27 Chain rule

For single variables the chain rule tells us how to differentiate a composition of functions, i.e. \( \frac{df(g(u))}{du} = f'(g(u))g'(u) \). We can understand the use of the word chain by thinking of computing the function in steps:

\[
\begin{align*}
  u & \leadsto x = g(u) \leadsto y = f(x) = f(g(u)).
\end{align*}
\]

This shows a chain of variables:
- \( u \) is independent.
- \( x \) is intermediate and depends on \( u \).
- \( y \) is dependent and depends on \( x \), so also depends on \( u \).

With this notation the chain rule can be written:

\[
\frac{dy}{du} = \frac{dy}{dx} \cdot \frac{dx}{du}.
\]

The chain rule for multivariable functions \( w = f(x, y) \) is similar. The chain can be written

\[
\begin{align*}
  (u, v) & \leadsto (x, y) = (x(u, v), y(u, v)) \leadsto w = f(x, y) = f(x(u, v), y(u, v)).
\end{align*}
\]

Here:
- \( u, v \) are independent variables.
- \( x, y \) are intermediate variables. They depend on \( u, v \).
- \( w \) is the dependent variable. It depends on \( x, y \), so it also depends on \( u, v \).

In this topic we will develop the multivariable chain rule. That is, we will learn to compute \( \frac{\partial w}{\partial u} \) in terms of \( \frac{\partial w}{\partial x}, \frac{\partial w}{\partial y}, \frac{\partial x}{\partial u} \) etc.

27.1 Review: tangent plane approximation formula

The key to understanding the chain rule is the tangent plane approximation formula. That is, if \( w = f(x, y) \) then

\[
\Delta w \approx \left. \frac{\partial w}{\partial x} \right|_0 \Delta x + \left. \frac{\partial w}{\partial y} \right|_0 \Delta y. \tag{1}
\]

27.2 Single variable approximation and chain rule

For warmup, we’ll derive the single variable chain rule from the tangent line approximation. If \( y = f(x) \) we have the approximation formula

\[
\Delta y \approx \frac{dy}{dx} \Delta x.
\]

If \( x \) is a function of \( u \) then we can divide the approximation formula by \( \Delta u \) to get

\[
\frac{\Delta y}{\Delta u} \approx \frac{df}{dx} \cdot \frac{\Delta x}{\Delta u}.
\]
In the limit this becomes the familiar chain rule: \( \frac{dy}{du} = \frac{dy}{dx} \cdot \frac{dx}{du} \).

**Example 27.1.** Suppose \( f(x) = x^3 \) and \( x(t) = \sin(t) \) compute \( \frac{df}{dt} \). Do this two ways  
(i) using the chain rule before replacing \( x \) by \( \sin(t) \)  
(ii) first replacing \( x \) by \( \sin(t) \)  

**answer:**  
(i) \( \frac{df}{dx} = 3x^2 \), \( \frac{dx}{dt} = \cos(t) \), so \( \frac{df}{dt} = \frac{df}{dx} \cdot \frac{dx}{dt} = 3x^2 \cos(t) = 3\sin^2(t) \cos(t) \).  
(ii) \( f(t) = x^3 = \sin^3(t) \), so \( \frac{df}{dt} = 3\sin^2(t) \cos(t) \).  

In this example, the two calculations are essentially the same. For more complicated problems it can be much easier to use the chain rule before substituting.

### 27.3 Multivariable chain rule

Again, suppose \( w = f(x, y) \) and \( x = x(u, v) \), \( y = y(u, v) \). So,  
\( u, v \) are independent variables  
\( x, y \) are intermediate variables  
\( w \) is the dependent variable.

**Chain rule.** The partial derivative of \( w \) with respect to \( u \) (with \( v \) held constant) is  
\[
\frac{\partial w}{\partial u} = \frac{\partial w}{\partial x} \cdot \frac{\partial x}{\partial u} + \frac{\partial w}{\partial y} \cdot \frac{\partial y}{\partial u}.
\]  
(2)

Likewise, the partial derivative of \( w \) with respect to \( v \) (with \( u \) held constant) is  
\[
\frac{\partial w}{\partial v} = \frac{\partial w}{\partial x} \cdot \frac{\partial x}{\partial v} + \frac{\partial w}{\partial y} \cdot \frac{\partial y}{\partial v}.
\]  
(3)

**Note.** With all the variables floating around we should be careful to say which variables are changing and which are held constant.

**Proof.** We start with the approximation formula in Equation 1 and divide by \( \Delta u \):  
\[
\frac{\Delta w}{\Delta u} = \frac{\partial w}{\partial x} \cdot \frac{\Delta x}{\Delta u} + \frac{\partial w}{\partial y} \cdot \frac{\Delta y}{\Delta u}
\]

Letting \( \Delta u \to 0 \) gives the chain rule for \( \frac{\partial w}{\partial u} \) in Equation 2. Likewise, we can prove Equation 3.

**Example 27.2.** Given \( w = x^2 + y^2 + x \), \( x = u^2v \), \( y = uv^2 \) find \( \frac{\partial w}{\partial u} \), and \( \frac{\partial w}{\partial v} \).  

**answer:** We first remind you that \( u, v \) are independent variables, \( x, y \) are intermediate variables, and \( w \) is a dependent variable.

