Multivariate chain rule

Donald Campbell

Summary

The multivariate chain rule is used in calculus to differentiate a function when its variables depend on other variables. It shows how the change in one variable affects the whole function by considering how the intermediate variables change.

Before reading this guide, it is recommended that you read Guide: Introduction to partial differentiation.

The **multivariate chain rule** allows you to compute derivatives of composite functions (functions of functions) involving multiple variables. It allows you to track how changes in **independent variables** influence a final quantity through intermediate variables called **dependent variables**.

Dependent and independent variables

When a function **depends** on a specific variable, this means that the function is written with the variable in its rule. For example, the function $f(x, y) = x^2 + y^2$ depends on x and y as the function is written in terms of x and y.

A **multivariate (or multivariable) function** is a function that depends on more than one variable.

Consider a multivariate function $w = f(x, y, z) = 2x^2 + xy + 3xz$ where

$$\begin{aligned} x(s,t) &= s^2 - t \\ y(s,t) &= s^2 + t \\ z(s,t) &= 2s + t^2 \end{aligned}$$

The variables x, y and z are referred to as **dependent variables** as they are functions that depend on (can be written in terms of) the variables s and t. Even though x, y and z are referred to as dependent variables, they are actually functions of the variables s and t. In this case, the dependent variables can be written in functional form as $x = x(s, t)$, $y = y(s, t)$ and $z = z(s, t)$ to emphasize dependency.

The variables s and t are referred to as **independent variables** as they do not depend on any other variables.

Occasionally, dependent variables may be referred to as **intermediate variables** since they connect the independent variables s and t to the final quantity w.

Functions of one dependent variable with one independent variable (recap)

When you have a function $y = f(x)$ where x depends on another variable t, the chain rule tells you how to differentiate y with respect to t.

Chain rule for functions of a single independent variable

Let $y = f(x)$ where f is a differentiable function of x. Also let $x = x(t)$ be a differentiable function of t.
The function y can then be differentiated with respect to t according to

$$\frac{\mathrm{d}y}{\mathrm{d}t} = \frac{\mathrm{d}y}{\mathrm{d}x}\frac{\mathrm{d}x}{\mathrm{d}t}$$

Important

It is important to use d here instead of ∂ since x and y are functions of a single independent variable t. In contrast, when working with functions of multiple variables, you would use a curly ∂.

It is useful to see how this works through an example. The chain rule helps you to differentiate composite functions by breaking them into an outer and inner function. You need to differentiate each part separately and then multiply the results.

Example 1

Let $y = f(x) = x^5$ where $x = 3t^2 + 2$. Use the chain rule to find $\frac{\mathrm{d}y}{\mathrm{d}t}$.
First, calculate the derivatives required for the chain rule.

$$\frac{\mathrm{d}y}{\mathrm{d}x} = 5x^4$$
$$\frac{\mathrm{d}x}{\mathrm{d}t} = 6t$$

Apply the chain rule then substitute for $x = 3t^2 + 2$.

$$\frac{\mathrm{d}y}{\mathrm{d}t} = \frac{\mathrm{d}y}{\mathrm{d}x}\frac{\mathrm{d}x}{\mathrm{d}t}$$
$$\frac{\mathrm{d}y}{\mathrm{d}t} = (5x^4)(6t)$$
$$\frac{\mathrm{d}y}{\mathrm{d}t} = \left(5(3t^2+2)^4\right)(6t)$$
$$\frac{\mathrm{d}y}{\mathrm{d}t} = 30t(3t^2+2)^4$$

The chain rule also extends to functions of two dependent variables.

Functions of two dependent variables with one independent variable

Suppose you have a function $z = f(x, y)$.

Here, z depends on two dependent variables x and y, however both x and y are functions of a single independent variable t.

$$x = x(t) \quad \text{and} \quad y = y(t)$$

This means that, even though z initially looks like a function of two dependent variables x and y, it really depends on a single independent variable t, and you could write z entirely in terms of t.

$$z = f\left(x(t), y(t)\right)$$

The goal is to find $\frac{\mathrm{d}z}{\mathrm{d}t}$.
You cannot differentiate $z = f(x, y)$ directly with respect to t because z depends intermedi-

ately on both x and y. Instead, you need to use the multivariate chain rule.

