3.4 The Chain Rule

You will recall from Calculus I that we apply the chain rule when we have the composition of two functions, for example when computing \( \frac{df}{dx} \). The chain rule applies in similar situations when dealing with functions of several variables. For example, \( f(x,y) \) is a function of two variables. However, we may be computing the partials of \( f(u,v) \) where both \( u \) and \( v \) are functions of one or more variables. We look at different cases.

3.4.1 Partials of \( f(x,y) \) where \( x \) and \( y \) are functions of one other variable \( t \)

Before we start, let us remind the reader that if a variable depends on several variables then the derivatives are partial derivatives and we will use the partial derivative notation as in \( \frac{\partial z}{\partial x} \) but if a variable depends only on one other variable, we will use the notation from calculus of functions of one variable as in \( \frac{dx}{dt} \).

Now, suppose we have \( z = f(x,y) \), \( x = g(t) \), \( y = h(t) \). Then, in this case we can also think of \( z \) as a function of \( t \). So, there is only one derivative to compute, \( \frac{dz}{dt} \). It can be computed as follows:

\[
\frac{dz}{dt} = \frac{\partial z}{\partial x} \frac{dx}{dt} + \frac{\partial z}{\partial y} \frac{dy}{dt} \tag{3.1}
\]

To remember this formula, we can use a tree structure shown in figure 3.6 as follows.

1. Draw the tree from top to bottom. Assuming \( z = f(x,y) \), \( x = g(t) \), \( y = h(t) \), we start with \( z \).
2. From \( z \), we draw a branch for each variable \( z \) depends on, \( x \) and \( y \) in this case, so we draw two branches.
3. From each of these variables, we repeat the procedure, that is draw a branch for each variable it depends on. Both \( x \) and \( y \) only depend on \( t \), so we draw one branch from each.

The tree is interpreted as follows. Since \( z \) is ultimately a function of \( t \), look at all the paths from \( z \) to \( t \).

1. Multiply all the partials which appear on a path.
2. Add all these products collected on each path.

In our case, we obtain the following: There are two paths from \( z \) to \( t \). The first one is \( z \to x \to t \) which gives \( \frac{\partial z}{\partial x} \frac{dx}{dt} \). The second path is \( z \to y \to t \) which gives \( \frac{\partial z}{\partial y} \frac{dy}{dt} \). Adding these gives \( \frac{\partial z}{\partial x} \frac{dx}{dt} + \frac{\partial z}{\partial y} \frac{dy}{dt} \) which is \( \frac{dz}{dt} \).
Similarly, if \( w = f(x, y, z) \) and \( x, y, z \) are functions of \( t \), then the corresponding tree structure is shown in figure 3.7.

Again, \( w \) is ultimately a function of \( t \). So, there is only one derivative to compute, \( \frac{dw}{dt} \). Using the interpretation outlines above, we obtain the following formula:

\[
\frac{dw}{dt} = \frac{\partial w}{\partial x} \frac{dx}{dt} + \frac{\partial w}{\partial y} \frac{dy}{dt} + \frac{\partial w}{\partial z} \frac{dz}{dt}
\]

and so on.

**Example 244** Find \( \frac{du}{dt} \) if \( u = x^2 - y^2 \) and \( x = t^2 - 1 \), \( y = 3 \sin \pi t \).

\[
\frac{du}{dt} = \frac{\partial u}{\partial x} \frac{dx}{dt} + \frac{\partial u}{\partial y} \frac{dy}{dt}
\]

\[
= 2x2t - 2y3\pi \cos \pi t
\]

\[
= 4xt - 6\pi y \cos \pi t
\]

\[
= 4 (t^2 - 1) t - 6\pi (3 \sin \pi t) \cos \pi t
\]

\[
= 4t^3 - 4t - 18\pi \sin \pi t \cos \pi t
\]

### 3.4.2 Partials of \( f(x, y) \) where \( x \) and \( y \) are functions of two variables \( s \) and \( t \)

We can use the same tree structure as above to compute partial derivatives in this case. Suppose we have \( z = f(x, y) \), \( x = g(s, t) \), \( y = h(s, t) \). So, in this case, \( z \) is a function of \( s \) and \( t \). So, there are two partial derivatives to compute: \( \frac{\partial z}{\partial s} \) and \( \frac{\partial z}{\partial t} \). The corresponding tree structure is shown in figure 3.8.
3.4. THE CHAIN RULE

Figure 3.7: $w = f(x(t), y(t), z(t))$

Figure 3.8: $z = f(x(s, t), y(s, t))$
The first partials of $z$ with respect to $x$ or $t$ are computed as follows:

\[
\frac{\partial z}{\partial s} = \frac{\partial z}{\partial x} \frac{\partial x}{\partial s} + \frac{\partial z}{\partial y} \frac{\partial y}{\partial s} \quad (3.2)
\]
\[
\frac{\partial z}{\partial t} = \frac{\partial z}{\partial x} \frac{\partial x}{\partial t} + \frac{\partial z}{\partial y} \frac{\partial y}{\partial t} \quad (3.3)
\]

