Chapter 8
Introducing the chain rule

“Don’t be a fool, you gotta be cool; learn to use that chain rule!”
(modified from) Bard Ermentrout, Pittsburgh, 1990’s

So far, examples were purposefully chosen to focus on power functions, polynomial, and rational functions that are relatively easy to differentiate. We now introduce the differentiation rule that opens up our repertoire to more elaborate examples that involve composite functions. We dedicate this chapter to the chain rule and its applications. Our first steps are to learn and understand what are composite functions, and how the chain rule is applied to their differentiation. After gaining confidence, we apply the new method to practical examples.

8.1 The chain rule

Section 8.1 Learning goals

1. Understand the concept of function composition and be able to express a composite function in terms of the underlying composed functions.

2. Understand the chain rule of differentiation and be able to use it to find the derivative of a composite function.

8.1.1 Function composition

Shown in Fig. 8.1 is an example of function composition: An independent variable, $x$, is used to evaluate a function, and the result, $u = f(x)$ then acts as an input to a second function, $g$. The final value is $y = g(u) = g(f(x))$. We refer to the two-step function operation as function composition.

Example 8.1 Consider the functions $f(x) = \sqrt{x}$ and $g(x) = x^2 + 1$. Determine the functions obtained by composing these, $h_1(x) = g(f(x))$ and $h_2(x) = f(g(x))$. 

139
Chapter 8. Introducing the chain rule

Solution: For $h_1$ we apply $f$ first, followed by $g$, so $h_1(x) = (\sqrt{x})^2 + 1 = x + 1$ (provided $x \geq 0$.) For $h_2$, the functions are applied in the reversed order so that $h_2(x) = \sqrt{x^2 + 1}$ (for any real $x$). We note that the domains of the two functions are slightly different. $h_1$ is only defined for $x \geq 0$ since $f(x)$ is not defined for negative $x$, whereas $h_2$ is defined for all $x$.

Example 8.2 Express the function $h(x) = 5(x^3 - x^2)^{10}$ as the composition of two simpler functions. What are the domains of each of the functions?

Solution: We can write this in terms of the two functions $f(x) = x^3 - x^2$ and $g(x) = 5x^{10}$. Then $h(x) = g(f(x))$.

8.1.2 The chain rule of differentiation

Given a composite function $y = f(g(x))$, we require a rule for differentiating $y$ with respect to $x$.

If $y = g(u)$ and $u = f(x)$ are both differentiable functions and $y = g(f(x))$ is the composite function, then the chain rule of differentiation states that

$$\frac{dy}{dx} = \frac{dy}{du} \frac{du}{dx}.$$

Informally, the chain rule states that the change in $y$ with respect to $x$ is a product of two rates of change: (1) the rate of change of $y$ with respect to its immediate input $u$, and (2) the rate of change of $u$ with respect to its input, $x$.

Why does it work this way? Although the derivative is not a simple quotient, we gain an intuitive grasp of the chain rule by writing

$$\frac{dy}{dx} \approx \frac{\Delta y}{\Delta x} = \frac{\Delta y}{\Delta u} \frac{\Delta u}{\Delta x}$$

then it is apparent that the “cancellation” of terms $\Delta u$ in numerator and denominator lead to the correct fraction on the left. The proof of the chain rule uses this essential idea, but care is taken to ensure that the quantity $\Delta u$ is nonzero, to avoid the embarrassment of dealing with the nonsensical ratio $0/0$.

Example 8.3 Compute the derivative of the function $h(x) = 5(x^3 - x^2)^{10}$. ■
8.1. The chain rule

Solution: We express the function as \( y = h(x) = 5u^{10} \) where \( u = (x^3 - x^2) \). Then

\[
\frac{dy}{dx} = \frac{dy}{du} \frac{du}{dx} = \left( \frac{d(5u^{10})}{du} \right) \left( \frac{d(x^3 - x^2)}{dx} \right) = 50u^9(3x^2 - 2x).
\]

Then, substituting for \( u \) leads to \( \frac{dy}{dx} = 50(x^3 - x^2)^9(3x^2 - 2x) \).

