The Mean Value Theorem

1 The Mean Value Theorem

We come now to one of the most important results in differential calculus: the Mean Value Theorem. Its usefulness is in its power in proving some of the big theorems of calculus. We will prove a few important results with it in this section.

The Mean Value Theorem says that the slope of the secant line of a function going through the endpoints of an interval on which the function is continuous and differentiable is equal to the slope of the tangent line to the function at some point between the endpoints. It asserts the existence of such a point but makes no attempt to determine what that point is, so it is much like the Intermediate Value Theorem in that respect.

First we will state and prove the following special case of the Mean Value Theorem which will then be used to prove the Mean Value Theorem in general.

Theorem 1.1 (Rolle's Theorem) *Suppose $f(x)$ is continuous on $[a, b]$ and differentiable on (a, b). If $f(a) = 0 = f(b)$, then there exists a point $c \in (a, b)$ such that $f'(c) = 0$.*

Proof. Suppose $f(x)$ is continuous on $[a, b]$, differentiable on (a, b) and $f(a) = 0 = f(b)$. We will prove the theorem using two cases. First, suppose that $f(x) > 0$ for some $x \in (a, b)$. Since $f(x)$ is continuous on $[a, b]$, there exists a point $c \in [a, b]$ for which $f(c)$ is the maximum value of f on $[a, b]$. Furthermore, $f(c) > 0$ implies $c \neq a$ and $c \neq b$, so $c \in (a, b)$ and so $f'(c) = 0$ because $f(x)$ is differentiable on (a, b).

Now suppose $f(x) \leq 0$ for all $x \in (a, b)$. Then either $f(x) = 0$ for all $x \in (a, b)$ in which case $f'(x) = 0$ for all $x \in (a, b)$, or else $f(x) < 0$ for some $x \in (a, b)$. Since $f(x)$ is continuous on $[a, b]$, we know that there is a point $c \in [a, b]$ for which $f(c)$ is the minimum value of $f(x)$ on $[a, b]$. Since $f(c)$ is the minimum on $[a, b]$ and $f(x) < 0$ for some $x \in (a, b)$, $f(c) < 0$. Consequently, $c \neq a$ and $c \neq b$, so $c \in (a, b)$ and therefore $f'(c) = 0$ because $f(x)$ is differentiable on (a, b). This proves the theorem. □

Corollary 1.2 *Suppose $f(x)$ is continuous on $[a, b]$ and differentiable on (a, b). If $f(a) = f(b)$, then there exists a point $c \in (a, b)$ such that $f'(c) = 0$.*

Proof. Suppose $f(a) = k = f(b)$. Let $h(x) = f(x) - k$. Then as a difference of continuous functions, $h(x)$ is continuous on $[a, b]$ and $h'(x) = f'(x)$ is differentiable on (a, b) since f is. Furthermore, $h(a) = f(a) - k = 0$ and $h(b) = f(b) - k = 0$. Therefore, by Rolle's Theorem, there exists a point $c \in (a, b)$ such that $0 = h'(c) = f'(c)$ as we wanted to show. □

Theorem 1.3 (Mean Value Theorem) *If $f(x)$ is continuous on $[a,b]$ and differentiable on (a,b), then there exists a point c in (a,b) such that*

$$f(b) - f(a) = f'(c)(b-a) \quad \Rightarrow \quad f'(c) = \frac{f(b)-f(a)}{b-a}.$$

Proof. Consider the function $h(x) = f(x)(b-a) - x(f(b) - f(a))$, noting that $h(x)$ is continuous on $[a,b]$ because $f(x)$ is. We compute

$$h'(x) = f'(x)(b-a) - (f(b) - f(a))$$

which is differentiable on (a,b) since $f(x)$ is. Note that

$$\begin{aligned} h(a) = bf(a) - af(a) - af(b) + af(a) &= bf(a) - af(b) \\ &= bf(b) - af(b) - bf(b) + bf(a) \\ &= f(b)(b-a) - b(f(b) - f(a)) = h(b). \end{aligned}$$

Therefore, by the corollary to Rolle's Theorem, there exists a point c in (a,b) such that

$$0 = h'(c) = f'(c)(b-a) - (f(b) - f(a)) \quad \Rightarrow \quad f(b) - f(a) = f'(c)(b-a)$$

which proves the Mean Value Theorem. □

In the special case that the function measures some quantity with respect to time (such as position, interest, population, etc.) the Mean Value Theorem says that the average rate of change of the quantity over some interval of time $[t_0, t_1]$ is equal to the instantaneous rate of change of the quantity at some time t between t_0 and t_1.

