Chapter 8: Applications of Derivatives

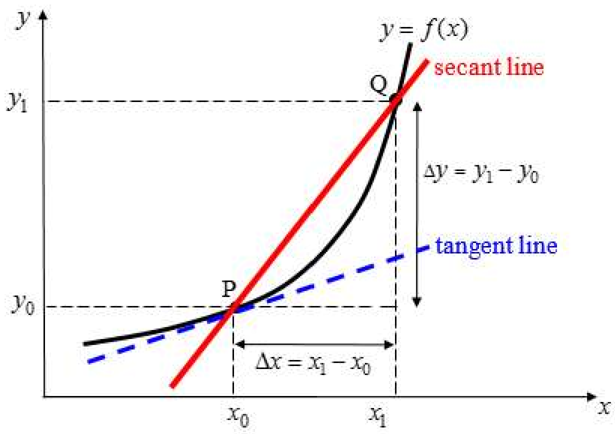
Learning Objectives:

1. *Interpret the derivative as the slope of the tangent to the graph of y=f(x) at a point.*
2. *Interpret the derivative as the instantaneous rate of change of y=f(x) at a point.*
3. *Apply the chain rule.*
4. *Identify increasing and decreasing functions.*
5. *Recognize the type of concavity of a function and identify points of inflection.*
6. *Use analytical techniques to find local extrema and if appropriate global extrema.*
7. *Use gradient descent to find local and global extrema.*

8.1 Instantaneous Rate of Change

Recall that the difference quotient $\dfrac{f(x_0+\Delta x)-f(x_0)}{\Delta x}$ gives us the slope of a secant line through two points on the graph of f. In the limit as $\Delta x \to 0$, we obtain the tangent of f at $x = x_0$. The slope of the tangent line at $x = x_0$ is then the value of the **derivative** at $x = x_0$. That is,

$$\text{Slope of tangent line is: } f'(x_0) \quad \text{or} \quad \left.\frac{dy}{dx}\right|_{x=x_0}$$



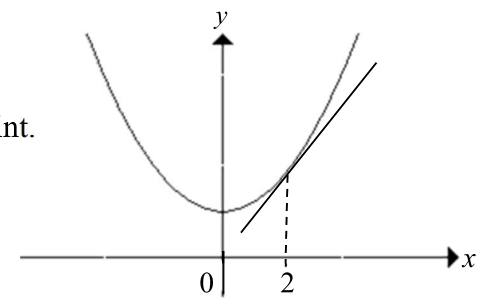
Alternatively, the derivative $f'(x_0)$ can also be thought of as the **instantaneous rate of change** of the function f at $x = x_0$.

Example 1.

Let $f(x) = 3x^2 + 1$.

(a) Calculate the slope of the tangent line at x = 2.

(b) Hence, find the equation of the tangent line at the given point.



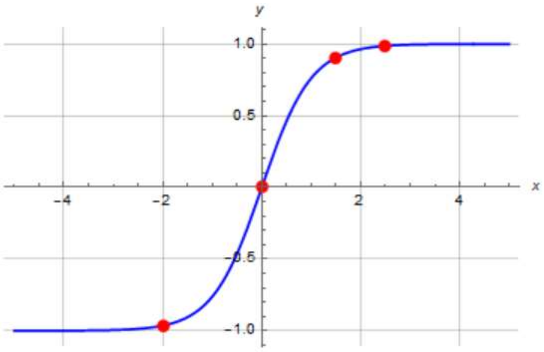
Example 2.

Activation functions in an artificial neural network help to introduce nonlinearity to the network. An example of an activation function is the hyperbolic tangent function given below.

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

(a) Find $f'(x)$.

(b) Hence find the slope of f at each point below.



8.2 Applications of the Chain Rule

Sometimes the rate of change of a function is related to the rate of change of another function. We will then use the Chain Rule, which was introduced in a previous chapter.

<u>Chain Rule</u>

If $y = f(g(x))$ and $u = g(x)$, then $\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx}$.

An alternative version is to use functional notations: $\frac{d}{dx}[f(g(x))] = f'(g(x)) \cdot g'(x)$

Example 3.

In an experiment, a spherical balloon is leaking air at a rate of 500 cm^3/s. How fast is the radius of this balloon shrinking when the radius is 10 cm? (Volume of a sphere: $v = \frac{4}{3}\pi r^3$)

Example 4.

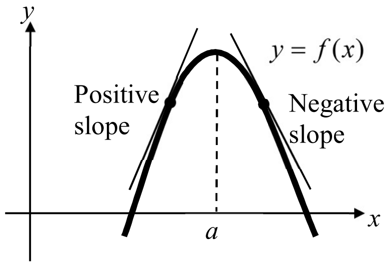
A square sheet of metal of side 8 cm is heated in a furnace. If each side expands at a rate of 0.2 mm/s, how fast is the area expanding after 5 seconds in the furnace?

8.3 Increasing & Decreasing Functions

A function f is said to be **increasing** when its graph rises as it goes from left to right. It is said to be **decreasing** when its graph falls as it goes from left to right.

The increasing/decreasing concept can be associated with the slope of the tangent line.

- In the figure below, we see that on the interval $(-\infty, a)$, the graph of f is rising, $f(x)$ is increasing, and the slope of the graph is positive, i.e. $f'(x) > 0$.
- On the interval (a, ∞), the graph of f is falling, $f(x)$ is decreasing, and the slope of the graph is negative, i.e. $f'(x) < 0$.



Example 5.