**Example 245** Let $z = \sin (x + y)$ where $x = 2st$ and $y = s^2 + t^2$. Find $\frac{\partial z}{\partial s}$ and $\frac{\partial z}{\partial t}$.

\[
\frac{\partial z}{\partial s} = \frac{\partial z}{\partial x} \frac{\partial x}{\partial s} + \frac{\partial z}{\partial y} \frac{\partial y}{\partial s} = \cos (x + y) (2t) + \cos (x + y) (2s) = 2t \cos (2st + s^2 + t^2) + 2s \cos (st + s^2 + t^2) = 2 \cos \left( (s + t)^2 \right) (s + t)
\]

and

\[
\frac{\partial z}{\partial t} = \frac{\partial z}{\partial x} \frac{\partial x}{\partial t} + \frac{\partial z}{\partial y} \frac{\partial y}{\partial t} = \cos (x + y) (2s) + \cos (x + y) (2t) = 2 \cos \left( (s + t)^2 \right) (s + t)
\]

**Example 246** Let $u = x^2 - 2xy + 2y^3$ where $x = s^2 \ln t$ and $y = 2st$. Find $\frac{\partial u}{\partial s}$ and $\frac{\partial u}{\partial t}$.

\[
\frac{\partial u}{\partial s} = \frac{\partial u}{\partial x} \frac{\partial x}{\partial s} + \frac{\partial u}{\partial y} \frac{\partial y}{\partial s} = (2x - 2y) (2s \ln t) + (-2x + 6y^2) (2t^3) = (2s^2 \ln t - 4st^3) (2s \ln t) + (-2s^2 \ln t + 24s^2 t^6) (2t^3)
\]

\[
\frac{\partial u}{\partial t} = \frac{\partial u}{\partial x} \frac{\partial x}{\partial t} + \frac{\partial u}{\partial y} \frac{\partial y}{\partial t} = (2s^2 \ln t - 4st^3) \left( \frac{s^2}{t} \right) + (-2s^2 \ln t + 24s^2 t^6) (6st^2)
\]

**Example 247** Given $z = f(x, y)$, $x = r^2 + s^2$ and $y = 2rs$ find $\frac{\partial z}{\partial r}$ and $\frac{\partial^2 z}{\partial r^2}$.
• **Computation of $\frac{\partial z}{\partial r}$:**

\[
\frac{\partial z}{\partial r} = \frac{\partial z}{\partial x} \frac{\partial x}{\partial r} + \frac{\partial z}{\partial y} \frac{\partial y}{\partial r} = 2r \frac{\partial z}{\partial x} + 2s \frac{\partial z}{\partial y}
\]

• **Computation of $\frac{\partial^2 z}{\partial r^2}$:**

\[
\frac{\partial^2 z}{\partial r^2} = \frac{\partial}{\partial r} \left( \frac{\partial z}{\partial r} \right) = \frac{\partial}{\partial r} \left( 2r \frac{\partial z}{\partial x} + 2s \frac{\partial z}{\partial y} \right) = 2 \left( \frac{\partial z}{\partial r} \right) \left( \frac{\partial z}{\partial x} \right) + 2r \frac{\partial}{\partial r} \left( \frac{\partial z}{\partial x} \right) + 2s \frac{\partial}{\partial r} \left( \frac{\partial z}{\partial y} \right) = 2 \frac{\partial z}{\partial x} + 2s \frac{\partial^2 z}{\partial y^2} + 2 \frac{\partial^2 z}{\partial x \partial r} + 2s \frac{\partial^2 z}{\partial x \partial y}
\]

We compute separately

\[
\frac{\partial}{\partial r} \left( \frac{\partial z}{\partial x} \right) = \frac{\partial}{\partial x} \left( \frac{\partial z}{\partial x} \right) \frac{\partial x}{\partial r} + \frac{\partial}{\partial y} \left( \frac{\partial z}{\partial x} \right) \frac{\partial y}{\partial r} = 2r \frac{\partial^2 z}{\partial x^2} + 2s \frac{\partial^2 z}{\partial y \partial x} \quad (3.5)
\]

and

\[
\frac{\partial}{\partial r} \left( \frac{\partial z}{\partial y} \right) = \frac{\partial}{\partial x} \left( \frac{\partial z}{\partial y} \right) \frac{\partial x}{\partial r} + \frac{\partial}{\partial y} \left( \frac{\partial z}{\partial y} \right) \frac{\partial y}{\partial r} = 2r \frac{\partial^2 z}{\partial x \partial y} + 2s \frac{\partial^2 z}{\partial y^2} \quad (3.6)
\]

If $f$ has continuous second partial derivatives, then combining Equations 3.4, 3.5 and 3.6 gives

\[
\frac{\partial^2 z}{\partial r^2} = 2 \frac{\partial z}{\partial x} + 4r^2 \frac{\partial^2 z}{\partial x^2} + 8rs \frac{\partial^2 z}{\partial y \partial x} + 4s^2 \frac{\partial^2 z}{\partial y^2}
\]