Example 8.4 Compute the derivative of the function \( y = f(x) = \sqrt{x^2 + a^2} \), where \( a \) is some positive real number.

Solution: This function can be considered as the composition of \( g(u) = \sqrt{u} = u^{1/2} \) and \( u(x) = x^2 + a^2 \). That is, we can write \( f(x) = g(h(x)) \). Then using the chain rule, we obtain

\[
\frac{dy}{dx} = \frac{1}{2} (x^2 + d^2)^{-1/2} \cdot 2x = \frac{x}{(x^2 + d^2)^{1/2}} = \frac{x}{\sqrt{x^2 + d^2}}.
\]

Example 8.5 Compute the derivative of the function

\[
y = f(x) = \frac{x}{\sqrt{x^2 + d^2}},
\]

where \( d \) is some positive real number.

Solution: We use both the quotient rule and the chain rule for this calculation.

\[
\frac{dy}{dx} = \frac{\left[ x \right]' \cdot \sqrt{x^2 + d^2} - \left[ \sqrt{x^2 + d^2} \right]' \cdot x}{(\sqrt{x^2 + d^2})^2}
\]

Here the \([\ldots]'\) denotes differentiation. Then

\[
\frac{dy}{dx} = \frac{1 \cdot \sqrt{x^2 + d^2} - \frac{1}{2} \cdot 2x \cdot (x^2 + d^2)^{-1/2} \cdot x}{(x^2 + d^2)}.
\]

We simplify algebraically by multiplying top and bottom by \((x^2 + d^2)^{1/2}\) and cancelling factors of 2 to obtain

\[
\frac{dy}{dx} = \frac{x^2 + d^2 - x^2}{(x^2 + d^2)^{1/2}(x^2 + d^2)} = \frac{d^2}{(x^2 + d^2)^{3/2}}.
\]

8.1.3 Interpreting the chain rule

The following intuitive examples may help to motivate why the chain rule is based on a product of two rates of change.

Example 8.6 (Pollution level in a lake) A species of fish is sensitive to pollutants in its lake. As humans settle and populate the area adjoining the lake, one may see a decline in the population of these fish due to increased levels of pollution. Quantify the rate at which the pollution level changes with time based on the pollution produced per human and the rate of increase of the human population.
**Solution:** The rate of fish population decline depends on the rate of change of the human population, and the rate of change in the pollution created per person. If either increases, the effect on the fish population increases. The chain rule implies that the net effect is a product of the two rates. Formally, for time in years, $x = f(t)$ the number of people at the lake in year $t$, and $p = g(x)$ the pollution created by $x$ people, the rate of change of the pollution $p$ over time is a product of $g'(x)$ and $f'(t)$:

$$\frac{dp}{dt} = \frac{dp}{dx} \frac{dx}{dt} = g'(x) f'(t).$$

**Example 8.7 (Population of carnivores, prey, and vegetation)** The population of large carnivores, $C$, on the African savannah depends on the population of gazelles that are their prey, $P$. The gazelle population, in turn, depends on the abundance of vegetation $V$, which depends on the amount of rain in a given year, $r$. Quantify the rate of change of the carnivore population with respect to the rainfall.