Multivariate chain rule for one independent variable

Let $z = f(x, y)$ where f is a differentiable function of x and y. Also let $x = x(t)$ and $y = y(t)$ be differentiable functions of t.
The function z can then be differentiated with respect to t according to

$$\frac{\mathrm{d}z}{\mathrm{d}t} = \frac{\partial z}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\partial z}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}t}$$

Important

The left-most derivative $\frac{\mathrm{d}z}{\mathrm{d}t}$ is written with a straight d as $z = z(t)$ is considered a function of t only at this stage. Here, z is written using the expressions for x and y in terms of t.
Derivatives such as $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$ are written with a curly ∂ as z is a function that depends on two variables x and y.
Derivatives such as $\frac{\mathrm{d}x}{\mathrm{d}t}$ and $\frac{\mathrm{d}y}{\mathrm{d}t}$ are written with a straight d as x is a function that depends only on a single variable t.

As t changes, both dependent variables x and y change as they are both functions of the independent variable t. This causes z to change as it is a function of variables x and y which are changing.

The total rate of change of z depends on:

- The product $\frac{\partial z}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}t}$ which is how sensitive z is to x, scaled by how sensitive x is to t.
- The product $\frac{\partial z}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}t}$ which is how sensitive z is to y, scaled by how sensitive y is to t.

The chain rule combines these two effects to get the required derivative $\frac{\mathrm{d}z}{\mathrm{d}t}$.

Example 2

Let $z = f(x, y) = x^3 + y^3$ where $x = 3t^2$ and $y = \cos(t)$. Use the multivariate chain rule to find $\frac{\mathrm{d}z}{\mathrm{d}t}$.
First, calculate the derivatives required for the multivariate chain rule.

$$\begin{aligned}
\frac{\partial z}{\partial x} &= 3x^2 \\
\frac{\partial z}{\partial y} &= 3y^2 \\
\frac{\mathrm{d}x}{\mathrm{d}t} &= 6t \\
\frac{\mathrm{d}y}{\mathrm{d}t} &= -\sin(t)
\end{aligned}$$

Apply the multivariate chain rule then substitute for $x = 3t^2$ and $y = \cos(t)$.

$$\begin{aligned}
\frac{\mathrm{d}z}{\mathrm{d}t} &= \frac{\partial z}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\partial z}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}t} \\
\frac{\mathrm{d}z}{\mathrm{d}t} &= (3x^2)(6t) + (3y^2)(-\sin(t)) \\
\frac{\mathrm{d}z}{\mathrm{d}t} &= 3(3t^2)^2(6t) + 3(\cos(t))^2(-\sin(t)) \\
\frac{\mathrm{d}z}{\mathrm{d}t} &= 162t^5 - 3\sin(t)\cos^2(t)
\end{aligned}$$

Important

You should make sure that your final answer is always expressed in terms of any independent variables. Your final answer should not include any intermediate variables.

Functions of two dependent variables with two independent variables

Suppose you have a function $z = f(x, y)$.

Here, z depends on two dependent variables x and y, however both x and y are functions of two independent variables s and t.

$$x = x(s, t) \quad \text{and} \quad y = y(s, t)$$

This means that, even though z initially looks like a function of two dependent variables x

and y, it really depends on two independent variables s and t, and you could write z entirely in terms of s and t.

$$z = f\left(x(s,t), y(s,t)\right)$$

Since there are now two independent variables, the goal is to find both $\dfrac{\partial z}{\partial s}$ and $\dfrac{\partial z}{\partial t}$.

You cannot differentiate $z = f(x,y)$ directly with respect to s or t because z depends intermediately on both x and y. Instead, you need to use the multivariate chain rule.

