportion of smokers in Ireland at about 20%.

Illustration

Suppose that we use a sample of size three to estimate the population mean ($\mu = 8$) and standard deviation ($\sigma = 2.19$) of the data[2]:

$$\{5, 7, 7, 10, 11\}.$$

[2]a far, far larger population size, and sample size small in comparison to it, would be far more realistic, but impossible to produce 'all' the numbers

Now there are $5\mathbf{C}3 = 10$ possible two-samples:

$(7, 10, 11)$, $(7, 10, 11)$, $(7, 7, 11)$, $(7, 7, 10)$, $(5, 10, 11)$, $(5, 7, 11)$, $(5, 7, 10)$, $(5, 7, 11)$, $(5, 7, 10)$, $(5, 7, 7)$

These have sample means

$$\approx 9.33,\ 9.33,\ 8.33,\ 8.00,\ 8.67,\ 7.67,\ 7.33,\ 7.67,\ 7.33,\ 6.33$$

The average of these[3] is 8... which is exactly the same as the population mean... so on average the sample mean is equal to the population mean so we have $\bar{x} \approx \mu$ because $\bar{x} = \mu$ *on average*.

Now let us calculate the standard deviations of the ten samples using the 'usual' formula:

$$\text{standard deviation} := \sqrt{\frac{\sum(x-\bar{x})^2}{n}} = \sqrt{\delta^2} \tag{3.10}$$

$$\approx \underbrace{1.70}_{\text{from }(7,10,11)},\ 1.70,\ 1.89,\ 1.41,\ 2.62,\ 2.49,\ 2.05,\ 2.49,\ 2.05,\ \underbrace{0.94}_{\text{from }(5,7,7)}.$$

These have an average of $\text{ave}(\sqrt{\delta^2}) = 1.94...$ but this *underestimates* the population standard deviation of ≈ 2.19. However if we use

$$s,\ \textit{sample}\text{ standard deviation} := \sqrt{\frac{\sum(x-\bar{x})^2}{n-1}} \tag{3.11}$$

on our ten samples we have '*sample*' standard deviations:

$$\underbrace{(7, 10, 11)}_{\bar{x}\approx 9.33} \to \sqrt{\frac{2.33^2 + 0.67^2 + 1.67^2}{3-1}} \approx 2.08,\ 2.08,\ 2.31,\ 1.73,\ 3.21,\ 3.06,\ 2.52,\ 3.06,\ 2.52,\ 1.15.$$

The average of these is ≈ 2.37 and is a much closer to $\sigma \approx 2.19$. As it happens there is no formula for a number k such that $k \approx \sigma$ in the sense that $k = \sigma$ on average. This s with the $n-1$ rather than n is about as good as we can do and we take $s \approx \sigma$ even though they aren't quite equal on average.

The sample means have what is known as a t-distribution (similar to the normal, bell-curve), and we might write $\bar{x} \sim t\left[\mu, \frac{\sigma}{\sqrt{n}}, n\right]$.

You may also come across the term *variance*. The variance is the square of the standard deviation. Similarly to the normal distribution, there is a table which, depending on the sample size, gives the percentage of the data which lies outside t standard deviations of the mean.

[3] up to rounding error

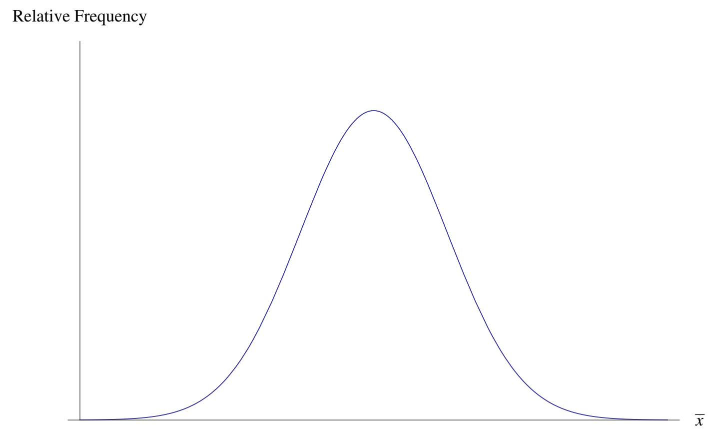


Figure 3.10: The most likely sample mean is equal to the population mean — the $\bar{x}$ on the x-axis are sample means. However, represented by the top and bottom tails, there are possibilities of choosing a sample with a particularly large or particularly small sample mean.

For example, suppose a sample of size $n = 10$ is taken from a population of standard deviation $\sigma \approx s$. Then 95% of the sample means $\bar{x}$ lie within 2.262 standard deviations of the (population mean):

$$\mu - 2.262\frac{\sigma}{\sqrt{10}} \leq \bar{x} \leq \mu + 2.262\frac{\sigma}{\sqrt{n}}.$$

This can also be presented as

$$\mu \pm t\frac{\sigma}{\sqrt{n}}.$$

t is a number of standard deviations — and depends on $n-1$ and a percentage p — the percentage of the $\bar{x}$ that are in this range.

Now we can be 95% confident that our sample mean is in this range:

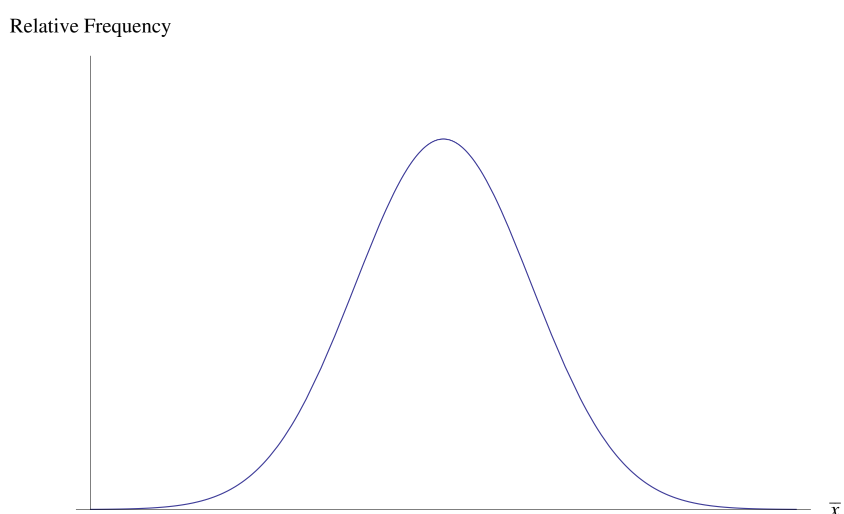


Figure 3.11: If the sample mean is within this distance of the population mean then, in turn, the population mean will be within this distance of the sample mean.

Note that in reality we have no idea what the standard deviation, σ, of the population is but it can be shown that the sample standard deviation provides a good estimate: $s \approx \sigma$.

Remark

This is a relatively subtle idea and although you can get by with a superficial understanding it would be better that you understand it a little deeper to help you with later studies. This analogy might help explain to you what a confidence interval is. Suppose that there is a laser pointing at a bullseye, and suppose that if you throw a dart 95% of the time you get within 10 cm of the bullseye:

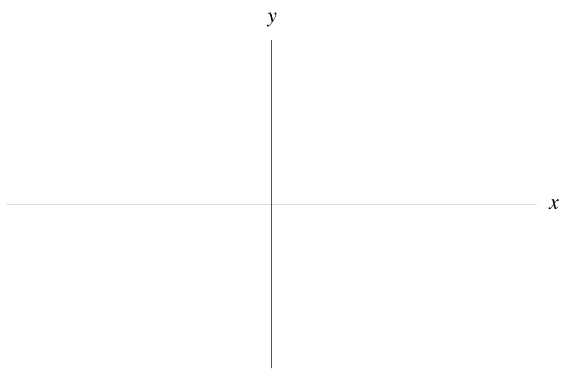


Figure 3.12: Suppose when you throw a dart that 95% of the time you get within 10 cm of the bull.

Formula: Confidence Interval for Small Sample Sizes

Suppose that a sample is taken from a population with a standard deviation of σ. If it is found that the sample mean is $\bar{x}$, then the following is the p% Confidence Interval for the population mean:

$$\bar{x} - t\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + t\frac{\sigma}{\sqrt{n}} \tag{3.12}$$

Here t depends on n and the level of confidence p, and is to be read from the t-distribution table. The population standard deviation, $\sigma \approx s$.

Note the interpretation:

> *On the p% confidence level, the population mean lies within the confidence interval.*

We are p% confident that the actual population mean is in this range.

Examples

1. **Summer 2012** In measuring the thickness (mm) of a certain material, the five values were recorded:

   $$9.98,\ 9.98,\ 10.01,\ 10.00,\ 10.03.$$

   Find a 99% confidence interval for the thickness.

   Solution: We must calculate the interval

   $$\bar{x} - t\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + t\frac{\sigma}{\sqrt{n}} \tag{3.13}$$

   We calculate the sample mean, $\bar{x}$:

   $$\bar{x} = \frac{9.98 + 9.98 + 10.01 + 10.00 + 10.03}{5} = 10.$$

We approximate $\sigma \approx s$. In the paper I gave the variance[4], and $s = \sqrt{s^2}$ but here we show how to calculate the sample standard deviation. We use

$$s = \sqrt{\frac{\sum_i (x_i - \bar{x})^2}{n-1}},$$

which will be given to you in a set of tables:

Now the t-value. Here we have $n = 5$ so the degrees of freedom are $n - 1 = 4$. For a 99% confidence interval we will want 1% in the tails... this yields the t-value $t = 4.604$:

$$\begin{aligned} \bar{x} - t\frac{\sigma}{\sqrt{n}} &\leq \mu \leq \bar{x} + t\frac{\sigma}{\sqrt{n}} \\ \Rightarrow 10 - 4.604\frac{0.02121}{\sqrt{5}} &\leq \mu \leq 10 + 4.604\frac{0.02121}{\sqrt{5}} \\ \Rightarrow 9.956 &\leq \mu \leq 10.04. \end{aligned}$$

We are 99% confident that μ is in the interval $[9.956, 10.04]$.

2. **Winter 2011** Six determinations were made about the density of a particular material

$$3.09,\ 3.11,\ 3.11,\ 3.11,\ 3.12,\ 3.12.$$

Find 99% confidence limits for the density.

Solution: We must calculate the interval

$$\bar{x} - t\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + t\frac{\sigma}{\sqrt{n}} \tag{3.14}$$

We calculate the sample mean, $\bar{x}$:

$$\bar{x} = \frac{3.09 + 3.11 + 3.11 + 3.11 + 3.12 + 3.12}{6} = 3.11.$$

We approximate $\sigma \approx s$ using

$$s = \underbrace{\sqrt{\frac{\sum_i (x_i - \bar{x})^2}{n-1}}}_{\text{in tables}} = \sqrt{\frac{\sum \delta_i^2}{n-1}},$$

[4] the variance is $s^2 \approx \sigma^2$

Now the t-value. Here we have $n = 6$ so the degrees of freedom are $n - 1 = 5$. For a 99% confidence interval we will want 1% in the tails... this yields the t-value $t = 4.032$:

$$\begin{aligned} \bar{x} - t\frac{\sigma}{\sqrt{n}} &\leq \mu \leq \bar{x} + t\frac{\sigma}{\sqrt{n}} \\ \Rightarrow 3.11 - 4.032\frac{0.01095}{\sqrt{6}} &\leq \mu \leq 3.11 + 4.032\frac{0.01095}{\sqrt{6}} \\ \Rightarrow 3.078 &\leq \mu \leq 3.142 \end{aligned}$$

We are 99% confident that μ is in the interval $[3.078, 3.142]$.

3.3.1 Confidence Intervals for Large Sample Sizes

In contrast to these confidence intervals we can generate far superior confidence intervals for larger sample sizes $n > 30$. We have the same analysis as before — except the t-distribution with a large n is well approximated by a normal distribution:

$$t[\mu, \sigma/\sqrt{n}, n] \underset{n\to\infty}{\longrightarrow} N[\mu, \sigma/\sqrt{n}] \text{ so that, roughly}$$

$$\bar{x} \sim N[\mu, \sigma/\sqrt{n}]. \tag{3.15}$$

The situation is that for $n > 30$, $t \approx z$ (which is like a t distribution with $n \to \infty$). This means that the sample means have a normal distribution with mean of μ, the population mean; and a standard deviation of $\sigma/\sqrt{n}$, where σ is the population standard deviation. When $n > 30$ and the data is not grouped, we cannot easily calculate the sample standard deviation and instead probably use a statistical software package. Following the analysis through yields:

Formula: Confidence Interval for Large Sample Sizes

Suppose that a sample of size $n > 30$ is taken from a population with a standard deviation of σ. If it is found that the sample mean is $\bar{x}$, then the following is the p% Confidence Interval for the population mean:

$$\bar{x} - z\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + z\frac{\sigma}{\sqrt{n}} \tag{3.16}$$

Here z depends on the level of confidence p (but not the sample size), and can be read from the bottom of the z-distribution tables. Sometimes this is written as

$$\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}},$$

$\alpha := 100\% - p\%$ is the complement of the confidence level and the $\alpha/2$ refers to half of this being above the upper confidence limit and half being below the lower confidence limit.

The z values for 90%, 95% and 99% are 1.645, 1.96 and $2.58 \approx 2.576$ (why)?

Remark

As mentioned above, the z values are the t values for $n \to \infty$. There is actually a line of values for $n \to \infty$ on your t tables and the z values for 90%, 95%, 98%, and 99% can be read straight from these tables.

1. **Autumn 2017** A road engineer measured and recorded the period between resurfacing of 50 randomly selected third class roads in one particular county. This study yielded a sample mean of 6.75 years and a sample standard deviation of 1.8 years. Hence calculate a 95% confidence interval for the (population) mean resurfacing period for this particular county.

