Chapter 3
Differentiation

Essential reading

(For full publication details, see Chapter 1.)

☞ Anthony and Biggs (1996) Chapters 6, 7 and 8.

Further reading

☞ Binmore and Davies (2001) Chapter 2, Sections 2.7–2.10 and Chapter 4, Sections 4.2 and 4.3.

☞ Booth (1998) Chapter 5, Modules 19 and 20.

☞ Bradley (2008) Chapter 6.

☞ Dowling (2000) Chapters 3 and 4.

3.1 Introduction

In this extremely important chapter we introduce the topic of calculus, one of the most useful and powerful techniques in applied mathematics. In this chapter we focus on the process of 'differentiation' of a function. The derivative (the result of differentiation) has numerous applications in economics and related fields. It provides a rigorous mathematical way to measure how fast a quantity is changing, and it also gives us the main technique for finding the maximum or minimum value of a function.

3.2 The definition and meaning of the derivative

The derivative is a measure of the instantaneous rate of change of a function[1] $f : \mathbb{R} \to \mathbb{R}$. The idea is to compare the value of the function at x with its value at $x + h$, where h is a small quantity. The change in the value of f is $f(x + h) - f(x)$, and this, when divided by the change h in the 'input', measures the average rate of change over the interval from x to $x + h$. Informally speaking, the *instantaneous* rate of change is the quantity this average rate of change approaches as h gets smaller and smaller.

A good analogy can be made with the speed of a car. Imagine that a car is driving along a straight road and that $f(x)$ represents its distance, in metres, from the starting

[1]See Anthony and Biggs (1996) Section 6.1.

point at time x, in seconds, from the start. (It would be more normal to use the symbol t as the variable here rather than x, but you will be aware from earlier that $f(x)$ and $f(t)$ convey the same information: it does not matter which symbol is used for the variable.) Let's suppose that at time $x = 10$, the distance $f(10)$ from the start is 150 and that at time $x = 11$, the distance from the start is 170. Then the average speed between times 10 and 11 is $(170 - 150)/(11 - 10) = 20$ metres per second. However, this need not be the same as the *instantaneous* speed at 10, since the car may accelerate or decelerate in the time interval from 10 to 11. Conceivably, then, the instantaneous speed at 10 could well be higher or lower than 20. To obtain better approximations to this instantaneous speed, we should measure average speed over smaller and smaller time intervals. In other words, we should measure the average speed from 10 to $10 + h$ and see what happens as h gets smaller and smaller. That is, we compute the limit of

$$\frac{f(10+h) - f(10)}{(10+h) - 10} = \frac{f(10+h) - f(10)}{h},$$

as h tends to 0.

We now give the definition of the derivative. The *derivative* (or instantaneous rate of change) of f at a number a (or 'at the point a') is the number which is the limit of

$$\frac{f(a+h) - f(a)}{h},$$

as h tends towards 0. It is not appropriate at this level to say formally what we mean by a limit, but the idea is quite simple: we say that $g(x)$ tends to the limit L as x tends to c if the distance between $g(x)$ and L can be made as small as we like provided x is sufficiently close to c. The derivative of f at a is denoted $f'(a)$. (We are assuming here that the limit exists: if it does not, then we say that the derivative does not exist at a. But we do not need to worry in this subject about the existence and non-existence of derivatives: these are matters for consideration in a more advanced course of study.) Now, if the derivative exists at all a, then for each a we have a derivative $f'(a)$ and we simply call the function f' the *derivative of f*. (Please don't be confused by this distinction between the derivative of f and the derivative of f at a point. For example, suppose that for each a, $f'(a) = 2a$. Then the derivative of f is the function f' given by $f'(x) = 2x$.)

Let's look at an example to make sure we understand the meaning of the derivative at a point a. (We will see soon that the type of numerical calculation we're about to undertake is not necessary in most cases, once we have learned some techniques for determining derivatives.)

Example 3.1 Suppose that $f(x) = 2^x$. Let's try to determine the derivative $f'(1)$ by working out the average rates of change

$$\frac{f(1+h) - f(1)}{h},$$

for successively smaller values of h. (As just mentioned, we will later see an easier way.) Table 3.1 shows some of the values of

$$\frac{f(1+h) - f(1)}{h} = \frac{2^{1+h} - 2}{h}.$$

h	$(2^{1+h} - 2)/h$
0.5	1.656854
0.1	1.435469
0.01	1.39111
0.001	1.386775
0.0001	1.386342

Table 3.1: Average rates of change for $f(x) = 2^x$.

It can be seen that as h approaches 0 these numbers seem to be approaching a number around 1.386. So we might guess that $f'(1)$ is around 1.386. (In fact, it turns out that the exact value of $f'(1)$ is $2 \ln 2 = 1.38629436\ldots$)

Activity 3.1 Calculate some more values of

$$\frac{f(1+h) - f(1)}{h},$$

for even smaller values of h.

A geometrical interpretation of the derivative can be given. The ratio

$$\frac{f(a+h) - f(a)}{h},$$

is the gradient of the line joining the points $(a, f(a))$ and $(a+h, f(a+h))$. As h tends to 0, this line becomes tangent to the curve at $(a, f(a))$; that is, it just touches the curve at that point. The derivative $f'(a)$ may therefore also be thought of as the gradient of the tangent to the curve $y = f(x)$ at the point $(a, f(a))$.

An alternative notation for $f'(x)$ is $\dfrac{df}{dx}$.

Derivatives can be calculated using the definition given above, in what is known as differentiation *from first principles*, but this is often cumbersome and you will not need to do this in an examination. We give one example by way of illustration, but we emphasise that you are not expected to carry out such calculations.

Example 3.2 Suppose $f(x) = x^2$. In order to work out the derivative we calculate as follows:

$$\frac{f(x+h) - f(x)}{h} = \frac{(x+h)^2 - x^2}{h} = \frac{(x^2 + 2xh + h^2) - x^2}{h} = 2x + h.$$

The first term is independent of h and the second term approaches 0 as h approaches 0, so the derivative is the function given by $f'(x) = 2x$.

3.3 Standard derivatives

In practice, to determine derivatives (that is, to *differentiate*), we have a set of *standard derivatives* together with rules for combining these. The standard derivatives (which you should memorise) are listed in Table 3.2.

$f(x)$	$f'(x)$
x^k	kx^{k-1}
e^x	e^x
$\ln x$	$1/x$
$\sin x$	$\cos x$
$\cos x$	$-\sin x$

Table 3.2: Standard derivatives.

We mentioned in Chapter 2 that the number e is very special. We can now see one reason why. We see from above that the derivative of the power function e^x is just itself, that is e^x. This is *not* the case for any other power function a^x. For example (as we shall see), the derivative of 2^x is *not* 2^x, but $2^x(\ln 2)$.

