The Chain Rule

Say we have one *initial* or *top-level* variable, and one or more *final* or *base-level* variables.

The initial variable is a *dependent* variable. Each of the final variables is an *independent* variable.

We draw a **tree diagram** where the initial variable appears at the *head* (or top level) of the tree, and all the final variables appear at the *foot* (or base level) of the tree.

Between the initial variable and the final variables, we will have one or more *layers* of *intermediate* variables. In theory, we could have any number of layers, but we will restrict our attention to problems where there are either *one* or *two* layers of intermediate variables.

At each layer of intermediate variables, we will have one or more intermediate variables.

Our tree diagram will have a separate *level* or *tier* for each *layer* of intermediate variables. All the variables for that layer appear at that level of the tree.

In general, the number of tiers in the tree diagram will be *two more than* the number of layers of intermediate variables.

In a problem where there is *one* layer of intermediate variables, we will have a **three-tier tree diagram**: The tiers are numbered from top to bottom. The first tier (i.e., the top tier) is for the initial variable. The second (or middle) tier is for the single layer of intermediate variables. The third tier (i.e., the bottom tier) is for the final variables.

In a problem where there are *two* layers of intermediate variables, we will have a **four-tier tree diagram**: The tiers are numbered from top to bottom. The first (top) tier is for the initial variable. The second tier is for the first layer of intermediate variables. The third tier is for the second layer of intermediate variables. The fourth (bottom) tier is for the final variables.

In the following discussion, the word "derivative" may refer to either an ordinary derivative or to a partial derivative.

To find the derivative of the initial variable with respect to any final variable, we employ the Chain Rule. This will give us a sum of various terms, where each term is a product of various factors.

- In a *three*-tier tree (i.e., where there is only *one* layer of intermediate variables), each term will be a product of *two* factors.
- In a *four*-tier tree (i.e., where there are *two* layers of intermediate variables), each term will be a product of *three* factors.

- In general, the number of factors in each product will be *one less than* the number of tiers in the tree, or *one more than* the number of layers of intermediate variables.

In a three-tier tree diagram, the number of terms in the sum is equal to the number of intermediate variables.

In a four-tier tree diagram, the number of terms in the sum is equal to the number of intermediate variables in the first layer *multiplied by* the number of intermediate variables in the second layer.

In either case (i.e., with either a three-tier tree or a four-tier tree), the number of terms in the sum is the number of *distinct pathways* in the tree *from* the initial variable *to* the specified final variable.

Example

Say we have a function $z = f(x,y) = x^3 + 7xy + 4y^2$. Suppose we have the parametric equations $x = 2t^2$ and $y = 5t + 1$.

We can substitute to obtain a formula for z in terms of t, as follows:
$z = (2t^2)^3 + 7(2t^2)(5t+1) + 4(5t+1)^2$
$z = 8t^6 + 14t^2(5t+1) + 4(25t^2 + 10t + 1)$
$z = 8t^6 + 70t^3 + 14t^2 + 100t^2 + 40t + 4$
$z = 8t^6 + 70t^3 + 114t^2 + 40t + 4$

We can now use the basic rules of differentiation to find the (ordinary) derivative of z with respect to t, giving us: $\frac{dz}{dt} = 48t^5 + 210t^2 + 228t + 40$.

Alternatively, it is possible to find $\frac{dz}{dt}$ without first expressing z directly in terms of t. This is based on computing four ordinary or partial derivatives:
- $\frac{\partial z}{\partial x} = 3x^2 + 7y$
- $\frac{\partial z}{\partial y} = 7x + 8y$
- $\frac{dx}{dt} = 4t$
- $\frac{dy}{dt} = 5$

We now multiply and add according to the following scheme:
$\frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt} = (3x^2 + 7y)(4t) + (7x + 8y)(5) = 12x^2 t + 28yt + 35x + 40y.$

This gives us $\frac{dz}{dt}$ in terms of x, y, and t. If we want the answer solely in terms of t. Hence, we substitute $2t^2$ in place of x and $5t+1$ in place of y. This gives us:
$12(2t^2)^2 t + 28(5t+1)t + 35(2t^2) + 40(5t+1) =$
$48t^5 + 140t^2 + 28t + 70t^2 + 200t + 40 =$

$48t^5 + 210t^2 + 228t + 40.$

If we are looking for the numerical value of $\frac{dz}{dt}$ based on a given numerical value of $t,$ then finding a general formula for $\frac{dz}{dt}$ purely in terms of t is unnecessary. The formula giving us $\frac{dz}{dt}$ in terms of $x,\ y,$ and t will work just as well. Given a numerical value for $t,$ say $t_0,$ we substitute that value into the equations expressing x and y in terms of $t,$ i.e., $x = x(t)$ and $y = y(t),$ giving us $x_0 = x(t_0)$ and $y_0 = y(t_0).$ We then substitute $t_0,\ x_0,$ and y_0 into the formula for $\frac{dz}{dt}$ and calculate the numerical result.

In the above example, find $\frac{dz}{dt}$ when $t = 3.$

- Since $x(t) = 2t^2,\ \ x(3) = 18.$
- Since $y(t) = 5t + 1,\ \ y(3) = 16.$
- Since $\frac{dz}{dt} = 12x^2t + 28yt + 35x + 40y,$ we get $12(18)^2(3) + 28(16)(3) + 35(18) + 40(16) = 14,278.$

This is the same answer we would get if we substituted 3 in place of t in the formula $48t^5 + 210t^2 + 228t + 40.$