Multivariate chain rule for two independent variables

Let $z = f(x,y)$ where f is a differentiable function of x and y. Also let $x = x(s,t)$ and $y = y(s,t)$ be differentiable functions of s and t.
The function z can then be differentiated with respect to s and t with

$$\frac{\partial z}{\partial s} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial s}$$
$$\frac{\partial z}{\partial t} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial t}$$

Important

The left-most derivatives $\dfrac{\partial z}{\partial s}$ and $\dfrac{\partial z}{\partial t}$ are written with a curly ∂ as $z = z(s,t)$ is considered a function of both s and t at this stage. Here, z is written using the expressions for x and y in terms of s and t.
Notice that the $\dfrac{\partial x}{\partial s}$, $\dfrac{\partial x}{\partial t}$, $\dfrac{\partial y}{\partial s}$ and $\dfrac{\partial y}{\partial t}$ derivatives are now written with a curly ∂ instead of a straight d as x and y are functions that now depend on two independent variables s and t. This requires partial derivatives instead of ordinary derivatives.

The multivariate chain rule works in two steps:

- To find the partial derivative of z with respect to s, combine the sensitivity of z to changes in x and y and the sensitivity of both x and y to changes in s.
- To find the partial derivative of z with respect to t, combine the sensitivity of z to changes in x and y and the sensitivity of both x and y to changes in t.

Example 3

Let $z = f(x, y) = x^2 + 3xy$ where $x = s + t^2$ and $y = s^2 - t$. Use the multivariate chain rule to find $\frac{\partial z}{\partial s}$ and $\frac{\partial z}{\partial t}$.
First, calculate the derivatives required for the multivariate chain rule.

$$\begin{aligned}
\frac{\partial z}{\partial x} &= 2x + 3y \\
\frac{\partial z}{\partial y} &= 3x \\
\frac{\partial x}{\partial s} &= 1 \\
\frac{\partial x}{\partial t} &= 2t \\
\frac{\partial y}{\partial s} &= 2s \\
\frac{\partial y}{\partial t} &= -1
\end{aligned}$$

Apply the multivariate chain rule equation then substitute for $x = s + t^2$ and $y = s^2 - t$.

Calculate the partial derivative of z with respect to s.

$$\begin{aligned}
\frac{\partial z}{\partial s} &= \frac{\partial z}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial s} \\
\frac{\partial z}{\partial s} &= (2x + 3y)(1) + (3x)(2s) \\
\frac{\partial z}{\partial s} &= \left(2(s + t^2) + 3(s^2 - t)\right)(1) + \left(3(s + t^2)\right)(2s) \\
\frac{\partial z}{\partial s} &= 9s^2 + 2t^2 + 6st^2 + 2s - 3t
\end{aligned}$$

Calculate the partial derivative of z with respect to t.

$$\begin{aligned}
\frac{\partial z}{\partial t} &= \frac{\partial z}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial t} \\
\frac{\partial z}{\partial t} &= (2x + 3y)(2t) + (3x)(-1) \\
\frac{\partial z}{\partial t} &= \left(2(s + t^2) + 3(s^2 - t)\right)(2t) + \left(3(s + t^2)\right)(-1) \\
\frac{\partial z}{\partial t} &= 4t^3 - 9t^2 + 6s^2t - 3s + 4st
\end{aligned}$$

Generalized multivariate chain rule

Suppose you have a function $w = f(x_1, x_2, \dots, x_n)$.

Here, w depends on n dependent variables called $x_1, x_2, \dots, x_n$ however each of these variables is a function of m independent variables $t_1, t_2, \dots, t_m$.

$$\begin{aligned} x_1 &= x_1(t_1, t_2, \dots, t_m) \\ x_2 &= x_2(t_1, t_2, \dots, t_m) \\ \vdots &= \vdots \\ x_n &= x_n(t_1, t_2, \dots, t_m) \end{aligned}$$

The previous cases investigated were:

$$\begin{array}{lllll} n = 1 & m = 1 & y = f(x) & x = x(t) & \\ n = 2 & m = 1 & z = f(x, y) & x = x(t) & y = y(t) \\ n = 2 & m = 2 & z = f(x, y) & x = x(s, t) & y = y(s, t) \end{array}$$

This multivariate chain rule can be generalized as follows.