These could also be stated in the 'd/dx' notation, as $\dfrac{d}{dx}(e^x) = e^x$ etc.

Example 3.3 The derivative of x^5 is $5x^4$.

Activity 3.2 What is the derivative of $\dfrac{1}{x}$?

3.4 Rules for calculating derivatives

To calculate the derivatives of functions other than the standard ones just given, it is useful to use the following rules.[2]

- *The sum rule:* If $h(x) = f(x) + g(x)$ then
$$h'(x) = f'(x) + g'(x).$$
- *The product rule:* If $h(x) = f(x)g(x)$ then
$$h'(x) = f'(x)g(x) + f(x)g'(x).$$
- *The quotient rule:* If $h(x) = f(x)/g(x)$ and $g(x) \neq 0$ then
$$h'(x) = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}.$$

[2] See Anthony and Biggs (1996) Section 6.2.

3

Example 3.4 Let $f(x) = x^3 e^x$. Then, by the product rule,

$$f'(x) = (x^3)'e^x + x^3(e^x)' = 3x^2e^x + x^3e^x.$$

Activity 3.3 Find the derivative of $x^2 \sin x$.

Activity 3.4 Find the derivative of $f(x) = (x^2 + 1) \ln x$.

Example 3.5 Let $f(x) = \dfrac{\ln x}{x}$. Then, by the quotient rule,

$$f'(x) = \frac{(1/x)x - (1)\ln x}{x^2} = \frac{1 - \ln x}{x^2}.$$

Activity 3.5 Determine the derivative of $\sin x/x$.

Another, very important, rule is the *composite function rule*,[3] or *chain rule*, which may be stated as follows:

- If $f(x) = s(r(x))$, then $f'(x) = s'(r(x))r'(x)$.

If you can write a function in this way, as the composition of s and r, then the composite function rule will tell you the derivative.

Example 3.6 Let $f(x) = \sqrt{x^3 + 2}$. Then $f(x) = s(r(x))$ where

$$s(x) = \sqrt{x} = x^{1/2} \quad \text{and} \quad r(x) = x^3 + 2.$$

Now, we have

$$s'(x) = \frac{1}{2}x^{1-1/2} = \frac{1}{2}x^{-1/2} \quad \text{and} \quad r'(x) = 3x^2,$$

so, by the composite function rule,

$$f'(x) = s'(r(x))r'(x) = \frac{1}{2}(x^3 + 2)^{-1/2}(3x^2) = \frac{3x^2}{2\sqrt{x^3 + 2}}.$$

Example 3.7 Suppose $f(x) = (ax + b)^n$. Then, by the composite function rule,

$$f'(x) = n(ax + b)^{n-1} \times (ax + b)' = an(ax + b)^{n-1}.$$

[3] See Anthony and Biggs (1996) Section 6.4.

Example 3.8 Suppose that $f(x) = \ln(g(x))$. Then, by the composite function rule,

$$f'(x) = \frac{1}{g(x)} g'(x) = \frac{g'(x)}{g(x)}.$$

(This result is often useful in integration, something we discuss later.)

Activity 3.6 Find the derivative of $(3x + 7)^{15}$.

Activity 3.7 Differentiate $f(x) = \sqrt{x^2 + 1}$.

Activity 3.8 Differentiate $g(x) = \ln(x^2 + 2x + 5)$.

Differentiation can sometimes be simplified by taking logarithms, as the following example demonstrates.

Example 3.9 We differentiate $f(x) = 2^x$ by observing that $\ln(f(x)) = \ln(2^x) = x \ln 2$. Now, the derivative of $\ln(f(x))$ is, by the composite function rule (chain rule) equal to $f'(x)/f(x)$, so, on differentiating both sides of $\ln(f(x)) = x \ln 2$, we obtain

$$\frac{f'(x)}{f(x)} = (x \ln 2)' = \ln 2,$$

and so

$$f'(x) = (\ln 2) f(x) = (\ln 2) 2^x.$$

In particular, $f'(1) = 2 \ln 2$, as alluded to in an earlier example in this chapter.

Activity 3.9 By taking logarithms first, find the derivative $f'(x)$ when $f(x) = x^x$.

3.5 Optimisation

Critical points

The derivative is very useful for finding the *maximum* or *minimum* value of a function — that is, for *optimisation*.[4] Recall that the derivative $f'(x)$ may be interpreted as a measure of the rate of change of f at x. It follows from this that we can tell whether a function is increasing or decreasing at a given point, simply by working out its derivative at that point.

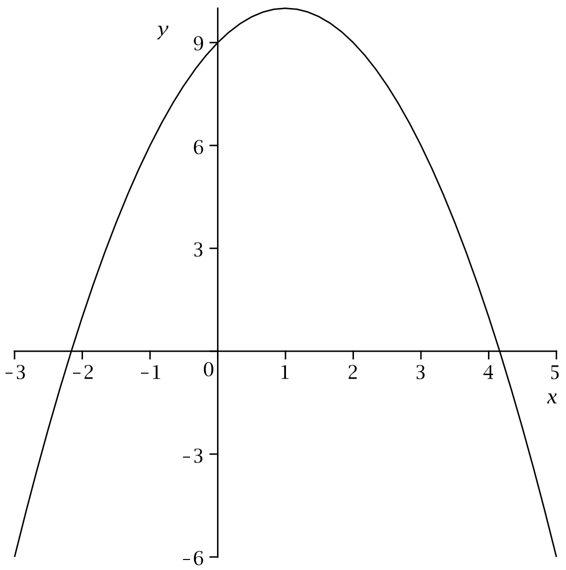
[4] See Anthony and Biggs (1996) Chapter 8.

- *If $f'(x) > 0$, then f is increasing at x.*
- *If $f'(x) < 0$, then f is decreasing at x.*

At a point c for which $f'(c) = 0$ the function f is neither increasing nor decreasing: in this case we say that c is a *critical point* (or *stationary point*) of f. It must be stressed that a function can have more than one kind of critical point. A critical point could be

- a *local maximum*, which is a c such that for all x close to c, $f(x) \leq f(c)$;
- a *local minimum*, which is a c such that for all x close to c, $f(x) \geq f(c)$; or
- an *inflexion point*, which is neither a local maximum nor a local minimum.

