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 maximum and minimum 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 global maximum or minimum, sometimes also called an 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 can 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 \( a \) and \( b \). We have not previously considered such points because we have not been interested in limiting a function to a small interval. An example should make this clear.

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

Since \( f'(1) = 2 \) would we not normally flag \( x = 1 \) as a point of interest, but it is clear from the graph that when \( f(x) \) is restricted to \([-2, 2]\) 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 \([-2, 2]\). If we try to compute the derivative of this new function we actually find that it does not have a derivative at \(-2 \) or \( 2 \). 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 undefined. It is simpler, however, simply to remember that we must always check the endpoints.

So the function \( g \), that is, \( f \) restricted to \([-2, 2]\), 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 A rectangular field is to be enclosed by a fence. The adjacent sides must be 100\( \times \)100 meters, and the function we want is the area of the field.

First we must translate this into a purely mathematical problem in which we want to maximize a quantity. In this case the quantity is \( A = l \cdot w \), where \( l \) and \( w \) are the dimensions of the field.

Figure 6.1.2

\[ f(x) = x^3 - x \]
to minimize is

$$f(x) = 2x + \frac{2100}{x}$$

since the perimeter is twice the length plus twice the width of the rectangle. Not all values of \( x \) make sense in this problem: lengths of sides of rectangles must be positive, so \( x > 0 \).

If \( x > 0 \) then so is \( 100/x \), so we need no second condition on \( x \).

We next find \( f'(x) \) and set it equal to zero: \( 0 = f'(x) = 2 - 2000/x^2 \). Solving \( f'(x) = 0 \) for \( x \) gives us \( x = 10 \). We are interested only in \( x > 0 \), so only the value \( x = 10 \) is of interest. Since \( f''(x) \) is defined everywhere on the interval \((0, \infty)\), there are no more critical values, and there are no endpoints. Is there a local maximum, minimum, or neither at \( x = 10 \)? The second derivative is \( f''(x) = 400x/x^3 \), and \( f''(10) > 0 \), so there is a local minimum. Since there is only one critical value, this is also the global minimum, so the \( x \) values, and there are no endpoints. Is there a local maximum, minimum, or neither at \( f(4) \)?

EXAMPLE 6.1.8 You want to sell a certain number \( n \) of items in order to maximize your profit. Market research tells you that if you set the price at \$1.50, you will be able to sell 5000 items, and for every 10 cents you lower the price below \$1.50, you will be able to sell another 1000 items. Suppose that your fixed costs ("start-up costs") total \$2000, and the per item cost of production ("marginal cost") is \$0.50. Find the price to set per item and the number of items sold in order to maximize profit, and also determine the maximum profit you can get.

The first step is to convert the problem into a function maximization problem. Since we want to maximize profit by setting the price per item, we should look for a function \( P(x) \) representing the profit when the price per item is \( x \). Profit is revenue minus costs, and revenue is number of items sold times the price per item, so we get \( P = ax - 2000 - 0.5bx \). The number of items sold is itself a function of \( x \), \( n = 5000 + 1000(1.5 - x)/0.10 \), because \( (1.5 - x)/0.10 \) is the number of multiples of 10 cents that the price is below \$1.50. Now we substitute for \( n \) in the profit function:

$$P(x) = (5000 + 1000(1.5 – x)/0.10)(x - 2000 - 0.5(5000 + 1000(1.5 – x)/0.10)) = -10000x^2 + 25000x - 12000$$

We want to know the maximum value of this function when \( x \) is between 0 and 1.5. The derivative is \( P'(x) = -20000x + 25000 \), which is zero when \( x = 1.25 \). Since \( P''(x) = -20000 < 0 \), there must be a local maximum at \( x = 1.25 \), and since this is the only critical value it must be a global maximum as well. (Alternatively, we could compute \( P(0) = -12000, P(1.25) = 3625, \) and \( P(1.5) = 3000 \) and note that \( P(1.25) \) is the maximum of these.) Thus the maximum profit is \$3625, attained when we set the price at \$1.25 and sell 7500 items.

EXAMPLE 6.1.9 Find the largest rectangle (that is, the rectangle with largest area) that fits inside the graph of the parabola \( y = x^2 \) below the line \( y = a \) (a is an unspecified constant value), with the top side of the rectangle on the horizontal line \( y = a \); see figure 6.1.3.)