Solution: (We have that $n > 30$ so we could use a z-distribution rather than a t-distribution) Our confidence interval is:

$$\bar{x} - t\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + t\frac{\sigma}{\sqrt{n}} \tag{3.17}$$

We have $\bar{x} = 6.75$, $\sigma \approx s = 1.8$, the sample standard deviation, and $n = 50$. All that remains is to find t... So we calculate

2. **Winter 2019** Consider a bespoke concrete dome, upon which more structures will rely. In order to estimate the structural strength of the dome, an engineer carried out an extensive survey, taking 50 small sample bores. The sample had a mean of 21.4 MPa and a sample standard deviation of 3.9 MPa. The structural engineer (SE) determined that at least 60% of the concrete would have to have a strength of more than 20 MPa.

   (a) Find a 95% confidence interval for the population mean, $[\mu_1, \mu_2]$.

   (b) Assuming that the population mean is equal to the sample mean, the population standard deviation equal to the sample standard deviation, and the compressive strength normally distributed, find the *characteristic compressive strength*, the strength S such that only 5% of the concrete has a compressive strength of less than S.

   (c) Assuming a population mean of μ_1, the population standard deviation equal to the sample standard deviation, and the compressive strength normally distributed, what proportion of the concrete has compressive strength of less than 20 MPa?

   (d) Hence, will the SE be happy with the dome?

   (a) Using $\overline{x} \pm t\frac{\sigma}{\sqrt{n}}$:

   (b) We draw:

   Therefore using $X = \mu + z\sigma$:

(c) We draw:

Therefore we find:

(d) Only approximately 53% of the concrete is stronger than 20 MPa:

Exercises: Make all calculations correct to four significant figures. *These answers have a health warning: take some of them with a pinch of salt. The thing is I wrote a little program to them but I didn't check all of them by hand as I didn't have the time. If you want clarification ask in the tutorial.* A few of them are a little bit off. Also if you use a z-distribution rather than a t-distribution then you will be off a little. Also if you don't use degrees of freedom $= n - 1$ on the t-tables.

1. **Winter 2016** A river engineer took 50 random samples from a river. This study yielded a sample mean of 86 ng/L and a sample standard deviation of 5 ng/L. Hence calculate a 95% confidence interval for the (population) mean pesticide level. **Ans:** $[84.58, 87.42]$

2. **Autumn 2020** Consider an old bridge, structural analysis of which suggests that if **more than** 20% of the concrete in a certain concrete component has a compressive strength of less than 20 MPa, that failure is extremely likely. The senior structural engineer (SE) suggests that the component can withstand a number of sample bores, and she determines that this number is five. Five sample bores are taken and measured to have compressive strengths of

$$25.3,\ 23.1,\ 24.4,\ 22.4,\ 19.8\,\text{MPa}.$$

   i. Find a 95% confidence interval for the population mean, $[\mu_1, \mu_2]$. **Ans:** $[20.38, 25.62]$ MPa.

   ii. The SE wants to be conservative, and so assumes that the population mean is the lower limit, μ_1. Assuming this, and the compressive strength normally distributed, find the *characteristic compressive strength*, the strength S such that only 5% of the concrete has a compressive strength of less than S. **Ans:** $S = 16.89$ MPa

   iii. Assuming a population mean of μ_1, and the compressive strength normally distributed, what proportion of the concrete has compressive strength of less than 20 MPa? **Ans:** 43%.

   iv. Hence, will the SE predict that the bridge will fail?

3. **Winter 2017** Following a spate of serious injuries on construction sites, Johnny, a consultant with the Health and Safety Authority, is carrying out a survey of the Irish construction sector. Johnny recorded the costs of five randomly selected accidents. They were:

$$€5,250,\ €6,000,\ €2,000,\ €7,000,\ €4,500.$$

   i. Calculate the sample mean. **Ans:** €4,950.
   ii. Calculate the sample standard deviation. **Ans:** €1,891
   iii. Hence use the t-table to calculate a 95% confidence interval for the (population) cost. **Ans:** [€2,602,€7,298]

4. **Autumn 2010**

   i. Five students measured a certain area and the following values were obtained.

$$1.05,\ 1.07,\ 1.07,\ 1.05,\ 1.06 \text{ hectares.}$$

   Find a 98% confidence interval for the actual area. **Ans**: [1.043,1.077]

   ii. If the measurements of this area by 50 students gave a mean value of 1.06 hectares with a standard deviation of 0.009 hectares, find a 98% confidence area for the actual area. **Ans**: [1.057,1.063]

---

5. **Autumn 2009**

   (i) Five students measured a certain area and the following values were obtained

$$1.05,\ 1.07,\ 1.07,\ 1.05,\ 1.06\,, \text{hectares.}$$

   Find a 98% confidence interval for the actual area. **Ans**: [1.043,1.077]

   (ii) If the measurements of this area by 50 students gave a mean value of 1.06 hectares with a standard deviation of 0.009 hectares, find a 98% confidence interval for the actual area. **Ans**: [1.057,1.063]

6. **Winter 2008** The compressive strength of concrete is subject to variation.

   (i) Five samples of a concrete mix were taken and the strength of concrete (N mm^{-2}) provided for these samples are recorded:

$$29.9,\ 29.9,\ 30.0,\ 30.1,\ 30.1.$$

   Set up a 95% confidence interval for the mean strength. **Ans**: [29.88,30.12]

   (ii) If a sample of 50 measurements gave a mean value of 30 N mm^{-2} with a standard deviation of 0.12 N mm^{-2}, set up a 95% confidence interval for the mean strength. **Ans**: [29.97,30.03]

3.4 Hypothesis Testing

Consider a manufacturing company which claims to produce metal rods that have a (population) mean length of 2 m. This is an assertion that has been made by the company about the population of produced goods. In this section this shall be referred to as H_0, the *null hypothesis* — the *status quo*:

$$H_0 : \ \mu = 2.$$

After using these rods for a period of time a construction company might say hang on; these rods aren't made to these specifications. They might put forward an *alternative hypothesis*, H_A. For example,

$$H_A : \ \mu \neq 2.$$

To settle this dispute the manufacturing company may take a random sample of n rods. From this they will find a sample mean $\bar{x}$ and standard deviation s. The question is: do the sample observations support H_0 or H_A? To decide this carry out the following steps:

1. State the H_0 and H_A for the situation under investigation.
2. On the assumption that H_0 is true, depending on the level of significance required, find an *acceptance interval* for the sample means.
3. If the sample mean falls outside this acceptance interval, reject H_0. Otherwise we cannot reject H_0.

Suppose for example, we are working with the 5% level of significance. Assuming H_0 is true, the probability that a sample mean lies outside the acceptance interval is 0.05. If in a random sample of that size, we find that the sample mean lies outside the acceptance interval this is good evidence in favour of rejecting H_0. Alternatively, if the sample mean lies inside the acceptance interval (but maybe quite far from the assumed mean), then we can say that 'rare events' happen and there is no strong reason to reject H_0 at this point.

3.4.1 Two types of Errors in Hypothesis Testing

Whenever we have to choose between two alternatives, there are two distinct types of error that may occur:

I: Reject H_0 when H_0 is true.

II: Cannot reject H_0 when H_A is true.

Type I errors are considered more serious then Type II. On the one hand, Type I errors can be seen as 'throwing the baby out with the bath water' but an analogy with criminal law shows why. In this context, the level of significance can be seen as the risk at which the decision-maker is willing to make a Type I error. The p value of a Hypothesis test is the probability of making a Type I error given that the null hypothesis is true. It is also called the *significance*.

Example: Murder Trial

In a murder trial, the null hypothesis (innocent until proven guilty) is given by:

The defendant did not murder the deceased: innocent (H_0)

The alternate hypothesis is

The defendant did murder the deceased: guilty (H_A)

Implicit in H_0 is a range of possibilities that are consistent with H_0. It is up to the prosecuting council to produce evidence which lies outside this range of possibilities (e.g. the probability of the defendant being innocent AND owning the gun is small). If the prosecuting council succeed in producing such evidence (similar to a sample mean outside the acceptance interval), the jury reject H_0 and the man is convicted. A Type I error here is convicting an innocent man — which is considered worse than a Type II: acquitting a guilty man. If you think about it, the level of significance is the probability of a Type I error.

For small sample sizes we should use a t-distribution but as before when $n \geq 30$ you can use a z-distribution. As mentioned in the previous sections, the z values for 90%, 95%, 98%, 99% and 99.8% can be read from the t-tables ($n \to \infty$).

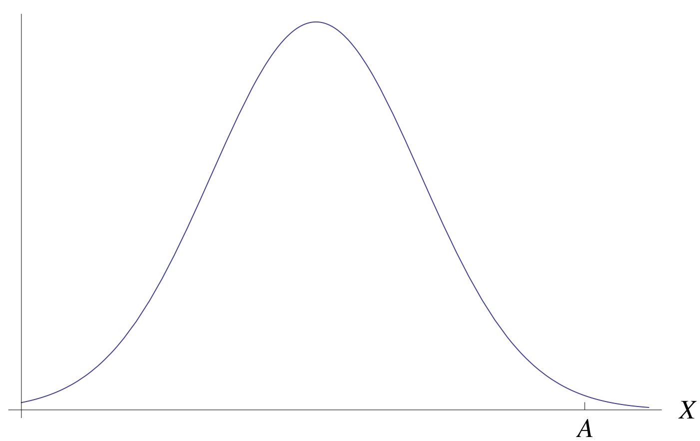


Figure 3.13: In hypothesis testing, we assume a status quo, a *null hypothesis* — e.g. the mean is equal to μ: we assume statistics about the data. Then we do a sample and calculate a sample mean. Finally we ask ourselves: is this sample mean consistent with my assumptions about the data? If it is we cannot reject the null hypothesis. If however we get a sample mean that is not consistent with our assumptions we reject our assumptions — the statistics are not as we believe them to be and we reject this null hypothesis.

Examples

1. **Autumn 2017** Test, at the 0.05 level of significance, the claim of local politician Harry Reilly-Jay that the average resurfacing period in this particular county is the same as the national average of 6 years, given that a sample of size 50 yielded a sample mean of 6.75 years and a sample standard deviation of 1.8 years. Specify the null and alternative hypothesis.

Solution: First we state the null and alternate hypothesis:

H_0: $\mu = 6$

H_A: $\mu \neq 6$

Now we set up the 95% acceptance interval for the sample means:

$$\mu - t\frac{\sigma}{\sqrt{n}} \leq \bar{x} \leq \mu + t\frac{\sigma}{\sqrt{n}}. \tag{3.18}$$

Here we want a 95% interval so we have $t = 2.009$. We take $\mu = 6$, $\sigma \approx s = 1.8$ and $n = 50$:

Finally 6.75 lies outside this acceptance interval so we *can* reject H_0 at the 0.05 level of significance.

2. **Winter 2019** *Cork Steelworks Limited* produce 8 m long steel rods used in construction. A customer of CS Ltd. claims that the rods are not 8 m long. The Quality Engineer replied to the customer saying that the rods have a mean-average length of 8 m: a fact the customer disputes. In light of this complaint, the QE takes a sample of thirty rods, measures their lengths, and calculates the sample mean as 7.97 m with a standard deviation of 0.04 m. The statistics calculated might possibly suggest that the average length of the rods is not 8 m, but the QE sets up her analysis in such away that this conclusion will be made incorrectly with a probability of only 5%. The QE comes up with an interval $[7.985, 8.015]$.

   (a) Is this *Sampling* or *Hypothesis Testing*?
   ANSWER PART ii. OR PART iii. AND NOT BOTH PARTS

   (b) If this is *Sampling*

      i. The sample mean is a random variable. In general, what is its mean?
      ii. What is the confidence level of this sampling process?
      iii. Reproduce the calculations of the QE to produce $[7.985, 8.015]$.
      iv. What is the conclusion of the sampling process?

   (c) If this is *Hypothesis Testing*

      i. What are the null and alternative hypotheses?
      ii. What is the p-value?
      iii. Reproduce the calculations of the QE to produce $[7.985, 8.015]$.
      iv. What is the conclusion of the hypothesis test?

   Solution: This is hypothesis testing. We state the null and alternate hypothesis:

   H_0: $\mu = 8$ m

   H_A: $\mu \neq 8$ m

   The p-value is 5%.

Now we set up the 95% acceptance interval for the sample means:

Here we want a 95% interval so we have $t =$. We take $\mu = 8$, $\sigma \approx s = 0.04$ and $n = 30$:

Assuming H_0, 95% of the sample means lie in this acceptance interval. Finally 7.97 m does *not* lie in this acceptance interval so we reject H_0 at the 0.05 level of significance.

3.4.2 "They're Small! They're Big!"

Suppose that a customer buying metal rods says that the rods are (on (mean) average) longer than the claimed length... or a cement distributor suspecting that the bags are (on (mean) average) lighter than the claimed mass. In these situations you can think of the alternative hypothesis as a claim (but not a null hypothesis). The null hypothesis is then the *negation* of the claim. For example, for 2 m rods "They're Long!"

H_0: $\mu \leq 2$ m

H_A: $\mu > 2$ m

For 25 kg bags, "They're Light!"

H_0: $\mu \geq 25$

H_A: $\mu < 25$

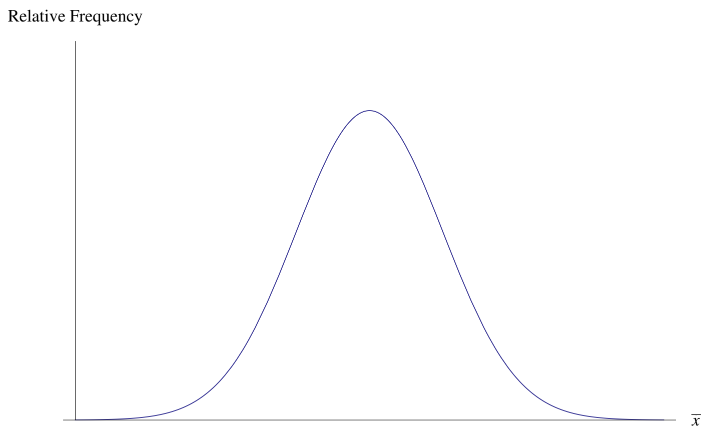


Figure 3.14: Suppose the (customer) claim is that the (mean average) length of rods is greater than the (producer claim) of μ. Then the $(1-p)\%$ acceptance region for sample means of size n for this claim is $\overline{x} \leq \mu + t\frac{\sigma}{n}$.

