

# Chain Rule

In this activity, we introduce the multi-variable chain rule, and we use it to compute derivatives of compositions of functions.

## The Chain Rule

Let's begin by recalling the chain rule from single variable calculus. If we have differentiable functions  $f$  and  $g$ , then we can compute the derivative of the composition as  $(f \circ g)'(x) = f'(g(x))g'(x)$ .

**Example 1.** Let  $f(x) = \sin(x)$  and  $g(x) = x^2$ . Then  $f'(x) = \cos(x)$ ,  $g'(x) = 2x$ , and we can differentiate the composition  $(f \circ g)(x) = \sin(x^2)$  using the chain rule:

$$\begin{aligned}(f \circ g)'(x) &= f'(g(x))g'(x) \\ &= \cos(g(x)) \cdot 2x \\ &= \cos(x^2) \cdot 2x.\end{aligned}$$

We can also use the chain rule to differentiate the composition  $(g \circ f)(x) = \sin^2(x)$ .

$$\begin{aligned}(g \circ f)'(x) &= g'(f(x))f'(x) \\ &= 2(f(x)) \cos(x) \\ &= 2 \sin(x) \cos(x)\end{aligned}$$

The multi-variable chain rule is similar, with the derivative matrix taking the place of the single variable derivative, so that the chain rule will involve matrix multiplication. We also need to pay extra attention to whether the composition of functions is even defined.

**Theorem 1.** Suppose  $\mathbf{f} : Y \subset \mathbb{R}^p \rightarrow \mathbb{R}^n$  and  $\mathbf{g} : X \subset \mathbb{R}^m \rightarrow \mathbb{R}^p$  are defined on open sets  $Y \subset \mathbb{R}^p$  and  $X \subset \mathbb{R}^m$ , respectively. Suppose that  $\mathbf{g}(X) \subset Y$ , so the image of  $\mathbf{g}$  is contained in the domain of  $\mathbf{f}$ . Suppose further that  $\mathbf{g}$  is differentiable at some point  $\mathbf{x}_0 \in X$ , and that  $\mathbf{f}$  is differentiable at  $\mathbf{y}_0 = \mathbf{g}(\mathbf{x}_0) \in Y$ .

Then the composition  $\mathbf{f} \circ \mathbf{g}$  is differentiable at  $\mathbf{x}_0$ , and

$$\begin{aligned}D(\mathbf{f} \circ \mathbf{g})(\mathbf{x}_0) &= D\mathbf{f}(\mathbf{y}_0)D\mathbf{g}(\mathbf{x}_0) \\ &= D\mathbf{f}(\mathbf{g}(\mathbf{x}_0))D\mathbf{g}(\mathbf{x}_0).\end{aligned}$$

---

Learning outcomes:  
Author(s):

## Chain Rule

Although the conditions sound complicated, essentially they're just requiring that all of the derivatives mentioned actually exist. Note the similarities to the single variable chain rule.

**Proof** super great proof of chain rule



## A Special Case

$\mathbb{R} \rightarrow \mathbb{R}^n \rightarrow \mathbb{R}$

## Examples

Examples of using the chain rule, including polar coords conversion.

## Conclusion