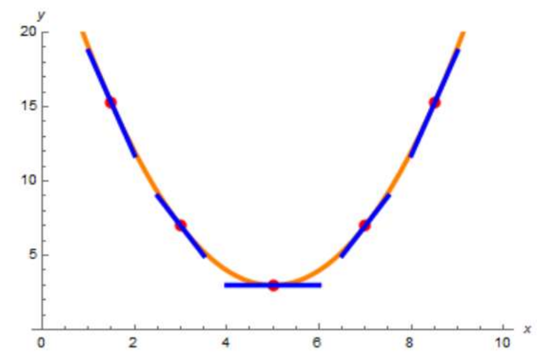
Find the interval on which the function $f(x) = 4x - x^2$ is increasing and the interval on which it is decreasing.

8.4 Concavity & Points of Inflection

Recall that the sign of $f'(x)$ tells us whether f is increasing or decreasing on an interval, but it does not tell us about the **concavity** of the curve of f on that interval. For this we have to look at the second derivative $f''(x)$ of the function.

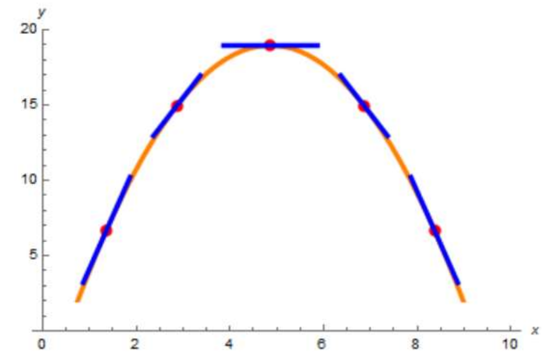
A function f is said to be **concave up** on an interval if its tangent lines have increasing slopes on that interval, that is, $f''(x) > 0$ in that interval.
Similarly, a function f is said to be **concave down** on an interval if its tangent lines have decreasing slopes on that interval, that is, $f''(x) < 0$ in that interval.

(a)



Graph is above tangent lines
$\Rightarrow$ f is **concave up** on interval

(b)



Graph is below tangent lines
$\Rightarrow$ f is **concave down** on interval

Notice that if f is concave up on an interval, its graph lies above the tangent lines on that interval. Similarly, if f is concave down on an interval, its graph lies below the tangent lines on that interval.
A point where the concavity changes (from up to down or down to up) is called a **point of inflection**.

Example 6.

Let $f(x) = 3x^4 - 4x^3$. Compute $f'(x)$ and $f''(x)$.

(a) Fill in the sign analysis in the tables below. (The first rows have been completed for you.)

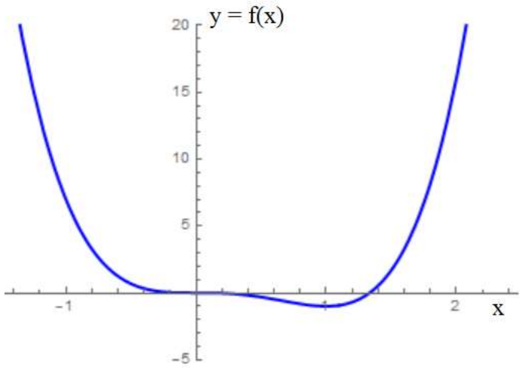
(b) In the tables, determine the intervals on which f is increasing, decreasing, concave up, and concave down.

Interval	$12x^2(x-1)$	$f'(x)$	Conclusion
$x < 0$	$(+)(-)$	$-$	f is decreasing on $(-\infty, 1]$
$0 < x < 1$			
$x > 1$			

Interval	$12x(3x-2)$	$f''(x)$	Conclusion
$x < 0$	$(-)(-)$	$+$	f is concave up on $(-\infty, 0)$
$0 < x < \frac{2}{3}$			
$x > \frac{2}{3}$			

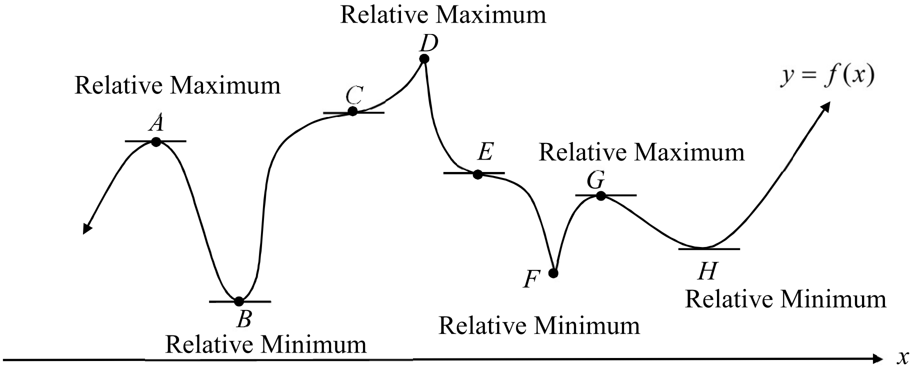
(c) Identify the points of inflection.

(d) Are your conclusions consistent with the graph of f below?



8.5 Local Extrema

When the graph of a continuous function changes from rising to falling, a relative (or local) **maximum** occurs. When the graph changes from falling to rising, a relative (or local) **minimum** occurs.



The graph will be "rounded" at the local maxima or minima if $f'(x)=0$. They are points *A*, *B*, *G* and *H* in the figure above.

