6

Chapter 6 Applications of the Derivative

6.1 Optimization

Many important applied problems involve finding the best way to accomplish some task. Often this involves finding the maximum or minimum value of some function: the minimum time to make a certain journey, the minimum cost for doing a task, the maximum power that can be generated by a device, and so on. Many of these problems can be solved by finding the appropriate function and then using techniques of calculus to find the maximum or the minimum value required.

Generally such a problem will have the following mathematical form: Find the largest (or smallest) value of \( f(x) \) when \( a \leq x \leq b \). Sometimes \( a \) or \( b \) are infinite, but frequently the real world imposes some constraint on the values that \( x \) may have.

Such a problem differs in two ways from the local minimum and maximum problems we encountered when graphing functions: We are interested only in the function between \( a \) and \( b \), and we want to know the largest or smallest value that \( f(x) \) takes on, not merely values that are the largest or smallest in a small interval. That is, we seek not a local maximum or minimum but a \textit{global} maximum or minimum, sometimes also called an \textit{absolute} maximum or minimum.

Any global maximum or minimum must of course be a local maximum or minimum. If we find all possible local extrema, then the global maximum, if it exists, must be the largest of the local maxima and the global minimum, if it exists, must be the smallest of the local minima. We already know where local extrema occur: only at those points at which \( f'(x) \) is zero or undefined. Actually, there are two additional points at which a maximum or minimum can occur if the endpoints \( a \) and \( b \) are not infinite, namely, at a

EXAMPLE 6.1.5 Let \( f(x) = x^2 \) restricted to \([-2, 1]\). The graph of \( f(x) = x^2 \), restricted to \([-2, 1]\), is shown in figure 6.1.1. We compute \( f(x) = 2x \), which is zero at \( x = 0 \) and is always defined.

Since \( f'(1) = 2 \) we would not normally flag \( x = 1 \) as a point of interest, but it is clear from the graph that when \( f(x) \) is restricted to \([-2, 1]\) there is a local maximum at \( x = 1 \). Likewise we would not normally pay attention to \( x = -2 \), but since we have truncated \( f \) at \(-2\) we have introduced a new local maximum there as well. In a technical sense nothing new is going on here: When we truncate \( f \) we actually create a new function, let's call it \( g \), that is defined only on the interval \([-2, 1]\). If we try to compute the derivative of this new function we actually find that it does not have a derivative at \(-2 \) or \( 1 \) Why? Because to compute the derivative at \( 1 \) we must compute the limit

\[
\lim_{\Delta x \to 0} \frac{g(1 + \Delta x) - g(1)}{\Delta x}
\]

This limit does not exist because when \( \Delta x > 0 \), \( g(1 + \Delta x) \) is not defined. It is simpler, however, simply to remember that we must always check the endpoints.

So the function \( g \), that is, \( f \) restricted to \([-2, 1]\), has one critical value and two finite endpoints, any of which might be the global maximum or minimum. We could first determine which of these are local maximum or minimum points (or neither); then the largest local maximum must be the global maximum and the smallest local minimum must be the global minimum. It is usually easier, however, to compute the value of \( f \) at every point at which the global maximum or minimum might occur; the largest of these is the global maximum, the smallest is the global minimum.

So we compute \( f(-2) = 4, f(0) = 0 \), \( f(1) = 1 \). The global maximum is 4 at \( x = -2 \) and the global minimum is 0 at \( x = 0 \).

EXAMPLE 6.1.6 (or smallest) value of \( f(x) = x^2 \) on the interval \([-2, 1]\), shown in figure 6.1.1. We compute \( f'(x) = 2x \), which is zero at \( x = 0 \) and is always defined.

Since \( f'(1) = 2 \) we would not normally flag \( x = 1 \) as a point of interest, but it is clear from the graph that when \( f(x) \) is restricted to \([-2, 1]\) there is a local maximum at \( x = 1 \). Likewise we would not normally pay attention to \( x = -2 \), but since we have truncated \( f \) at \(-2\) we have introduced a new local maximum there as well. In a technical sense nothing new is going on here: When we truncate \( f \) we actually create a new function, let's call it \( g \), that is defined only on the interval \([-2, 1]\). If we try to compute the derivative of this new function we actually find that it does not have a derivative at \(-2 \) or \( 1 \). Why? Because to compute the derivative at \( 1 \) we must compute the limit

\[
\lim_{\Delta x \to 0} \frac{g(1 + \Delta x) - g(1)}{\Delta x}
\]

This limit does not exist because when \( \Delta x > 0 \), \( g(1 + \Delta x) \) is not defined. It is simpler, however, simply to remember that we must always check the endpoints.

So the function \( g \), that is, \( f \) restricted to \([-2, 1]\), has one critical value and two finite endpoints, any of which might be the global maximum or minimum. We could first determine which of these are local maximum or minimum points (or neither); then the largest local maximum must be the global maximum and the smallest local minimum must be the global minimum. It is usually easier, however, to compute the value of \( f \) at every point at which the global maximum or minimum might occur; the largest of these is the global maximum, the smallest is the global minimum.

So we compute \( f(-2) = 4, f(0) = 0 \), \( f(1) = 1 \). The global maximum is 4 at \( x = -2 \) and the global minimum is 0 at \( x = 0 \).

EXAMPLE 6.1.1 Find the maximum and minimum values of \( f(x) = x^2 \) on the interval \([-2, 1]\), shown in figure 6.1.1. We compute \( f'(x) = 2x \), which is zero at \( x = 0 \) and is always defined.

Since \( f'(1) = 2 \) we would not normally flag \( x = 1 \) as a point of interest, but it is clear from the graph that when \( f(x) \) is restricted to \([-2, 1]\) there is a local maximum at \( x = 1 \). Likewise we would not normally pay attention to \( x = -2 \), but since we have truncated \( f \) at \(-2\) we have introduced a new local maximum there as well. In a technical sense nothing new is going on here: When we truncate \( f \) we actually create a new function, let's call it \( g \), that is defined only on the interval \([-2, 1]\). If we try to compute the derivative of this new function we actually find that it does not have a derivative at \(-2 \) or \( 1 \). Why? Because to compute the derivative at \( 1 \) we must compute the limit

\[
\lim_{\Delta x \to 0} \frac{g(1 + \Delta x) - g(1)}{\Delta x}
\]

This limit does not exist because when \( \Delta x > 0 \), \( g(1 + \Delta x) \) is not defined. It is simpler, however, simply to remember that we must always check the endpoints.

