Chain rule

Now we will formulate the chain rule when there is more than one independent variable.

We suppose w is a function of x, y and that x, y are functions of u, v. That is,

$$w = f(x, y) \text{ and } x = x(u, v),\ y = y(u, v).$$

The use of the term chain comes because to compute w we need to do a chain of computations

$$(u, v) \to (x, y) \to w.$$

We will say w is a *dependent* variable, u and v are *independent* variables and x and y are *intermediate* variables.

Since w is a function of x and y it has partial derivatives $\dfrac{\partial w}{\partial x}$ and $\dfrac{\partial w}{\partial y}$.

Since, ultimately, w is a function of u and v we can also compute the partial derivatives $\dfrac{\partial w}{\partial u}$ and $\dfrac{\partial w}{\partial v}$. The chain rule relates these derivatives by the following formulas.

$$\begin{aligned}
\frac{\partial w}{\partial u} &= \frac{\partial w}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial u} \\
\frac{\partial w}{\partial v} &= \frac{\partial w}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial v}.
\end{aligned}$$

Example: Given $w = x^2y + y^2 + x$, $\quad x = u^2v$, $\quad y = uv^2 \quad$ find $\dfrac{\partial w}{\partial u}$.

<u>**Answer:**</u> First we compute

$$\frac{\partial w}{\partial x} = 2xy + 1,\ \ \frac{\partial w}{\partial y} = x^2 + 2y,\ \ \frac{\partial x}{\partial u} = 2uv,\ \ \frac{\partial y}{\partial u} = v^2,\ \ \frac{\partial x}{\partial v} = u^2,\ \ \frac{\partial y}{\partial v} = 2uv.$$

The chain rule then implies

$$\begin{aligned}
\frac{\partial w}{\partial u} &= \frac{\partial w}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial u} \\
&= (2xy + 1)2uv + (x^2 + 2y)v^2 \\
\frac{\partial w}{\partial v} &= \frac{\partial w}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial v} \\
&= (2xy + 1)u^2 + (x^2 + 2y)2uv.
\end{aligned}$$

Often, it is okay to leave the variables mixed together. If, for example, you wanted to compute $\frac{\partial w}{\partial u}$ when $(u, v) = (1, 2)$ all you have to do is compute x and y and use these values, along with u, v, in the formula for $\dfrac{\partial w}{\partial u}$.

$$x = 2,\ y = 4\ \Rightarrow\ \frac{\partial w}{\partial u} = (5)(4) + (12)(4) = 68.$$

If you actually need the derivatives expressed in just the variables u and v then you would have to substitute for x, y and z.

Proof of the chain rule:
Just as before our argument starts with the tangent approximation at the point (x_0, y_0).

$$\Delta w \approx \left.\frac{\partial w}{\partial x}\right|_o \Delta x + \left.\frac{\partial w}{\partial y}\right|_o \Delta y.$$

$$\Delta x \approx \frac{\partial x}{\partial u}\Delta u + \frac{\partial x}{\partial v}\Delta v \qquad \Delta y \approx \frac{\partial y}{\partial u}\Delta u + \frac{\partial y}{\partial v}\Delta v$$

$$\Rightarrow \Delta w \approx \frac{\partial w}{\partial x}\left(\frac{\partial x}{\partial u}\Delta u + \frac{\partial x}{\partial v}\Delta v\right) + \frac{\partial w}{\partial y}\left(\frac{\partial y}{\partial u}\Delta u + \frac{\partial y}{\partial v}\Delta v\right)$$

$$= \left(\frac{\partial w}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial u}\right)\Delta u + \left(\frac{\partial w}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial v}\right)\Delta v$$

Now hold v constant and divide by Δu to get

$$\frac{\Delta w}{\Delta u} \approx \left.\frac{\partial w}{\partial x}\right|_o \frac{\Delta x}{\Delta u} + \left.\frac{\partial w}{\partial y}\right|_o \frac{\Delta y}{\Delta u}.$$

$$\frac{x(u_0+\Delta u, v_0) - x(u_0, v_0)}{\Delta u} \xrightarrow{\Delta u \to 0} \left.\frac{d}{du} x(u, v_0)\right|_{u_0} = \left.\frac{\partial x}{\partial u}\right|_{(u_0, v_0)}$$

Finally, letting $\Delta u \to 0$ gives the chain rule for $\dfrac{\partial w}{\partial u}$.

Ambiguous notation

Often you have to figure out the dependent and independent variables from context.

Thermodynamics is a big player here. It has, for example, the variables P, T, V, U, S. and *any* two can be taken to be independent and the others are functions of those two.

We will do more with this topic in the future.