![Figure 8.2. An example in which the population of carnivores, $C = h(P) = P^2$ depends on prey $P$, while the prey depend on vegetation $P = f(V) = 2V$, and the vegetation depends on rainfall $V = g(r) = r^{1/2}$.](image)

**Solution:** We can express these dependencies through functions; for instance, we could write $V = g(r)$, $P = f(V)$ and $C = h(P)$, where we understand that $g, f, h$ are some functions (resulting from measurement or data collection on the savanna). As one example, shown in Figure 8.2,

$$C = h(P) = P^2, \quad P = f(V) = 2V, \quad V = g(r) = r^{1/2}.$$

A drought that decreases rainfall will also decrease the abundance of vegetation. This will decrease the gazelle population, and eventually affect the population of carnivores. The rate of change in the carnivores population with respect to the rainfall, $dC/dr$, according to the chain rule, would be

$$\frac{dC}{dr} = \frac{dC}{dP} \frac{dP}{dV} \frac{dV}{dr}.$$

Computing all necessary derivatives,

$$\frac{dV}{dr} = \frac{1}{2} r^{-1/2}, \quad \frac{dP}{dV} = 2, \quad \frac{dC}{dP} = 2P,$$

so that

$$\frac{dC}{dr} = \frac{dC}{dP} \frac{dP}{dV} \frac{dV}{dr} = \frac{1}{2} r^{-1/2} (2P) = \frac{2P}{r^{1/2}}.$$
8.1. The chain rule

Using the fact that $V = r^{1/2}$ and $P = 2V$, we obtain

$$\frac{dC}{dr} = \frac{2P}{V} = \frac{2(2V)}{V} = 4.$$

An alternate approach for this example (which is often not possible in more complicated cases) is to express the number of carnivores explicitly in terms of rainfall, using the fact that $C = h(P) = h(f(V)) = h(f(g(r)))$. Then

$$C = P^2 = (2V)^2 = 4V^2 = 4(r^{1/2})^2 = 4r.$$  \[\Rightarrow\]  $$\frac{dC}{dr} = 4.$$

We can see that our two answers agree.

**Example 8.8 (Budget for coffee)** The budget for coffee depends on the number of cups consumed per day and on the price per cup. The total budget changes if the price goes up or if the consumption goes up (e.g. during late nights preparing for midterm exams). Quantify the rate at which your budget for coffee would change if both consumption and price change.

**Solution:** The total rate of change of the coffee budget is a product of the change in the price and the change in the consumption. For $t$ time in days, $x = f(t)$ the number of cups of coffee consumed, and $y = g(x)$ the price for $x$ cups of coffee, we obtain

$$\frac{dy}{dt} = \frac{dy}{dx} \frac{dx}{dt} = g'(x)f'(t).$$

**Example 8.9 (Earth’s temperature and greenhouse gases)** In Exercise 21 of Chapter 1, we found that the temperature of the Earth depends on the albedo $a$ (fraction of incoming radiation energy reflected) according to the formula

$$T = \left(\frac{(1-a)S}{\epsilon \sigma}\right)^{1/4}. \quad (8.1)$$

Suppose that the albedo $a$ depends on the level of greenhouse gases $G$ so that $da/dG$ is known. If this is the only quantity that depends on $G$, determine how the temperature would change as the level of greenhouse gases $G$ increases.

**Solution:** The information provided specifies that $T$ depends on the level of greenhouse gases via the chain of dependencies $G \rightarrow a \rightarrow T$. Let us write

$$T = \left(\frac{S}{\epsilon \sigma}\right)^{1/4} (1-a)^{1/4}$$

In this problem the quantities $S, \epsilon, \sigma$ are all constants, so it simplifies calculation to write the function in the form shown above. According to the chain rule,

$$\frac{dT}{dG} = \frac{dT}{da} \frac{da}{dG}.$$
We are given \( da/dG \) and we can compute \( dT/da \). Hence, we find that

\[
\frac{dT}{dG} = \left( \frac{S}{e \sigma} \right)^{1/4} \frac{d}{da} \left[ (1 - a)^{1/4} \right] \frac{da}{dG} = \left( \frac{S}{e \sigma} \right)^{1/4} \frac{1}{4} \left( 1 - a \right)^{(1/4) - 1} \cdot (-1) \frac{da}{dG}.
\]

Rearranging leads to

\[
\frac{dT}{dG} = -\frac{1}{4} \left( \frac{S}{e \sigma} \right)^{1/4} (1 - a)^{-3/4} \frac{da}{dG}.
\]

In general, greenhouse gases affect both the Earth’s albedo \( a \) and its emissivity \( \epsilon \). We generalize our results in Exercise 2.