To start, we use the given formulas to compute partial derivatives:  
\[
\frac{\partial w}{\partial x} = 2xy + 1, \quad \frac{\partial w}{\partial y} = x^2 + 2y, \quad \frac{\partial x}{\partial u} = 2uv, \quad \frac{\partial y}{\partial u} = v^2, \quad \frac{\partial x}{\partial v} = u^2, \quad \frac{\partial y}{\partial v} = 2uv.
\]
So
\[
\frac{\partial w}{\partial u} = \frac{\partial w}{\partial x} \cdot \frac{\partial x}{\partial u} + \frac{\partial w}{\partial y} \cdot \frac{\partial y}{\partial u} = (2xy + 1)(2uv) + (x^{2} + 2y)(v^{2})
\]
\[
= (2u^{2}v \cdot uv^{2} + 1)(2uv) + (u^{4}v^{2} + 2uv^{2})(v^{2})
\]
\[
= 5u^{4}v^{4} + 2uv + 2uv^{4}
\]

This case is simple enough that we can check our answer directly.

\[
w = x^{2}y + y^{2} + x = u^{5}v^{4} + u^{2}v^{4} + u^{2}v \Rightarrow \frac{\partial w}{\partial u} = 5u^{4}v^{4} + 2uv + 2uv^{4}.
\]

Notes. 1. The boxed answers above are usually sufficient. There is often no reason to substitute for \(x\) and \(y\).

2. This case was simple enough to check directly. More frequently, it will be much easier to use the chain rule as shown that to first find \(w\) in terms of \(u\) and \(v\) and then compute the partial derivatives.

**Example 27.3.** Suppose \(z = x^{2}y^{3} + x^{4}y^{5}\) with \(x = u^{2}v, \ y = u^{2} + v^{2}\) find \(\frac{\partial z}{\partial u}\).

**answer:** Use the chain rule:

\[
\frac{\partial z}{\partial u} = \frac{\partial z}{\partial x} \cdot \frac{\partial x}{\partial u} + \frac{\partial z}{\partial y} \cdot \frac{\partial y}{\partial u} = (2xy^{3} + 4x^{3}y^{3})(2uv) + (3x^{2}y^{2} + 5x^{4}y^{4})(2u).
\]

**27.4 A special case example**

**Example 27.4.** The temperature in space varies and is given by the function \(T(x, y, z)\). An ant crawls along a wire. Its path is described by \(\mathbf{r}(t) = \langle x(t), y(t), z(t) \rangle\). What is the rate of change of temperature experienced by the ant. (You might not care but the ant certainly does.)

Suppose that \(T(x, y, z) = x^{2} + y^{3} + z^{4}\), and \((x(t), y(t), z(t)) = (\cos(t), t, \sin(t))\).

**answer:** The chain in this case is

\[
t \leadsto (x, y, z) \leadsto T
\]

Since \(T(t)\) is a function of one variable we use normal derivative notation

\[
\frac{dT}{dt} = \frac{\partial w}{\partial x} \cdot \frac{dx}{dt} + \frac{\partial T}{\partial y} \cdot \frac{dy}{dt} + \frac{\partial T}{\partial z} \cdot \frac{dz}{dt}
\]
\[
= 2x \cdot (-\sin(t)) + 3y^{2} \cdot 1 + 4z^{3} \cdot \cos(t).
\]

We leave our answer in this implicit (‘mixed’) form. If needed it could be written out all in terms of \(t\).
27.4.1 Writing in terms of the gradient

In the previous example, we can write the curve in vector form \( \mathbf{r}(t) = (x(t), y(t), z(t)) \). Also, as a function of \((x, y, z)\) the gradient of \( T \) is

\[
\nabla T(x, y, z) = \begin{bmatrix} \frac{\partial T}{\partial x} & \frac{\partial T}{\partial y} & \frac{\partial T}{\partial z} \end{bmatrix}.
\]

Thus, the derivative \( \frac{dT}{dt} \) can be written

\[
\frac{dT}{dt} = \frac{\partial T}{\partial x} \frac{dx}{dt} + \frac{\partial T}{\partial y} \frac{dy}{dt} + \frac{\partial T}{\partial z} \frac{dz}{dt} = \begin{bmatrix} \frac{\partial T}{\partial x} & \frac{\partial T}{\partial y} & \frac{\partial T}{\partial z} \end{bmatrix} \cdot \begin{bmatrix} \frac{dx}{dt} & \frac{dy}{dt} & \frac{dz}{dt} \end{bmatrix} = \nabla T \cdot \frac{d\mathbf{r}}{dt}.
\]

27.4.2 The gradient is perpendicular to the level surfaces revisited

We showed this in an earlier topic. If you look at it again you’ll see that the argument there hinged on the approximation formula. We’ll show this now using the chain rule –of course, the chain rule hinged on the approximation formula.

Claim. Suppose \( w = f(x, y, z) \) and \( P_0 \) is on the level surface \( f(x, y, z) = c \). The \( \nabla f(P_0) \) is perpendicular to this level surface.

Proof. Draw any curve on the level surface \( \mathbf{r}(t) = (x(t), y(t), z(t)) \) such that \( \mathbf{r}(0) = P_0 \). Since the curve is on the level surface we have

\[
f(x(t), y(t), z(t)) = c
\]

So, the derivative with respect to \( t \) is 0. That is,

\[
\frac{df}{dt} = \nabla f(P_0) \cdot \mathbf{r}'(0) = 0.
\]

We have shown that the gradient \( \nabla w(P_0) \) is perpendicular to any curve on the level surface through \( P_0 \). QED

27.5 Ambiguous notation

Often you have to figure out the dependent and independent variables from context. Thermodynamics is a big culprit here:

Variables: \( P, T, V, U, S \). Any two can be taken to be independent and the others are functions of those two.

We will do more with this in the future.