Differentiability

A multivariable function $\boldsymbol{f}\colon \mathbb{R}^n \to \mathbb{R}^m$ is said to be differentiable at $\boldsymbol{a} \in \mathbb{R}^n$ if all its partial derivatives $\frac{\partial \boldsymbol{f}}{\partial x_i}$ exist *and* the local tangent plane $\boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{a}) + \nabla \boldsymbol{f}(\boldsymbol{a}) \cdot (\boldsymbol{x} - \boldsymbol{a})$ is a "good" approximation of the function:

$$\lim_{\boldsymbol{x}\to\boldsymbol{a}} = \frac{\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{h}(\boldsymbol{x})}{|\boldsymbol{x} - \boldsymbol{a}|} = 0.$$

A nice little shortcut, because we're college math students and not hardcore real analysts, is that a function is differentiable if its partial derivatives are all continuous.

Chain Rule

The chain rule deals with the *composition* of two functions. In one-variable calculus, the chain rule looks something like, for some function $f = g \circ h$ (that is, $f(x) = g(h(x))$),

$$\frac{\mathrm{d}f}{\mathrm{d}x} = \frac{\mathrm{d}f}{\mathrm{d}h}\frac{\mathrm{d}h}{\mathrm{d}x} = \frac{\mathrm{d}g}{\mathrm{d}h}\frac{\mathrm{d}h}{\mathrm{d}x} = g'(h(x))h'(x).$$

For multi-variable functions (D, the chain rule takes on a matrix form:

$$D(\boldsymbol{g} \circ \boldsymbol{h}) = D(\boldsymbol{g})D(\boldsymbol{h}).$$

Example Consider $\boldsymbol{f}\colon \mathbb{R}^3 \to \mathbb{R}^2$ given by

$$\boldsymbol{f}(x_1, x_2, x_3) = (x_1 - x_2, x_1x_2x_3),$$

and $\boldsymbol{x}\colon \mathbb{R}^2 \to \mathbb{R}^3$ given by

$$\boldsymbol{x}(t_1, t_2) = (t_1t_2, t_1^2, t_2^2).$$

Context Here's how the mapping looks:

$$\mathbb{R}^2 \xrightarrow{\boldsymbol{x}} \mathbb{R}^3 \xrightarrow{\boldsymbol{f}} \mathbb{R}^2,$$

or

$$(t_1, t_2) \overset{\boldsymbol{x}}{\mapsto} (\underbrace{x_1(t_1, t_2)}_{t_1t_2}, \underbrace{x_2(t_1, t_2)}_{t_1^2}, \underbrace{x_3(t_1, t_2)}_{t_2^2}) \overset{\boldsymbol{f}}{\mapsto} (\underbrace{f_1(x_1, x_2, x_3)}_{x_1 - x_2}, \underbrace{f_2(x_1, x_2, x_3)}_{x_1x_2x_3}).$$

Notice the overall (composed) mapping is then

$$\mathbb{R}^2 \xrightarrow{\boldsymbol{f}\circ\boldsymbol{x}} \mathbb{R}^2.$$

Solution (method 1): tedious substitution Find the derivative of the composition of f and x,

$$D(\boldsymbol{f} \circ \boldsymbol{x})(\boldsymbol{t}).$$

We can simply plug in the composed function

$$(\boldsymbol{f} \circ \boldsymbol{x})(\boldsymbol{t}) = (x_1 - x_2, x_1x_2x_3) = (t_1t_2 - t_1^2, (t_1t_2)(t_1^2)(t_2^2)) = (t_1t_2 - t_1^2, t_1^3t_2^3)$$

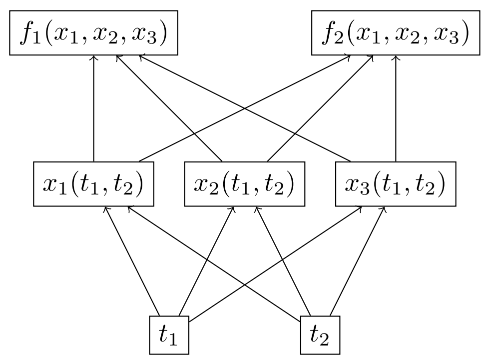
and take the derivative:

$$D(\boldsymbol{f} \circ \boldsymbol{x})(\boldsymbol{t}) = \begin{pmatrix} \partial_{t_1}\left(t_1t_2 - t_1^2\right) & \partial_{t_2}\left(t_1t_2 - t_1^2\right) \\ \partial_{t_1}\left(t_1^3t_2^3\right) & \partial_{t_2}\left(t_1^3t_2^3\right) \end{pmatrix} = \begin{pmatrix} t_2 - 2t_1 & t_1 \\ 3t_1^2t_2^3 & 3t_1^3t_2^2 \end{pmatrix}.$$