So the function \( g \), that is, \( f \) restricted to \([-2, 1]\), has one critical value and two finite endpoints, any of which might be the global maximum or minimum. We could first determine which of these are local maximum or minimum points (or neither); then the largest local maximum must be the global maximum and the smallest local minimum must be the global minimum. It is usually easier, however, to compute the value of \( f \) at every point at which the global maximum or minimum might occur; the largest of these is the global maximum, the smallest is the global minimum.

So we compute \( f(-2) = 4, f(0) = 0 \), \( f(1) = 1 \). The global maximum is 4 at \( x = -2 \) and the global minimum is 0 at \( x = 0 \).

EXAMPLE 6.1.2 Find the maximum value of \( f(x) = x^2 \) on the interval \([0, 4]\). First note that \( f'(x) = 2x + 4 \) when \( x = 2 \), and \( f(2) = 1 \). Next observe that \( f'(x) \) is defined for all \( x \), so there are no other critical values. Finally, \( f(0) = -3 \) and \( f(4) = -1 \). The largest value of \( f(x) \) on the interval \([0, 4]\) is \( f(2) = 1 \).

EXAMPLE 6.1.4 Let \( f(x) = x^3 + 4x - 3 \). Find the maximum value of \( f(x) \) on the interval \([-1, 1]\). First note that \( f'(x) = 3x^2 + 4 \) when \( x = 2 \), but \( f(2) = 1 \). Next observe that \( f'(x) \) is defined for all \( x \), so there are no other critical values. Finally, \( f(0) = -3 \) and \( f(4) = -1 \). The largest value of \( f(x) \) on the interval \([-1, 1]\) is \( f(1) = 0 \).

EXAMPLE 6.1.5 Find the maximum and minimum values of the function \( f(x) = 7 + |x - 2| \) for \( x \) between 1 and 4 inclusive. The derivative \( f'(x) \) is never zero, but \( f'(x) \) is undefined at \( x = 2 \), so we compute \( f(2) = 7 \). Checking the end points we get \( f(1) = 8 \) and \( f(4) = 9 \). The smallest of these numbers is \( f(2) = 7 \), which is, therefore, the minimum value of \( f(x) \) on the interval \( 1 \leq x \leq 4 \), and the maximum is \( f(4) = 9 \).

EXAMPLE 6.1.6 Find all local maxima and minima for \( f(x) = x^3 - x \), and determine whether there is a global maximum or minimum on the open interval \((-2, 2)\). In example 5.1.2 we found a local maximum at \((-\sqrt{3}/3, 2\sqrt{3}/9)\) and a local minimum at \((\sqrt{3}/3, -2\sqrt{3}/9)\). Since the endpoints are not in the interval \((-2, 2)\) they cannot be con-
6.1 Optimization 107

6.1 Optimization 109

EXAMPLE 6.1.10 If you fit the largest possible cone inside a sphere, what fraction of the volume of the sphere is occupied by the cone? (Here by “cone” we mean a right circular cone, i.e., a cone for which the base is perpendicular to the axis of symmetry, and for which the cross-section cut perpendicular to the axis of symmetry at any point is a circle.)

Let \( R \) be the radius of the sphere, and let \( r \) and \( h \) be the base radius and height of the cone inside the sphere. What we want to maximize is the volume of the cone: \( \pi r^2 h/3 \). Here \( R \) is a fixed value, but \( r \) and \( h \) can vary. Namely, we could choose \( r \) to be as large as possible—equal to \( R \)—by taking the height equal to \( R \); or we could make the cone’s height \( h \) larger at the expense of making \( r \) a little less than \( R \). See the cross-section depicted in Figure 6.1.5.

EXAMPLE 6.1.11 You are making cylindrical containers to contain a given volume. Suppose that the top and bottom are made of a material that is \( N \) times as expensive (cost per unit area) as the material used for the lateral side of the cylinder. Find (in terms of \( N \)) the ratio of height to base radius of the cylinder that minimizes the cost of making the containers.

Let us first choose letters to represent various things: \( h \) for the height, \( r \) for the base radius, \( V \) for the volume of the cylinder, and \( c \) for the cost per unit area of the lateral side of the cylinder; \( V \) and \( c \) are constants, and \( h \) and \( r \) are variables. Now we can write the cost of materials:

\[
\pi r^2 h + c(V/h) + N(2\pi r^2).
\]

Again we have two variables; the relationship is provided by the fixed volume of the cylinder: \( V = \pi r^2 h \). We use this relationship to eliminate \( h \) (we could eliminate \( r \), but it’s a little easier if we eliminate \( h \), which appears in only one place in the above formula for cost).

The result is

\[
f(r) = 2\pi r V/\pi r^2 + 2Nc\sqrt{V}/r = 2V/r + 2Nc/r^2.
\]

We want to know the minimum value of this function when \( r \) is in \((0, \infty)\). We now set \( f'(r) = -2V/r^2 + 4Ncr \), giving \( r = \sqrt{V/(2Nc)} \). Since \( f''(r) = 4V/r^3 + 4Nc/3r \) is positive when \( r \) is positive, there is a local minimum at the critical value, and hence a global minimum since there is only one critical value.

Finally, since \( h = V/(\pi r^2) \),

\[
\frac{h}{r} = \frac{V}{\pi r^2} = \frac{V}{\pi(2Nc/r)} = 2N,
\]

so the minimum cost occurs when the height \( h \) is 2 times the radius. If, for example, there is no difference in the cost of materials, the height is twice the radius (or the height is equal to the diameter).
EXAMPLE 6.1.12  Suppose you want to reach a point A that is located across the sand from a nearby road (see figure 6.1.5). Suppose that the road is straight, and h is the distance from A to the closest point C on the road. Let x be your speed on the road, and let w, which is less than x, be your speed on the sand. Right now you are at the point D, which is r distance a from A. At what point B should you turn off the road and head across the sand in order to minimize your travel time to A? Let b be the distance short of C where you turn off, i.e., the distance from B to C. We want to minimize the total travel time. Recall that when traveling at constant velocity, time is distance divided by velocity.