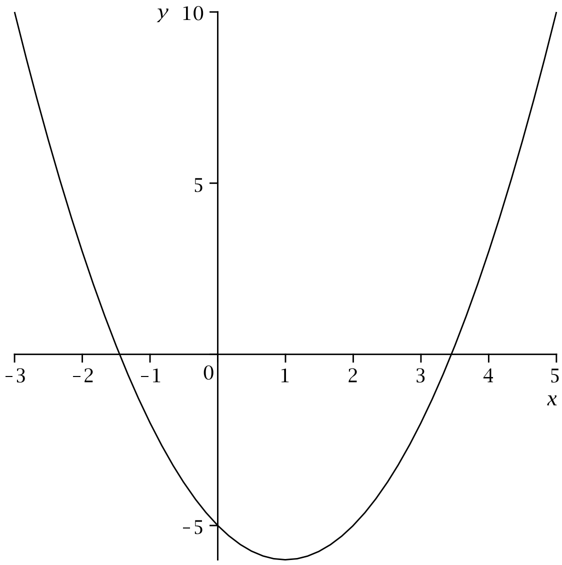
In the first of the following three figures, $c = 1$ is a local maximum of the function whose graph is sketched, and in the second $c = 1$ is a local minimum of the function sketched. In the third figure, $c = 1$ is a critical point, but not a maximum or a minimum; in other words, it is an inflexion point.

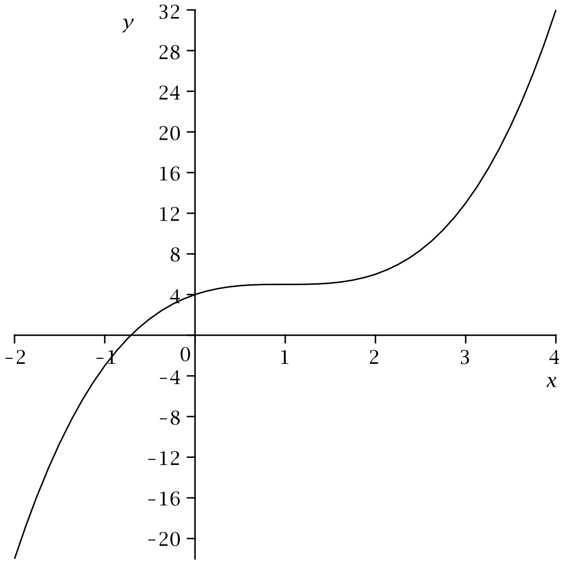


Deciding the nature of a given critical point

We can decide the nature of a given critical point by considering what happens to the derivative f' in a region around the critical point. Suppose, for example, that c is a critical point of f and that f' is positive for values just less than c, zero at c, and negative for values just greater than c. Then, f is increasing just before c and decreasing just after c, so c is a local maximum. Similarly, if the derivative f' changed sign from negative to positive around the point c then we can deduce that c is a local minimum. At an inflexion point, the derivative would not change sign: it would be either non-negative on each side of the critical point, or non-positive on each side. Thus a critical point can be *classified* by considering the sign of f' on either side of the point.

There is another way of classifying critical points. Let's think about a local maximum point c as described above. Note that the derivative f' is decreasing at c (since it goes





from positive, through 0, to negative), so the derivative of the derivative f' is negative at c. We call the derivative of f' the *second derivative* of f, and denote it by $f''(x)$ or

$$\frac{d^2 f}{dx^2}.$$

In other words, then, $f''(c) < 0$. It can be proved that a *sufficient* condition for f to have a local maximum at a critical point c is $f''(c) < 0$. (By saying that this is a 'sufficient condition', we mean that if $f''(c) < 0$ then c is a maximum; you should understand that even if c is a maximum, it need not be the case that $f''(c) < 0$.) There is a similar condition for a minimum, in which the corresponding condition is $f''(c) > 0$. Summarising, we have,

- *if $f'(a) = 0$ and $f''(a) < 0$, then $x = a$ is a local maximum of f;*
- *if $f'(b) = 0$ and $f''(b) > 0$, then $x = b$ is a local minimum of f.*

These observations together form the *second-order conditions* for the nature of a critical point. If a critical point c is an inflexion point, then the condition $f''(c) = 0$ must hold (since the point is neither a local maximum nor a local minimum). However, as mentioned above, if f'' is zero at a critical point then we *cannot* conclude that the point is an inflexion point. For example, if $f(x) = x^6$ then $f''(0) = 0$, but f does not have an inflexion point at 0; it has a local minimum there.

Example 3.10 Let's find the critical points of the function

$$f(x) = 2x^3 - 9x^2 + 1,$$

and determine the natures of these points. The derivative is

$$f'(x) = 6x^2 - 18x = 6x(x - 3).$$

The solutions to $f'(x) = 0$ are 0 and 3 and these are therefore the critical points. To determine their nature we could examine the sign of f' in the vicinity of each point, or we could check the sign of $f''(x)$ at each. For completeness of exposition, we shall do both here, but in practice you only need to carry out one of these tests.

First, let's examine the sign of $f'(x)$ in the vicinity of $x = 0$. We have $f'(x) = 6x(x - 3)$, which is positive for $x < 0$ (since it is then the product of two negative numbers). For x just greater than 0, $x > 0$ and $x - 3 < 0$, so that $f'(x) < 0$. (Note: we are interested only in the signs of $f'(x)$ *just* to either side of the critical point, in its immediate vicinity). Thus, at $x = 0$, f' changes sign from positive to negative and hence $x = 0$ is a local maximum. Now for the other critical point. When x is just less than 3, $6x(x - 3) < 0$ and when $x > 3$, $6x(x - 3) > 0$; thus, since f' changes from negative to positive around the point, $x = 3$ is a local minimum.

Alternatively, we note that $f''(x) = 12x - 18$. Since $f''(0) < 0$, $x = 0$ is a local maximum. Since $f''(3) > 0$, $x = 3$ is a local minimum.

Identifying local and global maxima

Now we turn to the problem of optimisation. Suppose we want to find the maximum value of a function $f(x)$. Such a wish only makes sense if the function has a maximum value; in other words, it does not take unboundedly large values. This value will occur at a local maximum point, but there may be several local maximum points. The *global maximum* is where the function attains its absolute maximum value (if such a value exists) and we can think of the local maximum points as giving the maximum value of the function in their vicinity. It should be emphasised that not all functions will have a global maximum. For instance, the function $f(x) = 2x^3 - 9x^2 + 1$ considered in the example above has no global maximum because the values $f(x)$ get increasingly large, without bound, for large positive values of x. Even though this function does, as we have seen, possess a local maximum, it does not have a global maximum.

If f does indeed have a global maximum, then we can find it as follows. We proceed by determining all the local maximum points of f, using the techniques outlined above, and then we calculate the corresponding values $f(x)$ and compare these to find the largest. (Of course, if there is only one local maximum, then it is the global maximum.) The analogous procedure is carried out if we want to find the global minimum value (if the function has one): we find the minimum points and, among these, find which gives the smallest value of f. These techniques are, like many other things in this subject, best illustrated by examples.