We want to find the maximum value of some function \( A(x) \) representing area. Perhaps the hardest part of this problem is deciding what \( x \) should represent. The lower right corner of the rectangle is at \((x, x^2)\), and once this is chosen the rectangle is completely determined. So we can let the \( x \) in \( A(x) \) be the \( x \) of the parabola \( f(x) = x^2 \). Then the area is \( A(x) = (2x)(a - x^2) - 2x^2 + 2ax \). We want the maximum value of \( A(x) \) when \( x \) is in \([0, \sqrt{a}]\). (You might object to allowing \( x = 0 \) or \( x = a \), since then the “rectangle” has either no width or no height, so is not “really” a rectangle. But the problem is somewhat easier if we simply allow such rectangles, which have zero area.)

Setting \( 0 = A'(x) = -4x^2 + 2ax \) we get \( x = \sqrt{a/2} \) as the only critical value. Testing this and the two endpoints, we have \( A(0) = A(\sqrt{a/2}) = 0 \) and \( A(\sqrt{a})/2) = (4/9)\sqrt{a^3/2} \). The maximum area thus occurs when the rectangle has dimensions \( 2\sqrt{a/3} \times (2/3)a \).

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:

$$c(2\pi rh) + N(2\pi r^2)$$

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

$$h = \frac{2rV}{\pi r^2} + \frac{2N\pi r^2}{V} = \frac{2V}{\pi r} + \frac{2N\pi r^2}{V}$$

We want to know the minimum value of this function when \( r \) is in \((0, \infty)\). We now set \( 0 = f'(r) = -2V/r^2 + 4N\pi c r \), giving \( r = \sqrt{V/(2N\pi)} \). Since \( f''(r) = 4V/r^3 + 4N\pi c \) 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 \( h = V/(\pi r^2) \),

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

so the minimum cost occurs when the height \( h \) is \( 2N \) 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 $b$ is the distance from $A$ to the closest point $C$ on the road. Let $v$ be your speed on the road, and let $w$, which is less than $v$, be your speed on the sand. Right now you are at the point $D$, which is a 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 $x$ 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 $770$ at speed $v$, and then the distance $b$ at speed $w$. Since $770 = a - x$ and, by the Pythagorean theorem, $770 = \sqrt{a^2 + b^2}$, the total time for the trip is

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

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

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

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 $w\sqrt{a^2 + b^2}$ is either in this interval or not. If it is, we can use the second derivative to test it:

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

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

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

and

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

So the upshot is this: If you start farther away from $C$ than $v\sqrt{a^2 + b^2}/w$, then you always want to cut across the sand when you are a distance $w\sqrt{a^2 + b^2}/v$ from point $C$. If you start closer than this to $C$, you should cut directly across the sand.

Summary—Steps to solve an optimization problem.

1. Decide what the variables are and what the constants are, draw a diagram if appropriate, understand clearly what it is that is to be maximized or minimized.
2. Write a formula for the function for which you wish to find the maximum or minimum.
3. Express that formula in terms of only one variable, that is, in the form $f(x)$.
4. Set $f'(x) = 0$ and solve. Check all critical values and endpoints to determine the extreme value.

Exercises 6.1.

1. Let $f(x) = \frac{1 + x - x^2}{2}$. If $x < 0$, $f(x)$ is a straight line which is always below $y = 1/2$. As $x$ approaches 1 from the left $f(x)$ approaches $\infty$. As $x$ approaches 2 from the left $f(x)$ approaches $-\infty$. If $x > 2$, $f(x)$ is a straight line which is always above $y = 1/2$. At $x = 1$, $f(x)$ has a vertical tangent line and at $x = 2$, $f(x)$ has a horizontal tangent line.

2. Find the dimensions of the rectangle of largest area having fixed perimeter 100.

3. Find the dimensions of the rectangle of largest area having fixed perimeter $P$.

4. A box with square base and no top is to hold a volume 100. Find the dimensions of the box that requires the least material for the five sides. Also find the ratio of height to side of the base.

5. A box with square base is to hold a volume 200. The bottom and top are formed by folding in flaps from all four sides, so that the bottom and top consist of two layers of cardboard. Find the dimensions of the box that requires the least material. Also find the ratio of height to side of the base. (This ratio will not involve $V$.)