### 8.2 The chain rule applied to optimization problems

Armed with the chain rule, we can now differentiate a wider variety of functions, and address problems that were not tractable with power, product, or quotient rule alone. We return to optimization problems in where derivatives require use of the chain rule.

#### Section 8.2 Learning goals

1. Read and follow the derivation of each optimization model.
2. Be able to carry out the calculations of derivatives appearing in the problems (using the chain rule)
3. Using optimization, find each critical point and identify its type.
4. Understand and be able to explain the interpretation of the mathematical results.

#### 8.2.1 Shortest path from food to nest

Ants are good mathematicians! They are able to find the shortest route connecting their nest to a food source. But how do they do it? Each ant secretes a chemical pheromone that other ants tend to follow. This marks up the trail that they use, and recruits nest-mates to food sources. The pheromone (chemical message for marking a route) evaporates after a while, so that, for a given number of foraging ants, a longer trail will have a less concentrated chemical marking than a shorter trail. This means that whenever a shorter route is found, the ants will favor it. After some time, this leads to selection of the shortest possible trail.

Shown in Fig. 8.3 is a common laboratory test scenario: ants in an artificial nest are offered two equivalent food sources. We ask what is the shortest total path connecting nest and sources.\(^{33}\)

**Example 8.10 (Minimizing the total path length)** Use the diagram to determine the length of the shortest path that connects the nest to both food sources. Assume that \( d << D \). \( \Box \)

\(^{33}\)This is a simplified version of the problem that the ants are solving.
8.2. The chain rule applied to optimization problems

Food Nest Food Nest Food Nest
(a) (b) (c)

Figure 8.3. Three ways to connect the ants’ nest to two food sources, showing (a) a V-shaped, (b) T-shaped, and (c) Y-shaped paths.

Solution: We first consider two possibilities for paths connecting nest to food (1) a V-shaped path and (2) a T-shaped path. Then for a for a V-shaped path the total length is $2\sqrt{D^2 + d^2}$, whereas for the T-shaped path, it is $D + 2d$. Now consider a third possibility, of a Y-shaped path. This means that the ants start to walk straight ahead and then veer off to the food after a while. All three possibilities are shown in Fig. 8.3.

Calculations are easiest if we denote the distance from the nest to the Y-junction as $D - x$, so that $x$ is distance shown in the diagram. The length of the Y-shaped path is then

$$L = L(x) = (D - x) + 2\sqrt{d^2 + x^2}. \quad (8.2)$$

Observe that when $x = 0$, then $L_T = D + 2d$, which corresponds to the T-shaped path length, whereas when $x = D$ then $L_V = 2\sqrt{d^2 + D^2}$ which V-shaped path length. Thus in this problem, we have $0 \leq x \leq D$ as the appropriate domain. We have determined the values of $L$ at the two domain endpoints.

To find the minimal path length, we look for critical points of the function $L(x)$. Differentiating (see Examples 8.4 and 8.5), we find

$$L'(x) = -1 + 2 \frac{x}{\sqrt{x^2 + d^2}}, \quad L''(x) = 2 \frac{d^2}{(x^2 + d^2)^{3/2}} > 0.$$

Then critical points occur when

$$L'(x) = 0 \quad \Rightarrow \quad -1 + 2 \frac{x}{\sqrt{x^2 + d^2}} = 0.$$

Simplifying leads to

$$\sqrt{x^2 + d^2} = 2x \quad \Rightarrow \quad x^2 + d^2 = 4x^2 \quad \Rightarrow \quad 3x^2 = d^2 \quad \Rightarrow \quad x = \frac{d}{\sqrt{3}}.$$

To determine the kind of critical point, we note that the second derivative is positive and the critical point is thus a local maximum.
Chapter 8. Introducing the chain rule

To determine the actual length of the path, we substitute \( x = d/\sqrt{3} \) into the function \( L(x) \) and obtain (after simplification, see Exercise 3)

\[
L = L(x) = D + \sqrt{3}d.
\]

The final result is summarized in Fig. 8.4. The shortest path is Y-shaped, with \( x = d/\sqrt{3} \). The ants march straight for a distance \( D - (d/\sqrt{3}) \), and then their trail branches to the right and left towards the food sources.