For example, suppose you drive the 60 miles to Atlanta in 1 hour. Then your average speed was 60 mph and so by the Mean Value Theorem, your speedometer (which measures the speed of the car at each moment in time) had to have read 60 mph at some point in your trip to Atlanta.

We are now ready for some important consequences of the Mean Value Theorem. First, we have a converse to the theorem that the derivative of a constant function is 0.

Corollary 1.4 *Suppose $f(x)$ is continuous on $[a,b]$ and differentiable on (a,b). If $f'(x) = 0$ for all x in (a,b), then there is a constant C such that $f(x) = C$ for all x in $[a,b]$.*

Proof. Let x_0 be a point in $(a,b]$. Then $f(x)$ is continuous on $[a,x_0]$ and differentiable on (a,x_0). By the Mean Value Theorem, there exists $c \in (a,x_0)$ such that $f'(c)(x_0 - a) = f(x_0) - f(a)$. But by our hypothesis, $f'(c) = 0$ and so $f(x_0) - f(a) = 0$ which implies that $f(x_0) = f(a)$. We chose $x_0 \in (a,b]$ arbitrarily so this argument holds for all $x \in (a,b]$; i.e., $f(x) = f(a)$ for all x in $(a,b]$. Thus, defining $C := f(a)$ we have that $f(x) = C$ for all x in $[a,b]$. □

A consequence of the previous corollary is the next result. It will be of vital importance to us when we discuss antiderivatives and differential equations.

Corollary 1.5 *Suppose $f(x)$ and $g(x)$ are continuous on $[a,b]$ and differentiable on (a,b). If $f'(x) = g'(x)$ for all x in (a,b), then there is a constant C such that $f(x) = g(x) + C$ for all x in $[a,b]$.*

Proof. Consider the function $h(x) = f(x) - g(x)$. Since $f(x)$ and $g(x)$ are continuous on $[a, b]$, so is $h(x)$. Also, since $h'(x) = f'(x) - g'(x)$, $h(x)$ is differentiable on (a, b) since both $f(x)$ and $g(x)$ are. But $f'(x) = g'(x)$ for all x in (a, b) and so $h'(x) = 0$ for all x in (a, b). By the previous corollary, there is a constant C such $C = h(x) = f(x) - g(x)$ for all $x \in [a, b]$. That is, $f(x) = g(x) + C$ for all x in $[a, b]$. □

Examples.

1. Suppose we know that $f'(x) = 3\sqrt{x}$ and that $f(0) = 4$. We can determine what $f(x)$ is by doing the following. First, we inspect $f'(x)$ and ask the question "what function, if we differentiate it, will give $3\sqrt{x}$." We note that $\sqrt{x} = x^{1/2}$ and so by the power rule, we need a function like $x^{3/2}$. It is not quite this function though because its derivative is $\frac{3}{2}x^{1/2}$. But that is just $1/2$ of $f'(x)$ and so if we take $g(x) = 2x^{3/2}$ we will get $g'(x) = f'(x)$. Note that for any $a, b \in \mathbb{R}$ with $0 < a < b$ we have that $g(x)$ is continuous on $[a, b]$ and differentiable on (a, b). Therefore, by the previous corollary, there is a constant C such that $f(x) = g(x) + C$. We know that
$$4 = f(0) = g(0) + C = 0 + C = C.$$
Therefore, we must have $f(x) = 2x^{3/2} + 4$ for all x in the interval $[a, b]$. Note that this implies that $f(x) = 2x^{3/2} + 4$ for all $x \in [0, \infty)$ since any interval $[a, b]$ with $0 < a < b$ works.