Example 3.11 To find the maximum value of the function $f(x) = xe^{-x^2}$, we first calculate the derivative, using the product rule to get

$$f'(x) = e^{-x^2} - (2x)xe^{-x^2} = e^{-x^2}(1 - 2x^2).$$

There are two solutions of $f'(x) = 0$, namely $x = 1/\sqrt{2}$ and $x = -1/\sqrt{2}$. (Note that e^{-x^2} is never equal to 0.) In other words, these values of x give the critical points, or stationary points. To determine their nature we could examine the sign of f' in the vicinity of each point, or we could check the sign of $f''(x)$ at each. For completeness of exposition, we shall do both here, but in practice you only need to use one method.

First, let's examine the sign of $f'(x)$ as x goes from just less than $-1/\sqrt{2}$ to just greater than $-1/\sqrt{2}$. For $x < -1/\sqrt{2}$, $1 - 2x^2 < 0$ and so $f'(x) < 0$, while for x just greater than $-1/\sqrt{2}$, $1 - 2x^2 > 0$ and $f'(x) > 0$. It follows that $-1/\sqrt{2}$ is a local minimum. In a similar way, one can check — and you should do this — that for x just less than $1/\sqrt{2}$, the derivative is positive and for x just greater than $1/\sqrt{2}$, the derivative is negative, so that we may deduce $1/\sqrt{2}$ is a local maximum.

Alternatively, we may calculate $f''(x)$, using the product rule to get

$$f''(x) = -2xe^{-x^2}(1 - 2x^2) - 4xe^{-x^2}.$$

Now, $f''(-1/\sqrt{2}) > 0$, so this point is a local minimum, and $f''(1/\sqrt{2}) < 0$, so this point is a local maximum.

Now, we are trying to find the maximum value of f. This is when $x = 1/\sqrt{2}$, and the maximum value is

$$f\left(\frac{1}{\sqrt{2}}\right) = \left(\frac{1}{\sqrt{2}}\right)e^{-1/2} = \frac{1}{\sqrt{2e}}.$$

(Note: this function does indeed have a global maximum and a global minimum. This might not be obvious, but it follows from that fact that for very large positive x or very 'large' negative x, xe^{-x^2} is extremely small in size.)

Activity 3.10 Find the critical points of $f(x) = x^3 - 6x^2 + 11x - 6$ and classify the nature of each such point (that is, determine whether the point is a local maximum, local minimum, or inflexion).

If we are trying to find the maximum value of a function $f(x)$ on an interval $[a, b]$, then it will occur either at a or at b, or at a critical point c in between a and b. Suppose, for instance, there was just one critical point c in the open interval (a, b), and that this was a local maximum. To be sure that it gives the maximum value on the interval, we should compare the value of the function at c with the values at a and b. To sum up, it is possible, when maximising on an interval, that the maximum value is actually at an end-point of the interval, and we should check whether this is so. (The same argument applies to minimising.)

3.6 Curve sketching

Another useful application of differentiation is in curve sketching. The aim in sketching the curve described by an equation $y = f(x)$ is to indicate the behaviour of the curve and the coordinates of key points. Curve sketching is a very different business from simply plotting a few points and joining them up: there's no room in this subject for such unsophisticated methods, and such 'plotting' is an inadequate substitute for proper curve sketching!

Given the equation $y = f(x)$ of a curve we wish to sketch, we have to determine key information about the curve. The main questions we should ask are as follows. Where does the curve cross the x-axis (if at all)? Where does it cross the y-axis? Where are the critical points (or stationary points, if you prefer that name)? What are the natures of the critical points? What is the behaviour of the curve for large positive values of x and 'large' negative values of x (where 'large' means large in absolute value)?

To outline a general technique, we take these in turn.

- **Where it crosses the x-axis:** The x-axis has equation $y = 0$ and the curve has equation $y = f(x)$, so the curve crosses the x-axis at the points $(x, 0)$ for which $f(x) = 0$. Thus we solve the equation $f(x) = 0$. This may have many solutions or none at all. (For instance, if $f(x) = \sin x$ there are infinitely many solutions, whereas if $f(x) = x^2 + 1$ there are none.)
- **Where it crosses the y-axis:** The y-axis has equation $x = 0$ and the curve has equation $y = f(x)$, so the curve crosses the y-axis at the single point $(0, f(0))$.
- **Finding the critical points:** We've seen how to do this already. We solve the equation $f'(x) = 0$.
- **The natures of the critical points:** This means determining whether each one is a local maximum, local minimum, or inflexion point, and the methods for doing this have been discussed earlier in this chapter.
- **Limiting behaviour:** We have to determine what happens to $f(x)$ as x tends to infinity and as x tends to minus infinity; in other words, we have to ask how $f(x)$ behaves for x far to the right on the x-axis and for x far to the left on the negative side of the axis.

As far as the last point is concerned, there are two standard results here which are useful.

First, the behaviour of a polynomial function is determined solely by its leading term, the one with the highest power of x. This term dominates for x of large absolute value. A useful observation is that if n is even, then

$$x^n \to \infty \text{ as } x \to \infty \text{ and also as } x \to -\infty,$$

while if n is odd,

$$f(x) \to \infty \text{ as } x \to \infty \text{ and } f(x) \to -\infty \text{ as } x \to -\infty.$$

(To say, for example, that $f(x) \to \infty$ as $x \to \infty$ means that the values of $f(x)$ are, for x large enough, greater than any value we want. For example, it means that there is some number X such that for all $x > X$, $f(x) > 1000000$; and that, for some value Y, we have $f(x) > 100000000$ for all $x > Y$, and so on. In words, we say that '$f(x)$ tends to infinity as x tends to infinity'.) Thus, for example, if $f(x) = -x^3 + 5x^2 - 7x + 2$, then we examine the leading term, $-x^3$. As $x \to \infty$, this tends to $-\infty$ and as $x \to -\infty$ it tends to ∞. So this is the behaviour of f.

Secondly, whenever we have a function which is the product of an exponential and a power, the exponential dominates. Thus, for example, $x^2 e^{-x} \to 0$ as $x \to \infty$ (even though $x^2 \to \infty$).