The graph will be "pointed" at the local maxima or minima when $f'(x)$ is not defined. They are points *D* and *F*. The slope of the tangent line is not defined at such points as the tangent line is either vertical or does not exist. We say that f is **not differentiable** at *D* and *F*.

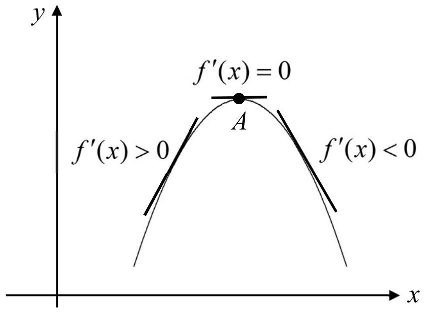
Critical points are those points $(x, f(x))$ on $y=f(x)$ where $f'(x)=0$ or at which $f(x)$ is not differentiable. The critical points where $f'(x)=0$ are called **stationary points** of $f(x)$.

Note: It does not imply that every critical point produces a local maximum or minimum. In the figure above, points *C* and *E* are critical points (the slope is zero), but the function does not have a local maximum or minimum at either of these points. In fact, points *C* and *E* are **Points of Inflection.**

8.5.1 First-Derivative Test

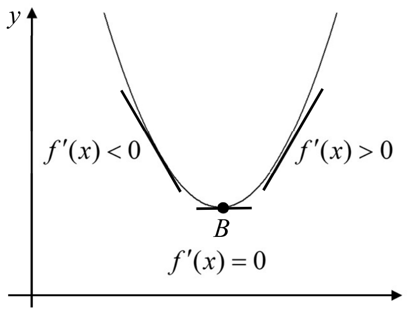
From the discussion of increasing and decreasing values for f, we see that the derivative changes sign from positive to negative when f passes through a local maximum point and from negative to positive when f passes through a local minimum point.

A relative (or local) maximum



In passing through the point A, $f'(x)$ decreases from positive to negative.

A relative (or local) minimum



In passing through the point B, $f'(x)$ increases from negative to positive.

Type of points	Sign of $f'(x)$ immediately before and after		
	Before stationary point	Stationary point	After stationary point
Maximum	**(+)**	**(0)**	**(–)**
Minimum	**(–)**	**(0)**	**(+)**

We may perform the following steps to determine the maximum or minimum points of a function $y = f(x)$:

(a) Differentiate the function $f(x)$.

(b) Let $f'(x)=0$ and solve for the values(s) of x (stationary points).

(c) Look at the slope of the tangent on either side of the stationary point(s). We must choose the test points close to the stationary point(s). If the test points are too far away, the curve might have already changed direction.

(d) If the slope of tangent is positive to the left of a stationary point and negative to the right, then the stationary point is a maximum. The reverse is true for a minimum point.

Example 7.

Find all the stationary points of the following functions and determine if each of the stationary points is a local maximum or minimum, or a point of inflection.

(a) $f(x) = 3x^4 - 1$ (b) $f(x) = 3x^4 + 4x^3$

8.5.2 Second-Derivative Test

It is apparent that a curve is concave down at a maximum point and concave up at a minimum point. Therefore, at $x = c$, where c is a critical point of f, we have:

1. A **local minimum** point at $(c,\ f(c))$ if $f''(c) > 0$
2. A **local maximum** point at $(c,\ f(c))$ if $f''(c) < 0$
3. There is no conclusion from this test if $f''(c) = 0$ or if $f''(c)$ is undefined. (Instead, we must use the first-derivative test)

Example 8.

Find all the stationary points of the function $f(x) = x^3 - 3x^2 - 9x + 2$ and determine if each of the stationary points is local maximum or minimum.

The above gives us a technique for finding local maxima and minima of a function f by finding values of x for which either $f'(x) = 0$ or $f'(x)$ does not exist. The question now is how about finding the *global* maximum and *global* minimum of f?

In the case where we have a *continuous* function f over a finite closed interval $[a,b]$, the ensuing theorem guarantees the existence of a global maximum and global minimum.

8.5.3 Theorem

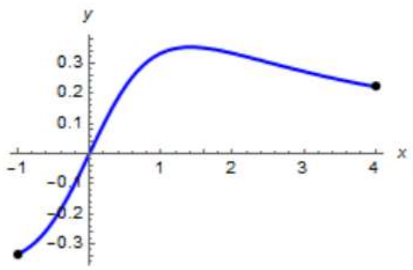
Let f be a continuous function on the interval $[a,b]$. The function f has both a global maximum and a global minimum on $[a,b]$.

To find the global maximum and global minimum of f, we do the following:

- Find all values of x in $[a,b]$, for which
    - $f'(x)=0$, or
    - $f'(x)$ does not exist, or
    - $x=a$ or $x=b$
- Evaluate $f(x)$ for each x above. The largest (or smallest) of these values is the largest (or smallest) value of f for values of x in $[a,b]$.

Example 9.

Let $f(x)=\dfrac{x}{x^2+2}$. Find the global maximum and minimum of f over the closed interval $[-1,4]$, and state where those values occur. Check if your results are consistent with the graph of f below.

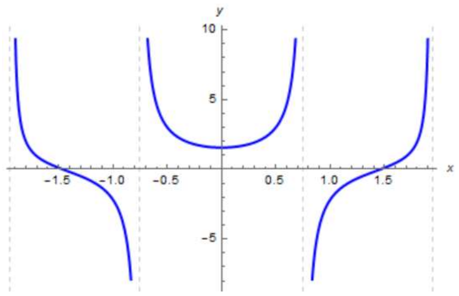


We now know that if f is *continuous* on a *finite closed* interval $[a,b]$, then the global extrema of f occur either at the endpoints of $[a,b]$ or at critical points inside $[a,b]$. In general, however, there is no guarantee that a function will actually have a global maximum or global minimum on an open interval.

