The Chain Rule

Suppose $f : \mathbb{R}^p \to \mathbb{R}^n$ and $g : \mathbb{R}^n \to \mathbb{R}^m$ are functions of several variables, where the number of outputs of f equals the number of inputs of g. You can "chain" f and g together to make the composite function $g \circ f$:

$$\mathbb{R}^p \xrightarrow{f} \mathbb{R}^n \xrightarrow{g} \mathbb{R}^m$$

That is, $(g \circ f)(x) = g(f(x))$.

The derivative of $g \circ f$ is given by the **Chain Rule**. It is exactly what you'd expect, based on your experience with functions of one variable.

Theorem. Suppose $f : \mathbb{R}^p \to \mathbb{R}^n$ is differentiable at c, and $g : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at $f(c)$. Then $g \circ f$ is differentiable at c, and

$$D(g \circ f)(c) = Dg[f(c)] \circ Df(c). \quad \square$$

In fact, Dg can be represented by an $m \times n$ matrix, while Df can be represented by an $n \times p$ matrix. The product on the right is the product of two matrices; it makes sense, because the n columns of Dg are compatible with the n rows of Df.

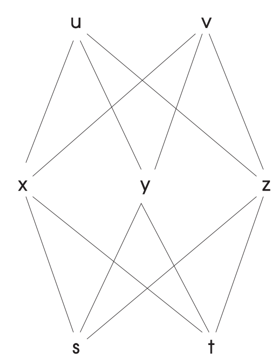
Example. Suppose

$$(x, y, z) = f(s, t) = (s^2 - t^2, s^2 + t^2, st), \quad (u, v) = g(x, y, z) = (2xy - 3yz, xz).$$

(a) Use the Chain Rule to compute $D(g \cdot f)(s, t)$.

(b) Find $\dfrac{\partial u}{\partial s}$ and $\dfrac{\partial v}{\partial s}$.

(a) Here is a picture which shows the dependencies of the variables:



For example, changing s causes x, y, and z to change, which in turn causes u and v to change.

First, compute Df and Dg:

$$Df = \begin{bmatrix} 2s & -2t \\ 2s & 2t \\ t & s \end{bmatrix}, \quad Dg = \begin{bmatrix} 2y & 2x - 3z & -3y \\ z & 0 & x \end{bmatrix}.$$

Next, multiply to obtain $D(g \circ f)(s, t)$, being careful to put Dg on the *left*:

$$D(g \circ f)(s, t) = \begin{bmatrix} 2y & 2x - 3z & -3y \\ z & 0 & x \end{bmatrix} \begin{bmatrix} 2s & -2t \\ 2s & 2t \\ t & s \end{bmatrix} = \begin{bmatrix} 4ys + 4xs - 6zs - 3yt & -4yt + 4xt - 6zt - 3ys \\ 2sz + xt & -2zt + xs \end{bmatrix}.$$

If you wish, you can substitute

$$x = s^2 - t^2, \quad y = s^2 + t^2, \quad z = st.$$

This gives

$$D(g \circ f)(s,t) = \begin{bmatrix} 8s^3 - 9s^2t + 3t^3 & -3s^3 + 9st^2 - 8t^3 \\ 3s^2t + t^3 & s^3 + 3st^2 \end{bmatrix}. \quad \Box$$

Note: This kind of substitution becomes messy when the functions are at all complicated, so I'll often leave the derivative as a product of matrices with "different variables",

(b) Here is how to interpret the matrix for $D(g \circ f)(s,t)$. The composite function is $(u,v) = (g \cdot f)(s,t)$. Therefore,

$$D(g \circ f)(s,t) = \begin{bmatrix} \dfrac{\partial u}{\partial s} & \dfrac{\partial u}{\partial t} \\ \dfrac{\partial v}{\partial s} & \dfrac{\partial v}{\partial t} \end{bmatrix}.$$

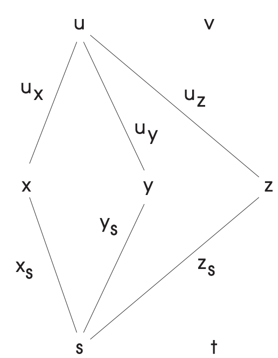
So, for example,

$$\frac{\partial v}{\partial s} = 3s^2t + t^3.$$

I'll check this directly.

$$v = xz = (s^2 + t^2)(st) = s^3t + st^3, \quad \text{so} \quad \frac{\partial v}{\partial s} = 3s^2t + t^3.$$

Alternatively, if all you need is one of the partials (say $\dfrac{\partial u}{\partial s}$), you can use the variable dependency picture to get the formula. Consider all paths in the picture from u to s. Label each path with the corresponding partial derivative. For example, the path from u to y is labelled with $u_y = \dfrac{\partial u}{\partial y}$.