Example 3.12 Let's do a really easy example. Consider the quadratic function $f(x) = 2x^2 - 7x + 5$. We already know a lot about sketching such curves (from the previous chapter), but let's apply the scheme suggested above. This curve crosses the x-axis when $2x^2 - 7x + 5 = 0$. The solutions to this equation (which can be found by using the formula or by factorising) are $x = 1$ and $x = 5/2$. The curve crosses the y-axis at $(0, 5)$. The derivative is $f'(x) = 4x - 7$, so there is a critical point at $x = 7/4$. The second derivative is $f''(x) = 4$, which is positive, so this critical point is a minimum. The value of f at the critical point is $f(7/4) = -9/8$. As $x \to \infty$, $f(x) \to \infty$ and as $x \to -\infty$, $f(x) \to \infty$. From this it follows that the graph of f is as in 3.1

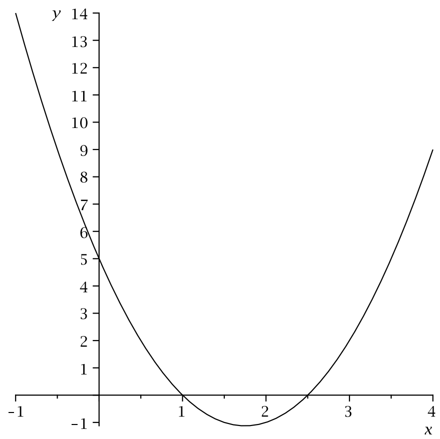
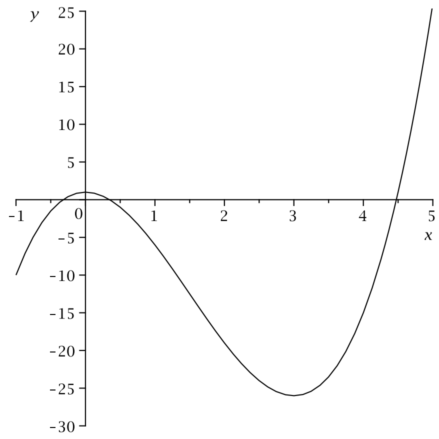


Figure 3.1: Graph of the quadratic function $f(x) = 2x^2 - 7x + 5$.

Example 3.13 We considered the function $f(x) = 2x^3 - 9x^2 + 1$ earlier. We saw that it has a local maximum at $x = 0$ and a local minimum at $x = 3$. The corresponding values of $f(x)$ are $f(0) = 1$ and $f(3) = -26$. The curve crosses the y-axis when $y = f(0) = 1$. It crosses the x-axis when $2x^3 - 9x^2 + 1 = 0$. Now, this is not an easy equation to solve! However, we can get some idea of the points where it crosses the x-axis by considering the shape of the curve. Note that as $x \to \infty$, $f(x) \to \infty$ and as $x \to -\infty$, $f(x) \to -\infty$. Also, we have $f(0) > 0$ and $f(3) < 0$. These observations imply that the graph must cross the x-axis somewhere to the left of 0 (since it must move from negative y-values to a positive y-value), it must cross again somewhere between 0 and 3 (since $f(0) > 0$ and $f(3) < 0$) and it must cross again at some point greater than 3 (because $f(x) \to \infty$ as $x \to \infty$ and hence must be positive from some point).

We therefore have the following sketch (in which I have shown the correct x-axis crossings):



Activity 3.11 Sketch the graph of $f(x) = x^3 - 6x^2 + 11x - 6$. (Note: this is the function considered in Activity 3.10.)

3.7 Marginals

We now turn our attention to economic applications of the derivative. In this section we consider 'marginals' and in the next section we approach the problem of profit maximisation using the derivative. Suppose that a firm manufactures chocolate bars and knows that in order to produce q chocolate bars it will have to pay out $C(q)$ dollars in wages, materials, overheads and so on. We say that C is the firm's *Cost function*. This is often called the *Total Cost*, and we shall often use the corresponding notation TC. The cost $TC(0)$ of producing no units (which is generally positive since a firm has certain costs in merely existing) is called the *Fixed Cost*, sometimes denoted FC. The difference between the cost and the fixed cost is known as the *Variable Cost*, VC. Other important measures are the *Average Cost*, defined by $AC = TC/q$ and the *Average Variable Cost* $AVC = VC/q$. We have the relationship

$$TC = FC + VC.$$

An increase in production by one chocolate bar is relatively small, and may be described as 'marginal'. The corresponding increase in total cost is $TC(q+1) - TC(q)$.

Now, we know that the derivative of a function f at a point a is the limit of

$$\frac{f(a+h) - f(a)}{h},$$

as h tends to 0. This means that, if h is small, then this quantity is approximately equal to $f'(a)$. Hence, for small h,

$$f(a+h) - f(a) \simeq hf'(a),$$

where '$\simeq$' means 'is approximately equal to'.

If the production level q is large, so that 1 unit is small compared with q, then we may take f to be the total cost function $TC(q)$ and take $h = 1$ to see that the cost incurred in producing one extra item, namely $TC(q+1) - TC(q)$, is given approximately by

$$TC(q+1) - TC(q) \simeq 1 \times (TC)'(q) = (TC)'(q).$$

It is for this reason that we *define* the *Marginal Cost function* to be the derivative of the total cost function TC. This marginal cost is often denoted MC. The marginal cost should not be confused with the average cost, which is given by $AC(q) = TC(q)/q$. In general, the marginal cost and the average cost are different.

In the traditional language of economics, the derivative of a function F is often referred to as the marginal of F. For example, if $TR(q)$ is the total revenue function, which describes the total revenue the firm makes when selling q items, then its derivative is called the *Marginal Revenue*, denoted MR.

Activity 3.12 A firm has total cost function

$$TC(q) = 50000 + 25q + 0.001q^2.$$

Find the fixed cost FC, and the marginal cost MC. What is the marginal cost when the output is 100? What is the marginal cost when the output is 10000?

3.8 Profit maximisation

The revenue or total revenue, TR, a firm makes is simply the amount of money generated by selling the good it manufactures. In general, this is the price of the good times the number of units sold. To calculate it for a specific firm, we need more information (as we shall see below). We have denoted the total cost function by TC. In this notation, the *profit*, Π, is the total revenue minus the total cost,

$$\Pi = TR - TC.$$

If we consider revenue and cost as functions of q, then $\Pi = \Pi(q)$ is given as a function of q by

$$\Pi(q) = TR(q) - TC(q).$$

To find which value, or values, of q give a maximum profit, we look for critical points of Π by solving $\Pi'(q) = 0$. But, since

$$\Pi'(q) = (TR)'(q) - (TC)'(q),$$

this means that the optimal value of q satisfies $(TR)'(q) = (TC)'(q)$. In other words, to maximise profit, marginal revenue equals marginal cost.[5]

The firm is said to be a *monopoly* if it is the only supplier of the good it manufactures. This means that if the firm manufactures q units of its good, then the selling price at

[5] See Anthony and Biggs (1996) Sections 8.1, 9.2 and 9.3 for a general discussion.

equilibrium is given by the inverse demand function, $p = p^D(q)$. Why is this? Well, the inverse demand function $p^D(q)$ tells us what price the consumers will be willing to pay to buy a total of q units of the good. But if the firm is a monopoly and it produces q units, then it is only these q units that are on the market. In other words, the 'q' in the inverse demand function (the *total* amount of the good on the market) is the same as the 'q' that the firm produces. So the selling price is defined by the inverse demand function, as a function of the production level q of the firm. The revenue is then given, as a function of q, by $TR(q) = qp^D(q)$.