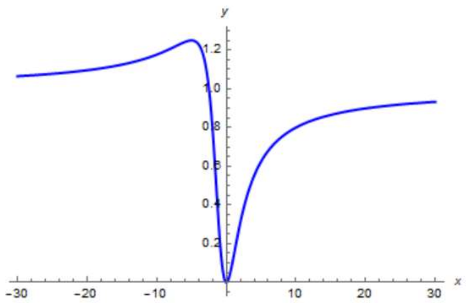
Example 10.

Discuss whether each function given below has any global extrema.

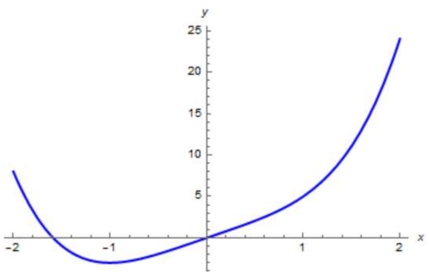
(a)



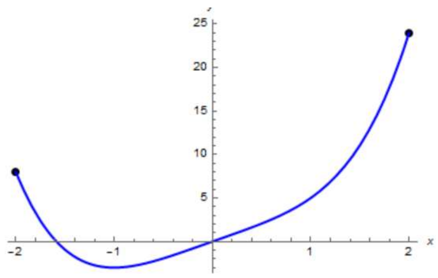
(b)



(c)



(d)



8.6 Optimization

Optimization is among one of the most important techniques in machine learning. We learned in the above analytical techniques to find the maximum and minimum values of a function. However, in many machine learning problems, where we have to deal with functions of several variables, the analytical approach may not be practical, and a numerical approach such as the gradient descent has to be used instead.

To gain an intuitive understanding of this algorithm, we consider **Univariate Gradient Descent**, where we try to locate the local minimum of a differentiable function of a single variable.

Suppose we have a univariate function given by: $f(x) = 4x^2 - 4x + 3$

Its first derivative is $f'(x) = 8x - 4$.

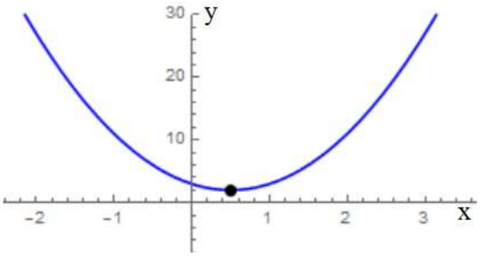
If we let $f'(x) = 0$, then $8x - 4 = 0 \quad \Rightarrow x = \frac{1}{2}$

And, $y = 4\left(\frac{1}{2}\right)^2 - 4\left(\frac{1}{2}\right) + 3 = 2$

Also, second derivative $f''(x) = 8 \quad \Rightarrow f''(\frac{1}{2}) = 8 > 0$

Thus, by the second derivative test, $\left(\frac{1}{2}, 2\right)$ is a local minimum.

In fact, the graph of $f(x) = 4x^2 - 4x + 3$ is as follows:



Univariate Gradient Descent

We now try to find the minimum value of $f(x) = 4x^2 - 4x + 3$, but this time using gradient descent. We start with an initial value of x_0, and then update the values of *x* repeatedly using the following update rule:

$$\boxed{x_{n+1} = x_n - \alpha f'(x_n)}$$

Here, α is called the **learning rate**.

The gradient descent algorithm can be summarized as follows:

<u>Algorithm</u>

Step 0: Set a learning rate $\alpha > 0$ and an initial point $x = x_0$ and compute $f(x_0)$.

Step 1: At *n*-th point $x = x_n$, compute $f'(x_n)$.

Step 2: Update to the $(n+1)th$ point, $x_{n+1} = x_n - \alpha f'(x_n)$ and compute $f(x_{n+1})$.

Step 3: Repeat steps 1 and 2 until a stopping criterion is reached.

There are a few ways to stop the iterations. Here, we choose to stop when the difference between two *x*-values is less than a user-defined threshold.

For the example above, let us choose to start with $x_0 = 2$ and a learning rate $\alpha = 0.1$.

Suppose also that we choose a threshold (also called epsilon) of 0.001, meaning that we stop the iterations when the difference between two *x*-values is less than epsilon.

Recall that $\dfrac{df}{dx} = f'(x) = 8x - 4$.