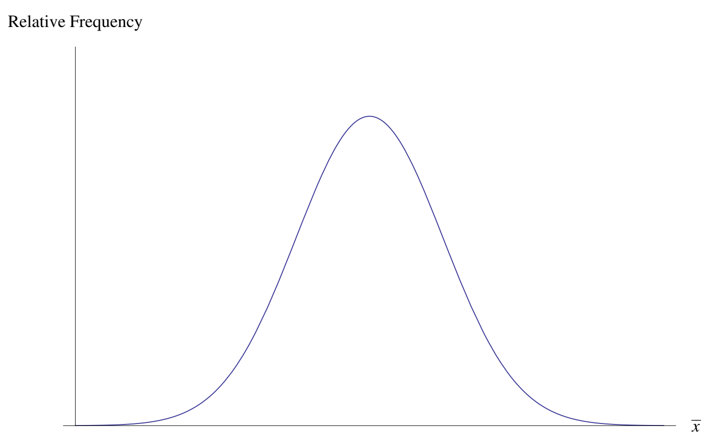


Figure 3.15: Suppose the (distributor) claim is that the (mean average) mass of cement bags is less than the (producer claim) of μ. Then the $(1-p)\%$ acceptance region for sample means of size n for this claim is $\overline{x} \geq \mu - t\frac{\sigma}{n}$.

After this the only thing to do is to make sure that you using the correct t value.

Example

A river engineer took 50 random samples from a river to test for pesticide level. This study yielded a sample mean of 86 ng/L and a sample standard deviation of 5 ng/L. Test the claim that the average pesticide level in the river is greater than 80 ng/L at the 0.05 level of significance. Specify the null and alternative hypothesis.

Solution: First we state the null and alternate hypothesis:

H_0: $\mu \leq 80$

H_A: $\mu > 80$

Let us draw a picture:

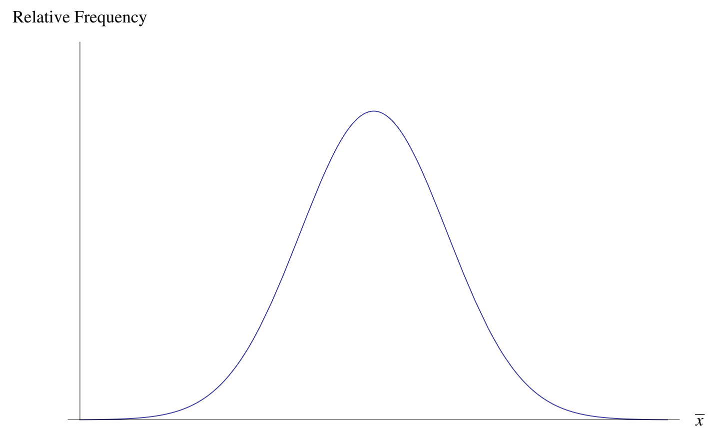


Figure 3.16: The claim is that the (mean average) pesticide level is more than 80 ng/L. We draw the acceptance region.

Now we set up the 95% acceptance interval for the sample means:

$$\bar{x} \leq \mu + t\frac{\sigma}{\sqrt{n}}. \tag{3.19}$$

Here we want a 95% one-sided interval so we have $t = 1.676$. We take $\mu = 80$, $\sigma \approx s = 5$ and $n = 50$:

Finally 86 lies outside this acceptance interval so we can reject H_0 at the 0.05 level of significance.

Exercises:

1. **Autumn 2012** The breaking strength of a certain metal is claimed to be 43.55 kN m^{-2} but a customer claims they are not this strong. A sample of 100 specimens gave a mean value of 43.6 kN m^{-2} with a standard deviation of 0.9 kN m^{-2}. At the 0.02 level of significance test the hypothesis that the mean is equal to 43.55 kN m^{-2} **Ans:** [43.34,43.76] cannot reject.

2. **Winter 2008** In the past it has been found that the mean life of components was 150 hours. The process of production was changed and a sample of 50 items yielded a mean value of 155 hours with a standard deviation of 12 hours. At the 5% level of significance, test that the mean has changed. **Ans:** [146.7,153.3] can reject.

3. **Autumn 2020** *Cork Brickworks Limited* hope to produce very strong bricks for specialist builders. They have established a manufacturing process and now they need to test how strong the bricks are in terms of crushing strength. The Quality Engineer takes a sample of 100 bricks. The QE finds a sample mean of 148.2 kg/cm^2, with a standard deviation of 5.3 kg/cm^2. The QE asks the CEO how much of an assurance does she want to give potential customers about the average crushing strength. The CEO responds that she wants to be 99% sure about the claims the company are making. The QE subsequently files a report containing the interval $[146.8, 149.6]$.

   i. Is this *Sampling* or *Hypothesis Testing*?
   [ANSWER PART ii. OR PART iii. AND NOT BOTH PARTS]

   ii. If this is *Sampling*

      A. The sample mean is a random variable. In general, what is its mean?
      B. What is the confidence level of this sampling process?
      C. Reproduce the calculations of the QE to produce the confidence interval $[146.8, 149.6]$.
      D. What is the conclusion of the sampling process?

   iii. If this is *Hypothesis Testing*

      A. What are the null and alternative hypotheses?
      B. What is the p-value?
      C. Reproduce the calculations of the QE to produce the acceptance region $[146.8, 149.6]$.
      D. What is the conclusion of the hypothesis test?

4. **Summer 2012** Concrete lintels are produced by a machine and it is claimed that the mean length is 2.5 m. A sample of 50 of these lintels gave a sample mean of 2.48 m with a sample standard deviation, $s \approx 0.06325$ m. At the 0.01 level of significance, test that the claim above is not true. **Ans:** [2.476,2.524] cannot reject.

5. **Autumn 2008** The lengths of blocks produced were assumed to be normally distributed with a mean value of 450 mm and with a standard deviation of 1.4 mm. A new type of mould was introduced and a sample of 100 blocks gave a mean length of 450.3 mm. At the 0.05 level of significance, test that the mean length has not changed. **Ans:** [449.7,450.27] can reject.

Chapter Checklist

1. What range of values can $\mathbb{P}[A]$ take?
   $0 \leq \mathbb{P}[A] \leq 1$.

2. Under what condition is $\mathbb{P}[A \text{ or } B] = \mathbb{P}[A] + \mathbb{P}[B]$?
   Events A and B are *mutually exclusive*.

3. Under what condition is $\mathbb{P}[A \text{ and } B] = \mathbb{P}[A] \cdot \mathbb{P}[B]$?
   Events A and B are *independent*.

4. Under what condition is $\mathbb{P}[\text{not } A] = 1 - \mathbb{P}[A]$?!
   All conditions.

5. What kind of data/random variable has a *normal distribution*?
   Roughly, symmetric about a single dominant average, bell-shaped.

6. In what sense is there only one kind of normal data?
   All probabilities involved normal X can be calculated using standard normal $z \sim N[0, 1]$.

7. Why do we sample?
   To estimate *population* statistics.

8. What is a *confidence interval* for a population mean?
   A range of values that we believe the population mean lies in (with a certain confidence).

9. Sample means have a t-distribution but for $n > 30$, $t \approx t_\infty = z$.

10. Can you describe hypothesis testing to the man on the street?
    Murder Trial. H_0: innocent, H_1: not innocent. Acceptance interval is range of possibilities consistent with H_0. Job of prosecution to show that evidence lies outside acceptance interval.

Chapter 4

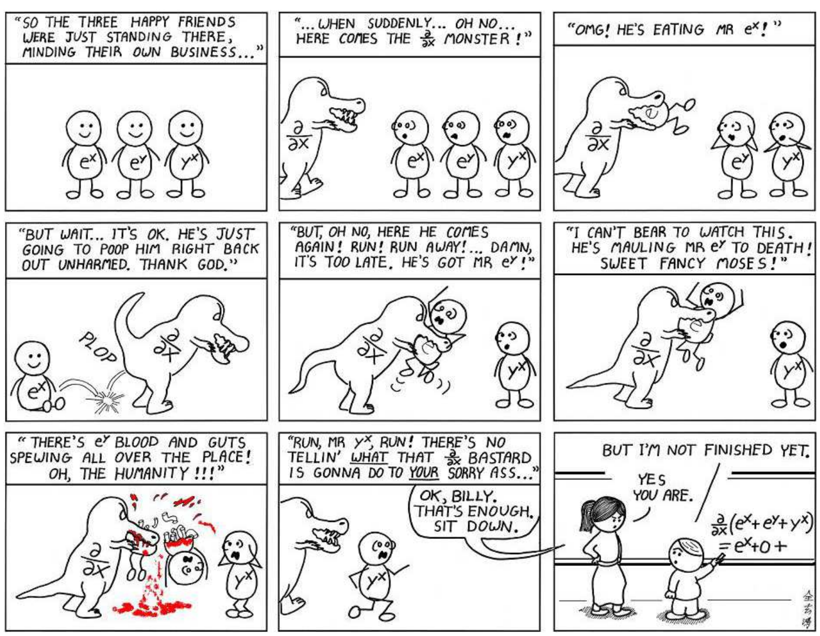
Taylor Series

The journey for an education starts with a childhood question[1].

David L. Finn

Although this may seem a paradox, all exact science is dominated by the idea of approximation.

Bertrand Russell

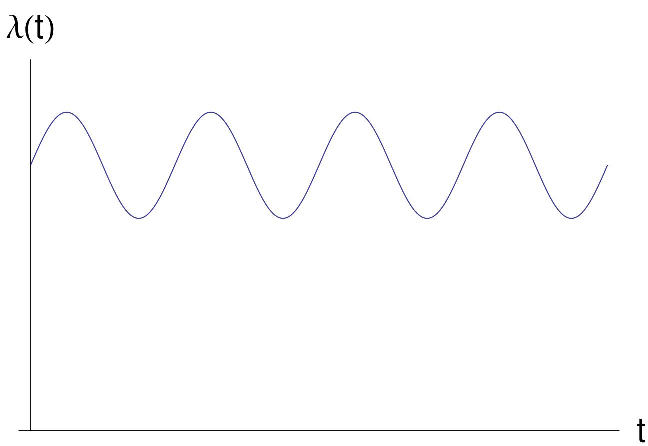


[1]the question in this case is: how does a calculator work?

4.0.1 Motivation

Figure 4.1: Top Gear dropped a VW Beetle from a height of 1 mile and it spun in the air as it fell.

If we are trying to formulate a model for the fall of this car we would have to try and account for the way the roll of the car means that the coefficient of the drag term ($\lambda v(t)$) varies between its maximum and minimum in a wave-like way:



A function with this behaviour is:

$$\lambda(t) = \frac{1}{2}(M+m) + \frac{1}{2}(M-m)\sin\omega t \tag{4.1}$$

where M and m are the maximum and minimum of $\lambda(t)$ and ω is a constant related to the angular frequency. Then the equation of motion is of the form:

$$\begin{aligned} ma &= \sum F(t) \\ \Rightarrow m\frac{d^2x}{dt^2} &= +mg - \frac{dx}{dt}\left(\frac{1}{2}(M+m) - \frac{1}{2}(M-m)\sin(\omega t)\right) \end{aligned}$$

Neither the method of undetermined coefficients (in MATH7021) nor any other straightforward method I know of solves this differential equation.

Unfortunately this is typical, and for many systems for which a differential equation may be drawn, it may be impossible to solve the equations. There are a number of numerical techniques which can give approximate answers. However if we are participating in some industrial project with millions spent on it we don't want to be chancing our arms on any old estimate or guess. *Approximation Theory* aims to control these errors. Suppose we have a Differential Equation with solution $y(x)$. An approximate solution $y_h(x)$ to the equation — depending on a *step-size* h — can be found using some numerical method. If the approximation method is sufficiently 'nice' we may be able to come up with a measure of the error, e.g:

$$\varepsilon_h = \max_{x\in[0,L]} |V(x) - V_h(x)|.$$

Some classes of problem are even nicer in that with increasing computational power we can develop increasingly approximate solutions with decreasing errors ($\varepsilon_h \to 0$). Even nicer still from a mathematical point of view if we can find a sequence of approximations with errors decreasing to zero.In this case we say that the sequence of approximations *converges*.

In this module we took a first foray into the approximation theory of numerical methods by estimating the solutions of differential equations in Section 2.4.

Calculus Review V: Higher Derivatives and the Product, Quotient & Chain Rules

When you differentiate a function twice you get the *second derivative*, e.g.

This generalises in the obvious way. The issue here is that we need to differentiate a function $f(x)$. The derivatives of functions are given in the function catalogue. The question is how do we find the derivative of

— sums/differences

— scalar multiples

— products

— quotients

— powers/roots/compositions

4.0.2 "Rules of Differentiation"

The answer is by the Sum, Scalar, Product, Quotient & Chain Rules which need to be well understood to do well with Taylor Series. They are not really rules but theorems/facts that describe how we should differentiate sums, products, compositions, etc.

“Rules of Differentiation”

Suppose that $u(x)$ and $v(x)$ are functions, $n \in \mathbb{Q}$ a fraction and $k \in \mathbb{R}$ a real number. Suppose also that $\frac{d}{dx}$ is the *differential operator* (i.e. it means differentiate), and $u'(x)$ and $v'(x)$ are the derivatives of $u(x)$ and $v(x)$ respectively. *Then*

$$\frac{d}{dx}(u(x) \pm v(x)) = u'(x) \pm v'(x) \quad \text{[Sum Rule]}$$

$$\frac{d}{dx}(k \cdot u(x)) = k \cdot u'(x) \quad \text{[Scalar Rule]}$$

$$\frac{d}{dx}(u(x) \cdot v(x)) = u(x) \cdot v'(x) + v(x) \cdot u'(x) \quad \text{[Product Rule]}$$

$$\frac{d}{dx}\left(\frac{u(x)}{v(x)}\right) = \frac{v(x) \cdot u'(x) - u(x) \cdot v'(x)}{[v(x)]^2} \quad \text{[Quotient Rule]}$$

$$\frac{d}{dx}x^n = nx^{n-1} \quad \text{[Power Rule]}$$

$$\frac{d}{dx}(u(v(x))) = u'(v(x)) \cdot v'(x) \quad \text{[Chain Rule]}$$

Remark

The Difference Rule is a corollary of the Sum and Scalar Rules. The Quotient Rule can be seen as a corollary of the Product and Chain Rules. The Power Rule handles roots. These formulas are all in the tables — the functions are called u and v but they are the same ideas.
Exercises

1. Differentiate the following

(a) $y = \frac{1}{x^2} - 6x + 4$	(g) $y = x \ln x$
(b) $y = xe^{2x}$	(h) $y = (\sin x)^2$
(c) $y = \sin 2x - \cos 4x$	(i) $y = e^x \sin x$
(d) $y = (2x+1)^3(x+2)$	(j) $y = \frac{6x+1}{x-4}$
(e) $y = 4x \sin x$	(k) $y = \frac{x^2}{x+3}$
(f) $y = \ln(x^2+1)$	(l) $y = \frac{e^x}{x-2}$.