6. You have 100 feet 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?

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 any. However, if you set the price of an item at $1$ and sell 500 items for each dollar below $10$ that you set the price. Suppose your fixed costs total $3000$ and your marginal cost is $82$ per item. What is the most profit you can make?

10. Find the area of the largest rectangle that fits inside a semicircle of radius 10 (one side of the rectangle is along the diameter of the semicircle). The area of the rectangle is $2ab$, so, by the Pythagorean theorem, $a^2 + b^2 = 100$. Since $a$ and $b$ are positive, the maximum area is $50\sqrt{2}$ when $a = b$. The area of the rectangle is $50\sqrt{2}$.

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 $2(2r)^2 = 8r^2$ 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 $2(2r)^2 = 8r^2$ 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 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 $A$ and $x$.)

17. In example 6.1.12, what happens if $w \geq v$ (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.

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?

126 Chapter 6 Applications of the Derivative

but it is difficult to determine which of these is smaller by direct comparison. If, as is likely in practice, we know the values of $v$, $w$, and $b$, then it is easy to determine this.

With a little cleverness, however, we can determine the minimum in general. We have seen that $f'(x)$ is always positive, so the derivative $f''(x)$ is always increasing. We know that if $w\sqrt{a^2 + b^2}$ the derivative is zero, so for values of $a$ less than that critical value, the derivative is negative. This means that if $0 < f(a)$, so the minimum occurs when $x = a$.

So the upshot is this: If you start farther away from $C$ than $w\sqrt{a^2 + b^2}/v$, then you always want to cut across the sand when you are a distance $w\sqrt{a^2 + b^2}/v$ from point $C$. If you start closer than this to $C$, you should cut directly across the sand.

Figure 6.1.6 Cutting a beam.
30. 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 placed so that the pull away from $A$ towards $B$ is minimal. Here assume that the force from each charge is proportional to the square of the distance 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 30 affected if the cost per item for the $x$ items, instead of being simply $2x$, decreases below $2x$ 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 $v_r$ 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 surface of 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 calculus, you instinctively choose the quickest route (shown in the figure) and save the child. Our purpose is to derive a relation between the angles $\theta$ your 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 $a, b, c, v, v_r$). 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. 

![Figure 6.2.1 Wading pool rescue.](image_url)

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, $dx/dt$—and we want to find the other rate $dy/dt$ 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:

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

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 $\frac{dx}{dt}$ to get $\frac{dy}{dt}$.

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 $\frac{dx}{dt} = 3$. At the same time, how fast is the $y$ coordinate changing? 

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

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 $\frac{dx}{dt}$), and then solving for $\frac{dy}{dt}$.

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 $\frac{dy}{dt}$ 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, $\frac{dx}{dt} = 500$. The distance between you and the plane is $y$; it is $\frac{dy}{dt}$ that we wish to know. By the Pythagorean Theorem we know that $x^2 + y^2 = 50^2$.

Taking the derivative:

$$2x \frac{dx}{dt} + 2y \frac{dy}{dt} = 0$$

Thus, $\frac{dy}{dt} = -\frac{x^2}{y} \frac{dx}{dt}$.

EXAMPLE 6.2.3 You are sailing a spherical balloon at the rate of 7 cm/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 the equation $V = 4\pi r^3/3$. Taking the derivative of both sides gives $dV/dt = 4\pi r^2 \frac{dr}{dt}$. We now substitute the values we know at the instant in question: $V = 4\pi (4)^3/3 = 7$$\frac{4\pi}{3}(16) \text{ cm/sec}$. 

EXAMPLE 6.2.4 Water is poured into a conical container at the rate of 10 cm$^3$/sec. The cone points directly down, and it has a height of 30 cm and a base radius of 10 cm; see figure 6.2.2. 

How fast is the water level rising when the water is 4 cm deep (at its deepest point)?

The water forms a conical shape within the big cone; its height and base radius and volume are all increasing as water is poured into the container. This means that we actually have three things varying with time: the water level $h$ (the height of the cone of water), the radius $r$ of the circular top of water (the base radius of the cone of water), and the volume of water $V$. The volume of a cone is given by $V = \pi r^2 h/3$. We know $dV/dt$, and we want $dh/dt$. At first something seems to be wrong: we have a third variable $r$ whose rate we don’t know.