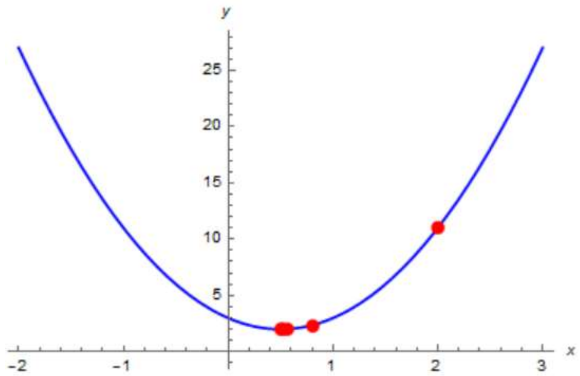
The Python code for gradient descent can be implemented as follows:

```python
x = 2 # Starting value of x
rate = 0.1 # Set learning rate
epsilon = 0.001 # Stop algorithm when absolute difference between 2 consecutive x-values is less than epsilon
diff = 1 # difference between 2 consecutive iterates
max_iter = 1000 # set maximum number of iterations
iter = 1 # iterations counter
f = lambda x: 4*x**2-4*x+3
deriv = lambda x: 8*x-4 # derivative of f

# Now Gradient Descent

while diff > epsilon and iter < max_iter:
    x_new = x - rate * deriv(x)
    print("Iteration ", iter, ": x-value is: ", x_new,"f(x) is: ", f(x_new) )
    diff = abs(x_new - x)
    iter = iter + 1
    x = x_new

print("The local minimum occurs at: ", x)
```

The output is as follows:

```
Iteration  1 : x-value is:  0.7999999999999998 f(x) is:  2.3599999999999994
Iteration  2 : x-value is:  0.5599999999999999 f(x) is:  2.0144
Iteration  3 : x-value is:  0.512 f(x) is:  2.0005759999999997
Iteration  4 : x-value is:  0.5024 f(x) is:  2.0000230400000003
Iteration  5 : x-value is:  0.50048 f(x) is:  2.0000009216
Iteration  6 : x-value is:  0.500096 f(x) is:  2.000000036864
The local minimum occurs at:  0.500096
```

Notice that from the fifth to the sixth iteration, x decreases by $0.50048 - 0.500096 = 0.000384$, which is less than epsilon. Hence, the iterations converge and the minimum value of f is 2, which occurs when x is 0.5.



The learning rate α determines how fast or slow x moves. If α is chosen to be too large, we might miss the minimum and the algorithm may not converge. On the other hand, if α is too small, we may require many iterations to reach convergence and the convergence is slow.

Example 11.

(The role of the learning rate)

Repeat the Example with $f(x) = 4x^2 - 4x + 3$ above, starting with $x_0 = 2$, but this time with a learning rate of $\alpha = 0.001$. Choose an epsilon value of 0.000001 and set the maximum number of iterations to be 2000. Comment on your results.

Python partial output:

```
Iteration  1166 : x-value is:  0.5001284402052271 f(x) is:  2.000000065987545
Iteration  1167 : x-value is:  0.5001274126835853 f(x) is:  2.0000000649359677
Iteration  1168 : x-value is:  0.5001263933821166 f(x) is:  2.0000000639011484
Iteration  1169 : x-value is:  0.5001253822350596 f(x) is:  2.0000000628828194
Iteration  1170 : x-value is:  0.5001243791771792 f(x) is:  2.0000000618807188
Iteration  1171 : x-value is:  0.5001233841437618 f(x) is:  2.000000060894588
The local minimum occurs at:  0.5001233841437618
```

In addition to the learning rate, the initial value of x from which we start the iterations of the gradient descent algorithm plays a very important role in the algorithm as well.

Example 12.

(The role of different initial values)

Let $f(x) = x^5 - 30x^3 + 50x$. Use gradient descent to find the minimum of f with each of the following:

(a) Start with $x_0 = 0$.

(b) Start with $x_0 = 2$.

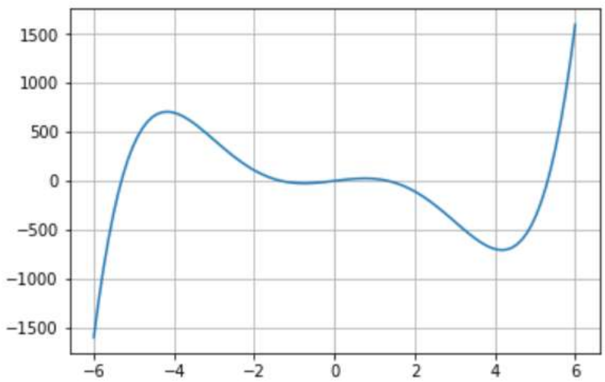
In each, use a learning rate of 0.001, epsilon 0.001, and set the maximum number of iterations to 1000. Comment on your results.

(a) Python partial output:

```
Iteration  35 : x-value is:  -0.7445901957441673 f(x) is:  -25.07402921017795
Iteration  36 : x-value is:  -0.74622976289979 f(x) is:  -25.07654823664344
Iteration  37 : x-value is:  -0.7476629245308597 f(x) is:  -25.078472707263586
Iteration  38 : x-value is:  -0.7489153423005702 f(x) is:  -25.079942218023618
Iteration  39 : x-value is:  -0.75000956449125 f(x) is:  -25.081063835228058
Iteration  40 : x-value is:  -0.7509653852286368 f(x) is:  -25.081919593264473
The minimum is:  -25.081919593264473
It occurs at:  -0.7509653852286368
```

(b) Python partial output:

```
Iteration  9 : x-value is:  4.1562117648945645 f(x) is:  -705.8431528389654
Iteration  10 : x-value is:  4.168908422393485 f(x) is:  -705.9486285945052
Iteration  11 : x-value is:  4.172803397193701 f(x) is:  -705.958491051829
Iteration  12 : x-value is:  4.173970434098405 f(x) is:  -705.9593747061026
Iteration  13 : x-value is:  4.174317556156385 f(x) is:  -705.9594528362722
The minimum is:  -705.9594528362722
It occurs at:  4.174317556156385
```



Tutorial 8

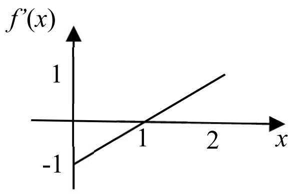
Section A: Multiple Choice Questions

1. Let $y = f(x)$ and $f''(0) = 4$. At $x = 0$, the function $f(x)$ ______________.

   (a) is equal to 0 (b) is equal to 1

   (c) is concave down (d) is concave up

2. The diagram of $f'(x)$ over $0 \le x \le 2$ is given below.



   At $x = 1$, $f(x)$ has _____________.