Ans: (a) $-6 - \frac{2}{x^3}$, (b) $2x \cdot e^{2x} + e^{2x}$, (c) $2\cos(2x) + 4\sin(4x)$, (d) $6(x+2)(2x+1)^2 + (2x+1)^3$, (e) $4x\cos x + 4\sin x$, (f) $\frac{2x}{x^2+1}$, (g) $1 + \ln x$, (h) $2\sin x \cos x$, (i) $e^x \cdot \cos x + e^x \cdot \sin x$, (j) $-\frac{25}{(x-4)^2}$, (k) $\frac{x(x+6)}{(x+3)^2}$, (l) $\frac{e^x(x-3)}{(x-2)^2}$.

2. Find the second derivatives of the following:

(i) $\frac{6}{x^2}$ (ii) $(6-5x)^5$ (iii) $\ln x$ (iv) $\sin(2x)$.

4.1 Maclaurin & Taylor Series

4.1.1 Power Series

How does a calculator work? Addition & Multiplication are fairly easy to programme but how about finding $\ln 2$? However no calculator can calculate $\ln 2$ exactly — it is an *irrational number* which means that it has a non-repeating decimal expansion. Therefore calculators must use approximations. Here we describe some methods that we might do this.

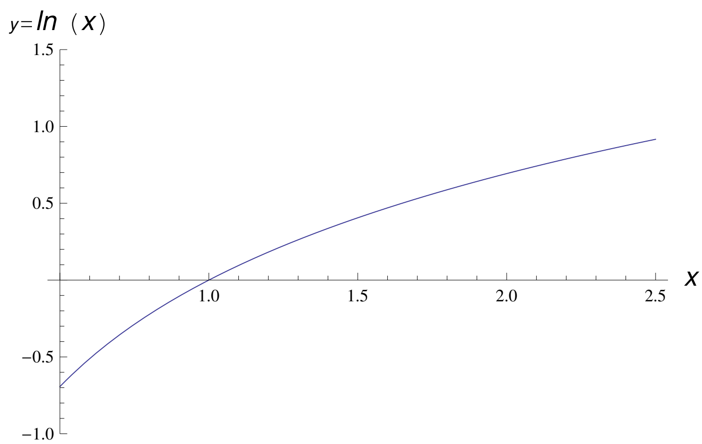


Figure 4.2: Smooth functions such as $y = \ln x$ admit good linear approximations.

To do this we pick a point close to $x = 2$, whose log we *do* know, for example $x = 1$. Now we draw the tangent to the curve at $x = 1$. To get the equation of the tangent to the curve we need a point (x_1, y_1) and a slope. We have $(x_1, y_1) = (1, 0)$ so all we need is a slope.

So therefore we have for points near $x = 1$ we have

Now instead of fitting a line to the curve we could fit a quadratic or more complicated *polynomials* to the curve[2] and all are somewhat easy for a calculator as they are just combinations of addition and multiplication, e.g.

$$\ln x \approx x - 1 - \frac{1}{2}(x-1)^2.$$

In general, all of these polynomial approximations are just that... approximations. However would it be possible to have a series of polynomial approximations to a non-polynomial function that converges?

[2]as in Chapter 1

The first clue that this might be possible was the following fact:

Geometric Series

For $|x| < 1$, we have the following

Now this formula is exact in the sense that no matter how close to $\dfrac{1}{1-x}$ as you want to get, you can get that close by taking enough terms of the 'infinite polynomial'. For example to using only five terms with $x = 0.2$:

$$\frac{1}{1-0.2} = 1.25 \approx 1.2499968 = 1 + 0.2 + 0.2^2 + 0.2^3 + 0.2^4.$$

Formally, such infinite sums of powers of x are called *Power/Taylor series*

In general, power series are only valid for some values of x. The example above only works for $|x| < 1$ (why??). Functions that can be somewhere-well-approximated by power series are called *analytic functions*. Now there are examples functions which are *not* well-behaved, but the ones of any use to ye are either well-enough understood in their original form[3] or else not of much use. We examine in particular two classes of power series.

4.1.2 Maclaurin Series

Above we used a line near $x = 1$ to approximate $\ln x$. The reason we used $x = 1$ was because we knew all about the function at that point: its value and slope. As a rule of thumb, we also tend to know all about functions near $x = 0$ and the Maclaurin Series makes use of this.

Given a function $y = f(x)$ the first thing we assume is that it is analytic near $x = 0$... i.e. a power series is going to work. So we assume that we can write

The problem therefore is to find the values of the coefficients $a_0, a_1, a_2, \dots$. This is actually an infinite number of problems but we can use differentiation to help. If the left-hand function is equal to the right-hand function then they must be equal for all values of x... and their first derivatives, and second derivatives and in fact all of their derivatives.

Let us see this with, say $f(x) = \sin x \overset{!}{=} a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \cdots$. We use this to our advantage as follows. What is happening at $x = 0$:

Now differentiate both functions and examine what happens at $x = 0$:

[3] see Step Functions

Now repeat the trick:

Now once more shows that:

$$a_3 := -\frac{1}{6} = \frac{f'''(0)}{3!}.$$

With a bit of thought you will see that the coefficients a_n are given by

Maclaurin Series Formula

If $f(x)$ is analytic near $x = 0$ then we have

$$f(x) = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \frac{f'''(0)}{3!}x^3 + \cdots = \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!}x^k. \tag{4.2}$$

Therefore, if we *truncate*, to say three terms:

Examples

1. **Winter 2012** Consider the function $f(x) = \ln(\sec x)$.

   (i) Show that $f''(x) = \sec^2 x$.

   (ii) Find the first two non-zero terms of the Maclaurin series of $f(x)$.

   Solution: Note that $\sec x = \dfrac{1}{\cos x}$ and has derivative $\sec x \cdot \tan x$. We will also require the *Chain Rule*:

   (i) To differentiate we need the chain rule:

   To find $f''(x)$ we differentiate again. We see in the tables that $(\tan x)' = \sec^2 x$ so we are done.

   (ii) We know that the Maclaurin series is given by

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!}x^k = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \frac{f'''(0)}{3!}x^3 + \cdots.$$

We calculate

This means we also need the value of $f'''(0)$. We differentiate $\sec^2 x$ using the chain rule and evaluate at $x = 0$:

this is also equal to zero so we must differentiate again, this time requiring the product rule:

So we have that

$$\begin{aligned}\ln(\sec x) &\approx 0 + 0 \cdot x + \frac{1}{2!}x^2 + \frac{0}{3!}x^3 + \frac{2}{4!}x^4 \\ &= \frac{1}{2}x^2 + \frac{2}{4!}x^4 = \frac{1}{2}x^2 + \frac{1}{12}x^4\end{aligned}$$

This is certainly a good approximation for $x \approx 0$, as for high powers $x^k \to 0$.

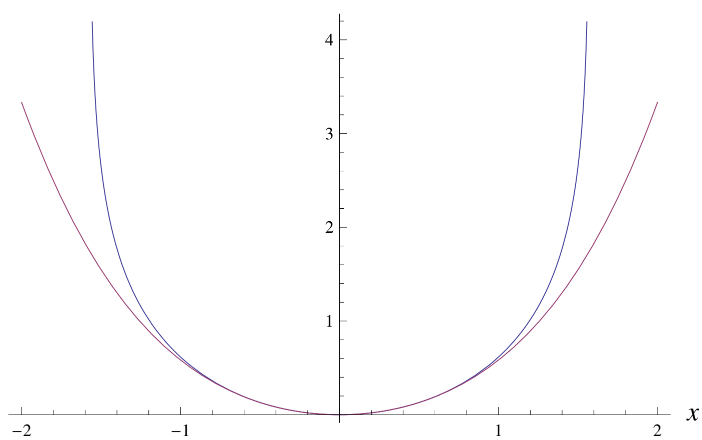


Figure 4.3: For small values of x, $\ln(\sec x)$ ('inside' curve) is well approximated by the degree four polynomial $x^2/2 + x^4/12$... certainly for $-1 < x < 1$.

2. **Autumn 2014**

   i. Find all terms up to x^3 in the Maclaurin Series of $f(x) = \ln(1+x)$.

   Solution: Using the formula from the tables (see the back of the manual)

   we calculate

   ii. Hence find correct to four significant figures an approximation to $\ln 1.1 = \ln(1+0.1)$.

   Solution: We use the Maclaurin Series at $x = 0.1$

   Remark: If you are asked to find the percentage error in an approximation of some quantity Q, note that the percentage error in approximating a true value Q_0 is given by $\dfrac{\Delta Q}{Q_0}$. In this case, with $Q_0 \approx 0.09531$:

$$\text{Percentage error} \approx \frac{|0.09533 - \ln 1.1|}{\ln 1.1} \approx 0.0002 = 0.02\%.$$

3. **Winter 2010** The maximum deflection $\delta_{\max}$ of a beam of span L is given by (exactly)

$$\delta_{\max} = \frac{WEI}{P}\left(\sec\left(\frac{mL}{2}\right) - 1\right), \text{ where } m^2 = \frac{P}{EI}.$$

   Simplify the formula for $\delta_{\max}$ using the Maclaurin Series[4]

$$\sec x \approx 1 + \frac{1}{2}x^2.$$

[4] deriving this was the first part

Solution: First we use the Maclaurin Series on $\sec \frac{mL}{2}$

Now we calculate

Remark: This assumes that $\dfrac{mL}{2} \approx 0$ (which it is because EI is large).

Worked Example

Find the first four terms of the Maclaurin Series of $f(x) = xe^x$.

Solution: The Maclaurin Series is just the Taylor Series about $x = 0$:

$$f(x) \approx f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \frac{f'''(0)}{3!}x^3. \tag{4.3}$$

Hence we have to calculate $f(0)$, $f'(0)$, etc. We have to realise that $f(x)$ is two functions multiplied together so when we differentiate we need the *product* rule:

$$f'(x) = \frac{d}{dx}\underbrace{x}_{u}\underbrace{e^x}_{v} = uv' + vu',$$

where e.g. v' is the derivative of v. We also might need the product rule in finding the second and third derivatives.

$$\begin{aligned}
f(0) &= (0)e^0 = 0 \\
\Rightarrow f'(x) &= xe^x + e^x(1) = xe^x + e^x|_{x=0} = 0e^0 + e^0 = 1 \\
\Rightarrow f''(x) &= \frac{d}{dx}(xe^x + e^x) = \frac{d}{dx}(xe^x) + \frac{d}{dx}e^x \\
&= (xe^x + e^x) + e^x = xe^x + 2e^x|_{x=0} = 0e^0 + 2e^0 = 2 \\
\Rightarrow f'''(x) &= \frac{d}{dx}(xe^x + 2e^x) = \frac{d}{dx}(xe^x) + 2\frac{d}{dx}e^x \\
&= xe^x + e^x + 2e^x = xe^x + 3e^x|_{x=0} = 0e^0 + 3e^0 = 3
\end{aligned}$$

where we used the fact that $\dfrac{d}{dx}x = 1$ and $\frac{d}{dx}e^x = e^x$ (indeed along with $e^0 = 1$ this is how you should *define* the exponential function).

Now we need to use the formula above and we write

$$\begin{aligned}
xe^x &\approx 0 + 1x + \frac{2}{2!}x^2 + \frac{3}{3!}x^3 = x + x^2 + \frac{3}{3 \cdot 2 \cdot 1}x^3 \\
&= x + x^2 + \frac{1}{2}x^3.
\end{aligned}$$

There is a quicker way to get there though. Find the Maclaurin Series of e^x, which because all the derivatives are e^x, and hence equal to one at $x = 0$ we find:

$$e^x \approx 1 + x + \frac{1}{2!}x^2$$
$$\Rightarrow xe^x \approx x\left(1 + x + \frac{1}{2}x^2\right) = x + x^2 + \frac{1}{2}x^3.$$

Exercises:

1. Find the first three non-vanishing terms of the Maclaurin Series of e^x. **Ans**: $1 + x + \frac{1}{2}x^2$.

2. **Winter 2019** Consider the function

$$f(x) = \sin x.$$

   (a) Find the first two non-zero terms of the Maclaurin Series of $f(x)$. **Ans**: $x - \frac{x^3}{3!}$.

   (b) Hence, find, correct to *four* significant figures, an approximation of $\sin 0.5$. **Ans**: $\sin(0.5) \approx 0.4792$.

   (c) Given that $\sin 0.5 \approx 0.4794$, find, correct to one significant figure, the percentage error in this approximation. **Ans**: 0.08%.