![Diagram](image)

**Figure 8.4.** (a) In the configuration for the shortest path we found that \( x = d/\sqrt{3} \). (b) The total length of the path \( L(x) \) as a function of \( x \) for \( D = 2, d = 1 \). The minimal path occurs when \( x = 1/\sqrt{3} \approx 0.577 \). The length of the shortest path is then \( L = D + \sqrt{3}d = 2 + \sqrt{3} \approx 3.73 \).

8.2.2 Food choice and attention

The example described in this section is based on research about animal behaviour. Several of its features are noteworthy. Aside from being an application of the chain rule, we encounter a surprise in some of our elementary calculations. The example also demonstrates that not every problem has an elegant or analytically simple solution. Indeed, we find that Newton’s method for approximating zeros of function proves useful. Finally, we see that general observations can provide insight not easily gained from specific cases.

**Paying attention**

Behavioural ecologist Reuven Dukas (McMaster U) studied choices that animals make when deciding which food to look for. The example below is based on his work with blue jays [7, 8, 9].

Many types of food are cryptic - hidden in the environment - and so, require time and attention to find. Some types of food are more easily detected, but other foods might
provide greater nourishment. The goal of this example is to find the best subdivision of time and attention between food types that maximizes the total energy gain.

**Setting up a model**

Suppose that there are two types of food available in the environment. Define notation as follows:

\[ x = \text{attention devoted to finding food of type 1} \]
\[ P(x) = \text{probability of finding the food given attention } x \]

We assume that \( 0 \leq x \leq 1 \), with \( x = 0 \) representing no attention and \( x = 1 \) full attention devoted to finding food type 1. Moreover, since \( P \) is a probability, \( 0 \leq P \leq 1 \). \( P = 0 \) means that the food is never found, and \( P = 1 \) means that the food is always found. About \( P \) we also assume that \( P(0) = 0, P(1) = 1 \), meaning that when no attention is paid \((x = 0)\) then there is no probability of finding the food \((P = 0)\), whereas if full attention paid \( x = 1 \) then there is always success \((P = 1)\).

Fig. 8.5 displays several hypothetical examples of \( P(x) \) versus \( x \). On the horizontal axis, we show the attention \( 0 \leq x \leq 1 \), and on the vertical axis, we show the probability of success at finding food, \( 0 \leq P \leq 1 \). All these curves share features we have assumed: full success for full attention, and no success for no attention. However, the curves differ in values at intermediate levels of attention.

![Figure 8.5](image.png)

**Figure 8.5.** The probability, \( P(x) \), of finding a food depends on the level of attention \( x \) devoted to finding that food. Here \( 0 \leq x \leq 1 \), with \( x = 1 \) being “full attention” devoted to the task. We show possible curves for four types of foods, some easier to find than others.

**Questions:**

1. What is the difference between foods of type 1 and 4?
2. Which food is easier to find, type 3 or type 4?
3. What role is played by the concavity of the curve?
Observe that concave down curves such as 3 and 4 rise rapidly, indicating that the probability of finding food increases a lot just by increasing the attention by a little: These represent foods that are relatively easy to find. Other curves (1 and 2) are concave up, indicating that much more attention is needed to gain appreciable increase in the probability of success: these represent foods that are harder to find, or more cryptic. The concavity of the curves carries this important information about relative ease or difficulty in finding a given food type.