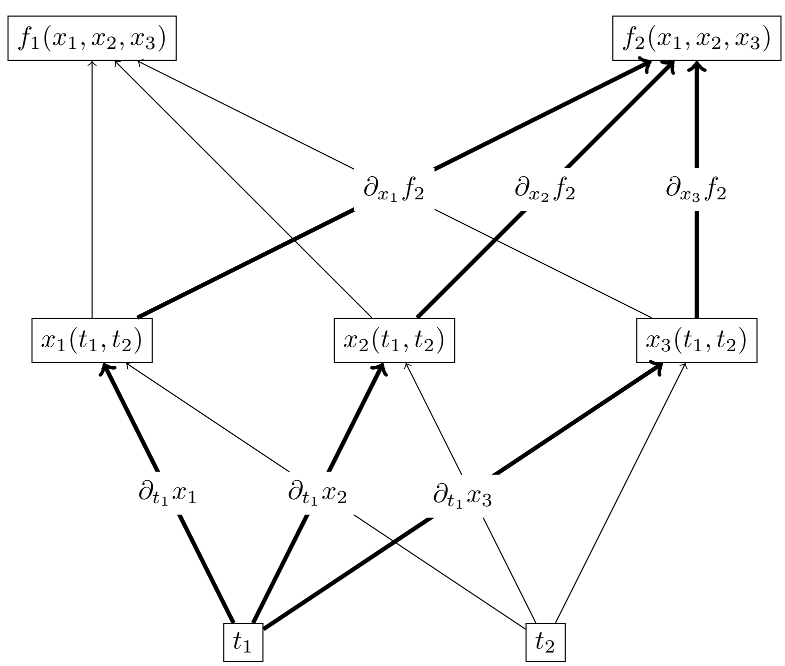
Solution (method 2): chain rule (*better!!!*) Alternatively, apply the chain rule:

$$\begin{aligned}
D(\boldsymbol{f} \circ \boldsymbol{x})(\boldsymbol{t}) &= D\boldsymbol{f}(\boldsymbol{x})D\boldsymbol{x}(\boldsymbol{t}) \\
&= \begin{pmatrix} \partial_{x_1}(x_1 - x_2) & \partial_{x_2}(x_1 - x_2) & \partial_{x_3}(x_1 - x_2) \\ \partial_{x_1}(x_1x_2x_3) & \partial_{x_2}(x_1x_2x_3) & \partial_{x_3}(x_1x_2x_3) \end{pmatrix} \begin{pmatrix} \partial_{t_1}(t_1t_2) & \partial_{t_2}(t_1t_2) \\ \partial_{t_1}(t_1^2) & \partial_{t_2}(t_1^2) \\ \partial_{t_1}(t_2^2) & \partial_{t_2}(t_2^2) \end{pmatrix} \\
&= \begin{pmatrix} 1 & -1 & 0 \\ x_2x_3 & x_1x_3 & x_1x_2 \end{pmatrix} \begin{pmatrix} t_2 & t_1 \\ 2t_1 & 0 \\ 0 & 2t_2 \end{pmatrix} \\
&= \begin{pmatrix} t_2 - 2t_1 & t_1 \\ x_2x_3t_2 + 2x_1x_3t_1 & x_2x_3t_1 + 2x_1x_2t_2 \end{pmatrix} \\
&= \begin{pmatrix} t_2 - 2t_1 & t_1 \\ (t_1^2)(t_2^2)t_2 + 2(t_1t_2)(t_2^2)t_1 & (t_1^2)(t_2^2)t_1 + 2(t_1t_2)(t_1^2)t_2 \end{pmatrix} \\
&= \begin{pmatrix} t_2 - 2t_1 & t_1 \\ 3t_1^2t_2^3 & 3t_1^3t_2^2 \end{pmatrix}.
\end{aligned}$$

Solution (method 3?): cool diagrams



The partial derivative of the function's i-th component f_i with respect to some input t_j can be found by tracing all "paths" on the diagram from t_j to f_i. For example, the partial derivative of f_2 with respect to t_1 is found by



$$\partial_{t_1} f_2 = \partial_{x_1} f_2 \cdot \partial_{t_1} x_1 + \partial_{x_2} f_2 \cdot \partial_{t_1} x_2 + \partial_{x_3} f_2 \cdot \partial_{t_1} x_3$$

Why does it work? Turns out, this is just another way to do, or *visualize*, derivative matrix multiplication.

