# 10. Statement of Chain Rule

Bookmark this page

Lecture due Oct 5, 2021 20:30 IST

**Synthesize**

**How to state the chain rule in general?**

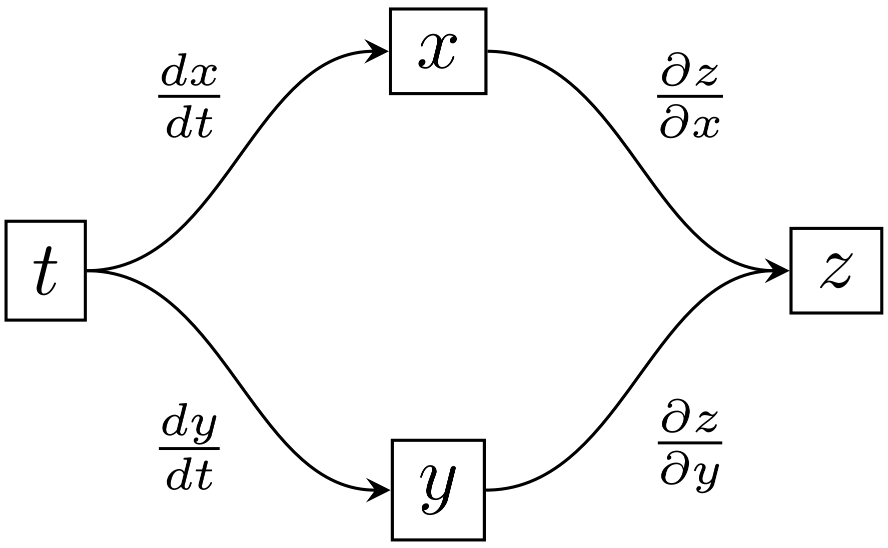
There are many "chain rules" in multivariable calculus, because of the many different possibilities for the number of input/output variables. Let's look at the statement in some of the most common cases that we already covered. At the end of this page we have included the most generalized statement of the chain rule.

**Example: From 1 variable to 2 variables to 1 variable**

Let's look at how the chain rule manifests when we have a quantity $z$ that depends on two variables, $x$ and $y$, which each depend on a single variable, $t$. The chain rule says that the (single-variable) derivative of $z$ is given by

$$\frac{dz}{dt} = \frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt} \tag{6.205}$$

How to remember this formula? First, we imagine the following diagram which shows the dependencies between the variables.




Then for each one of the paths from $t$ to $z$, we have a term in the formula for $\frac{dz}{dt}$. By adding up the relevant partial derivatives, we obtain the total expression for $\frac{dz}{dt}$.

We have seen an example of this situation, where there was an output $V = xy^2$ and $x$ and $y$ each depended on a variable $t$.

**Example: From 2 variables to 2 variables to 1 variable**

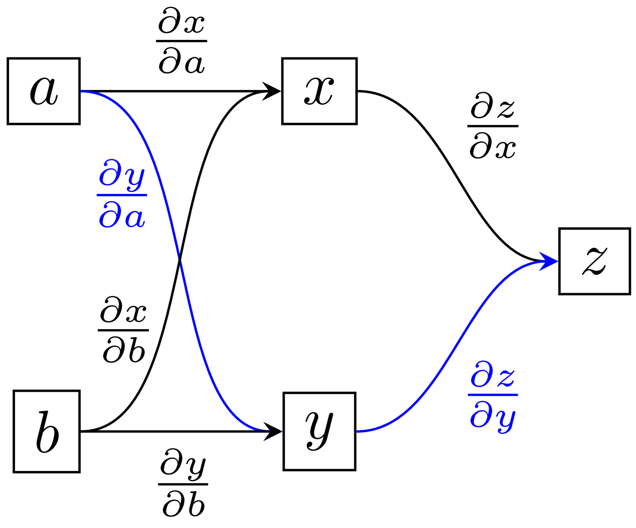
Now let's look at how the chain rule manifests when we have a quantity $z$ that depends on two variables, $x$ and $y$, which each depend on two variables $a$ and $b$. The chain rule says that the partial derivatives of $z$ with respect to $a$ and $b$ are given by:

$$\frac{\partial z}{\partial a} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial a} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial a} \tag{6.206}$$

$$\frac{\partial z}{\partial b} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial b} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial b} \tag{6.207}$$

We have highlighted the $\frac{\partial z}{\partial y}\frac{\partial y}{\partial a}$ term in blue to emphasize its connection to the following diagram:

Calculator Hide Notes




The term $\frac{\partial z}{\partial y}\frac{\partial y}{\partial a}$ arises because of the highlighted path from $a$ to $z$.

In general we see the same pattern as before: to obtain the partial derivative of an output variable with respect to one of the input variables, we sum up each of the paths from input to output and multiply the corresponding partial derivatives.

We have seen an example of this situation, where there was an output $V = xy^2$ and $x$ and $y$ each depended on two variables $a$ and $b$.

**More generally: From $n$ variables to $m$ variables to $1$ variable**

Both of the special cases stated above are examples of the following more general statement of the chain rule.

> **Theorem** (Chain Rule) Suppose $z$ is a quantity that depends on $m$ variables, $y_1, \ldots, y_m$ and each of the $y$'s depends on the $n$ variables $x_1, \ldots, x_n$. Then the derivatives of $z$ are given by:
>
> $$\text{For } 1 \leq i \leq n, \qquad \frac{\partial z}{\partial x_i} = \sum_{j=1}^{m} \frac{\partial z}{\partial y_j}\frac{\partial y_j}{\partial x_i} \tag{6.208}$$

When the theorem is written this way, it is often not clear how to apply it to a given problem. But you can always draw a diagram such as the ones pictured above, and then the theorem just says that the derivative of the output, with respect to one of the inputs, is found by summing up each of the $m$ paths from the input to the output, with each path weighted by the product of the appropriate partial derivatives.