For the purposes of this example, we consider two food types with attentions \( x \) and \( y \) fully divided between them so that \( x + y = 1 \). Suppose that the relative nutritional values of the foods are 1 and \( N \). Let \( P_1(x) \) and \( P_2(y) \) be probabilities of finding food of type 1 and of 2 given that we spend attention \( x \) or \( y \) in looking for the given type. Then the total nutritional value gained by splitting up attention between the two foods is:

\[
V(x) = P_1(x) + N P_2(y) = P_1(x) + N P_2(1 - x).
\]

**Example 8.11 (\( P_1 \) and \( P_2 \) as power function with integer powers:)** Consider the case that the probability of finding the food types is given by the simple power functions,

\[
P_1(x) = x^2, \quad P_2(y) = y^3.
\]

(These functions satisfy \( P(0) = 0, P(1) = 1 \), in accordance with Figure 8.5.) Further, suppose that both foods are equally nutritious so \( N = 1 \). Find the optimal \( V(x) \).

**Solution:** The total nutritional value in this case is

\[
V(x) = P_1(x) + N P_2(1 - x) = x^2 + (1 - x)^3.
\]

We look for a maximum value of \( V \): Using the chain rule to differentiate, we find that

\[
V'(x) = 2x + 3(1 - x)^2(-1), \quad V''(x) = 2 - 3(2)(1 - x)(-1) = 2 + 6(1 - x).
\]

We observe that a negative factor \((-1)\) comes from applying the chain rule to the factor \((1 - x)^3\). Setting \( V'(x) = 0 \) we get

\[
2x + 3(1 - x)^2(-1) = 0. \quad \Rightarrow \quad -3x^2 + 8x - 3 = 0 \quad \Rightarrow \quad x = \frac{4 \pm \sqrt{7}}{3} \approx 0.4514, 2.21.
\]

Since the attention must take on a value in \( 0 \leq x \leq 1 \), we must reject the second of the two solutions. It would appear that the animal may benefit most by spending a fraction 0.4514 of its attention on food type 1 and the rest on type 2.

However, to confirm our speculation, we must check whether the critical point is a maximum.

Observing the second derivative, and recalling that \( x \leq 1 \), we note that the second derivative is positive for all values of \( x \)! This is unfortunate, as it signifies a local minimum! The animal gains least by splitting up its attention between the foods in this case. Indeed, from Figure 8.6(a), we see that the most gain occurs at either \( x = 0 \) (only food of type 2 sought) or \( x = 1 \) (only food of type 1 sought). Again we observe the importance of checking for the type of critical point before drawing hasty conclusions.
8.2. The chain rule applied to optimization problems

The chain rule applied to optimization problems

Figure 8.6. (a) Figure for Example 8.11 and (b) for Example 8.12. In (a) the probabilities of finding foods of types 1 and 2 are concave up power functions, whereas in (b) both functions are concave down. As a result there is a local maximum for the nutritional value in (b) but not in (a).

Example 8.12 (Fractional-power functions for $P_1, P_2$) As a second example, consider the case that the probability of finding the food types is given by the concave down power functions,

$$P_1(x) = x^{1/2}, \quad P_2(y) = y^{1/3}$$

and both foods are equally nutritious ($N = 1$). Find the optimal food value $V(x)$.

Solution: These functions also satisfy $P(0) = 0, P(1) = 1$, in accordance with the sketches shown in Figure 8.5. Then

$$V(x) = P_1(x) + P_2(1 - x) = \sqrt{x} + (1 - x)^{(1/3)},$$

$$V'(x) = \frac{1}{2\sqrt{x}} - \frac{1}{3(1.0 - x)^{(2/3)}},$$

$$V''(x) = -\frac{1}{4x^{(3/2)}} - \frac{2}{9(1.0 - x)^{(5/3)}}.$$

At this point we would like to proceed to solve $V'(x) = 0$ to find the critical point. Unfortunately, this problem, while seemingly routine, turns out to be algebraically nasty. However, rather than despair, we seek an approximate solution to the problem, for which Newton’s Method proves ideal, as shown next.