Generalized multivariate chain rule

Let $w = f(x_1, x_2, \dots, x_n)$ where f is a differentiable function of $x_1, x_2, \dots, x_n$. Also let the following dependent variables be differentiable functions of $t_1, t_2, \dots, t_m$.

$$\begin{aligned} x_1 &= x_1(t_1, t_2, \dots, t_m) \\ x_2 &= x_2(t_1, t_2, \dots, t_m) \\ \vdots &= \vdots \\ x_n &= x_n(t_1, t_2, \dots, t_m) \end{aligned}$$

The function w can then be differentiated with respect to $t_1, t_2, \dots, t_m$ with

$$\begin{aligned} \frac{\partial w}{\partial t_1} &= \frac{\partial w}{\partial x_1}\frac{\partial x_1}{\partial t_1} + \frac{\partial w}{\partial x_2}\frac{\partial x_2}{\partial t_1} + \dots + \frac{\partial w}{\partial x_n}\frac{\partial x_n}{\partial t_1} \\ \frac{\partial w}{\partial t_2} &= \frac{\partial w}{\partial x_1}\frac{\partial x_1}{\partial t_2} + \frac{\partial w}{\partial x_2}\frac{\partial x_2}{\partial t_2} + \dots + \frac{\partial w}{\partial x_n}\frac{\partial x_n}{\partial t_2} \\ \vdots &= \vdots \\ \frac{\partial w}{\partial t_m} &= \frac{\partial w}{\partial x_1}\frac{\partial x_1}{\partial t_m} + \frac{\partial w}{\partial x_2}\frac{\partial x_2}{\partial t_m} + \dots + \frac{\partial w}{\partial x_n}\frac{\partial x_n}{\partial t_m} \end{aligned}$$

! Important

A particular case exists where the variables x_i each depend on only one independent variable $x_1 = x_1(t), x_2 = x_2(t), \dots, x_n = x_n(t)$. In this case, a straight d should be used instead of a curly ∂.

$$\frac{\mathrm{d}w}{\mathrm{d}t} = \frac{\partial w}{\partial x_1}\frac{\mathrm{d}x_1}{\mathrm{d}t} + \frac{\partial w}{\partial x_2}\frac{\mathrm{d}x_2}{\mathrm{d}t} + \dots + \frac{\partial w}{\partial x_n}\frac{\mathrm{d}x_n}{\mathrm{d}t}$$

Similarly, if there is only one dependent variable x, then a straight d should be used instead of a curly ∂ for the derivatives of w with respect to x.

Example 4

Consider the case with $n = 3$ dependent variables and $m = 2$ independent variables. Let $w = f(x, y, z) = x^2 - 3xz + 2xy$ where

$$x(s,t) = s^2 - t$$
$$y(s,t) = s\cos(t)$$
$$z(s,t) = t^2 - s$$

Use the multivariate chain rule to find $\frac{\partial w}{\partial s}$ and $\frac{\partial w}{\partial t}$.
The following multivariate chain rule equations will be useful in determining these partial derivatives for w. These have been derived from the generalized chain rule equation above.

$$\frac{\partial w}{\partial s} = \frac{\partial w}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial s} + \frac{\partial w}{\partial z}\frac{\partial z}{\partial s}$$
$$\frac{\partial w}{\partial t} = \frac{\partial w}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial w}{\partial z}\frac{\partial z}{\partial t}$$

First, calculate the derivatives required for the multivariate chain rule.

$$\frac{\partial w}{\partial x} = 2x - 3z + 2y$$
$$\frac{\partial w}{\partial y} = 2x$$
$$\frac{\partial w}{\partial z} = -3x$$
$$\frac{\partial x}{\partial s} = 2s$$
$$\frac{\partial x}{\partial t} = -1$$

Tree representation of the multivariate chain rule

There is a useful way for determining the required form of the multivariate chain rule without needing to remember the generalized version above. The following diagram is a **tree representation** of the dependencies between variables in a multivariate function.