Figure 4.4: $\sin x \approx x$

3. **Winter 2015**

   (a) Find all terms up to x^2 of the Maclaurin series for $f(x) = \ln(1 + x)$ about $x = 0$. **Ans:** $\ln(1 + x) \approx x - \frac{1}{2}x^2$

   (b) Hence approximate $\ln(1.5) = \ln(1 + 0.5)$ without using a calculator. **Ans:** 0.375

   (c) Use your calculator to find $\ln(1.5)$ correct to three significant figures and calculate the percentage error in the Maclaurin series approximation. **Ans:** 0.405 and 7.407%

4. **Autumn 2020** Suppose that the loading on a beam of span 5 m is given by:

$$w(x) = 16 \cdot \cos\left(\frac{x^2}{12}\right).$$

   i. Show carefully that the first two non-zero terms of the Maclaurin Series of $f(x) = \cos x$ gives

$$\cos x \approx 1 - \frac{x^2}{2}.$$

   ii. Hence, by substituting $x \to \dfrac{x^2}{12}$ into this Maclaurin Series, and multiplying by 16, show that for $x \approx 0$,

$$w(x) \approx 16 - \frac{x^4}{18}.$$

5. **Winter 2016** It can be shown that if a point load of magnitude P is exerted at the top of a column of length L, Young's Modulus E and second moment of area I, at a distance e from the central axis, that the horizontal deflection of the beam is given by

$$\delta_{\text{max}} = e \cdot \left(\sec\left(\sqrt{\frac{P}{EI}} \cdot \frac{L}{2}\right) - 1\right).$$

   i. Show carefully that the first two non-zero terms of the Maclaurin series for $f(x) = \sec(x)$ are given by:

$$f(x) = 1 + \frac{1}{2}x^2 + \cdots.$$

   ii. If $\sqrt{\dfrac{P}{EI}} \cdot \dfrac{L}{2} \ll 1$, then $\sec\left(\sqrt{\dfrac{P}{EI}} \cdot \dfrac{L}{2}\right)$ can be approximated using the first two non-zero terms of its Maclaurin Series:

$$\sec x \approx 1 + \frac{1}{2}x^2.$$

   Making this assumption, show that

$$\delta_{\text{max}} \approx \frac{ePL^2}{8EI}.$$

6. **Autumn 2018** Suppose that for a load of P applied to a beam of length L, Young's Modulus E and second moment of area I, that the deflection of the beam is given by

$$\delta_{\text{max}} = \frac{1 - \cos(kL)}{\cos(kL)}, \text{ where } k = \sqrt{\frac{P}{EI}}.$$

   (a) Show that

$$\delta_{\text{max}} = \sec(kL) - 1.$$

   (b) If $kL \ll 1$, then $\sec(kL)$ can be approximated using the first two non-zero terms of its Maclaurin Series (this question also required the student to derive this Maclaurin Series):

$$\sec x \approx 1 + \frac{1}{2}x^2.$$

   Making this assumption, show that

$$\delta_{\text{max}} = \sec(kL) - 1 \approx \frac{PL^2}{2EI}.$$

4.1.3 Taylor Series

What is the Maclaurin Series of $\ln x$? OK first of all what is $\ln 0$... This example shows that not all functions are analytic near $x = 0$. However, suppose that a function $y = f(x)$ is analytic near a point $x = a$? It turns out that if we write

then an identical analysis to that above yields the following formula:

Taylor Series Formula

If $f(x)$ is analytic near $x = a$ ($x \approx a$) then we have

$$\underbrace{f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f'''(a)}{3!}(x-a)^3 + \cdots}_{\text{on exam paper}} = \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!}(x-a)^k. \quad (4.4)$$

Remark

The Maclaurin Series is just the Taylor Series expansion about $a = 0$.

Winter 2013

Find the first three terms of the Taylor Series of $f(x) = \sec x$ about the point $a = \dfrac{\pi}{4}$.

[NOTE: $\sec x = 1/\cos x$, $\sin(\pi/4) = \cos(\pi/4) = 1/\sqrt{2}$, $\tan(\pi/4) = 1$]

Solution: About the point $x = \pi/4$ means near the point $a = \pi/4$. We calculate

Now using the formula:

which is valid for x near $\pi/4$.

Exercises:

1. **Autumn 2010** Find the first three terms of the Taylor Series of $f(x) = \sec x$ about the point $x = \dfrac{\pi}{4}$. [NOTE: $\sec x = 1/\cos x$, $\sin(\pi/4) = \cos(\pi/4) = 1/\sqrt{2}$, $\tan(\pi/4) = 1$] **Ans**: $\sqrt{2} + \sqrt{2}\left(x - \dfrac{\pi}{4}\right) + \dfrac{3}{\sqrt{2}}\left(x - \dfrac{\pi}{4}\right)^2$.

2. **Winter 2011** Find the first three terms of the Taylor Series of $f(x) = \tan x$ about the point $x = \dfrac{\pi}{4}$. [NOTE: $\sin(\pi/4) = \cos(\pi/4) = 1/\sqrt{2}$, $\tan(\pi/4) = 1$] **Ans**: $1 + 2\left(x - \dfrac{\pi}{4}\right) + 2\left(x - \dfrac{\pi}{4}\right)^2$.

3. **Autumn 2012** Find the first three terms of the Taylor Series of $f(x) = \ln(\sec x)$ about the point $x = \dfrac{\pi}{4}$. [NOTE: $\sin(\pi/4) = \cos(\pi/4) = 1/\sqrt{2}$, $\tan(\pi/4) = 1$] **Ans**: $\dfrac{\ln 2}{2} + \left(x - \dfrac{\pi}{4}\right) + \left(x - \dfrac{\pi}{4}\right)^2$.

---

4. **Autumn 2009** Find the first three terms of the Taylor Series of $f(x) = \ln(\cos x)$ about the point $x = \dfrac{\pi}{4}$. [NOTE: $\sec x = 1/\cos x$, $\sin(\pi/4) = \cos(\pi/4) = 1/\sqrt{2}$, $\tan(\pi/4) = 1$] **Ans**: $-\dfrac{\ln 2}{2} - \left(x - \dfrac{\pi}{4}\right) - \left(x - \dfrac{\pi}{4}\right)^2$.

5. **Winter 2008** Find the first three terms of the Taylor Series of $f(x) = \ln(\sec x + \tan x)$ about the point $x = \dfrac{\pi}{4}$. [NOTE: $\sec x = 1/\cos x$, $\sin(\pi/4) = \cos(\pi/4) = 1/\sqrt{2}$, $\tan(\pi/4) = 1$] **Ans**: $\ln(1 + \sqrt{2}) + \sqrt{2}\left(x - \dfrac{\pi}{4}\right) + \dfrac{1}{\sqrt{2}}\left(x - \dfrac{\pi}{4}\right)^2$.

Partial Differentiation

Functions of Several Variables: Surfaces

We can think of functions as expressing a relationship between variables. We write $y = f(x)$ or even $y = y(x)$ if the variable y depends on the variable x.
Examples of functions seen in engineering include:

1. The area of a square, A, depends on the side-length, s; and we write $A = A(s)$. Indeed $A(s) = s^2$.

2. The height of a projectile, h, depends on how much time, t, has elapsed since projection and we write $h = h(t)$.

3. The yearly sales of a product, S, depends on the number of years the product has been on the market, t; and we write $S = S(t)$.

4. The temperature of a component, θ, depends on the time it has been used, t; and we write $\theta = \theta(t)$.

5. For beams of constant E and I, we have that the load per unit length, w; the shearing force, V; the bending moment, M; the slope, y'; and the deflection, y; ALL depend on the distance from one end of the beam, x; and we write $w = w(x)$, $V = V(x)$, $M = M(x)$, $y' = y'(x)$ and $y = y(x)$.

6. The volume of a sphere depends on the radius:

$$V = V(r) = \frac{4}{3}\pi r^3.$$

Consider the bending moment on a light beam of span 6 m, subject to a constant/udl load :

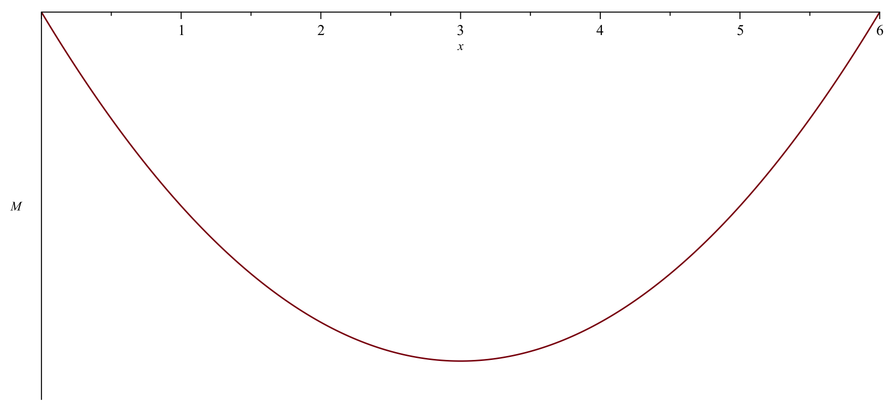


Figure 4.5: $M = M(x)$ is a functions of a single variable, x: the distance along the beam.

Many functions in engineering, physics and mathematics tie together more than two variables. For example, the bending moment on a square plate depends on two variables, x and y:

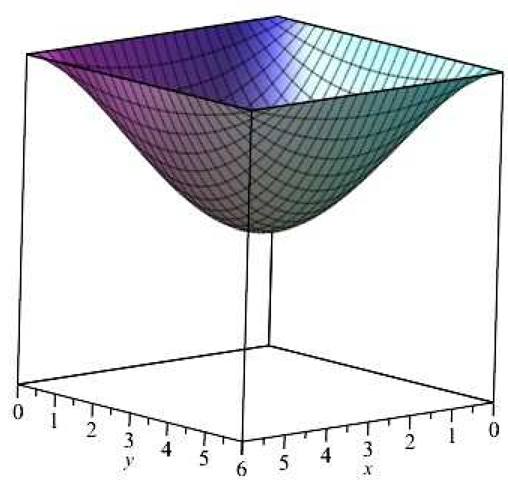


Figure 4.6: $M = M(x, y)$ is a functions of *two* variable, x and y: the coordinates of the plate.

Another example, the second moment of area, I, of a rectangle of height h and base b about the x axis is given by:

$$I = I(b, h) = \frac{bh^3}{12}.$$

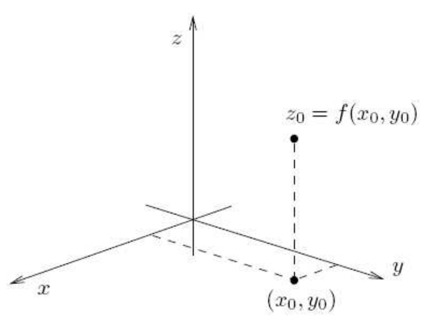
If we vary any two of these then the behaviour of the third can be calculated.

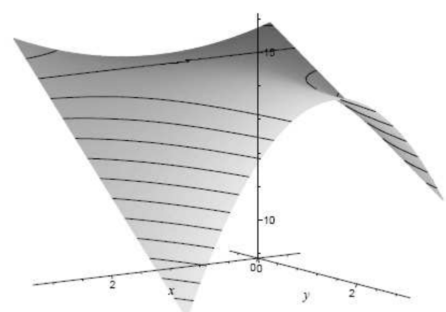
How I varies as we change h and b is easy to see from the above (e.g. when $b \to 2b$, $I \to 2I$; and when $h \to 2h$, $I \to 8h$), but we want to adapt the tools of one-variable calculus to help us investigate functions of more than one variable. For the most part we shall concentrate on functions of two variables such as $z = x^2 + y^2$ or $z = x \sin(y + e^x)$. Some examples of functions of two or more variables:

1. The area of a rectangle, A, depends on the width, w and length l; and we write $A = A(w, l)$. Indeed $A(w, l) = wl$.
2. For beams, we have that the slope, y'; and the deflection, y; both depend on the distance from one end of the beam, x, the Young's Modulus of the material, E, and the second moment of area of the beam, I; and we write $y' = y'(x, E, I)$ and $y = y(x, E, I)$.
3. The volume of a cylinder depends on the radius, r, and height, h:

$$V = V(r, h) = \pi r^2 h.$$

Graphically $z = f(x, y)$ describes a surface in 3D space — varying the x- and y-coordinates gives the z-coordinate, producing the surface:





4.1.4 Linear Operators

In first year, and Chapter 2, we get away with the simple words *differentiate* and *anti-differentiate*. In MATH6040 and MATH7019 we need to be a lot more precise with our notation. In particular, we need to understand what $\dfrac{d}{dx}$ means.

A *function* associates to an input x a unique output $f(x)$:

$$x \mapsto f(x).$$

In first year, the inputs and outputs were both real numbers but as maths gets more technical we need to allow for more complicated objects for inputs and outputs. So rather than just assuming that a function f maps numbers to numbers we might write:

$$f : \mathbb{R} \to \mathbb{R}.$$

Here $\mathbb{R}$ stands for the set of real numbers and this signifies that f sends real numbers to real numbers. If we write that g is a function and

$$g : A \to B,$$

it means that A is the collection/set of inputs and B is the collection/set of outputs.

Now if we take the collection of all possible functions from real numbers to real numbers[5], and denote it by $F(\mathbb{R})$ (you can kinda think as this as the set of all 2D graphs), then we can consider a function as a possible input to another function...

I have in my mind a function

$$\varphi : F(\mathbb{R}) \to F(\mathbb{R}).$$

Some 'values' of φ are:

$$\begin{aligned} \varphi((f(x) =)2) &= 0 \\ \varphi(3x-2) &= 3 \\ \varphi(x^2) &= 2x \\ \varphi(\sin x) &= \cos x \\ \varphi(e^x) &= e^x \\ \varphi(\ln x) &= \frac{1}{x}. \end{aligned}$$

What function is φ? It is a perfectly good function because it associates to each input a unique output.

