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*](introtopartialdifferentiation.qmd)*.*

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 depends on and as the function is written in terms of and .

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

Consider a multivariate function where

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

The variables and 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 and to the final quantity .

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

When you have a function where depends on another variable , the chain rule tells you how to differentiate with respect to .

|  |
| --- |
| Chain rule for functions of a single independent variable |
| Let where is a differentiable function of . Also let be a differentiable function of .  The function can then be differentiated with respect to according to |

|  |
| --- |
| Important |
| It is important to use here instead of since and are functions of a single independent variable . 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 where . Use the chain rule to find .  First, calculate the derivatives required for the chain rule.  Apply the chain rule then substitute for . |

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 .

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

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

The goal is to find .

You cannot differentiate directly with respect to because depends intermediately on both and . Instead, you need to use the multivariate chain rule.

|  |
| --- |
| Multivariate chain rule for one independent variable |
| Let where is a differentiable function of and . Also let and be differentiable functions of .  The function can then be differentiated with respect to according to |

|  |
| --- |
| Important |
| The left-most derivative is written with a straight as is considered a function of only at this stage. Here, is written using the expressions for and in terms of .  Derivatives such as and are written with a curly as is a function that depends on two variables and .  Derivatives such as and are written with a straight as is a function that depends only on a single variable . |

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

The total rate of change of depends on:

* The product which is how sensitive is to , scaled by how sensitive is to .
* The product which is how sensitive is to , scaled by how sensitive is to .

The chain rule combines these two effects to get the required derivative .

|  |  |
| --- | --- |
|  | **Example 2**  Let where and . Use the multivariate chain rule to find .  First, calculate the derivatives required for the multivariate chain rule.  Apply the multivariate chain rule then substitute for and . |

|  |
| --- |
| 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 .

Here, depends on two dependent variables and , however both and are functions of two independent variables and .

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

Since there are now two independent variables, the goal is to find both and .

You cannot differentiate directly with respect to or because depends intermediately on both and . Instead, you need to use the multivariate chain rule.

|  |
| --- |
| Multivariate chain rule for two independent variables |
| Let where is a differentiable function of and . Also let and be differentiable functions of and .  The function can then be differentiated with respect to and with |

|  |
| --- |
| Important |
| The left-most derivatives and are written with a curly as is considered a function of both and at this stage. Here, is written using the expressions for and in terms of and .  Notice that the , , and derivatives are now written with a curly instead of a straight as and are functions that now depend on two independent variables and . This requires partial derivatives instead of ordinary derivatives. |

The multivariate chain rule works in two steps:

* To find the partial derivative of with respect to , combine the sensitivity of to changes in and and the sensitivity of both and to changes in .
* To find the partial derivative of with respect to , combine the sensitivity of to changes in and and the sensitivity of both and to changes in .

|  |  |
| --- | --- |
|  | **Example 3**  Let where and . Use the multivariate chain rule to find and .  First, calculate the derivatives required for the multivariate chain rule.  Apply the multivariate chain rule equation then substitute for and .  Calculate the partial derivative of with respect to .  Calculate the partial derivative of with respect to . |

# Generalized multivariate chain rule

Suppose you have a function .

Here, depends on dependent variables called however each of these variables is a function of independent variables .

The previous cases investigated were:

This multivariate chain rule can be generalized as follows.

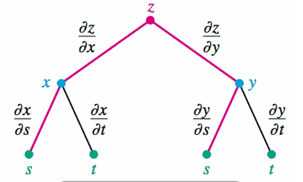
|  |
| --- |
| Generalized multivariate chain rule |
| Let where is a differentiable function of . Also let the following dependent variables be differentiable functions of .  The function can then be differentiated with respect to with |

|  |
| --- |
| Important |
| A particular case exists where the variables each depend on only one independent variable . In this case, a straight should be used instead of a curly .  Similarly, if there is only one dependent variable , then a straight should be used instead of a curly for the derivatives of with respect to . |

|  |  |
| --- | --- |
|  | **Example 4**  Consider the case with dependent variables and independent variables.  Let where  Use the multivariate chain rule to find and .  The following multivariate chain rule equations will be useful in determining these partial derivatives for . These have been derived from the generalized chain rule equation above.  First, calculate the derivatives required for the multivariate chain rule.  Apply the multivariate chain rule equation then substitute for , and .  Calculate the partial derivative of with respect to .  Calculate the partial derivative of with respect to . |

# 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.



Tree diagram for the multivariate chain rule.

At the top of the tree is the function , which depends on three intermediate variables , and . Each of these three intermediate variables depend on independent variables and . The tree shows all possible paths from to the independent variables and .

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 with respect to the relevant independent variable.

To find , follow all paths leading from to .

* Path via :
* Path via :
* Path via :
* Combine all terms:

To find , follow all paths leading from to .

* Path via :
* Path via :
* Path via :
* Combine all terms:

|  |
| --- |
| Important |
| Remember that if a variable has only a single dependency on another variable (no separate branch), a straight 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 where and .

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 .

Find the derivative of with respect to .

Find the derivative of with respect to .

Find the derivative of with respect to .

Find the derivative of with respect to .

# Further reading

[For more questions on the subject, please go to Questions: Multivariate chain rule.](../questions/qs-multivariatechainrule.qmd)

[For a way to differentiate implicit functions of more than one variable, please see Guide: Multivariate implicit differentiation.](multivariateimplicitdifferentiation.qmd)

## 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.](https://creativecommons.org/licenses/by-nc-sa/4.0/?ref=chooser-v1)