If $f$ does not have continuous partial derivatives, then we cannot combine the mixed partials, and we get an expression slightly more complicated:

\[
\frac{\partial^2 z}{\partial r^2} = 2 \frac{\partial z}{\partial x} + 4r^2 \frac{\partial^2 z}{\partial x^2} + 4rs \frac{\partial^2 z}{\partial y \partial x} + 4s \frac{\partial^2 z}{\partial x \partial y} + 4s^2 \frac{\partial^2 z}{\partial y^2}
\]
3.4.3 General Case

Suppose $f$ is a function of $n$ variables $x_1, x_2, \ldots, x_n$ and each $x_i$ is in turn a function of $m$ variables $t_1, t_2, \ldots, t_m$. $f$ is then a function of $m$ variables. The $m$ first partials of $f$ are:

$$\frac{\partial f}{\partial t_i} = \frac{\partial f}{\partial x_1} \frac{\partial x_1}{\partial t_i} + \frac{\partial f}{\partial x_2} \frac{\partial x_2}{\partial t_i} + \ldots + \frac{\partial f}{\partial x_n} \frac{\partial x_n}{\partial t_i}$$

for each $i = 1, 2, \ldots, m$.

3.4.4 Implicit Differentiation

The chain rule can be used to derive a simpler method for finding the derivative of an implicitly defined function.

$y$ defined implicitly as a function of $x$ in a relation of the form $F(x, y) = 0$

Suppose that $F(x, y) = 0$ defines $y$ as an implicit function of $x$ we will call $y = f(x)$. We wish to find $\frac{dy}{dx}$. We do so by differentiating both sides of $F(x, y) = 0$ with respect to $x$. The right side is easy to differentiate and we will skip the details here! To differentiate the left side with respect to $x$, $F(x, y)$, we will use the chain rule, remembering that $F(x, y) = F(x, f(x))$. So, $F$ is ultimately a function of $x$. The corresponding tree structure is shown in figure 3.9. We obtain:
3.4. THE CHAIN RULE

Figure 3.10: $F(x, y, f(x, y))$

\[
\frac{\partial F}{\partial x} \frac{dx}{dx} + \frac{\partial F}{\partial y} \frac{dy}{dx} = 0 \\
\frac{\partial F}{\partial x} + \frac{\partial F}{\partial y} \frac{dy}{dx} = 0 \\
\frac{dy}{dx} = -\frac{\partial F}{\partial y} = \frac{F_x}{F_y}
\]

Example 248 Find $\frac{dy}{dx}$ if $2xy - y^3 + 1 - x - 2y = 0$.

Using the notation above, $F(x, y) = 2xy - y^3 + 1 - x - 2y$

\[
\frac{dy}{dx} = \frac{F_x}{F_y} \\
= \frac{-2y - 1}{2x - 3y^2 - 2}
\]

$z$ defined implicitly as a function of $x$ and $y$ in a relation of the form $F(x, y, z) = 0$

Suppose that $F(x, y, z) = 0$ defines $z$ as an implicit function of $x$ and $y$ that is $F(x, y, z) = F(x, y, f(x, y))$. We can use a tree structure to find the partial derivatives. It is shown in figure 3.10. The dotted lines indicate that if $x$ were a function of $y$, it is where we would have drawn a branch. Since $x$ and $y$ are the independent variables, they do not depend on each other. So, $x$ is not a function of $y$ and $y$ is not a function of $x$.

If we differentiate each side with respect to $x$, using the tree structure, we
get
\[ \frac{\partial F}{\partial x} \frac{\partial x}{\partial x} + \frac{\partial F}{\partial z} \frac{\partial z}{\partial x} = 0 \]

Now, \( \frac{\partial x}{\partial x} = 1 \) so, we get
\[ \frac{\partial F}{\partial x} + \frac{\partial F}{\partial z} \frac{\partial z}{\partial x} = 0 \]

That is
\[ \frac{\partial z}{\partial x} = -\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial z}} = -\frac{F_x}{F_z} \]

Similarly, we can show that
\[ \frac{\partial z}{\partial y} = -\frac{\frac{\partial F}{\partial y}}{\frac{\partial F}{\partial z}} = -\frac{F_y}{F_z} \]

**Example 249** Find \( \frac{\partial z}{\partial x} \) if \( z \) is defined implicitly by
\[ x^3 + y^3 + z^3 + 6xyz = 1 \]

Using the formula above with \( F(x,y,z) = x^3 + y^3 + z^3 + 6xyz - 1 \), we see that
\[ \frac{\partial z}{\partial x} = -\frac{F_x}{F_z} = -\frac{3x^2 + 6yz}{3z^2 + 6xy} = -\frac{x^2 + 2yz}{z^2 + 2xy} \]

Compare this solution with the solution of example 234 on page 134.

### 3.4.5 Assignment

Do odd# 1 - 33 at the end of 11.4 in your book.