You travel the distance $\sqrt{a^2 + b^2}$ at speed v, and then the distance $\sqrt{b^2 + w^2}$ at speed w. Since $\sqrt{a^2 + b^2} = a - x$ and, by the Pythagorean theorem, $\sqrt{b^2 + w^2} = \sqrt{\frac{2w^2}{x} + b^2}$; the total time for the trip is

$$f(x) = \frac{a - x}{v\sqrt{\frac{2w^2}{x} + b^2}} + \frac{\sqrt{\frac{2w^2}{x} + b^2}}{w}$$

We want to find the minimum value of f when x is between 0 and a. As usual we set $f'(x) = 0$ and solve for x:

$$0 = f'(x) = \frac{1 - \frac{x}{v^2} + \frac{\sqrt{\frac{2w^2}{x} + b^2}}{v\sqrt{\frac{2w^2}{x} + b^2}}}{\sqrt{\frac{2w^2}{x} + b^2}} + \frac{x^2}{2w^2} - \frac{1}{2}$$

$$x = \frac{v^2}{\sqrt{2w^2}} - b$$

Notice that a does not appear in the last expression, but a is not irrelevant, since we are interested only in critical values that are in [0, a], and $v\sqrt{\frac{2w^2}{x} + b^2}$ is either in this interval or not. If it is, we can use the second derivative to test it:

$$f''(x) = -\frac{v^4}{(2w^2)^{\frac{3}{2}}} + \frac{2v\sqrt{\frac{2w^2}{x} + b^2}}{2w^2} + \frac{x^2}{2w^2}$$

Since this is always positive there is a local minimum at the critical point, and so it is a global minimum as well.

If the critical value is not in [0, a] it is larger than a. In this case the minimum must occur at one of the endpoints. We can compute

$$f(0) = \frac{b}{w}$$
$$f(a) = \frac{\sqrt{\frac{2w^2}{a} + b^2}}{w}$$

You want to make cylindrical containers to hold 1 liter (1000 cubic centimeters) using the least amount of construction material. The side is made from a rectangular piece of material, and this can be done with no material wasted. However, the top and bottom are cut from squares of side 2r, so that $2\pi r^2 = 2\pi r^2$ of material is needed (rather than $2\pi r^2$, which is the total area of the top and bottom). Find the dimensions of the container using the least amount of material, and also find the ratio of height to radius for this container.

15. You want to make cylindrical containers of a given volume V using the least amount of construction material. The side is made from a rectangular piece of material, and this can be done with no material wasted. However, the top and bottom are cut from squares of side 2r, so that $2\pi r^2 = 2\pi r^2$ of material is needed (rather than $2\pi r^2$, which is the total area of the top and bottom). Find the dimensions of the container using the least amount of material, and also find the ratio of height to radius for this container.

16. Given a right circular cone, you can put an upside-down cone inside it so that its vertex is at the center of the base of the larger cone and its base is parallel to the base of the larger cone. If you choose the upside-down cone to have the largest possible volume, what fraction of the volume of the larger cone does it occupy? (Let H and R be the height and base radius of the larger cone, and let h and r be the height and base radius of the smaller cone. Hint: Use similar triangles to get an equation relating H and r.)

17. In example 6.1.12, what happens if w ≤ x (i.e., your speed on sand is at least your speed on the road)?

18. A container holding a fixed volume is being made in the shape of a cylinder with a hemispherical top. (The hemispherical top has the same radius as the cylinder.) Find the ratio of height to radius of the cylinder which minimizes the cost of the container if (a) the cost per unit area of the top is twice as great as the cost per unit area of the side; and the container is made with no bottom; (b) the same as in (a), except that the container is made with a circular bottom, for which the cost per unit area is 1.5 times the cost per unit area of the side; (c) the cost per unit area of side is 1 meter per square meter; (d) the container is made with a circular bottom, for which the cost per unit area is 1.5 times the cost per unit area of the side.

19. A piece of cardboard is 1 meter by 1.2 meter. A square is to be cut from each corner and the sides folded up to make an open-top box. What are the dimensions of the box with maximum possible volume?

6.1 Optimization 113

8. You have 1 foot of fence to make a rectangular play area alongside the wall of your house. The wall of the house bounds one side. What is the largest size possible (in square feet) for the play area?

9. Marketing tells you that if you set the price of an item at $10 then you will be unable to sell it, but that you can sell 500 items for each dollar below $10 that you set the price. Suppose your fixed costs total $3000, and your marginal cost is $2 per item. What is the most profit you can make?

10. Find the areas of the largest rectangle that fits inside a semicircle of radius 10 (one side of the rectangle is along the diameter of the semicircle).

11. Find the area of the largest rectangle that fits inside a semicircle of radius r (one side of the rectangle is along the diameter of the semicircle).

12. For a cylinder with surface area 50, including the top and the bottom, find the ratio of height to base radius that maximizes the volume.

13. For a cylinder with given surface area $S$, including the top and the bottom, find the ratio of height to base radius that maximizes the volume.

14. You want to make cylindrical containers to hold 1 liter (1000 cubic centimeters) using the least amount of construction material. The side is made from a rectangular piece of material, and this can be done with no material wasted. However, the top and bottom are cut from squares of side 2r, so that $2\pi r^2 = 2\pi r^2$ of material is needed (rather than $2\pi r^2$, which is the total area of the top and bottom). Find the dimensions of the container using the least amount of material, and also find the ratio of height to radius for this container.

15. You want to make cylindrical containers of a given volume V using the least amount of construction material. The side is made from a rectangular piece of material, and this can be done with no material wasted. However, the top and bottom are cut from squares of side 2r, so that $2\pi r^2 = 2\pi r^2$ of material is needed (rather than $2\pi r^2$, which is the total area of the top and bottom). Find the optimal ratio of height to radius.

16. Given a right circular cone, you can put an upside-down cone inside it so that its vertex is at the center of the base of the larger cone and its base is parallel to the base of the larger cone. If you choose the upside-down cone to have the largest possible volume, what fraction of the volume of the larger cone does it occupy? (Let H and R be the height and base radius of the larger cone, and let h and r be the height and base radius of the smaller cone. Hint: Use similar triangles to get an equation relating H and r.)