But the dimensions of the cone of water must have the same proportions as those of the container. That is, because of similar triangles, $r/h = 10/30$ so $r = h/3$. Now we can eliminate $r$ from the problem entirely: $V = \pi (h^3)/3 h^3 = \pi h^3/27$. We take the derivative of both sides and plug in $h = 4$ and $dV/dt = 10$, obtaining $10 = (3\pi x^2/27)(dh/dt)$. Thus, $dh/dt = 90/(16\pi) \text{ cm/sec}$. 

130 Chapter 6 Applications of the Derivative

![Figure 6.2.2 Conical water tank.](image_url)

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 $P$ at the end of the rope, and let $Q$ be the point of attachment at the other end. Suppose that the swing is directly below $Q$ at time $t = 0$, and is being pushed by someone who walks at 6 ft/sec from left to right. Find (a) how fast the swing is rising after 1 sec; (b) the angular speed of the rope in deg/sec after 1 sec.

We start out by asking: What is the geometric quantity whose rate of change we know, and what is the geometric quantity whose rate of change we’re being asked about? Note that the person pushing the swing is moving horizontally at a rate we know. In other words, the horizontal coordinate of $P$ is increasing at 6 ft/sec. In the xy-plane let us make the convenient choice of putting the origin at the location of $P$ at time $t = 0$, i.e., a distance 10 directly below the point of attachment. Then the rate we know is $\frac{dx}{dt}$, and in part (a) the rate we want is $\frac{dy}{dt}$ (the rate at which $P$ is rising). In part (b) the rate we want is $\frac{d\theta}{dt}$, where $\theta$ stands for the angle in radians through which the swing has swung from the vertical. (Actually, since we want our answer in deg/sec, at the end we must convert $\frac{d\theta}{dt}$ from rad/sec by multiplying by $180/\pi$.)

(a) From the diagram we see that we have a right triangle whose legs are $x = 10$ and $y$, and whose hypotenuse is 10. Hence $x^2 + (10 - y)^2 = 100$. Taking the derivative of both sides we obtain: $2x \cdot 10 - 2(10 - y) \frac{dy}{dt} = 0$. We now look at what we know after 1 second, namely $x = 6$ (because $x$ started at 0 and has been increasing at the rate of 6 ft/sec for 1 sec), $y = 2$ (because we get $10 - y = 8$ from the Pythagorean theorem applied to the triangle with hypotenuse 10 and leg 6), and $\theta = 6$. Putting these values gives us $2 \cdot 6 - 2 \cdot 2 \cdot \frac{dy}{dt} = 0$, from which we can easily solve for $\frac{dy}{dt} = 4.5$ ft/sec.

(b) Here our two variables are $x$ and $\theta$, so we want to use the same right triangle as in part (a), but this time relate $\theta$ to $x$. Since the hypotenuse is constant (equal to 10), the best way to do this is to use the sine: $\sin \theta = x/10$. Taking derivatives we obtain
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 108 km/hr. How fast is the distance between the two cars changing?

Figure 6.2.4 Cars moving apart.

\[
\text{Distance} = \sqrt{10^2 + 15^2} = 18.75 \text{ km}
\]

\[
\frac{\text{dr}}{\text{dt}} = \frac{10 \times 80 + 15 \times 108}{18.75} = 123.6 \text{ km/hr}
\]

Do example 6.2.6 assuming that 8 seconds ago car A started from rest at 55 km/hr. At what time is the tip of his shadow moving? At what rate is his shadow lengthening?

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 5.0 miles 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 180 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 light shines from the top of a pole 20 m high. A ball is falling 10 meters from the pole, casting a shadow on a building 30 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 it shadow moving?

Figure 6.2.7 Falling ball.

\[
\text{Distance} = \sqrt{20^2 + 30^2} = 36.5 \text{ m}
\]

\[
\frac{\text{dr}}{\text{dt}} = \frac{10 \times 80 + 15 \times 108}{18.75} = 123.6 \text{ km/hr}
\]

\[
\frac{\text{dr}}{\text{dt}} = \frac{18.75 \times 123.6}{10^2 + 15^2} = 23.5 \text{ km/hr}
\]

19. Do example 6.2.6 assuming that the angle between the two roads is 120° instead of 90° (that is, the “north–south” road actually goes in a somewhat northerly direction from \( P \)). Recall the law of cosines: \( c^2 = a^2 + b^2 - 2ab \cos \theta \).