This view can be very confusing with a function an input to another function so we can use the term *operator* (or *transform*) to describe a function which takes as input a function and produces as output a unique function. Now, being more specific, φ is the *differentiate with respect to x* operator. We shall denote it by $\frac{d}{dx}$. Therefore whenever you see $\frac{d}{dx}$ it means differentiate with respect to x:

$$\frac{d}{dx}\left(e^{\sin x}\right) = e^{\sin x} \cdot \cos x.$$

Similarly we have $\frac{d}{dt}$, $\frac{d}{dr}$ operators, etc.:

$$\begin{aligned} \frac{d}{dt}\left(t^2 \ln t\right) &= t^2 \cdot \frac{1}{t} + \ln t \cdot 2t = t + 2t \cdot \ln t \\ \frac{d}{dr}\left(\frac{4}{3}\pi r^3\right) &= \frac{4}{3}\pi \cdot 3r^2 = 4\pi r^2. \end{aligned}$$

Can anybody figure out what these operators $\mathcal{I}$ and $\mathcal{S}$ ar doing:

$$\begin{aligned} \mathcal{I}(f(x) = 2) &= 2x + C \\ \mathcal{S}(f(x) = 2) &= 0 \\ \mathcal{I}(\sin x) &= -\cos x + C \\ \mathcal{S}(\sin x) &= -\sin x \\ \mathcal{I}(x^4) &= \frac{x^5}{5} + C \\ \mathcal{S}(x^4) &= 12x^2 \end{aligned}$$

[5]and the following ignores many technicalities

The first operator, $\mathcal{I}$ is the anti-differentiate with respect to x operator:

$$\mathcal{I} = \int \cdot \, dx.$$

Strictly, speaking, $\int \cdot \, dx$ does not produce a single function as an output but a family of functions... this might have been addressed by your first year lecturer. The second operator, $\mathcal{S}$, is the 'differentiate with respect to x twice' operator:

$$\mathcal{S} = \frac{d}{dx}\frac{d}{dx} = \left(\frac{d}{dx}\right)^2 = \frac{d^2}{dx^2}.$$

In this chapter we will encounter a number of *differential operators*... and later in the chapter we will see functions that depend not one numeric input but two or more: these will yield the *partial derivative operators*:

$$\frac{\partial}{\partial x} \qquad \text{and} \qquad \frac{\partial}{\partial y},$$

as well as the products of these:

$$\frac{\partial}{\partial x}\frac{\partial}{\partial y} = \frac{\partial^2}{\partial x \partial y}, \qquad \text{,etc..}$$

In MATH7021 you will do a lot of work with the *Laplace Transform* operator, denoted $\mathcal{L}$; and if you go into Level 8 you should see the *Fourier Transform* operator, denoted by $\mathcal{F}$.

Partial Derivatives

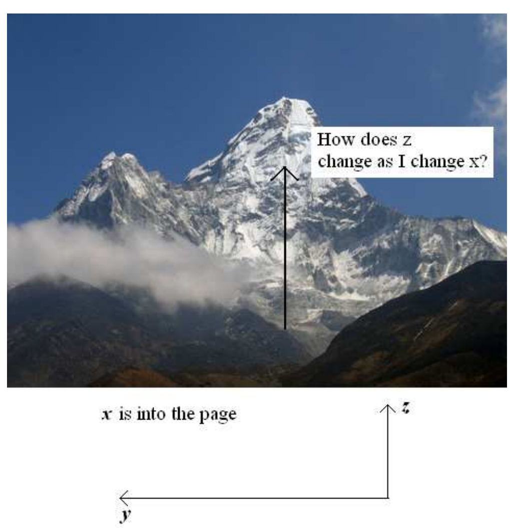


Figure 4.7: What is the rate of change in z as I keep y constant?

If we were to look at this from side-on:

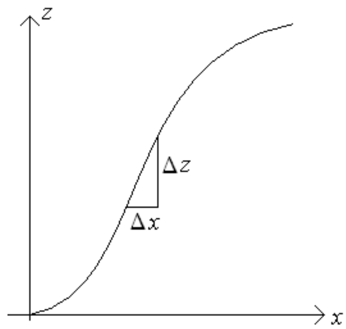


Figure 4.8: When y is a constant z can be considered a function of x only.

In general we have that $z = f(x, y)$; but if $y = b$ is fixed (constant):

Which is also the slope of the tangent to f at x. Hence the rate of change of $f(x, y)$ with respect to x at $x = a$ when y is fixed at $y = b$ is the slope of the surface in the x-direction.

Example

Let $z = f(x, y) = x^3 + x^2y^3 - 2y^3$. What is the rate of change[6] of z with respect to x when $y = 2$?

Solution:

More generally, we fix $y = y$, *some* constant, and define

as *the partial derivative of f with respect to x.* We define the partial derivative of f with respect to y in exactly the same way.

Example

What are the partial derivatives of

$$z = x^2 + xy^5 - 6x^3y + y^4$$

with respect to x and y respectively?

Solution:

[6]the same as the slope in the x-direction.

There are many alternative notations for partial derivatives. For instance, instead of $\frac{\partial f}{\partial x}$ we can write f_x or f_1. In fact,

$$\frac{\partial f}{\partial x} \equiv \frac{\partial z}{\partial x} \equiv f_x$$
$$\frac{\partial f}{\partial y} \equiv \frac{\partial z}{\partial y} \equiv f_y$$

We have $z = f(x, y)$ — z depends on x AND y. This is comparison to the single-variable case where $y = f(x)$ — y depends on x. To compute partial derivatives, all we have to do is remember that the partial derivative of a function with respect to x is the same as the *ordinary* derivative of the function g of a single variable that we get by keeping y fixed. Thus we have the following:

1. To find f_x, regard y as a constant and differentiate f with respect to x.
2. To find f_y, regard x as a constant and differentiate f with respect to y.

Using this technique we can make use of known results from one-variable theory such as the product, quotient and chain rules.
Exercises: Find all the first order derivatives of the following functions:

$$(i)\ f(x,y) = x^3 - 4xy^2 + y^4 \qquad (ii)\ f(x,y) = x^2e^y - 4y$$
$$(iii)\ f(x,y) = x^2 \sin xy - 3y^2 \quad (iv)\ f(x,y,z) = 3x \sin y + 4x^3y^2z$$

Solutions:

(i)

$$\frac{\partial f}{\partial x} = 3x^2 - 4y^2(1) + 0 = 3x^2 - 4y^2.$$

$$\frac{\partial f}{\partial y} = -4x(2y) + 4y^3 = 4y^4 - 8xy.$$

(ii)

$$\frac{\partial f}{\partial x} = e^y(2x) + 0 = 2xe^y.$$

$$\frac{\partial f}{\partial y} = x^2(e^y) - 4 = x^2e^y - 4.$$

(iii) This one needs a product and a chain rule for f_x and a chain rule for f_y.

$$\begin{aligned}\frac{\partial f}{\partial x} &= x^2 \times \frac{\partial \sin xy}{\partial x} + \sin xy \times \frac{\partial x^2}{x} + 0 \\ &= x^2 \times \cos xy \times \frac{\partial xy}{\partial x} + \sin xy \times 2x \\ &= x^2 \cos xy \times y + 2x \sin xy = x^2 y \cos xy + 2x \sin xy.\end{aligned}$$

$$\begin{aligned}\frac{\partial f}{\partial y} &= x^2 \times \cos xy \times \frac{\partial xy}{\partial y} - 6y \\ &= x^2 \times \cos xy \times x - 6y = x^3 \cos xy - 6y.\end{aligned}$$

(iv)

$$\frac{\partial f}{\partial x} = \sin y(3) + 4y^2z(3x^2) = 3(\sin y + 4x^2y^2z).$$

$$\frac{\partial f}{\partial y} = 3x(\cos y) + 4x^3z(2y) = 3x\cos y + 8x^3yz.$$

$$\frac{\partial f}{\partial z} = 0 + 4x^3y^2(1) = 4x^3y^2.$$

4.2 Differentials & Applications to Rounding Error

Differentials

Using the fact that many of the functions $y = f(x)$ that we see in engineering are analytic/*nice* and hence can be written as Taylor series, actually gives us a way of approximating the change in (the calculation of) f, Δf, when (the measurement) x is changed by Δx. We do this by first writing the Taylor Series for $y = f(x)$ about some point $x = a$:

where $\mathcal{O}((x-a)^2)$ just means that the terms beyond this point are powers of $(x-a)^2$ or greater. Now suppose that x changes from $x = a$ to $x = a + \Delta x$ so that we have $x - a = \Delta x$ so that our expression reads:

$$\underbrace{f(x)}_{\text{true}} = \underbrace{f(a)}_{\text{calculated}} + \underbrace{f'(a)\cdot\Delta x + \overbrace{\mathcal{O}((\Delta x)^2)}^{\approx 0}}_{\text{error}}$$

In context:

- a is a measurement,
- x is the true value of the quantity being measured,
- $\Delta x = x - a$ is the error in the measurement.

Now the next story is that for our later application of this business, to error analysis, we will be interested in changes in x, Δx, that are *small*. For example, consider the sine of an angle of elevation, x, $\sin x$. Suppose now we are interested in a Taylor Series of $\sin x$ about $x = \dfrac{\pi}{6} = 30^\circ$ where the change in x, $\Delta x = 0.1$. Now the higher powers of Δx are getting smaller and smaller and we say for small errors that:

$$\underbrace{f(a+\Delta x)}_{\text{true}} = f(a) + f'(a)\Delta x + \underbrace{\mathcal{O}((\Delta x)^2)}_{\approx 0} \approx \underbrace{f(a)}_{\text{calculated}} + \underbrace{f'(a)\Delta x}_{\approx \Delta f},$$

so approximately, when x goes from $x = a$ to $x = a + \Delta x$, the value of f goes from $f = f(a)$ to $f + \Delta f \approx f(a) + f'(a)\Delta x$ so that

$$\text{error in calculation} \ \approx\ \text{derivative at measurement} \ \times\ \text{measurement error}.$$

We call this the differential, df:

$$df := f'(a)\,\Delta x,$$

and this $df \approx \Delta f$, the change in f due to the change Δx in x. For a differentiable function of two variables $z = f(x, y)$, we might once again be interested in estimating the change in f when x changes to $x + \Delta x$ and y changes to $y + \Delta y$.

Propagation of Rounding Errors

In this section we see the effect of rounding error on your calculation. The lesson I want ye to learn here is to keep as many decimal places in your *intermediate* calculations as possible — only round when you are presenting your answer. If writing down intermediate steps you can write rounded figures but keep the full precision on your calculator. You saw in MATH6040 how:

Further Remark: Multivariable Taylor Series

Similarly to the way we can write a function of a single variable as a Taylor Series (for x near a):

$$\begin{aligned} f(x) &= f(a) + f'(a)(x-a) + \mathcal{O}((x-a)^2) \\ \Rightarrow f(x) &\approx f(a) + f'(a)(x-a). \end{aligned}$$

we can also write a function of *several* variables using a power series. For example, for (x, y) near (a, b)

$$f(x, y) = f(a, b) + \frac{\partial f}{\partial x}(x-a) + \frac{\partial f}{\partial y}(y-a) + \text{higher order terms}$$

$$\Rightarrow f(x, y) \approx f(a, b) + \frac{\partial f}{\partial x}(x-a) + \frac{\partial f}{\partial y}(y-a).$$

Therefore suppose $x = a + \Delta x \Rightarrow x - a = \Delta x$ and $y = b + \Delta y \Rightarrow y - b = \Delta y$. Then

$$\underbrace{f(a+\Delta x, b+\Delta y)}_{\text{true}} = f(a,b) + \frac{\partial f}{\partial x}\Delta x + \frac{\partial f}{\partial y}\Delta y + \underbrace{\text{higher order terms}}_{\approx 0} \approx \underbrace{f(a,b)}_{\text{calculated}} + \underbrace{\frac{\partial f}{\partial x}\Delta x + \frac{\partial f}{\partial y}\Delta y}_{\approx \Delta f},$$

A similar principle applies to rounding errors. Think of rounding errors as errors in measurement: how do they affect, and cause errors in, calculations? Suppose we have a formula for a property P related to two other properties A and B by:

Now suppose we have values A_0 and B_0, but we round them to say one decimal place. Then we have associated errors $\Delta A = 0.05$ and $\Delta B = 0.05$. We can now keep track of the errors in the *calculation* of P due to these rounding errors. The maximum absolute value of the differential $|dP|_{\max}$, gives an estimate of this:

The reason that we use absolute values is because we don't want rounding errors to cancel each other out: we want to allow for the worst case scenario. Typically we will have a quantity $P = P(A, B)$

with given values A_0 & B_0 rounded to so many decimal places, and our job will be to present the calculation of P as

$$P = P(A_0, B_0) \pm \Delta P.$$

It is good practise to

1. round the rounding error ΔP to *one significant figure* (because $|dP|_{\max} \approx \Delta P$ is rough)
2. match the number of decimal places/precision of P with ΔP

As engineers, all measurements should also come with units and therefore errors also. There will be marks deducted in this question for missing units or incorrect rounding/precision.

Examples

1. If a loaded beam has a bending moment (measured in kN m if x is in metres) given by

$$M(x) = \frac{1}{5}x^3 - 9x^2 + 40x,$$

estimate the error in the calculation of the bending moment if we round to one decimal place at

(a) $x = 1$ m.

(b) the location of the maximum bending moment, $x_{\max} = 15 - 5\frac{\sqrt{57}}{3} \approx 2.4$ m.

Solution: We can calculate that $M(1) = 31.2$ and $M(2.4) \approx 46.9248$

(a) If we are rounding to one decimal place, then the error in x, $\Delta x = 0.05$. Now we have

Now we differentiate to find

so that

$$M(1) = (31 \pm 1) \text{ kN m}$$

(b) Again we have $\Delta M \approx \left|\dfrac{\partial M}{\partial x}\right| \Delta x$. However at the maximum we have $\dfrac{\partial M}{\partial x} = 0$ so that even at $x = 2.4$ the error will be small:

Hence rounding is safe in this case.

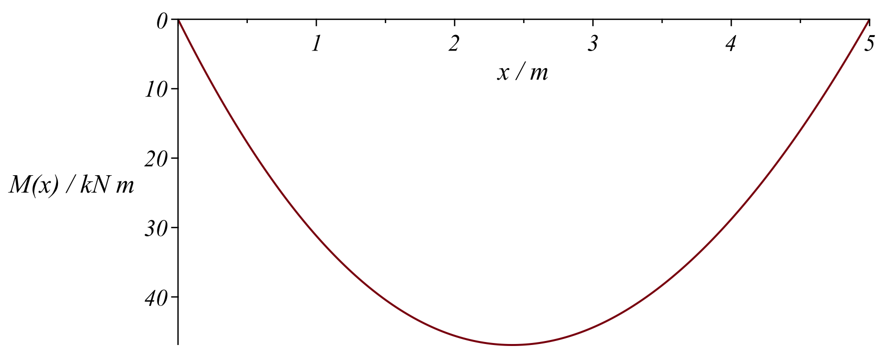


Figure 4.9: A small derivative means that rounding is safe(r).