17. In example 6.1.12, what happens if w ≥ x (i.e., your speed on sand is at least your speed on the road)?

18. A container holding a fixed volume is being made in the shape of a cylinder with a hemispherical top. (The hemispherical top has the same radius as the cylinder.) Find the ratio of height to radius of the cylinder which minimizes the cost of the container if (a) the cost per unit area of the top is twice as great as the cost per unit area of the side; and the container is made with no bottom; (b) the same as in (a), except that the container is made with a circular bottom, for which the cost per unit area is 1.5 times the cost per unit area of the side; (c) the cost per unit area of side is 1 meter per square meter; (d) the container is made with a circular bottom, for which the cost per unit area is 1.5 times the cost per unit area of the side.

19. A piece of cardboard is 1 meter by 1.2 meter. A square is to be cut from each corner and the sides folded up to make an open-top box. What are the dimensions of the box with maximum possible volume?
If you fit the cone with the largest possible surface area (lateral area plus area of base) into a sphere, what percent of the volume of the sphere is occupied by the cone? 

31. Two electrical charges, one a positive charge A of magnitude a and the other a negative charge B of magnitude b, are located a distance c apart. A positively charged particle P is situated on the line between A and B. Find where P should be put so that the pull away from A towards B is minimal. Here assume that the force from each charge is proportional to the strength of the source and inversely proportional to the square of the distance from the source.

32. Find the fraction of the area of a triangle that is occupied by the largest rectangle that can be drawn in the triangle (with one of its sides along a side of the triangle). Show that this fraction does not depend on the dimensions of the given triangle.

33. How are your answers to Problem 9 affected if (a) the cast per item for the x items, instead of being simply $s$, decrease below $s$ in proportion to $x$ (because of economy of scale and volume discounts) by 1 cent for each 25 items produced? 

34. You are standing near the side of a large wading pool of uniform depth when you see a child in trouble. You can run at a speed $v$ on land and at a slower speed $v_2$ in the water. Your perpendicular distance from the side of the pool is $a$, the child's perpendicular distance is $b$, and the distance along the pool between the closest point to you and the closest point to the child is $c$ (see the figure below). Without stopping to do any calculations, you instinctively choose the quickest route (shown in the figure) and save the child. Our purpose is to derive a relation between the angle $\theta$ you path makes with the perpendicular to the side of the pool when you're on land, and the angle $\theta$ your path makes with the perpendicular when you're in the water. To do this, let $x$ be the distance between the closest point to you at the side of the pool and the point where you enter the water. Write the total time you run (on land and in the water) in terms of $x$ (and also the constants $a, b, c, v, v_2$). Then set the derivative equal to zero. The result, called "Snell's law" or the "law of refraction," also governs the bending of light when it goes into water.

6.2 RELATED RATES

Suppose we have two variables $x$ and $y$ (in most problems the letters will be different, but for now let’s use $x$ and $y$) which are both changing with time. A "related rates" problem is a problem in which we know one of the rates of change at a given instant — say,

$$\frac{dx}{dt} = 7$$

— and we want to find the other rate $\frac{dy}{dt}$ at that instant. (The use of $x$ to mean $dx/dt$ goes back to Newton and is still used for this purpose, especially by physicists.) If $y$ is written in terms of $x$, i.e., $y = f(x)$, then this is easy to do using the chain rule:

$$\frac{dy}{dt} = \frac{dy}{dx} \cdot \frac{dx}{dt} \frac{dx}{dt}$$

That is, find the derivative of $f(x)$, plug in the value of $x$ at the instant in question, and multiply by the given value of $x = dx/dt$ to get $y = dy/dt$.

EXAMPLE 6.2.1 Suppose an object is moving along a path described by $y = x^2$, that is, it is moving on a parabolic path. At a particular time, say $t = 5$, the $x$ coordinate is 6 and we measure the speed at which the $x$ coordinate of the object is changing and find that $dx/dt = 3$. At the same time, how fast is the $y$ coordinate changing?

Using the chain rule, $dy/dt = 2x \cdot dx/dt$. At $t = 5$ we know that $x = 6$ and $dx/dt = 3$, so $dy/dt = 2 \cdot 6 \cdot 3 = 36$.

In many cases, particularly interesting ones, $x$ and $y$ will be related in some other way, for example $x = f(y)$, or $F(x, y) = k$, or perhaps $F(x, y) = G(x, y)$, where $F(x, y)$ and $G(x, y)$ are expressions involving both variables. In all cases, you can solve the related rates problem by taking the derivative of both sides, plugging in all the known values (namely, $x$, $y$, and $\dot{x}$), and then solving for $\dot{y}$.

To summarize, here are the steps in doing a related rates problem:

1. Decide what the two variables are.
2. Find an equation relating them.
3. Take $d/dt$ of both sides.
4. Plug in all known values at the instant in question.
5. Solve for the unknown rate.

EXAMPLE 6.2.2 A plane is flying directly away from you at 500 mph at an altitude of 3 miles. How fast is the plane’s distance from you increasing at the moment when the plane is flying over a point on the ground 4 miles from you?

To see what’s going on, we first draw a schematic representation of the situation, as in figure 6.2.1.

Because the plane is in level flight directly away from you, the rate at which $x$ changes is the speed of the plane, $dx/dt = 500$. The distance between you and the plane is $y$; it is $dy/dt$ that we wish to know. By the Pythagorean theorem we know that $x^2 + y^2 = 100$.

EXAMPLE 6.2.3 A swing consists of a board at the end of a 10 ft long rope. Think of the board as a point

Changing pool rescue.

To see what's going on, we first draw a schematic representation of the situation, as in figure 6.2.1.

Taking the derivative:

$$2\dot{x} = 2y \dot{y}$$

We are interested in the time at which $x = 4$; at this time we know that $4^2 + 9 = y^2$, so $y = 5$. Putting together all the information we get

$$2(4)(500) = 2(5))$$

Thus, $\dot{y} = 400$ mph.

EXAMPLE 6.2.4 Water is poured into a conical container at the rate of 10 cm$^3$/sec. How fast is its radius increasing when the radius is 4 cm? 