A plotting program used to graph $V(x)$ in Figure 8.6(b) demonstrates that there is a maximum in $0 \leq x \leq 1$, i.e for attention split between finding both foods. We further see from $V''(x)$ that the second derivative is negative for all values of $x$ in the interval, indicating that we have obtained a local maximum, as expected.
Applying Newton's method to finding the critical point

Example 8.13 Use Newton’s Method to find the critical point for the function \( V(x) \) in Example 8.12.

Solution: Let \( f(x) = V'(x) \). Then finding the critical point of \( V(x) \) reduces to finding the zero of the function \( V'(x) \), i.e., solving \( f(x) = 0 \). This is precisely the type of problem that Newton’s Method addresses, as discussed in Section 5.4. Since the interval of interest is \( 0 \leq x \leq 1 \), we start with an initial “guess” for the critical point at \( x_0 = 0.5 \), midway along this interval. Then, according to Newton’s method, the improved guess would be

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

and, repeating this, at the \( k \)'th stage,

\[
x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}.
\]

To use this method, we must carefully note that

\[
f(x) = V'(x) = \frac{1}{2 \sqrt{x}} - \frac{1}{3 (1.0 - x)^{2/3}},
\]

\[
f'(x) = V''(x) = -\frac{1}{4 x^{3/2}} - \frac{2}{9 (1.0 - x)^{5/2}}.
\]

Thus, we might use a spreadsheet in which cells A1 stores our initial guess, whereas B1, C1, and D1 store the values of \( f(x) \), \( f'(x) \) and \( x_0 - f(x_0)/f'(x_0) \). In the typical syntax of spreadsheets, this might read something like the following.

A1 0.5  
B1 = (1/(2*SQRT(A1)) - 1/(3*(1-A1)^(2/3)))  
C1 = (-1/(4*A1^(3/2)) - 2/(9*(1-A1)^(5/2)))  
D1 = A1 - B1/C1

Applying this idea, and repeating the calculation by dragging the values to successive rows would lead to iterated approximations as follows.

\[
x_0 = 0.50000, x_1 = 0.59061, \text{and thereafter, successive values 0.60816, 0.61473, 0.61751, 0.61875, 0.61931, 0.61956, 0.61968, 0.61973, 0.61976, 0.61977, 0.61977, 0.61977, ...}
\]

Thus, we see that the values converge to the location of the critical point, \( x = 0.61977 \) (and \( y = 1 - x = 0.38022 \)) within the interval of interest.

Epilogue

While the conclusions drawn above were disappointing in one specific case, it is not always true that concentrating all one’s attention on one type is optimal. We can examine the
problem in more generality to find when the opposite conclusion might be satisfied. In the general case, the value gained is

\[ V(x) = P_1(x) + N P_2(1 - x). \]

A critical point occurs when

\[ V'(x) = \frac{d}{dx}[P_1(x) + N P_2(1 - x)] = P'_1(x) + N P'_2(1 - x)(-1) = 0. \]

(By now you realize where the extra term \((-1)\) comes from the Chain Rule.) Suppose we have found a value of \(x\) in \(0 < x < 1\) at which this is satisfied. We then examine the second derivative:

\[ V''(x) = \frac{d}{dx}[V'(x)] = \frac{d}{dx}[P'_1(x) - N P'_2(1 - x)]
= P''_1(x) - N P''_2(1 - x)(-1) = P''_1(x) + N P''_2(1 - x). \]

The concavity of the function \(V\) is thus related to the concavity of the two functions \(P_1(x)\) and \(P_2(1 - x)\). If these are concave down (e.g. as in food types 3 or 4 in Figure 8.5), then \(V''(x) < 0\) and a local maximum will occur at any critical point found by our differentiation. Another way of stating this observation is: if both food types are relatively easy to find, one can gain most benefit by splitting up the attention between the two. Otherwise, if both are hard to find, then it is best to look for only one at a time.
Exercises