Example 3.14 A monopoly has cost function $TC(q) = 1000 + 2q + 0.06q^2$ and its demand curve has equation $q + 10p = 500$. What value of q maximises the profit?

To answer this, we first have to determine the revenue as a function of q. Since the firm is a monopoly, we know that $TR(q) = qp^D(q)$. From the equation for the demand curve, $q + 10p = 500$, we obtain

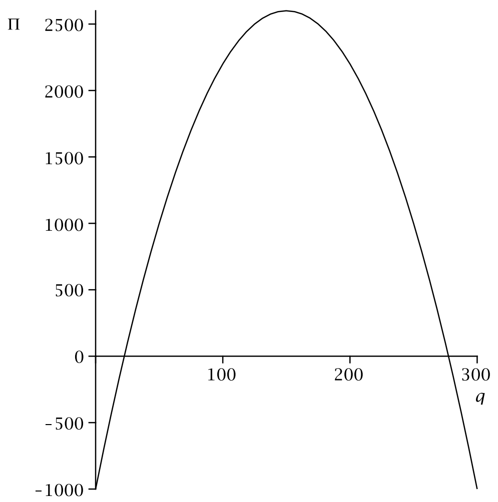
$$p = 50 - 0.1q \quad \text{so} \quad p^D(q) = 50 - 0.1q \quad \text{and} \quad R(q) = q(50 - 0.1q).$$

The profit is therefore given by

$$\begin{aligned} \Pi(q) &= TR(q) - TC(q) \\ &= q(50 - 0.1q) - (1000 + 2q + 0.06q^2) \\ &= 48q - 0.16q^2 - 1000. \end{aligned}$$

The equation $\Pi'(q) = 0$ is $48 - 0.32q = 0$, which has solution $q = 150$. To verify that this does indeed give a maximum profit, we note that $\Pi''(q) = -0.32 < 0$.

Consider this last example a little further. The profit function $\Pi(q)$ is a quadratic function with a negative q^2 term, so we well know what it looks like: its graph will be as follows.



Notice that the profit is negative to start with (because there is no revenue when nothing is produced, but there is a cost of producing nothing, namely the fixed cost of

1000). As production is increased, profit starts to rise, and becomes positive. The point at which profit just starts to become positive (that is, where it first equals 0), is called the breakeven point. So, the breakeven point is the smallest positive value of q such that $\Pi(q) = 0$. When the firm is producing the breakeven quantity, it is breaking even in the sense that its revenue matches its costs. In this specific example, we can calculate the breakeven point by solving the equation

$$-0.16q^2 + 48q - 1000 = 0.$$

This has the solutions 22.524 and 277.475. It is clearly the first of these that we want, as the second, higher, value of q is where the profit, having increased to a maximum and then decreased, becomes 0 again. So the breakeven point is 22.524.

> **Activity 3.13** A firm is a monopoly with cost function
>
> $$TC(q) = q + 0.02q^2.$$
>
> The demand equation for its product is $q + 20p = 300$. Work out (a) the inverse demand function; (b) the profit function; (c) the optimal value q_m and the maximum profit; (d) the corresponding price.

Learning outcomes

At the end of this chapter and the relevant reading, you should be able to:

- explain what is meant by the derivative
- state the standard derivatives
- calculate derivatives using sum, product, quotient, and composite function (chain) rules
- calculate derivatives by taking logarithms
- establish the nature of the critical/stationary points of a function
- use the derivative to help sketch functions
- explain and use the terminology surrounding 'marginals' in economics, and be able to find fixed costs and marginal costs, given a total cost function
- explain what is meant by the breakeven point and be able to determine this
- make use of the derivative in order to minimise or maximise functions, including profit functions

You do *not* need to be able to differentiate from first principles (that is, by using the formal definition of the derivative).

Sample examination/practice questions

Question 3.1

Differentiate $y = (1 + 2x - e^x)$ and find when the derivative is zero.

Question 3.2

Differentiate the following functions.

(a) $y = x^3 + \exp(3x^2)$.

(b) $y = \dfrac{3x+5}{x^2+3x+2}$.

(c) $y = \ln(3x^2) + 3x + \dfrac{1}{\sqrt{1+x}}$.

Question 3.3

Assume that the price/demand relationship for a particular good is given by

$$p = 10 - 0.005q$$

where p is the price (\$) per unit and q is the demand per unit of time. Also assume that the fixed costs are \$100 and the average variable cost per unit is $4 + 0.01q$.

(a) What is the maximum profit obtainable from this product?

(b) What level of production is required to break even?

(c) What are the marginal cost and marginal revenue functions?

Question 3.4

The demand function relating price p and quantity x, for a particular product, is given by

$$p = 5\exp(-x/2).$$

Find the amount of production, x, which will maximise revenue from selling the good, and state the value of the resulting revenue. Produce a rough sketch graph of the marginal revenue function for $0 \leq x \leq 6$.

Question 3.5

Suppose you have a nineteenth-century painting currently worth \$2000, and that its value will increase steadily at \$500 per year, so that the amount realised by selling the painting after t years will be $2000 + 500t$. An economic model shows that the optimum time to sell is the value of t for which the function

$$P(t) = (2000 + 500t)e^{-0.1t}$$

is maximised. Given this, find the optimum time to sell, and verify that it is optimal.

Question 3.6

A monopolist's average cost function is given by

$$AC = 10 + \frac{20}{Q} + Q.$$

Her demand equation is

$$P + 2Q = 20,$$

where P and Q are price and quantity, respectively. Find expressions for the total revenue and for the profit, as functions of Q. Determine the value of Q which maximises the total revenue. Determine also the value of Q maximising profit.

Question 3.7

A firm's average cost function is given by

$$\frac{300}{Q} - 10 + Q,$$

and the demand function is given by $Q + 5P = 850$, where P and Q are quantity and price, respectively.

Supposing that the firm is a monopoly, find expressions for the total revenue and for the profit, as functions of Q.

Determine the value of Q which maximises the total revenue and the value of Q which maximises profit.

Answers to activities

Feedback to activity 3.1

Taking $h = 0.00001$, for example,

$$\frac{2^{1+h} - 2}{h},$$

is 1.386299 and for $h = 0.000001$ it is 1.386295. These are even closer to the true value $2 \ln 2 = 1.38629436 \ldots$.

Feedback to activity 3.2

The function $1/x$ can be written as x^{-1}. Its derivative is therefore $(-1)x^{-1-1} = -x^{-2} = -1/x^2$.