Here the variables are the radius $r$ and the volume $V$. We know $dV/dt$, and we want $dr/dt$. The two variables are related by means of the equation $V = \frac{1}{3} \pi r^2 h$. Taking the derivative of both sides gives $dV/dt = \frac{1}{3} \pi r^2 \dot{h}$. We now substitute the values we know at the instant in question: $7 = 4 \pi \frac{1}{3}$, so $\dot{h} = \frac{7}{16\pi}$ cm/sec.

EXAMPLE 6.2.5 A swing consists of a board at the end of a 10 ft long rope. Think of the board as a point

Changing pool rescue.
We have seen that sometimes there are apparently more than two variables that change with time, but in reality there are just two, as the others can be expressed in terms of just two. But sometimes there really are several variables that change with time; as long as you know the rates of change of all but one of them you can find the rate of change of the remaining one. As in the case when there are just two variables, take the derivative of both sides of the equation relating all of the variables, and then substitute all of the known values and solve for the unknown rate.

**EXAMPLE 6.2.6** A road running north to south crosses a road going east to west at the point $P$. Car A is driving north along the first road, and car B is driving east along the second road. At a particular time car A is 10 kilometers to the north of $P$ and traveling at 80 km/hr; while car B is 15 kilometers to the east of $P$ and traveling at 100 km/hr. How fast is the distance between the two cars changing?

Let $a(t)$ be the distance of car A north of $P$ at time $t$, and $b(t)$ the distance of car B east of $P$ at time $t$, and let $c(t)$ be the distance from car A to car B at time $t$. By the Pythagorean Theorem, $c(t)^2 = a(t)^2 + b(t)^2$. Taking derivatives we get $2c(t)\frac{dc}{dt} = 2a(t)\frac{da}{dt} + 2b(t)\frac{db}{dt}$,

so

\[
\frac{dc}{dt} = \frac{a(t)b(t)}{c(t)} \frac{db}{dt} - \frac{b(t)a(t)}{c(t)} \frac{da}{dt}.
\]

Substituting known values we get:

\[
\frac{dc}{dt} = \frac{10 \times 15}{\sqrt{(10)^2 + (15)^2}} (\frac{10}{3}) - \frac{15 \times 10}{\sqrt{(10)^2 + (15)^2}} (\frac{-8}{3}) = \frac{125}{17} \approx 7.35
\]

at the time of interest.

Notice how this problem differs from example 6.2.2. In both cases we started with the Pythagorean Theorem and took derivatives on both sides. However, in example 6.2.2 one of the sides was a constant (the altitude of the plane), and so the derivative of the square of that side of the triangle was simply zero. In this example, on the other hand, all three sides of the right triangle are variables, even though we are interested in a specific value of each side of the triangle (namely, when the sides have lengths 10 and 15). Make sure that you understand at the start of the problem what are the variables and what are the constants.

**Exercises 6.2.**

1. A cylindrical tank standing upright (with one circular base on the ground) has radius 20 cm. How fast does the water level in the tank drop when the water is being drained at 25 cm$^3$/sec? 

2. A cylindrical tank standing upright (with one circular base on the ground) has radius 1 meter. How fast does the water level in the tank drop when the water is being drained at 3 liters per second? 

3. A ladder 13 meters long rests on horizontal ground and leans against a vertical wall. The foot of the ladder is pulled away from the wall at the rate of 0.6 m/sec. How fast is the top sliding down the wall when the foot of the ladder is 5 m from the wall? 

4. A ladder 13 meters long rests on horizontal ground and leans against a vertical wall. The top of the ladder is being pulled up the wall at 0.1 meters per second. How fast is the foot of the ladder approaching the wall when the foot of the ladder is 5 m from the wall? 

5. A rotating beacon is located 2 miles out in the water. Let $A$ be the point on the shore that is closest to the beacon. As the beacon rotates at 10 rev/min, the beam of light sweeps down the shore once each time it revolves. Assume that the shore is straight. How fast is the point where the beam hits the shore moving at an instant when the beam is lighting up a point 2 miles along the shore from the point $A$? 

6. A baseball diamond is a square 90 ft on a side. A player runs from first base to second base at 15 ft/sec. At what rate is the player’s distance from third base decreasing when she is half way from first to second base? 

12. **Chapter 6 Applications of the Derivative**

13. A man 1.8 meters tall walks at the rate of 1 meter per second toward a streetlight that is 4 meters off the ground. At what rate is his shadow shortening? 

14. A woman 5 ft tall walks at the rate of 3.5 ft/sec away from a streetlight that is 12 ft above the ground. At what rate is the tip of her shadow moving? At what rate is her shadow lengthening? 

15. A man 1.8 meters tall walks at the rate of 1 meter per second toward a streetlight that is 4 meters above the ground. At what rate is the tip of his shadow moving? At what rate is his shadow shortening? 

16. A police helicopter is flying at 150 mph at a constant altitude of 0.5 mile above a straight road. The pilot uses radar to determine that an oncoming car is at a distance of exactly 1 mile from the helicopter, and that this distance is decreasing at 100 mph. Find the speed of the car. 

17. A police helicopter is flying at 200 kilometers per hour at a constant altitude of 1 km above a straight road. The pilot uses radar to determine that an oncoming car is at a distance of exactly 2 kilometers from the helicopter, and that this distance is decreasing at 250 kph. Find the speed of the car. 

18. A baseball is hit by a 20 m/s pitch in the air. A ball is falling 10 meters from the pole, casting a shadow on a building 10 meters away, as shown in figure 6.2.7. When the ball is 25 meters from the ground it is falling at 6 meters per second. How fast is its shadow moving? 

19. A woman of height 6 feet walks at a rate of 1.5 feet per second away from a streetlight of height 16 feet. If the woman has a constant speed of walking of 2 feet per second, how fast is the tip of her shadow moving? How fast is her shadow lengthening? 

20. A man 6 feet tall is walking at a rate of 2 feet per second away from a streetlight of height 10 feet. How fast is the woman’s shadow changing? How fast is her shadow lengthening? 

21. A man 6 feet tall is walking at a rate of 2 feet per second away from a streetlight of height 10 feet. How fast is the woman’s shadow changing? How fast is her shadow lengthening? 