   (a) a maximum value (b) a minimum value

   (c) a point of inflexion (d) none of the above

3. The curve $y = f(x)$ has a maximum point at $x = a$, which of the following is not true?

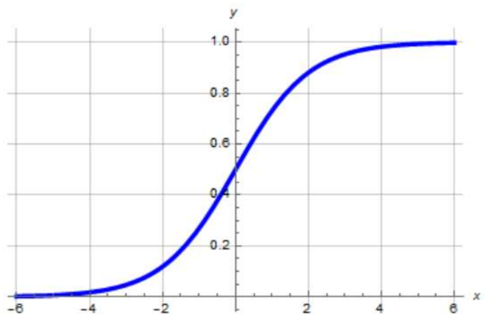
   (a) $\dfrac{dy}{dx}$ is an increasing function (b) $\dfrac{dy}{dx}$ is a decreasing function

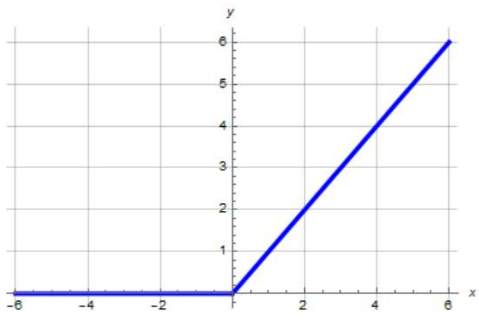
   (c) $\dfrac{d^2y}{dx^2} < 0$ at $x = a$ (d) $\dfrac{dy}{dx} > 0$ for $x < a$ and $\dfrac{dy}{dx} < 0$ for $x > a$

Section B

1. The Sigmoid and Rectified Linear Unit (ReLU) activation functions commonly used in machine learning and deep learning are given below. Find $f'(x)$ in each.

   (a) Sigmoid Function: $f(x) = \dfrac{1}{1+e^{-x}}$ (b) ReLU Function: $f(x) = \begin{cases} 0, & \text{for } x \le 0 \\ x, & \text{for } x > 0 \end{cases}$





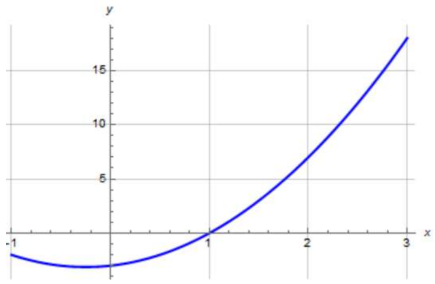
2. Find the equation of the tangent line to the curve of $y = 2x^3 - 5x^2 + 6x$ at $x = 1$.

3. Given the function $y = e^{\sin x + \cos x}$, find the gradient and the equation of the tangent line at $x = \dfrac{\pi}{2}$.

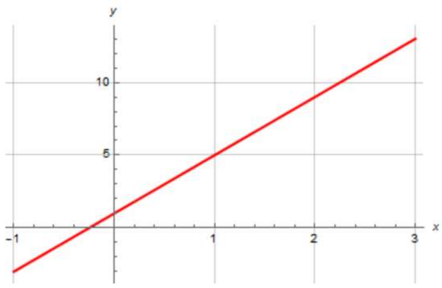
4. A company in the computer software business makes a monthly profit of $P = 150x - \dfrac{x^2}{100} - 60000$ for $0 \le x \le 5000$, where x is the number of units sold per month. Find the rate of change of the profit with respect to the number of units sold.

5. The graphs of $f(x)$, $f'(x)$, $g(x)$ and $g'(x)$ are given below.
Let $F(x) = f(g(x))$ and $G(x) = g(f(x))$. Find $F'(1)$ and $G'(1)$.

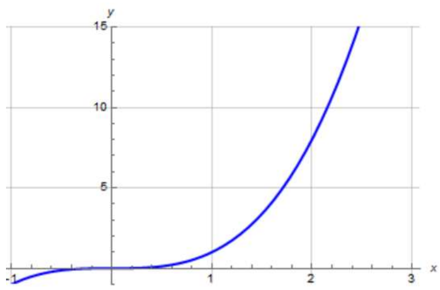
Graph of f



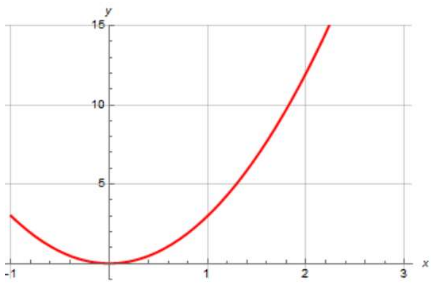
Graph of f'



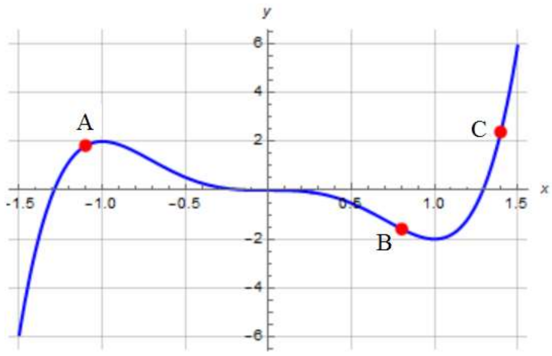
Graph of g



Graph of g'