Feedback to activity 3.3

By the product rule,

$$(x^2 \sin x)' = 2x \sin x + x^2 \cos x.$$

Feedback to activity 3.4

By the product rule,

$$((x^2 + 1) \ln x)' = 2x \ln x + (x^2 + 1)\frac{1}{x} = 2x \ln x + x + \frac{1}{x}.$$

Feedback to activity 3.5

By the quotient rule,

$$\left(\frac{\sin x}{x}\right)' = \frac{\cos x(x) - (\sin x)(1)}{x^2} = \frac{x \cos x - \sin x}{x^2}.$$

Feedback to activity 3.6

By the chain rule (composite function rule),

$$\left((3x+7)^{15}\right)' = 15(3x+7)^{14}(3) = 45(3x+7)^{14}.$$

Feedback to activity 3.7

$\sqrt{x^2+1} = (x^2+1)^{1/2}$. By the chain rule,

$$\left((x^2+1)^{1/2}\right)' = \frac{1}{2}(x^2+1)^{-1/2}(2x) = \frac{x}{\sqrt{x^2+1}}.$$

Feedback to activity 3.8

By the chain rule,

$$\left(\ln(x^2+2x+5)\right)' = \frac{1}{x^2+2x+5}(x^2+2x+5)' = \frac{2x+2}{x^2+2x+5}.$$

Feedback to activity 3.9

If $f(x) = x^x$ then $\ln f(x) = x \ln x$ and so

$$\frac{f'(x)}{f(x)} = x\left(\frac{1}{x}\right) + (1)\ln x = 1 + \ln x,$$

from which we obtain

$$f'(x) = (1+\ln x)f(x) = (1+\ln x)x^x.$$

Feedback to activity 3.10

The derivative of $f(x) = x^3 - 6x^2 + 11x - 6$ is $f'(x) = 3x^2 - 12x + 11$. The stationary points, the solutions to $3x^2 - 12x + 11 = 0$, are $(12-\sqrt{12})/6$ and $(12+\sqrt{12})/6$. The second derivative is $6x - 12$, which is negative at the first stationary point and positive at the second. We therefore have a local maximum at $(12-\sqrt{12})/6$ and a local minimum at $(12+\sqrt{12})/6$. The corresponding values of f are $2\sqrt{3}/9$ and $-2\sqrt{3}/9$. (Approximately, $2\sqrt{3}/9$ is 0.3849.)

Feedback to activity 3.11

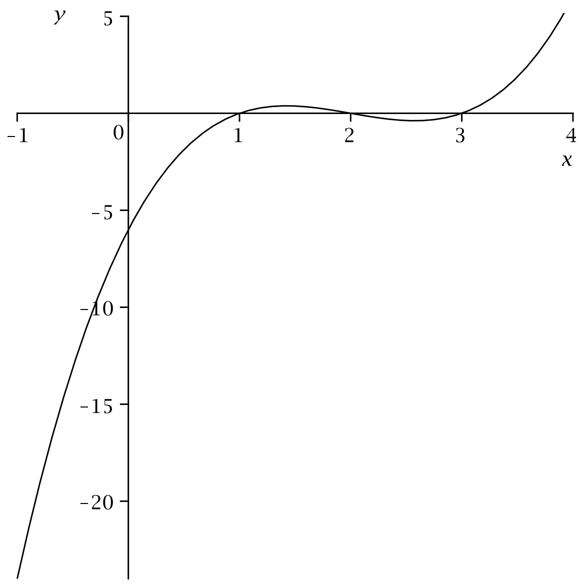
To sketch $f(x) = x^3 - 6x^2 + 11x - 6$, we first note from the previous activity that f has a local maximum at $(12-\sqrt{12})/6$ and a local minimum at $(12+\sqrt{12})/6$. The corresponding values of f are $2\sqrt{3}/9$ and $-2\sqrt{3}/9$. (Approximately, $2\sqrt{3}/9$ is 0.3849.) As x tends to infinity, so does $f(x)$ and as x tends to $-\infty$, $f(x) \to -\infty$. The curve crosses the y-axis at $(0,-6)$. To find where it crosses the x-axis, we have so solve $x^3 - 6x^2 + 11x - 6 = 0$. One solution to this is easily found (by guesswork) to be $x = 1$, and so $(x-1)$ is a factor. Therefore, for some numbers a, b, c,

$$x^3 - 6x^2 + 11x - 6 = (x-1)(ax^2 + bx + c).$$

Straight away, we see that $a = 1$ and $c = 6$. To find b, we could notice that the number of terms in x^2 is $-a + b$, which should be -6, so that $-1 + b = -6$ and $b = -5$. Hence

$$x^3 - 6x^2 + 11x - 6 = (x-1)(x^2 - 5x + 6) = (x-1)(x-2)(x-3)$$

and we see there are three solutions: $x = 1, 2, 3$. Piecing all this information together, we can sketch the curve as follows:



Feedback to activity 3.12

The total cost function is $50000 + 25q + 0.001q^2$. The fixed cost is obtained by setting $q = 0$, giving $FC = 50000$. The marginal cost is $MC = 25 + 0.002q$. The marginal cost is \$25.2 if the output is 100, but it rises to \$45 if the output is 10000.

Feedback to activity 3.13

(a) The inverse demand function is

$$p^D(q) = \frac{300 - q}{20} = 15 - 0.05q.$$

(b) The profit function is

$$\Pi(q) = qp^D(q) - TC(q) = q(15 - 0.05q) - (q + 0.02q^2) = 14q - 0.07q^2.$$

(c) We have $\Pi'(q) = 14 - 0.14q$, so $q = 100$ is a critical point. The second derivative of Π is $\Pi''(q) = -0.14$, which is negative, so the critical point is a local maximum. The value of the profit there is $\Pi(100) = 1400 - 700 = 700$, whereas $\Pi(0) = 0$ and $\Pi(200) = 0$. Since the maximum profit in the interval $[0, 200]$ must be either at a local maximum or an end-point, it follows that the maximum profit is 700, obtained when $q = 100$.

(d) The price when $q = 100$ is $p^D(100) = 15 - (0.05)(100) = 10$.

Answers to Sample examination/practice questions

Answer to question 3.1

The derivative is $\frac{dy}{dx} = 2 - e^x$, and this is equal to 0 when $e^x = 2$; that is, when $x = \ln 2$.