22. A light shines from the top of a pole 20 m high. An object is dropped from the same height from a point 10 m away, so that its height at time $t$ seconds is $h(t) = 20 - 9.8t^2/2$. How fast is the object’s shadow moving on the ground one second later?
6.3 Newton's Method

Suppose you have a function \( f(x) \), and you want to find \( x \) as accurately as possible where it crosses the \( x \)-axis; in other words, you want to solve \( f(x) = 0 \). Suppose you know of no way to find an exact solution by any algebraic procedure, but you are able to use an approximation, provided it can be made quite close to the true value. Newton’s method is a way to find a solution to the equation to as many decimal places as you want. It is what is called an “iterative procedure,” meaning that it can be repeated again and again to get an answer of greater and greater accuracy. Iterative procedures like Newton’s method are well suited to programming for a computer. Newton’s method uses the fact that the tangent line to a curve is a good approximation to the curve near the point of tangency.

**EXAMPLE 6.3.1** Approximate \( \sqrt{3} \). Since \( \sqrt{3} \) is a solution to \( x^2 = 3 \) or \( x^2 - 3 = 0 \), we use \( f(x) = x^2 - 3 \). We start by guessing something reasonably close to the true value; this is usually easy to do; let’s use \( x = 3 \). Now use the tangent line to the curve when \( x = 2 \) as an approximation to the curve, as shown in figure 6.3.1. Since \( f'(x) = 2x \), the slope of this tangent line is 4 and its equation is \( y = 4x - 7 \). The tangent line is quite close to \( f(x) \), so it crosses the \( x \)-axis near the point at which \( f(x) \) crosses, that is, near \( \sqrt{3} \). It is easy to find where the tangent line crosses the \( x \)-axis: solve \( 0 = 4x - 7 \) to get \( x = 7/4 = 1.75 \). This is certainly a better approximation than 2, but let us say not close enough. We can improve it by doing the same thing again: find the tangent line at \( x = 1.75 \), find where this new tangent line crosses the \( x \)-axis, and use that value as a better approximation. We can continue this indefinitely, though it gets a bit tedious. Let’s see if we can shortcut the process. Suppose the host approximation to the intercept we have so far is \( x_0 \). To find a better approximation we will always do the same thing: find the slope of the tangent line at \( x_0 \), find the equation of the tangent line, find the \( x \)-intercept. The slope is \( f'(x_0) \). The tangent line is \( y = (2x_0)(x - x_0) + (x_0^2 - 3) \), the point-slope formula for a line. Finally, the intercept is found by solving \( 0 = (2x_0)(x - x_0) + (x_0^2 - 3) \). With a little algebra this turns into \( x = (x_0^2 + 3)/(2x_0) \), this is the next approximation, which we naturally call \( x_{n+1} \). Instead of doing the whole tangent line computation every time we can simply use this formula to get as many approximations as we want. Starting with \( x_0 = 2 \), we get \( x_1 = (4 + 3)/(2(2)) = 7/4 \) (the same approximation we got above, of course), \( x_2 = (7/4 + 3)/(2(7/4)) = (7/4 + 3)/(7/2) = 56/35 \approx 1.6180, x_3 \approx 1.6180 \), and so on. This is still a bit tedious by hand, but with a calculator or, even better, a good computer program, it is quite easy to get many, many approximations. We might guess already that 1.73205 is accurate to two decimal places, and in fact it turns out that it is accurate to 5 places.

Let’s think about this process in more general terms. We want to approximate a solution to \( f(x) = 0 \). We start with a rough guess, which we call \( x_0 \). We use the tangent line to \( f(x) \) to get a new approximation that we hope will be closer to the true value. What is the equation of the tangent line when \( x = x_0 \)? The slope is \( f'(x_0) \) and the line goes through \((x_0, f(x_0))\), so the equation of the line is \( y = f'(x_0)(x - x_0) + f(x_0) \).

Now we find where this crosses the \( x \)-axis by substituting \( y = 0 \) and solving for \( x \):

\[
x = \frac{x_0 f'(x_0) - f(x_0)}{f'(x_0)} = x_0 - \frac{f(x_0)}{f'(x_0)}
\]

We will typically want to compute more than one of these improved approximations, so we number them consecutively; from \( x_0 \) we have computed \( x_1 \):

\[
x_1 = \frac{x_0 f'(x_0) - f(x_0)}{f'(x_0)} = x_0 - \frac{f(x_0)}{f'(x_0)},
\]

and in general from \( x_n \) we compute \( x_{n+1} \):

\[
x_{n+1} = \frac{x_n f'(x_n) - f(x_n)}{f'(x_n)} = x_n - \frac{f(x_n)}{f'(x_n)}
\]

**EXAMPLE 6.3.2** Returning to the previous example, \( f(x) = x^2 - 3 \). \( f'(x) = 2x \), and the formula becomes \( x_{n+1} = x_n - (x_n^2 - 3)/(2x_n) \), as before.

In practice, it is to say, if you need to approximate a value in the course of designing a bridge or a building or an airplane, you must have some confidence that the approximation you settle on is accurate enough. As a rule of thumb, once a certain number of decimal places stop changing from one approximation to the next it is likely that those decimal places are correct. Still, this may not be enough assurance, in which case we can test the result for accuracy.

**EXAMPLE 6.3.3** Find the \( x \) coordinate of the intersection of the curves \( y = 2x \) and \( y = \tan x \), accurate to three decimal places. To put this in the context of Newton’s method,
6.4 Linear Approximations

Newton’s method is one example of the usefulness of the tangent line as an approximation to a curve. Here we explore another such application.

Recall that the tangent line to \( f(x) \) at a point \( x = a \) is given by \( L(x) = f'(a)(x-a) + f(a) \). The tangent line in this context is also called the linear approximation to \( f \) at \( a \).

If \( f \) is differentiable at \( a \), then \( L \) is a good approximation of \( f \) so long as \( a \) is “not too far” from \( a \). Put another way, if \( f \) is differentiable at \( a \) then under a microscope \( f \) will look very much like a straight line. Figure 6.4.1 shows a tangent line to \( y = x^2 \) at three different magnifications.

If we want to approximate \( f(b) \), because computing it exactly is difficult, we can approximate the value using a linear approximation, provided that we can compute the tangent line at some close to \( b \).