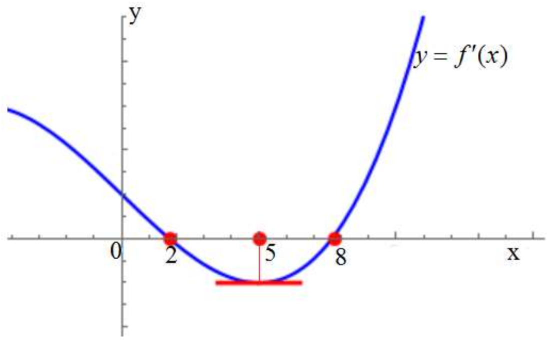
6. The graph of $y = f(x)$ is given below.



(a) Determine the signs of $\dfrac{dy}{dx}$ and $\dfrac{d^2y}{dx^2}$ at the points A, B and C.

(b) Locate all points at which f has a point of inflection.

7. The graph of $y = f'(x)$ is given below. State if each of the statements (a) to (f) is true or false. Explain.



(a) $f(-2) > f(0)$

(b) $f'(2) > 0$

(c) $f'(5) > 0$

(d) $f''(-2) > 0$

(e) $f''(9) > 0$

(f) f has a point of inflection at $x = 5$

8. The impedance Z ohms of a circuit in series with reactance $X\ \Omega$ is given by:

$$Z = \sqrt{16 + X^2}$$

If X decreases at 2 ohms/s, find the rate of change of Z when $X = 3$ ohms.

9. A ladder 3 metres long is leaning against the side of a wall. If the foot of the ladder slides away from the wall at a speed of 30 cm/s, how fast is the top end sliding down the wall when the foot of the ladder is 1 metre away from the wall?

10. Find all the relative maximum and minimum points of the following functions.

(a) $f(x) = x^3 + 9x^2 + 15x - 25$

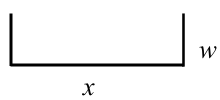
(b) $f(x) = x^3 - 3x^2 + 1$

(c) $f(x) = (x+1)^3$

(d) $h(x) = \dfrac{x^2 + 5x + 3}{x - 1}$

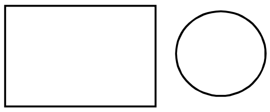
(e) $g(x) = \dfrac{x^2 - 2x + 4}{x - 2}$

11. A piece of wire 160mm is bent to form the shape of a rectangle without the top as shown in the diagram on the right. Find the width w and length x which gives the maximum cross-sectional area.



12. A window frame is made in the shape of a rectangle with a semi-circle of radius r (m) on top of the breadth of the rectangle. The breadth of the rectangle is the same as the diameter of the semi-circle. If the total area is to be 8 m^2, show that the perimeter P meters of the frame is $P = \frac{8}{r} + r\left(\frac{\pi}{2} + 2\right)$. Hence, find the minimum cost of producing the frame if 1 m costs $5.

13. A wire 3 m long, is cut into 2 pieces to form a rectangle and a circle. The rectangle has a length that is double of its breadth. Find the radius of the circle such that the total area of both the rectangle and the circle is minimum.



14. Find the global maximum and minimum values of f on the given closed interval, and state where those values occur.

(a) $f(x) = 8x - x^2$; $[0,6]$

(b) $f(x) = |1 - x|$; $[-1,2]$

Answers

Section A

1. (d) 2. (b) 3. (a)

Section B

1. (a) $f'(x) = \dfrac{e^{-x}}{\left(1+e^{-x}\right)^2}$ or $f(x)\left(1-f(x)\right)$ (b) $f'(x) = \begin{cases} 0, & \text{for } x<0 \\ 1, & \text{for } x>0 \\ \text{does not exist}, & \text{for } x=0 \end{cases}$

2. $y = 2x+1$

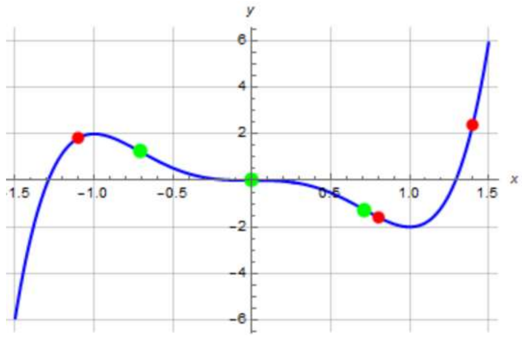
3. $y = e\left(1+\dfrac{\pi}{2}\right) - ex$ or $y = 6.988 - ex$

4. $\dfrac{dP}{dx} = 150 - \dfrac{x}{50}$

5. $F'(1) = 15$, $G'(1) = 0$

6. (a) $\left.\dfrac{dy}{dx}\right|_{x=-1.1} > 0$, $\left.\dfrac{dy}{dx}\right|_{x=0.8} < 0$, $\left.\dfrac{dy}{dx}\right|_{x=1.4} > 0$; $\left.\dfrac{d^2y}{dx^2}\right|_{x=-1.1} < 0$, $\left.\dfrac{d^2y}{dx^2}\right|_{x=0.8} > 0$, $\left.\dfrac{d^2y}{dx^2}\right|_{x=1.4} > 0$