20. Do example 6.2.6 assuming that car A is 300 meters north of \( P \), car B is 400 meters east of \( P \), both cars are going at constant speed toward \( P \), and the two cars will collide in 10 seconds.

21. Do example 6.2.6 assuming that 8 seconds ago car A started from rest at \( P \) and has been picking up speed at the steady rate of 5 m/sec², and 6 seconds after car A started car B passed \( P \) moving east at constant speed 60 m/sec.

22. Refering again to example 6.2.6, suppose that instead of car B an airplane is flying at speed 200 km/hr to the east of \( P \) at an altitude of 2 km, as depicted in figure 6.2.8. How fast is the distance between car and airplane changing?

23. Refering again to example 6.2.6, suppose that instead of car B an airplane is flying at speed 200 km/hr to the east of \( P \) at an altitude of 2 km, that it is gaining altitude at 10 km/hr. How fast is the distance between car and airplane changing?

24. A right triangle is inscribed in a circle of diameter 10 m. Express the area of the triangle as a function of the angle \( \theta \) between the hypotenuse and the diameter. Find the rate of change of the area with respect to \( \theta \). If the area is 6 square meters, find the measure of angle \( \theta \) and the rate of change of \( \theta \) with respect to time.
6.3 Newton's Method

Suppose you have a function \( f(x) \), and you want to find 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 could 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 \( x_0 = 2 \). Now take 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 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 \( x_1 \). To find a better approximation we will always do the same thing: find the slope of the tangent line at \( x_1 \), find the equation of the tangent line, find the x-intercept. The slope is \( 2x_1 \). The tangent line is \( y = (2x_1)(x - x_1) + (x_1^2 - 3) \), using the point-slope formula for a line. Finally, the intercept is found by solving \( 0 = (2x_1)(x - x_1) + (x_1^2 - 3) \). With a little algebra this turns into \( x = (x_1^2 + 3)/(2x_1) \), this is the next approximation, which we naturally call \( x_{1+1} \). 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 \( x_1 = 2 \), we get \( x_2 = (x_1^2 + 3)/(2x_1) = (4^2 + 3)/4 = 2.75 \) (the same approximation we got above, of course).\n
\[ x_3 = (x_2^2 + 3)/(2x_2) = (7.875 + 3)/5.5 = 1.73214, \] 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 \( y = f(x_0)(x - x_0) + f(x_0) \).

we note that we want to know where \( 2x = \tan x \) or \( f(x) = \tan x - 2x = 0 \). We compute \( f'(x) = \sec^2 x - 2 \) and set up the formula:

\[ x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)} \]

From the graph in figure 6.3.2 we guess \( x_0 = 1 \) as a starting point, then using the formula we compute:

\[ x_1 = 1.311047803, x_2 = 1.229292096, x_3 = 1.16765900, x_4 = 1.1652658, x_5 = 1.165561636. \] So we guess that the first three places are correct, but that is not the same as saying 1.165 is correct to three decimal places — 1.165 might be the correct, rounded approximation. How can we tell? We can substitute 1.165, 1.1655 and 1.166 into \( tan x - 2x \) and see which one gives a difference that is correct. Since the first two are negative and the third is positive, \( tan x - 2x \) crosses the x-axis between 1.165 and 1.166, so the correct value to three places is 1.166.

**Exercises 6.3.**

You may want to use this Sage worksheet.

1. Approximate the fifth root of 7, using \( x_0 = 1.5 \) as a first guess. Use Newton’s method to find \( x_{10} \) as your approximation.

2. Use Newton’s Method to approximate the cube root of 10 to 4 decimal places.

3. The function \( f(x) = x^3 - 3x - 2x = 0 \) has a root between 3 and 4, because \( f(3) = -3 \) and \( f(4) = 10 \). Approximate the root to two decimal places.