**What about more than one output?**

In each of the above examples, there was just one "output" variable $z$. If there are two or more output variables, say $z_1$ and $z_2$, then you can use the above theorem for each variable separately. The overall rule is the same: we get the derivative of a given output with respect to a given input by summing up all the paths from the chosen input to the chosen output, weighted by the product of the appropriate partial derivatives.

**How to draw the diagrams?**

The diagrams above are sometimes known as “dependency graphs." They encapsulate how each variable depends on the other variables. Once you have a dependency graph, it is straightforward to use the chain rule, since we can look at each of the paths in the diagram and write down a corresponding term for the derivative.

In a recitation video, you will see an example of drawing the dependency graph for a given situation. If you aren't sure how to draw the diagram, one strategy you could try is to start with the output variable, and ask yourself “what variables does this depend on?" By repeatedly asking this question until you reach the input variables, you can fill out the entire diagram.

**Generalized Chain Rule**

Calculator Hide Notes

We do not expect you to know the most general statement of the chain rule. But for those who are curious, see the following.

**Generalized Chain Rule**

In its most general form, the Chain Rule is best stated in terms of Jacobian matrices. Namely, if we have transformations $T$ and $W$ taking input vectors to output vectors:

$$\underbrace{\vec{x}}_{n \text{ variables}} \xrightarrow{T} \underbrace{\vec{y}}_{m \text{ variables}} \xrightarrow{W} \underbrace{\vec{z}}_{k \text{ variables}}$$

Then the **generalized chain rule** says the Jacobian matrix of the transform $W \circ T$ is given by the product of the Jacobian of $W$ and the Jacobian of $T$.

$$\textbf{(Generalized chain rule)} \qquad \text{Jacobian of } W \circ T = \text{Jacobian of } W \cdot \text{Jacobian of } T \qquad (6.209)$$

Letting $J_T$ stand for the Jacobian matrix of the transformation $T$ (at the appropriate point), the generalized chain rule says:

$$\underbrace{J_{W \circ T}}_{k \times n \text{ matrix}} = \underbrace{J_W}_{k \times m \text{ matrix}} \cdot \underbrace{J_T}_{m \times n \text{ matrix}} \qquad (6.210)$$

Unpacking the statement slightly, the generalized chain rule just says that the linear approximation of a composite transformation can be done one transformation at a time. In other words, to approximate the value of $\vec{z}$ for a given $\vec{x}$, we can first approximate the value of $\vec{y}$ (using the Jacobian of $T$) and then use this $\vec{y}$ to make an approximation for the resulting value of $\vec{z}$ (using the Jacobian of $W$). The matrix multiplication shows up in the chain rule because matrix multiplication corresponds to composition (applying one transformation, and then the other).

Another way of understanding the role of matrices is to recognize all of the formulas on this page as "hidden dot products." By scrutinizing each of these dot products, one can package them all into one formula, leading to the matrix product formulation above.

Hide

## Volume of sphere

2/2 points (graded)

Suppose the radius of a sphere is controlled by two parameters $a$ and $b$. Suppose that, when $a = 3$ and $b = 5$, we have $r = 5$, $\frac{\partial r}{\partial a} = -1$ and $\frac{\partial r}{\partial b} = 2$. What are the partial derivatives of the volume of the sphere when $a = 3$ and $b = 5$?

$\frac{\partial V}{\partial a} =$ -100*pi ✔ **Answer:** -100*pi

$\frac{\partial V}{\partial b} =$ 200*pi ✔ **Answer:** 200*pi

? INPUT HELP

**Solution:**

To find $\frac{\partial V}{\partial a}$ we ask the question "how does $V$ depend on $a$?" First, $V$ depends on $r$ via $\frac{dV}{dr}$, and then $r$ depends on $a$ via $\frac{\partial r}{\partial a}$. Therefore, the chain rule gives us:

$$\frac{\partial V}{\partial a} = \frac{dV}{dr}\frac{\partial r}{\partial a}$$

Calculator Hide Notes

$$\frac{\partial V}{\partial b} = \frac{dV}{dr}\frac{\partial r}{\partial b} \tag{6.212}$$

We find $\frac{dV}{dr}$ from $V = \frac{4}{3}\pi r^3$, so $\frac{dV}{dr} = 4\pi r^2$. We were given $r = 5$, so $\frac{dV}{dr} = 100\pi$. We were also given $\frac{\partial r}{\partial a} = -1$, so taking the product we obtain $\frac{\partial V}{\partial a} = -100\pi$. Similarly we obtain $\frac{\partial V}{\partial b} = 200\pi$.

In fact we did not need to know the values $a = 3$ and $b = 5$, since the entire question is about rates of change.

Submit You have used 1 of 5 attempts

Answers are displayed within the problem

## 10. Statement of Chain Rule

Hide Discussion

**Topic:** Unit 5: Curves and Surfaces / 10. Statement of Chain Rule

**Add a Post**

Show all posts by recent activity

There are no posts in this topic yet.

×

Previous Next



# edX

About
Affiliates
edX for Business
Open edX
Careers
News

# Legal

Terms of Service & Honor Code
Privacy Policy
Accessibility Policy
Trademark Policy
Sitemap

# Connect

Blog

Calculator Hide Notes