(b)



7. (a) False. $f'(x) > 0$ for x in $(-\infty, 2)$, so f is increasing in $(-\infty, 2) \Rightarrow f(-2) < f(0)$
   (b) False. The graph of $y = f'(x)$ cuts the x-axis at $x = 2$, so $f'(2) = 0$.
   (c) False. The graph of $y = f'(x)$ is below the x-axis at $x = 5$, so $f'(5) < 0$.
   (d) $f''(-2)$ refers to the slope of the tangent to $y = f'(x)$ at $x = -2$, which is negative. Hence $f''(-2) < 0$.
   (e) True, since the tangent to the curve of $y = f'(x)$ at $x = 9$ has a positive slope.
   (f) True, since $f''(x) < 0$ to the left of $x = 5$ and $f''(x) > 0$ to the right of $x = 5$. This means the concavity of f changes at $x = 5$.

8. −1.2 ohm/s

9. −10.607 cm/s

10. (a) (−5, 0) maximum , (−1, −32) minimum (b) (0, 1) maximum , (2, −3) minimum
    (c) (−1, 0) point of inflection (d) (−2, 1) maximum , (4, 13) minimum
    (e) (0, −2) maximum , (4, 6) minimum

11. 40 mm , 80 mm

12. $53.44

13. 0.196m

14. (a) Global maximum is 16, when $x = 4$; global minimum is 0, when $x = 0$
    (b) Global maximum is 2, when $x = -1$; global minimum is 0, when $x = 1$

Practical 8

Task 1

Use Matplotlib in Python to graph each of the functions f from (a) to (c).
Then, estimate the global maximum and minimum values of f (if any) on the given interval.
Also, state where those values occur.

(a) $f(x) = 4x^3 - 3x^4$ on the interval $(-\infty, +\infty)$

(b) $f(x) = (x^2 - 1)^2$ on the interval $(-\infty, +\infty)$

(c) $f(x) = 1 + \frac{1}{x}$ on the interval $(0, \infty)$

Task 2

Let $f(x) = x^4 - 3x^3 + 15$.
Implement gradient descent in Python to find the global minimum of f, and state where it occurs.

(a) Start the search at $x_0 = 6$ with the following settings:
learning rate $\alpha = 0.01$, epsilon = 0.00001, and maximum number of iterations = 10000

(b) Start the search at $x_0 = 6$ with the following settings:
learning rate $\alpha = 0.001$, epsilon = 0.00001, and maximum number of iterations = 10000

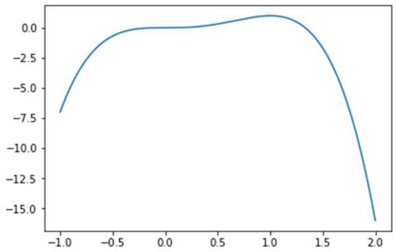
(c) Start the search at $x_0 = 10$ with the following settings:
learning rate $\alpha = 0.01$, epsilon = 0.00001, and maximum number of iterations = 10000

(d) Start the search at $x_0 = -2$ with the following settings:
learning rate $\alpha = 0.01$, epsilon = 0.00001, and maximum number of iterations = 10000

Answers

Task 1

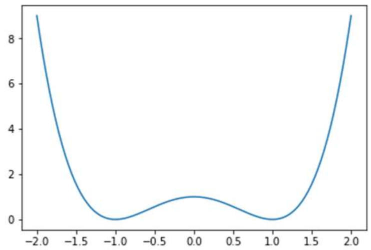
(a)



Global maximum = 1, at $x = 1$

No minimum

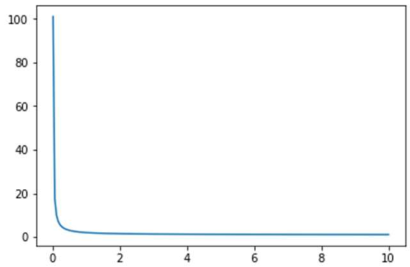
(b)



Minimum value = 0, at $x = \pm 1$

No maximum

(c)



No maximum or minimum

Task 2

(a)

```
Iteration  69 : x-value is:  2.2499556210437 f(x) is:  6.457031269940586
Iteration  70 : x-value is:  2.2499646074278457 f(x) is:  6.457031262682655
The minimum is:  6.457031262682655
It occurs at:  2.2499646074278457
```

(b)

```
Iteration  357 : x-value is:  2.2504880374611953 f(x) is:  6.457033662275709
Iteration  358 : x-value is:  2.250478150414891 f(x) is:  6.457033565512635
The minimum is:  6.457033565512635
It occurs at:  2.250478150414891
```

(c) Iterations fail to converge:

```
Iteration  1 : x-value is:  -21.0 f(x) is:  222279.0
Iteration  2 : x-value is:  389.13 f(x) is:  22751900903.504723
Iteration  3 : x-value is:  -2342899.03245888 f(x) is:  3.013109038267281e+25
Iteration  4 : x-value is:  5.144238880063777e+17 f(x) is:  7.003006184304912e+70
Iteration  5 : x-value is:  -5.44531958766121e+51 f(x) is:  8.792117731840386e+206
```

(d)

```
Iteration  1023 : x-value is:  -0.01050984024169965 f(x) is:  15.000003494848825
Iteration  1024 : x-value is:  -0.010499852699619663 f(x) is:  15.000003484883223
The minimum is:  15.000003484883223
It occurs at:  -0.010499852699619663
```