Answer to question 3.2

For the first, we have, using the composite function rule on the exponential term,

$$\frac{dy}{dx} = 3x^2 + 6x\exp(3x^2).$$

For the second, using the quotient rule,

$$\frac{dy}{dx} = \frac{3(x^2+3x+2)-(2x+3)(3x+5)}{(x^2+3x+2)^2} = \frac{-3x^2-10x-9}{(x^2+3x+2)^2}.$$

Lastly,

$$\begin{aligned}\frac{d}{dx}\left(\ln(3x^2)+3x+\frac{1}{\sqrt{1+x}}\right) &= \frac{d}{dx}\left(\ln(3x^2)+3x+(1+x)^{-1/2}\right)\\ &= \frac{6x}{3x^2}+3-\frac{1}{2}(1+x)^{-3/2}\\ &= \frac{2}{x}+3-\frac{1}{2(1+x)^{3/2}}.\end{aligned}$$

Answer to question 3.3

(a) The average variable cost is AVC$=4+0.01q$, so the variable cost is VC$=4q+0.01q^2$. So the total cost function is $TC = 4q+0.01q^2+FC$, where FC is the fixed cost; that is, $TC = 4q+0.01q^2+100$. Given that the firm is a monopoly, the revenue is $TR = pq = (10-0.005q)q = 10q-0.005q^2$. The profit function is

$$\Pi = TR - TC = 10q - 0.005q^2 - (4q+0.01q^2+100) = 6q - 0.015q^2 - 100.$$

To find the maximum, we solve $\Pi' = 0$, which is $6-0.03q = 0$, so $q = 200$. To check it gives a maximum, we check that the second derivative is negative, which is true since $\Pi'' = -0.03$.

(b) The breakeven point is the (least) value of q for which $\Pi(q) = 0$. So we solve $6q - 0.015q^2 - 100 = 0$. We can rewrite this as

$$0.015q^2 - 6q + 100 = 0,$$

which has solutions

$$\frac{6\pm\sqrt{6^2-4(0.015)(100)}}{2(0.015)} = \frac{6\pm\sqrt{36-6}}{0.03} = \frac{100}{3}(6\pm\sqrt{30}).$$

Since $\sqrt{30} < 6$ (because $30 < 6^2 = 36$), there are two positive solutions. The breakeven point is the smaller of these, which is $100(6-\sqrt{30})/3$.

(c) The marginal cost and marginal revenue functions are the derivatives of the total cost and total revenue. Thus,

$$MC = \frac{d}{dq}\left(4q+0.01q^2+100\right) = 4+0.02q$$

and

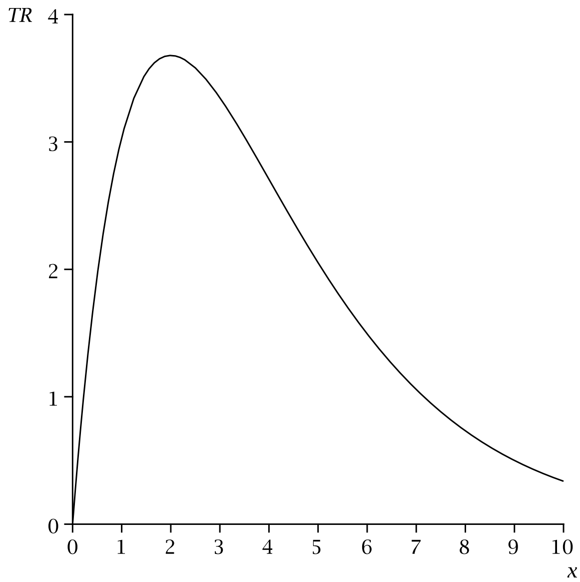
$$MR = \frac{d}{dq}\left(10q - 0.005q^2\right) = 10 - 0.01q.$$

Answer to question 3.4

The total revenue obtained from the sale of the good is simply (price times quantity) $TR = 5xe^{-x/2}$. To find the maximum, we differentiate:

$$(TR)' = 5e^{-x/2} + 5x\left(-\frac{1}{2}\right)e^{-x/2} = \frac{5}{2}e^{-x/2}\left(2 - x\right),$$

and this is zero only if $x = 2$. We can see that the derivative changes sign from positive to negative on passing through $x = 2$, and so this is a local maximum. The value of the revenue there is $5(2)e^{-2/2} = 10/e$. When $x = 0$, $TR = 0$ and there are no solutions to $TR = 0$, so the graph of TR will not cross the x-axis. Sketching the curve between 0 and 6, we obtain the following.



Answer to question 3.5

By routine application of the rules for differentiation we get

$$P'(t) = 500e^{-0.1t} + (2000 + 500t)(-0.1)e^{-0.1t} = e^{-0.1t}(300 - 50t).$$

Since this is zero when $t = 6$, that is a critical point of P. Differentiating again we get

$$P''(t) = (-0.1)e^{-0.1t}(300 - 50t) + e^{-0.1t}(-50) = e^{-0.1t}(5t - 80).$$

It follows that $P''(6) < 0$, so the critical point $t = 6$ is a local maximum.

The fact that $t = 6$ is indeed the maximum in $[0, \infty)$ can be verified by common-sense arguments. We know that $t = 6$ is the only critical point of $P(t)$, and that it is a local maximum. It follows that $P(t)$ must decrease steadily for $t > 6$, because if at any stage it started to increase again, it would have to pass through a critical point first. (Alternatively, to see that this local maximum is a global maximum, it could be noted that $P(t) \to 0$ as $t \to \infty$. This is because the exponential part $e^{-0.01t}$ tends to 0 and exponentials 'dominate' polynomials, so that even if we multiply this by $(2000 + 500t)$, the result still tends to 0.)

Answer to question 3.6

First, since the total cost is Q times the average cost, we have

$$TC = Q\left(10 + \frac{20}{Q} + Q\right) = 10Q + 20 + Q^2.$$

The monopolist's demand equation is $P + 2Q = 20$, so when the quantity produced is Q, the selling price will be $P = 20 - 2Q$ and hence the total revenue will be

$$TR = QP = Q(20 - 2Q) = 20Q - 2Q^2.$$

So the profit function is

$$\Pi(Q) = 20Q - 2Q^2 - (10Q + 20 + Q^2) = 10Q - 3Q^2 - 20.$$

To maximise TR, we set $(TR)' = 0$, which is $20 - 4Q = 0$, giving $Q = 5$. This does indeed maximise revenue because $(TR)'' = -4 < 0$. For the profit function, we have $\Pi'(Q) = 10 - 6Q$, and this is 0 when $Q = 5/3$. Again, this gives a maximum because $\Pi''(Q) = -6 < 0$.

Answer to question 3.7

The total cost is

$$TC = Q(AC) = Q\left(\frac{300}{Q} - 10 + Q\right) = 300 - 10Q + Q^2.$$

The inverse demand function is given by $P = 170 - Q/5$, so the total revenue is

$$TR = \left(170 - \frac{Q}{5}\right)Q = 170Q - 0.2Q^2.$$

The derivative $(TR)'$ is $170 - 0.4Q$, which is 0 when $Q = 425$, this giving a maximum because $(TR)'' = -0.4 < 0$. The profit function is

$$\begin{aligned}\Pi(Q) &= TR - TC = 170Q - 0.2Q^2 - (300 - 10Q + Q^2)\\ &= 180Q - 1.2Q^2 - 300.\end{aligned}$$

To maximise the profit, we set $\Pi' = 0$, which is $180 - 2.4Q = 0$, so $Q = 75$. This is a maximum because $\Pi''(Q) = -2.4 < 0$.