![Figure 6.4.1](image)

**EXAMPLE 6.4.1** Let \( f(x) = \sqrt{x} \). Then \( f'(x) = 1/2\sqrt{x} \). The linear approximation to \( f \) at \( x = 5 \) is \( L(x) = f'(5)(x-5) + f(5) \). As an immediate application we can approximate square roots of numbers near 9 by hand. To estimate \( \sqrt{10} \), we substitute 6 into the linear approximation instead of into \( f(x) \), so \( \sqrt{10} 

\approx \frac{1}{2}(\sqrt{9} + 1) \approx 3.17 \). This rounds to 3.17 while the square root of 10 is actually 3.16 to two decimal places, so this estimate is only accurate to one decimal place. This is not too surprising, as 10 is really not very close to 9; on the other hand, for many calculations, 3.2 would be accurate enough.

With modern calculators and computing software it may not appear necessary to use linear approximations. But in fact they are quite useful. In cases requiring an explicit numerical approximation, they allow us to get a quick rough estimate which can be used as a “reality check” on a more complex calculation. In some complex calculations involving numerical approximation, they allow us to get a quick rough estimate which can be used without serious loss of accuracy.

**EXAMPLE 6.4.2** Consider the trigonometric function \( \sin x \). Its linear approximation at \( x = 0 \) is simply \( L(x) = x \). When \( x \) is small this is quite a good approximation and is used frequently by engineers and scientists to simplify some calculations.

**DEFINITION 6.4.3** Let \( y = f(x) \) be a differentiable function. We define a new independent variable \( dx \), and a new dependent variable \( dy = f'(x)dx \). Notice that \( dy \) is a function both of \( x \) (since \( f'(x) \) is a function of \( x \)) and of \( dx \). We say that \( dx \) and \( dy \) are differentials.

We say \( \Delta y \) of the actual change in the function \( f \) between \( a \) and \( x \). Thus, \( dy \) can be used to approximate \( \Delta y \), the actual change in the function \( f \) between \( a \) and \( x \). This is exactly the approximation given by the tangent line:

\[
\Delta y = f'(a)dx 
\]

Thus, \( dy \) can be used to approximate \( \Delta y \), the actual change in the function \( f \) between \( a \) and \( x \).

While \( L(x) \) approximates \( f(x) \), \( dy \) approximates how \( f(x) \) has changed from \( f(a) \). Figure 6.4.2 illustrates the relationships.

![Figure 6.4.2](image)

**Exercises 6.4**

1. Let \( f(x) = x^2 \). At \( a = 1 \) and \( \Delta x = 1/2 \), what are \( \Delta y \) and \( dy \) ?

2. Let \( f(x) = \sqrt{x} \). At \( a = 1 \) and \( \Delta x = 1/10 \), what are \( \Delta y \) and \( dy \) ?

3. Let \( f(x) = \sin(2x) \). At \( a = \pi/4 \) and \( \Delta x = \pi/10 \), what are \( \Delta y \) and \( dy \) ?

4. Use differentials to estimate the amount of paint needed to apply a coat of paint 0.02 cm thick to a sphere with diameter 40 meters. (Recall that the volume of a sphere of radius \( r \) is \( V = (4/3)\pi r^3 \).) Notice that you are given that \( dr = 0.02 \).

5. Show in detail that the linear approximation of \( \sin x \) at \( x = 0 \) is \( L(x) = x \) and the linear approximation of \( \cos x \) at \( x = 0 \) is \( L(x) = 1 \).

6.5 The Mean Value Theorem

Here are two interesting questions involving derivatives:

1. Suppose two different functions have the same derivative; what can you say about the relationship between the two functions?

2. Suppose you drive a car from toll booth a on a toll road to another toll booth at an average speed of 70 miles per hour. What can be concluded about your average speed during the trip? In particular, did you exceed the 65 mile per hour speed limit?

While those sound very different, it turns out that the two problems are very closely related. We know that “speed” is really the derivative by a different name; let’s start by translating the second question into something that may be easier to visualize. Suppose that the function \( f(t) \) gives the position of your car at the toll road at time \( t \). Your change in position between one toll booth and the next is given by \( f(t_1) - f(t_0) \), assuming that at time \( t_0 \) you were at the first booth and at time \( t_1 \) you arrived at the second booth. Your average speed for the trip is \( \frac{f(t_1) - f(t_0)}{t_1 - t_0} \).

If we think about the graph of \( f(t) \), the average speed is the slope of the line that connects the two points \( (t_0, f(t_0)) \) and \( (t_1, f(t_1)) \). Your speed at any particular time \( t \) between \( t_0 \) and \( t_1 \) is \( f'(t) \), the slope of the curve. Now question (2) becomes a question about slope. In particular, if the slope between endpoints is 70, what can be said of the slopes at points between the endpoints?

As a general rule, when faced with a new problem it is often a good idea to examine one or more simplified versions of the problem, in the hope that this will lead to an understanding of the original problem. In this case, the problem in its “slope” form is somewhat easier to simplify than the original, but equivalent, problem.

Here is a special instance of the problem. Suppose that \( f(t_0) = f(t_1) \). Then the two endpoints have the same height and the slope of the line connecting the endpoints is zero. What can we say about the slope between the endpoints? It shouldn’t take much experimentation before you are convinced of the truth of this statement: Somewhere between \( t_0 \) and \( t_1 \) the slope is exactly zero, that is, somewhere between \( t_0 \) and \( t_1 \) the slope is equal to the slope of the line between the endpoints. This suggests that perhaps the same is true even if the endpoints are at different heights, and again a bit of experimentation will probably convince you that this is so. But we can do better than “experimentation” — we can prove that this is so.

We start with the simplified version:

**THEOREM 6.5.1 Rolle’s Theorem** Suppose that \( f(x) \) has a derivative on the interval \( [a, b] \), is continuous on the interval \( [a, b] \), and \( f(a) = f(b) \). Then at some value \( c \) in \( (a, b) \), \( f'(c) = 0 \).

**Proof.** We know that \( f(x) \) has a maximum and minimum value on \( [a, b] \) (because it is continuous), and we also know that the maximum and minimum must occur at an endpoint at a point at which the derivative is zero, or at a point where the derivative is undefined. Since the derivative is never undefined, that possibility is removed.