8.1. **Practicing the Chain Rule:** Use the chain rule to calculate the following derivatives

(a) \( y = f(x) = (x + 5)^5 \)

(b) \( y = f(x) = 4(x^2 + 5x - 1)^8 \)

(c) \( y = f(x) = (\sqrt{x} + 2x)^3 \)

8.2. **Earth’s temperature:** In this problem, we expand and generalize the results of Example 8.9. As before, let \( G \) denote the level of greenhouse gasses on Earth, and consider the relationship of temperature of the earth to the albedo \( a \) and the emissivity \( \epsilon \) given by Eqn. (8.1).

(a) Suppose that \( a \) is constant, but \( \epsilon \) depends on \( G \). Assume that \( d\epsilon/dG \) is given. Determine the rate of change of temperature with respect to the level of greenhouse gasses in this case.

(b) Suppose that both \( a \) and \( \epsilon \) depend on \( G \). Find \( dT/dG \) in this more general case. (Hint: the quotient rule as well as the chain rule will be needed in this case.)

8.3. **Shortest path from nest to food sources:**

(a) Use the first derivative test to verify that the value \( x = \frac{d}{\sqrt{3}} \) is a local minimum of the function \( L(x) \) given by Eqn (8.2)

(b) Show that the shortest path is \( L = D + \sqrt{3}d \).

(c) In Section 8.2.1 we assumed that \( d << D \), so that the food sources were close together relative to the distance from the nest. Now suppose that \( D = d/2 \). How would this change the solution to the problem?

8.4. **Geometry of the shortest ants’ path:** Use the results of Section 8.2.1 to show that in the shortest path, the angles between the branches of the Y-shaped path are all 120°. You may find it helpful to recall that \( \sin(30) = 1/2 \), \( \sin(60) = \sqrt{3}/2 \).

8.5. **More about the ant trail:** Consider the lengths of the V and T-shaped paths in the ant trail example of Section 8.2.1. We will refer to these as \( L_V \) and \( L_T \), and both depend on the distances \( d \) and \( D \) in Fig. 8.3.

(a) Write down the expressions for each of these functions.

(b) Suppose the distance \( D \) is fixed. How do the two lengths \( L_V \), \( L_T \) depend on the distance \( d \)? Use your sketching skills to draw a rough sketch of \( L_V(d) \), \( L_T(d) \).

(c) Use you sketch to determine whether there is a value of \( d \) for which the lengths \( L_V \) and \( L_T \) are the same.

8.6. **Divided attention:**

This problem is based on the material on food choice and attention described in Section 7.6. It is advisable to first read that section.

A bird in its natural habitat feeds on two kinds of seeds, whose nutritional values are 5 calories per seed of type 1 and 3 calories per seed of type 2. Both kinds of seeds are hidden among litter on the forest floor and have to be found. If the bird splits its
attention into a fraction $x_1$ searching for seed type 1 and a fraction $x_2$ searching for seed type 2, then its probability of finding 100 seeds of the given type is

$$P_1(x_1) = (x_1)^3, \quad P_2(x_2) = (x_2)^5.$$ 

Assume that the bird pays full attention to searching for seeds so that $x_1 + x_2 = 1$ where $0 \leq x_1 \leq 1$ and $0 \leq x_2 \leq 1$.

(a) Write down an expression for the total nutritional value $V$ gained by the bird when it splits its attention. Use the constraint on $x_1, x_2$ to eliminate one of these two variables. (For example, let $x = x_1$ and write $x_2$ in terms of $x_1$.)

(b) Find critical points of $V(x)$ and classify those points.

(c) Find absolute minima and maxima of $V(x)$ and use your results to explain what is the bird’s optimal strategy to maximize the nutritional value of the seeds it can find.