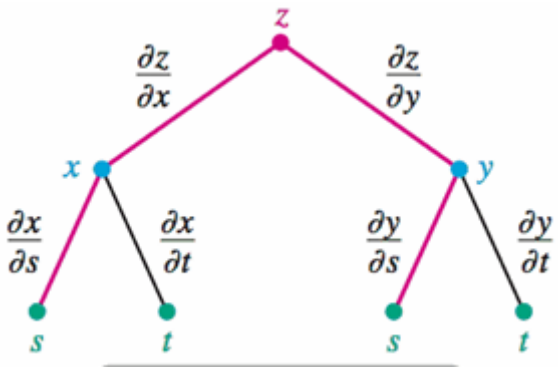

Figure 1: Tree diagram for the multivariate chain rule.

At the top of the tree is the function $w = f(x, y, z)$, which depends on three intermediate variables x, y and z. Each of these three intermediate variables depend on independent variables s and t. The tree shows all possible paths from w to the independent variables s and t.

Each path represents a **chain of differentiation** (a product of derivatives) when applying the chain rule. All paths are combined by summing these chains at the end to find the complete derivative of w with respect to the relevant independent variable.

To find $\dfrac{\partial w}{\partial s}$, follow all paths leading from z to s.

- Path via x: $\dfrac{\partial w}{\partial x}\dfrac{\partial x}{\partial s}$
- Path via y: $\dfrac{\partial w}{\partial y}\dfrac{\partial y}{\partial s}$
- Path via z: $\dfrac{\partial w}{\partial z}\dfrac{\partial z}{\partial s}$
- Combine all terms: $\dfrac{\partial w}{\partial s} = \dfrac{\partial w}{\partial x}\dfrac{\partial x}{\partial s} + \dfrac{\partial w}{\partial y}\dfrac{\partial y}{\partial s} + \dfrac{\partial w}{\partial z}\dfrac{\partial z}{\partial s}$

To find $\dfrac{\partial w}{\partial t}$, follow all paths leading from z to s.

- Path via x: $\dfrac{\partial w}{\partial x}\dfrac{\partial x}{\partial t}$

- Path via y: $\dfrac{\partial w}{\partial y}\dfrac{\partial y}{\partial t}$
- Path via z: $\dfrac{\partial w}{\partial z}\dfrac{\partial z}{\partial t}$
- Combine all terms: $\dfrac{\partial w}{\partial t} = \dfrac{\partial w}{\partial x}\dfrac{\partial x}{\partial t} + \dfrac{\partial w}{\partial y}\dfrac{\partial y}{\partial t} + \dfrac{\partial w}{\partial z}\dfrac{\partial z}{\partial t}$

> **! Important**
>
> Remember that if a variable has only a single dependency on another variable (no separate branch), a straight d should be used in the derivative instead of a curly ∂.

This visual aid can be useful as an alternative method to remembering the generalized equation.

Quick check problems

Consider the function $z = f(x, y) = xy^2 + 3x^2$ where $x = t^2 - \cos(2t)$ and $y = 2t + 5\sin(t)$.

How many dependent variables and independent variables are there?

Determine whether the following statement is true or false.

The required form of the multivariate chain rule is $\dfrac{\mathrm{d}z}{\mathrm{d}t} = \dfrac{\partial z}{\partial x}\dfrac{\mathrm{d}x}{\mathrm{d}t} + \dfrac{\partial z}{\partial y}\dfrac{\mathrm{d}y}{\mathrm{d}t}$.

Find the derivative of z with respect to x.

Find the derivative of x with respect to t.

Find the derivative of z with respect to y.

Find the derivative of y with respect to t.

Further reading

For more questions on the subject, please go to Questions: Multivariate chain rule.

For a way to differentiate implicit functions of more than one variable, please see Guide: Multivariate implicit differentiation.

Version history

v1.0: initial version created 05/25 by Donald Campbell as part of a University of St Andrews VIP project.

This work is licensed under CC BY-NC-SA 4.0.