If the maximum or minimum occurs at a point \( c \) other than an endpoint, where \( f'(c) = 0 \), then we have found the point we seek. Otherwise, the maximum and minimum both occur at an endpoint, and since the endpoints have the same height, the maximum and minimum are the same. This means that \( f(x) = f(a) = f(b) \) at every \( x \in [a, b] \), so the function is a horizontal line, and it has derivative zero everywhere in \( (a, b) \). Then we may choose any \( c \) at all to get \( f'(c) = 0 \).

Perhaps remarkably, this special case is all we need to prove the more general one as well.

**THEOREM 6.5.2 Mean Value Theorem** Suppose that \( f(x) \) has a derivative on the interval \( (a, b) \) and is continuous on the interval \( [a, b] \). Then at some value \( c \) in \( (a, b) \),

\[
f'(c) = \frac{f(b) - f(a)}{b - a}
\]

**Proof.** Let \( m = \frac{f(b) - f(a)}{b - a} \) and consider a new function \( g(x) = f(x) - m(x-a) - f(a) \). We know that \( g(x) \) has a derivative everywhere, since \( g'(x) = f'(x) - m \). We can compute \( g(a) = f(a) - m(a-a) - f(a) = 0 \) and

\[
g(b) - g(a) = m(b-a) - f(a) - f(b) = f(b) - f(a) = \frac{f(b) - f(a)}{b - a} - f(a)
\]

\[
= f(b) - f(a) - f(a) = 0.
\]
So the height of \( g(x) \) is the same at both endpoints. This means, by Rolle’s Theorem, that at some \( c \), \( g'(c) = 0 \). But we know that \( g'(c) = f'(c) - m \), so

\[
0 = f'(c) - m = f'(c) - \frac{f(b) - f(a)}{b - a}.
\]

which turns into

\[
f'(c) = \frac{f(b) - f(a)}{b - a},
\]

exactly what we want.

Returning to the original formulation of question (2), we see that if \( f(t) \) gives the position of your car at time \( t \), then the Mean Value Theorem says that at some time \( c \), \( f'(c) = \frac{f(b) - f(a)}{b - a} \), that is, at some time you must have been traveling at exactly your average speed for the trip, and that indeed you exceeded the speed limit.

Now let’s return to question (1). Suppose, for example, that two functions are known to have derivative equal to 5 everywhere, \( f'(x) = g'(x) = 5 \). It is easy to find such functions: \( 5x, \ 5x + 47, \ 5a - 132, \) etc. Are there other, more complicated, examples? No—the only functions that work are the “obvious” ones, namely, \( 5x \) plus some constant. How can we see that this is true?

Although “5” is a very simple derivative, let’s look at an even simpler one. Suppose that \( f'(x) = g'(x) = 0 \). Again we can find examples: \( f(x) = 0, \ f(x) = 47, \ f(x) = -511 \) all have \( f'(x) = 0 \). Are there non-constant functions \( f \) with derivative 0? No, and here’s why: Suppose that \( f(c) \) is not a constant function. This means that there are two points on the function with different heights, say \( f(a) \neq f(b) \). The Mean Value Theorem tells us that at some point \( c \), \( f'(c) = \frac{(f(b) - f(a))/(b - a)}{} \) \( \neq 0 \). So any non-constant function does not have a derivative that is zero everywhere; this is the same as saying that the only functions with zero derivative are the constant functions.

Let’s go back to the slightly less easy example: suppose that \( f'(x) = g'(x) = 5 \). Then

\[
(f(x) - g(x))' = f'(x) - g'(x) = 5 - 5 = 0.
\]

So what we discovered in the previous paragraph, we know that \( f(x) - g(x) = k \), for some constant \( k \). So any two functions with derivative 5 must differ by a constant; since \( 5x \) is known to work, the only other examples must look like \( 5x + k \).

Now we can extend this to more complicated functions, without any extra work. Suppose that \( f'(x) = g'(x) \). Then as before \( (f(x) - g(x))' = f'(x) - g'(x) = 0 \), so \( f(x) - g(x) = k \). Again this means that if we find just a single function \( g(x) \) with a certain derivative, then every other function with the same derivative must be of the form \( g(x) + k \).

EXAMPLE 6.5.3 Describe all functions that have derivative \( 5x - 3 \). It’s easy to find one: \( g(x) = (5/2)x^2 - 3x \) has \( g'(x) = 5x - 3 \). The only other functions with the same derivative are therefore of the form \( f(x) = (5/2)x^2 - 3x + k \).

EXAMPLE 6.5.5 Describe all functions with derivative \( \sin x \). One such function is \( -\cos x \); so all such functions have the form \( -\cos x + k \).

### Exercises 6.5.

1. Let \( f(x) = x^2 \). Find a value \( c \in (-1,2) \) so that \( f'(c) \) equals the slope between the endpoints of \( f(x) \) on \((-1,2) \).

2. Verify that \( f(x) = x/(x + 2) \) satisfies the hypotheses of the Mean Value Theorem on the interval \([-1,2) \) and then find all of the values, \( c \), that satisfy the conclusion of the theorem.

3. Verify that \( f(x) = 3x/(x + 7) \) satisfies the hypotheses of the Mean Value Theorem on the interval \([-2,6) \) and then find all of the values, \( c \), that satisfy the conclusion of the theorem.

4. Let \( f(x) = \tan x \). Show that \( f((\pi/2)) = 0 \) but there is no number \( c \in (\pi/2, \pi) \) such that \( f'(c) = 0 \). Why does this not contradict Rolle’s theorem?

5. Let \( f(x) = (x - 3)^2 \). Show that there is no value \( c \in (1,4) \) such that \( f'(c) = (f(4) - f(1))/(4-1) \). Why is this not a contradiction of the Mean Value Theorem?

6. Describe all functions with derivative \( x^2 + 47x - 5 \).

7. Describe all functions with derivative \( \sqrt{x}x \).

8. Show that the equation \( x^4 - 7x + 1 = 0 \) does not have more than two distinct real roots.

9. Let \( f \) be differentiable on \( R \). Suppose that \( f(x) \neq 0 \) for every \( x \). Prove that \( f \) has at most one real root.

10. Prove that for all real \( x \) and \( y \), \( \cos x - \cos y \leq |x - y| \). State and prove an analogous result involving sine.

11. Show that \( \sqrt{1 + x^2} \leq 1 + (x/2) \) if \( -1 < x < 0 \).