2. Can you determine what $f(x)$ must be if $f'(x) = 6e^{-3x}$ and $f(0) = 3$?

The final result of this section has to do with increasing and decreasing functions. First though we must define what we mean by a function increasing on a certain interval in mathematical language.

Definition 1.1 *Suppose $f(x)$ is a function with domain D. If there is an open interval (a, b) such that for all x_1, x_2 in (a, b) with $x_1 < x_2$ we have $f(x_1) < f(x_2)$, then we say that $f(x)$ is* ***increasing*** *on (a, b). Similarly, if there is an open interval (c, d) such that for all x_1, x_2 in (c, d) with $x_1 < x_2$ we have $f(x_1) > f(x_2)$, then we say that $f(x)$ is* ***decreasing*** *on (c, d).*

This final result says that the sign of the derivative tells us if a function is increasing or decreasing on an interval.

Corollary 1.6 *Suppose $f(x)$ is continuous on $[a, b]$ and differentiable on (a, b). If $f'(x) > 0$ for all x in (a, b), then $f(x)$ is increasing on (a, b). If $f'(x) < 0$ for all x in (a, b), then $f(x)$ is decreasing on (a, b).*

Proof. Suppose $f'(x) > 0$ for all x in (a, b). Let x_1, x_2 be in (a, b) with $x_1 < x_2$. Then $x_2 - x_1 > 0$ and so by the Mean Value Theorem there exists a point c in (x_1, x_2) such that
$$f(x_2) - f(x_1) = f'(c)(x_2 - x_1) > 0 \quad \Rightarrow \quad f(x_2) > f(x_1).$$
That is, $f(x)$ is increasing on (a, b). On the other hand, suppose $f'(x) < 0$ for all x in (a, b) and let x_1, x_2 be in (a, b) with $x_1 < x_2$. Then again $x_2 - x_1 > 0$ and by the Mean Value Theorem there exists a point c in (x_1, x_2) such that
$$f(x_2) - f(x_1) = f'(c)(x_2 - x_1) < 0 \quad \Rightarrow \quad f(x_2) < f(x_1)$$
showing that $f(x)$ is decreasing on (a, b). □

Examples.

1. We will determine the open intervals on which $f(x) = x^3 - 12x + 17$ is increasing and those on which it is decreasing. First, we compute

$$f'(x) = 3x^2 - 12 = 3(x^2 - 4) = 3(x - 2)(x + 2).$$

Now, $f'(x) = 0$ when $x = 2$ and when $x = -2$ and so we divide the real line into three segments with these cut points. On the interval $(-\infty, -2)$ we find that $f'(x) > 0$ so by the previous corollary f is increasing on this interval. For the interval $(-2, 2)$ we find that $f'(x) < 0$ and so f is decreasing there. Finally, on the interval $(2, \infty)$ we find $f'(x) > 0$ and so f is increasing on that interval.

2. Can you find the open intervals on which $g(x) = x^2e^{-2x}$ is increasing and those on which it is decreasing?

3. We will use the last corollary of the Mean Value Theorem together with the Intermediate Value Theorem to prove that the equation $e^{-x} = x - 1$ has exactly one solution in the interval $[1, 2]$. First we define a function $f(x) = e^{-x} - x + 1$ and note that solutions of the equation correspond to roots of the function. We evaluate $f(1) = e^{-1} - 1 + 1 = e^{-1} > 0$ and $f(2) = e^{-2} - 2 + 1 = \frac{1}{e^2} - 1 < 0$. Since the exponential function is continuous everywhere and sums of continuous functions are continuous, $f(x)$ is continuous on $[1, 2]$ and so by the intermediate value theorem, there exists $c \in (1, 2)$ for which $f(c) = 0$ (since $f(2) < 0 < f(1)$). The question now is whether or not there is more than one root of f in the interval $[1, 2]$. For this, we compute

$$f'(x) = -e^{-x} - 1 = -(e^{-x} + 1) < 0 \quad \text{for all } x \in [1, 2].$$

This means that $f(x)$ can only pass through the x-axis at most once in the interval $[1, 2]$ and consequently c is the only solution of the equation in the interval $[1, 2]$.