4. A rectangular piece of cardboard of dimensions 8 \( \times \) 10 is used to make an open-top box by cutting out a small square of side \( x \) from each corner and bending up the sides. (See exercise 20 in 6.1.) If \( x = 2 \), then the volume of the box is 24 \( \times \) 13 = 104. Use Newton's method to find a value of \( x \) for which the box has volume 100, accurate to 3 significant figures.
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 \( L(x) = f(a) + f'(a)(x-a) \). The tangent line in this context is also called the linear approximation to \( f(x) \) at \( a \).

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] The linear approximation to \( y = x^2 \).

**Example 6.4.1**

Let \( f(x) = \sqrt{1+x} \). Then \( f'(x) = \frac{1}{2\sqrt{1+x}} \). The linear approximation to \( f \) at \( x = 5 \) is \( L(x) = 1/(2\sqrt{6}) + (x-5)/6 - 3 \). 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 \( f(x) \), since \( \sqrt{10} \) is not too far from 9. We have \( \sqrt{10} \approx 3.17 \) and 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 computer 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 as a “reality check” on a more complex calculation.

**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 differen-

tials.

\[ \Delta x = x - a \] and \( \Delta y = f(x) - f(a) \). If \( x \) is near \( a \) then \( \Delta x \) is small. If we set \( dx = \Delta x \) then \( dy = f'(a)dx = \frac{dy}{dx} \Delta x = \Delta y \).

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:

\[ dy = f'(a)(x - a) - f(a) = f(x) - f(a) = L(x) - f(a). \]

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] Differentials.

**Exercises 6.4**

1. Let \( f(x) = x^3 \). If \( a = 1 \) and \( dx = \Delta x = 1/2 \), what are \( \Delta y \) and \( dy \)?
2. Let \( f(x) = \sqrt{x} \). If \( a = 1 \) and \( dx = \Delta x = 1/10 \), what are \( \Delta y \) and \( dy \)?
3. Let \( f(x) = \sin(x) \). If \( a = \pi \) and \( dx = \Delta x = \pi/100 \), 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 \( V = (4/3)\pi r^3 \).
5. Show in detail that the linear approximation of \( \sin(x) \) at \( x = 0 \) is \( E(x) = x \) and the linear approximation of \( \cos(x) \) at \( x = 0 \) is \( E(x) = 1 \).

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 to a toll road to another toll booth at an average speed of 70 miles per hour. What can be concluded about your actual 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 if the function \( f(t) \) gives the position of your car at time \( t \). Your change in position between one toll booth and the next is given by \( f(t_2) - f(t_1) \), assuming that at time \( t_1 \) you were at the first booth and at time \( t_2 \) you arrived at the second booth. Your average speed for the trip is \( (f(t_2) - f(t_1))/(t_2 - t_1) \). If we think about the graph of \( f(t) \), the average speed is the slope of the line that connects the two points \( (t_1, f(t_1)) \) and \( (t_2, f(t_2)) \). Your speed at any particular time \( t \) between \( t_1 \) and \( t_2 \) 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 \( f(t_1) = f(t_2) \). 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 \( t_1 \) and \( t_2 \) the slope is exactly zero, that is, somewhere between \( t_1 \) and \( t_2 \) 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 \([a, b] \) and is continuous on the interval \([a, b] \), 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 at which 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

\[ g'(b) = f(b) - m(b-a) - f(a) = f(b) - \frac{f(b) - f(a)}{b - a}(b-a) - f(a) = f(b) - (f(b) - f(a)) - f(a) = 0. \]
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 \( t \), \( f'(t) = 70 \), 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, 5x − 32, 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 \( f(x) \) 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 \( c \), \( f'(c) = (f(b) - f(a))/(b - a) \neq 0 \). 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 \( f(x) = f(a) + 5(x - a) \). Again we know that \( f(b) - g(a) = k \), 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 \( f(x) = (5/2)x^2 - 3x + k \).

Exercises 6.5.

1. Let \( f(x) = x^3 \). 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 \([1, 4] \) 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 \( f(x) = f(2\pi) = 0 \) 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 \( x^2 + 4x - 5 \).

7. Describe all functions with derivative \( x^3 - 4x^2 + 1 \).

8. Describe all functions with derivative \( x^3 - \frac{1}{2}x^2 \).

9. Describe all functions with derivative \( \sin(2x) \).

10. Show that the equation \( 6x^4 - 7x + 1 = 0 \) does not have more than two distinct real roots.

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

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

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