Now to get $\dfrac{\partial u}{\partial s}$, multiply along each path and add the results:

$$\frac{\partial u}{\partial s} = \frac{\partial u}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial s} + \frac{\partial u}{\partial z}\frac{\partial z}{\partial s} = (2y)(2s) + (2x - 3z)(2s) + (-3y)(t). \quad \Box$$

Example. Suppose that $(x,y) = f(s,t)$, $(u,v) = g(x,y)$, $f(0,1) = (2,-2)$, and

$$Df(0,1) = \begin{bmatrix} 1 & -1 \\ 0 & 2 \end{bmatrix}, \quad Df(2,-2) = \begin{bmatrix} 3 & -3 \\ 2 & 0 \end{bmatrix},$$

$$Dg(0,1) = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}, \quad Dg(2,-2) = \begin{bmatrix} 1 & 1 \\ 5 & -1 \end{bmatrix}.$$

Find $D(g \circ f)(0,1)$ and $\left.\dfrac{\partial v}{\partial s}\right|_{(0,1)}$.

First,

$D(g \circ f)(0,1) = Dg\left(f(0,1)\right) \circ Df(0,1) = Dg(2,-2) \circ Df(0,1) = \begin{bmatrix} 1 & 1 \\ 5 & -1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 5 & -7 \end{bmatrix}$.

Now

$$D(g \circ f)(s,t) = \begin{bmatrix} \dfrac{\partial u}{\partial s} & \dfrac{\partial u}{\partial t} \\ \dfrac{\partial v}{\partial s} & \dfrac{\partial v}{\partial t} \end{bmatrix}.$$

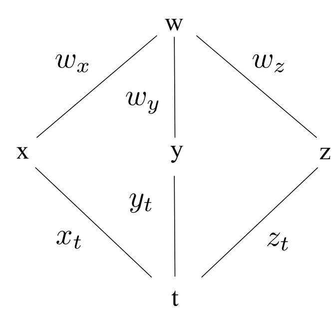
It follows that $\left.\dfrac{\partial v}{\partial s}\right|_{(0,1)} = 5$. □

Example. Suppose

$$w = 5x^2y - y^2z + \ln z.$$

$$x = \cos 6t, \quad y = \frac{2}{t} + 3, \quad z = e^{t^2}.$$

Find $\dfrac{dw}{dt}$.



$$\frac{dw}{dt} = \frac{\partial w}{\partial x}\frac{dx}{dt} + \frac{\partial w}{\partial y}\frac{dy}{dt} + \frac{\partial w}{\partial z}\frac{dz}{dt} = (10xy)(-6\sin 6t) + (5x^2 - 2yz)\left(-\frac{2}{t^2}\right) + \left(-y^2 + \frac{1}{z}\right)(2te^{t^2}). \quad □$$

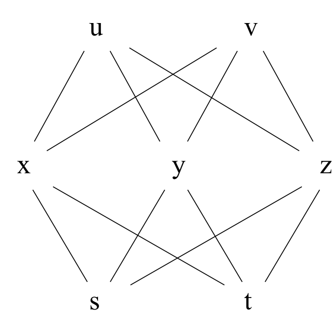
Note: I'm leaving the answer in terms of x, y, z, and t. If you really needed everything in terms of t, you could substitute using the x, y, and z-equations.

Example. Suppose $(u,v) = f(x,y,z)$ and $(x,y,z) = g(s,t)$ are defined by

$$u = x^2 + y^2 + z^2, \quad v = 5xyz,$$

$$x = \cos s \cos t, \quad y = \cos s \sin t, \quad z = \sin s.$$

Find $\dfrac{\partial u}{\partial s}$ and $\dfrac{\partial v}{\partial t}$.



$$\frac{\partial u}{\partial s} = \frac{\partial u}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial s} + \frac{\partial u}{\partial z}\frac{\partial z}{\partial s} = (2x)(-\sin s\cos t) + (2y)(-\sin s\sin t) + (2z)(\cos s).$$

$$\frac{\partial v}{\partial t} = \frac{\partial v}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial v}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial v}{\partial z}\frac{\partial z}{\partial t} = (5yz)(-\cos s\sin t) + (5xz)(\cos s\cos t) + (5xy)(0). \quad \square$$

Note: I'm leaving the answer in terms of x, y, z, s, and t. If you really needed everything in terms of s and t, you could substitute using the x, y, and z-equations.