2. **Autumn 2017** For a light beam of a particular length $L \approx 6.24$ m, simply supported under a load varying from 4 kN m^{-1} at $x = 0$ m to 16.48 kN m^{-1} at $x = L$ m, it can be shown that the bending moment, in kN m, at $x = 6$ m is given by:

$$M(a, b) = -216a - 72 + 6b.$$

where the *exact* values of a and b are $\dfrac{1}{3}$ and $\dfrac{280}{11}$ respectively.

i. Explain why if a and b are rounded to two decimal places that the rounding errors Δa and Δb are both equal to 0.005.

ii. Use differentials to estimate the rounding error in M, ΔM, caused by the rounding errors in a and b.

iii. Calculate the exact bending moment at $x = 6$, $M(1/3, 280/11)$.

iv. Calculate the approximate bending moment at $x = 6$, $M(0.33, 25.45)$, using $1/3 \approx 0.33$ and $280/11 \approx 25.45$.

Solution:

i. When rounding to two decimal places, the largest error is $0.005 \longrightarrow 0.01$.

ii. Using

$$\Delta M \approx |dM|_{\max} = \left|\frac{\partial M}{\partial a}\right| \Delta a + \left|\frac{\partial M}{\partial b}\right| \Delta b$$

iii. We have:

iv. We have:

Exercises: Recall, if $z = f(x, y)$, and $x = x_0 \pm \Delta x$, $y = y_0 \pm \Delta y$, then the calculation of z is $z_0 = f(x_0, y_0)$, and the error in this calculation is approximately:

$$\Delta z \approx \left|\frac{\partial z}{\partial x}\right| \Delta x + \left|\frac{\partial z}{\partial y}\right| \Delta y.$$

This approximation is rough so round this approximation to one significant figure. Round the calculation to the same precision/number of decimal places. Include units as appropriate.

1. **Autumn 2020** Suppose that a beam of length L and width w is uniformly supporting a mass of weight F. Then the stress induced is given by:

$$\sigma = \frac{F}{L \cdot w}.$$

Suppose an engineer uses the heavily rounded measurements $L = 5$ m, with a rounding error of 0.05 m; $F = 1400$ N, with a rounding error of 50 N; and $w = 0.4$ m, with a rounding error of 0.05 m.

i. Find the stress induced by the mass, σ_0, as calculated by the engineer. **Ans:** 700 N/m^2.

ii. Use

$$\Delta\sigma \approx |d\sigma|_{\max} = \left|\frac{\partial \sigma}{\partial F}\right| \Delta F + \left|\frac{\partial \sigma}{\partial L}\right| \Delta L + \left|\frac{\partial \sigma}{\partial w}\right| \Delta w,$$

to estimate the error in the calculation of σ due to the rounding errors. **Ans:** 100 N/m^2.

iii. Suppose that the weight is multiplied by a safety factor of 1.15. Is this safety factor large enough to 'cover' this error in the calculation of σ_0? Justify your answer.

2. **Winter 2019** Consider the quantity

$$Q = 200\pi.$$

Suppose rather than using a good approximation of π, an engineer uses $\pi \approx 3.14$ instead. This has a rounding error of 0.002.

(a) Use **differentials** to estimate the error in his calculation, ΔQ. Present your answer in the form $Q = Q_0 \pm \Delta Q$. **Ans:** $Q = 628.0 \pm 0.4$.

(b) Use a calculator to find the exact error in this calculation, $\Delta Q = |200\pi - 200(3.14)|$. **Ans:** $\Delta Q = 0.3$ rad.

(c) If Q is an angle measured in radians, $200\pi = 100 \times 2\pi$ represents 100 full rotations: $200\,\pi$ rad $\equiv 0$ rad. Where $Q_0 = 200 \times 3.14$, use part (a), or part (b), to estimate how far away, in degrees, the angle Q_0 is from 100 full rotations. [HINT: π rad $= 180°$] **Ans:** $\approx 17 - -23°$.

3. Suppose an engineer uses the formula $g(x) = \sqrt{x}$ with a value of $x \approx 16$ (correct to the nearest whole number).

(a) Explain why if $x \approx 16$ is correct to the nearest whole number, that the rounding error $\Delta x = 0.5$.

(b) Use **differentials** to estimate the error in their calculation, Δg. Present your answer in the form $g = g_0 \pm \Delta g$. **Ans:** $g = 4.00 \pm 0.06$

(c) Estimate the percentage error in this calculations. **Ans:** 1.5%

4. If a loaded beam has a bending moment, in kN m, given by

$$M(x) = x^3 - 2x^2 + 8x,$$

estimate the error in the calculation of the bending moment if we round to one decimal place at $x = 1$.

(a) Explain why if $x \approx 1$ is correct to one decimal place, that the rounding error $\Delta x = 0.05$.

(b) Use **differentials** to estimate the error in their calculation, ΔM. Present your answer in the form $M = M_0 \pm \Delta M$. **Ans:** $M = (7.0 \pm 0.4)$ kN m.

(c) Estimate the percentage error in this calculation. **Ans:** 5.7%.

5. **Winter 2018** An engineer wishes to calculate the area of a circle. The area of a circle is given by

$$A = \pi r^2.$$

He measures the radius to be 9.6 m with a rounding error of 5 cm = 0.05 m. Rather than using a good approximation to π, he uses 3.14 which has a rounding error of approximately 0.002.

(a) Use **differentials** to estimate the error in his calculation, ΔA, caused by these errors of measurement and of rounding. Present your answer in the form $A = A_0 \pm \Delta A$. **Ans**: $A = (289 \pm 3)$ m^2

(b) What is the percentage error in this calculation? **Ans:** 1.0%

---

6. Suppose an engineer uses the formula $f(x) = x^{10}$ with a value of $x \approx 10$ (correct to one decimal place).

(a) Explain why if $x \approx 10$ is correct to one decimal place, that the rounding error $\Delta x = 0.05$.

(b) Use **differentials** to estimate the error in their calculation, Δf. Present your answer in the form $f = f_0 \pm \Delta f$. **Ans:** $f = (10,000,000,000 \pm 500,000,000)$.

(c) Estimate the percentage error in this calculation. **Ans:** 5%

4.3 Chapter Summary

1. *Taylor Series* allow an approximation to a function of a single variable $y = f(x)$ near a point $x = a$. For $x \approx a$

$$f(x) \approx f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f'''(a)}{3!}(x-a)^3 + \cdots .$$

2. A *Maclaurin Series* is just an approximation to a function $y = f(x)$ near the point $a = 0$:

$$f(x) \approx f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \frac{f'''(0)}{3!}x^3 + \cdots .$$

3. *Three-Term-Taylor Method* allows approximations to $y(x_0 + kh) \approx y_k$ where $y(x)$ is the solution of a given differential equation

$$\frac{dy}{dx} = F(x, y)\,, \qquad y(x_0) = y_0.$$

The approximations $y_1, y_2, \ldots$ can be found using the iteration

$$y_{k+1} = y_k + h \cdot y'_k + \frac{h^2}{2} y''_k,$$

where $y'_k = y'(x_0 + kh)$ and $y''_k = y''(x_0 + kh)$.

4. *The Euler Method* is just a Two-Term-Taylor Method:

$$y_{k+1} = y_k + hy'_k.$$

5. *Partial Derivatives* are rates of change of functions of several variables. Equivalently slopes in *particular directions* on a surface.

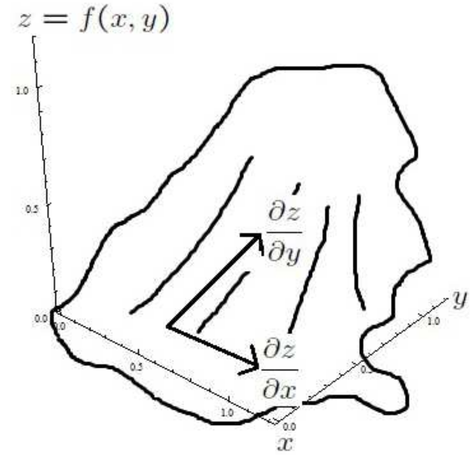


Figure 4.10: If we have a *surface*, then $\frac{\partial z}{\partial x}$ is the slope when we move in the x-direction. Similarly, $\frac{\partial z}{\partial y}$ is the slope when we move in the y-direction.

$$\frac{\partial z}{\partial x} = f_x = \left.\frac{dz}{dx}\right|_{y=\text{constant}}$$

$$\frac{\partial z}{\partial y} = f_z = \left.\frac{dz}{dy}\right|_{x=\text{constant}}$$

6. *Rounding Error Analysis using Differentials* is estimating the rounding error in the calculation of a function $z = f(x, y)$ due to rounding errors in the x and y. If $x = x_0 \pm \Delta x$ and $y = y_0 \pm \Delta y$ then

$$\Delta z \approx' dz' = \left|\frac{\partial f}{\partial x}\right|_{x_0,y_0} \cdot \Delta x + \left|\frac{\partial f}{\partial y}\right|_{x_0,y_0} \cdot \Delta y.$$

This formula is *not* in the tables.

Where $z_0 = f(x_0, y_0)$, present the calculation in the form

$$z_0 \pm \Delta z.$$

It is good practise to

(a) round the rounding error Δz to *one significant figure*

(b) match the number of decimal places/precision of z_0 with Δz

4.4 Function Catalogue

Constant Functions

1. **Definition** Let $k \in \mathbb{R}$. A *constant* function is of the form
$$f(x) = k.$$
An example of a constant function is $f(x) = 2$.

2. **Main Idea/Properties** A constant function outputs the same number for all inputs and so has a rate of change of zero.

3. **Derivative** The derivative of a constant (function) is zero:
$$\frac{d}{dx}(k) = 0.$$

Figure 4.11: The graph of a constant function. Note that the slope morryah the derivative is zero.

Lines

1. **Definition** Let $m \in \mathbb{R}$ and $c \in \mathbb{R}$. A *line of slope m and y-intercept c* is given by
$$f(x) = mx + c.$$
An example of a line is $f(x) = 3x - 2$.

2. **Main Idea/Properties** A line does exactly what it says on the tin. The slope/derivative of a line is a constant so the rate of change of a line is constant.

3. **Derivative** The derivative of a line is the slope:
$$\frac{d}{dx}(mx + c) = m.$$

Quadratics

1. **Definition** Let $a, b, c \in \mathbb{R}$. A quadratic is a function of the form
$$f(x) = ax^2 + bx + c.$$
An example of a quadratic is $f(x) = x^2 + 1$.

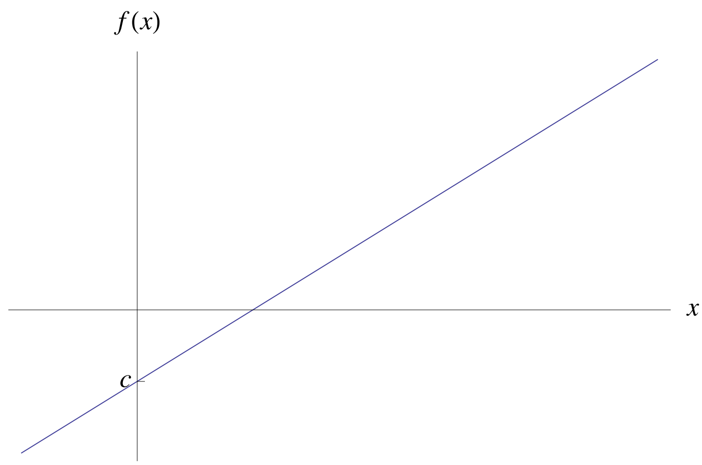


Figure 4.12: The graph of a line. Note that the slope morryah the derivative is constant.

2. **Main Idea/Properties** A quadratic either has a $\bigcup$ shape (when $a > 0$) or a $\bigcap$ shape (when $a < 0$). It has two *roots* given by the $\dfrac{-b \pm \sqrt{\cdots}}{2a}$ formula. If they are both *real* (when $b^2 - 4ac > 0$), then the graph cuts the x-axis at two points. The graph is symmetric about the max/min. Hence the max/min can be found by looking at $f'(x) = 0$ or else be found at the midpoint of the roots. If $b^2 - 4ac < 0$ then the roots contain a $\sqrt{(-)}$ — *complex roots.*

3. **Derivative** The derivative of a quadratic is a line!

$$\frac{d}{dx}(ax^2 + bx + c) = a(2x) + b(1) + 0 = 2ax + b.$$

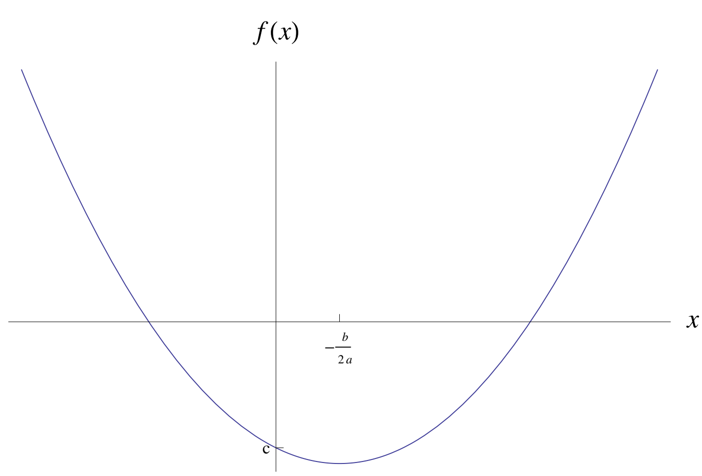


Figure 4.13: The graph of a $+x^2$ quadratic with $a > 0$. Note that the slope goes from negative to zero to positive — like a line. This quadratic has two real roots and the minimum occurs at $-\dfrac{b}{2a}$. At this point the tangent is horizontal. This point can be found by differentiating $ax^2 + bx + c$, e.g. getting the slope, and setting it equal to zero.

Polynomials

1. **Definition** Let $a_n, a_{n-1}, \dots, a_2, a_1, a_0 \in \mathbb{R}$. A *polynomial of degree n* is a function of the form

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_2 x^2 + a_1 x + a_0.$$

   As example of a polynomial is $f(x) = x^5 - 3x^3 + 2x^2 + 1$.

2. **Main Idea/Properties** A polynomial of degree n has n roots — some of which may be complex, some of which may be repeated. However if all the roots are real and distinct then the polynomial cuts the x-axis n times. The derivative of a polynomial of degree n is a polynomial of degree $n-1$;

$$\text{e.g.} \quad \frac{d}{dx}(x^5 - 3x^3 + 2x^2 + 1) = 5x^4 - 3(3x^2) + 2(2x) = 5x^4 - 9x^2 + 4x,$$

   which has potentially $n-1$ real roots and hence $n-1$ points where $f'(x) = 0$ — potentially $n-1$ turning points.

   As an example note that quadratics are degree two polynomials and have one turning point.

3. **Derivative** We differentiate a polynomial using the Sum, Scalar & Power Rules:

$$\frac{d}{dx}(ax^n) = a\frac{d}{dx}x^n = a(nx^{n-1}) = anx^{n-1}.$$

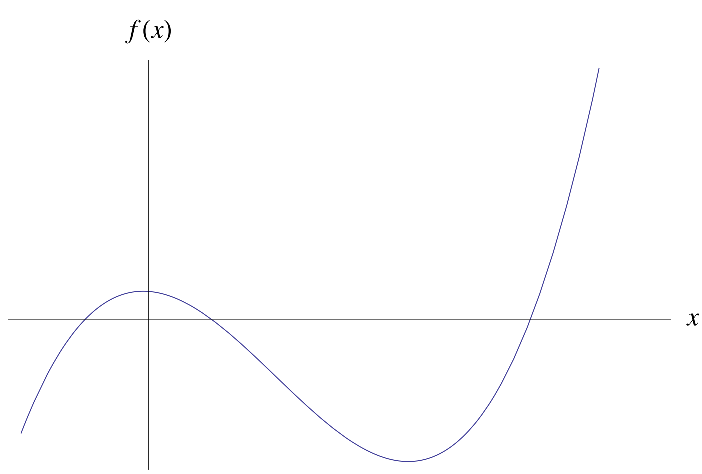


Figure 4.14: This is an example of a cubic: $ax^3 + bx^2 + cx + d$. Note that it has *three* real roots and *two* turning points because $f'(x) \sim x^2$. In some sense this is typical behaviour of polynomials.

Roots

1. **Definition** Let $n \in \mathbb{N}$. The *nth root function* is a function of the form:

$$f(x) = \sqrt[n]{x},$$

   the *positive n-th root* of x.

2. **Main Idea/Properties** We can show that if we define

$$x^{1/n} = \sqrt[n]{x}$$

then all of the theorems of indices and differentiation work properly with this definition and it turns out that $\sqrt[n]{x} = x^{1/n}$ written as a power can be differentiated using the Power Rule.

3. **Derivative** Using the power rule

$$\frac{d}{dx}\sqrt[n]{x} = \frac{d}{dx}x^{1/n} = \frac{1}{n}x^{1/n-1}.$$

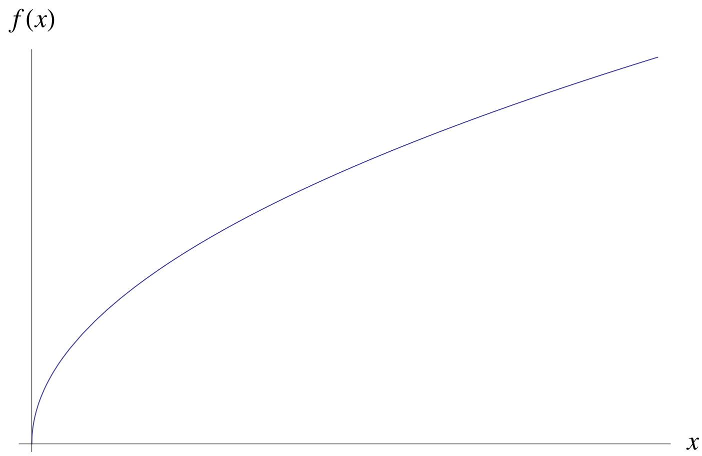


Figure 4.15: A plot of the square root function $f(x) = \sqrt{x}$. Note that roots are only defined for *positive* values of x... for example $\sqrt{-4}$ is not a real number.

Trigonometric

1. **Definition** Let $0 \leq \theta \leq 2\pi$ be an angle. Consider now the unit circle

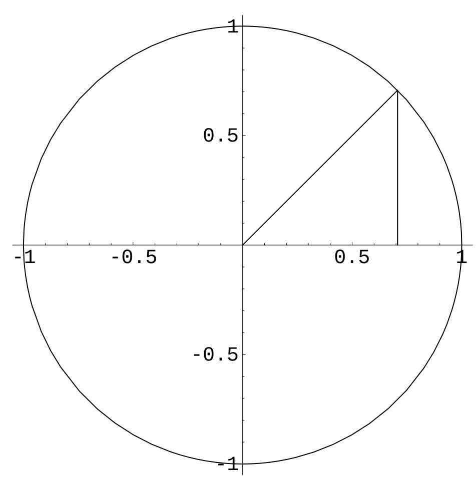


Figure 4.16: If the angle made with the positive x-axis is θ, then the coordinate of the point on the circle is $(\cos\theta, \sin\theta)$.

This defines cosine and sine for angles between 0 and 2π. The definition is extended by periodicity to the whole of the number line by

$$\cos(\theta + 2\pi) = \cos(\theta)$$
$$\sin(\theta + 2\pi) = \sin\theta$$

We also define

$$\tan\theta = \frac{\sin\theta}{\cos\theta}$$

2. **Main Idea/Properties** Sine and Cosine are waves that oscillate between ± 1. Sine begins at zero ($\sin 0 = 0$) while cosine begins at one ($\cos 0 = 1$). Apart from this they are very similar: the graph of sine is got by shifting the graph of cosine $\pi/2$ units to the right.

3. **Derivative** We can show that

$$\frac{d}{dx}\sin x = \cos x$$
$$\frac{d}{dx}\cos x = -\sin x.$$

Using the Quotient Rule we have

$$\frac{d}{dx}\tan x = \sec^2 x = (\sec x)^2,$$

where

$$\sec x := \frac{1}{\cos x}. \tag{4.5}$$

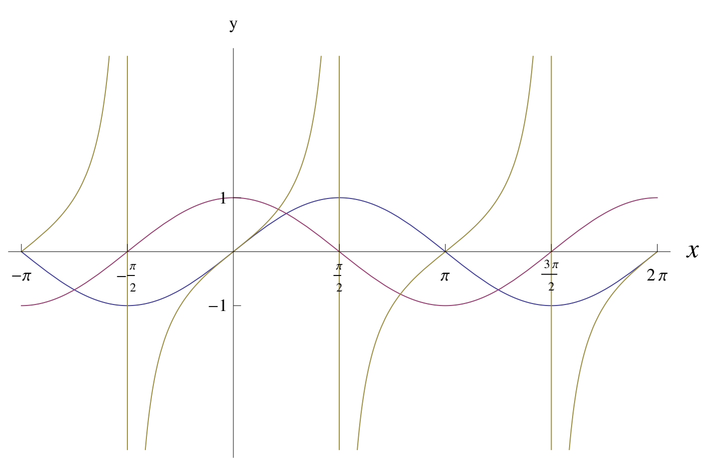


Figure 4.17: Note that $|\sin x|, |\cos x| \leq 1$ while $\tan x \to \pm\infty$ at $\pi/2$.

Inverse Trigonometric

1. **Definition** These functions are *inverse functions* of the trigonometric functions.

   Let $y \in [-\pi/2, \pi/2]$ and $x \in [-1, 1]$:
   $$y = \sin^{-1}(x) \Leftrightarrow x = \sin y. \tag{4.6}$$
   Let $y \in [-\pi/2, \pi/2]$:
   $$y = \tan^{-1}(x) \Leftrightarrow x = \tan y. \tag{4.7}$$

2. **Main Idea/Properties** These are the inverse functions of $\sin x$ and $\tan x$. For example
   $$\sin\theta = \frac{1}{2} \Rightarrow \sin^{-1}(\sin\theta) = \sin^{-1}(1/2) \Rightarrow \theta = \frac{\pi}{6}.$$
   In other words, $\sin^{-1}(x)$ asks for the angle — *between* $\pm\pi/2$ — which has a sine of x.

3. **Derivative** We can show that
   $$\frac{d}{dx}\sin^{-1}\left(\frac{x}{a}\right) = \frac{1}{\sqrt{a^2 - x^2}} \tag{4.8}$$
   and
   $$\frac{d}{dx}\tan^{-1}\left(\frac{x}{a}\right) = \frac{a}{a^2 + x^2}. \tag{4.9}$$

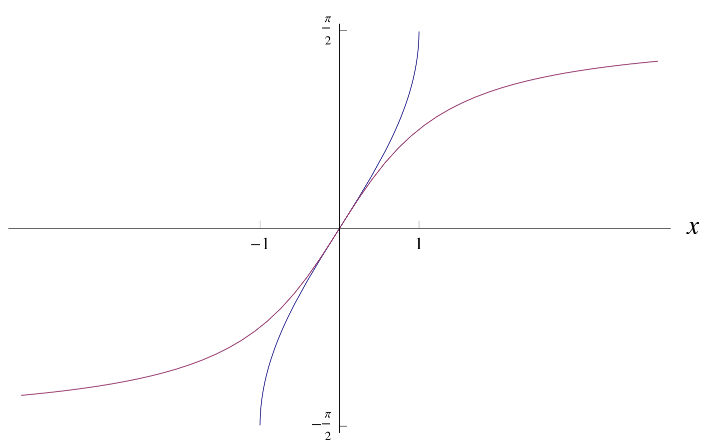


Figure 4.18: Inverse Sine can only take inputs between ± 1. However we have $\tan^{-1}(x) \to \pi/2$ as $x \to \infty$.

Exponential

1. **Definition** The *exponential function* can be defined as a power series

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots \tag{4.10}$$

2. **Main Idea/Properties** The exponential function is the unique function that is equal to its own derivative.

3. **Derivative** We have that

$$\frac{d}{dx} e^x = e^x, \tag{4.11}$$

and using the Chain Rule

$$\frac{d}{dx} e^{ax} = a \cdot e^{ax}. \tag{4.12}$$

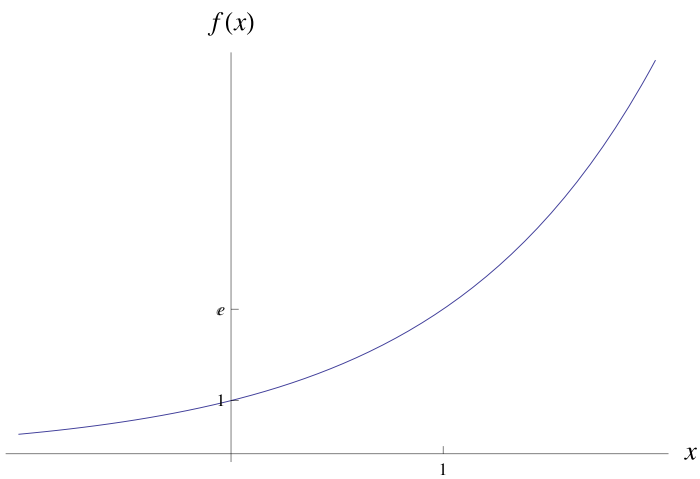


Figure 4.19: Note that $e \approx 2.7183$ is a constant. We have that $e^0 = 1$. $e^x \to \infty$ as $x \to +\infty$. When the input is negative, say $x = -N$, then $e^x = e^{-N} = \frac{1}{e^N} \to 0$ as $N \to \infty \Leftrightarrow x \to -\infty$.

Logarithmic

1. **Definition** The natural logarithm is the inverse function of e^x:

$$y = \ln x \Leftrightarrow x = e^y \tag{4.13}$$

2. **Main Idea/Properties** As $e^y > 0$, the natural logarithm can only take strictly positive inputs. They can be used to solve exponential equations:

$$e^x = 2 \Rightarrow \ln(e^x) = \ln 2 \Rightarrow x = \ln 2.$$

   Note that we have

$$\begin{aligned} \ln 1 &= 0 \\ \ln(xy) &= \ln x + \ln y \\ \ln(x^n) &= n \ln x. \end{aligned}$$

3. **Derivative** We can show that

$$\frac{d}{dx} \ln x = \frac{1}{x}.$$

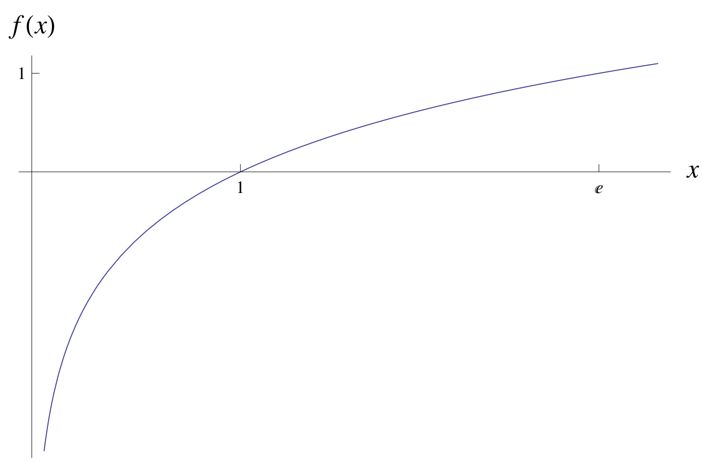


Figure 4.20: As $x \to 0$, $\ln x \to -\infty$; $\ln 1 = 0$ and $\ln x \to \infty$ — slowly — as $x \to \infty$.