Tracing the "paths" of each component is the same thing as component-wise matrix multiplication:

$$\begin{aligned}
D_{\boldsymbol{t}}\boldsymbol{f} &= D_{\boldsymbol{x}}\boldsymbol{f}\, D_{\boldsymbol{t}}\boldsymbol{x} \\
&= \begin{pmatrix} \partial_{x_1} f_1 & \partial_{x_2} f_1 & \partial_{x_3} f_1 \\ \boxed{\partial_{x_1} f_2 \quad \partial_{x_2} f_2 \quad \partial_{x_3} f_2} \end{pmatrix} \begin{pmatrix} \boxed{\begin{matrix}\partial_{t_1} x_1 \\ \partial_{t_1} x_2 \\ \partial_{t_1} x_3\end{matrix}} & \begin{matrix}\partial_{t_2} x_1 \\ \partial_{t_2} x_2 \\ \partial_{t_2} x_3\end{matrix} \end{pmatrix} \\
&= \begin{pmatrix} \partial_{x_1} f_1\, \partial_{t_1} x_1 + \partial_{x_2} f_1\, \partial_{t_1} x_2 + \partial_{x_3} f_1\, \partial_{t_1} x_3 & \partial_{x_1} f_1\, \partial_{t_2} x_1 + \partial_{x_2} f_1\, \partial_{t_2} x_2 + \partial_{x_3} f_1\, \partial_{t_2} x_3 \\ \boxed{\partial_{x_1} f_2\, \partial_{t_1} x_1 + \partial_{x_2} f_2\, \partial_{t_1} x_2 + \partial_{x_3} f_2\, \partial_{t_1} x_3} & \partial_{x_1} f_2\, \partial_{t_2} x_1 + \partial_{x_2} f_2\, \partial_{t_2} x_2 + \partial_{x_3} f_2\, \partial_{t_2} x_3 \end{pmatrix}
\end{aligned}$$

Implicit Derivatives

Consider some surface defined by the equation $F(x, y, z) = 0$. Implicitly take partial derivatives with respect to each variable, considering that z to be a function of x and y:

$$\begin{aligned}
\partial_x F + \partial_z F\, \partial_x z &= 0, \\
\partial_y F + \partial_z F\, \partial_y z &= 0.
\end{aligned}$$

Consequently,

$$\begin{aligned}
\partial_x z &= -\frac{\partial_x F}{\partial_z F}, \\
\partial_y z &= -\frac{\partial_y F}{\partial_z F}.
\end{aligned}$$

Example Consider the surface given by

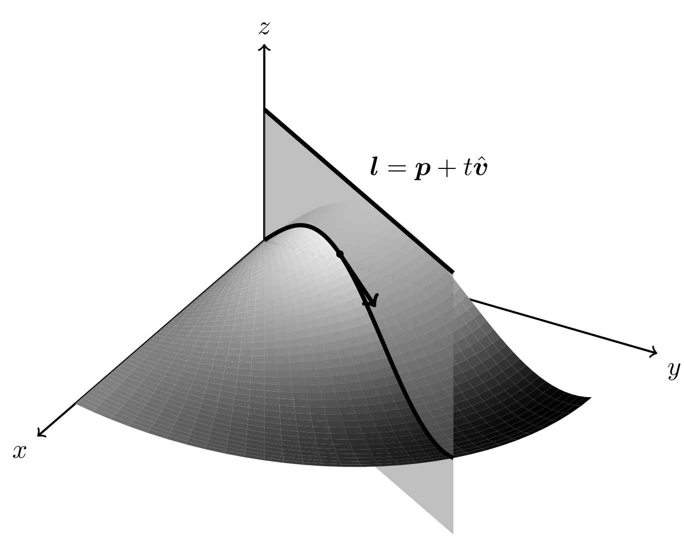
$$x^2 + y^2 + z^2 - 1 = 0,$$

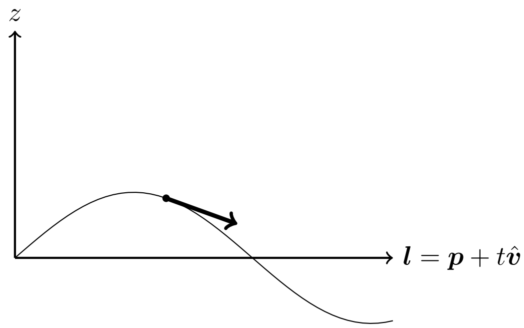
a sphere with radius 1. The partial derivatives of z with respect to x and y are then

$$\begin{aligned}
\partial_x z &= -\frac{\partial_x (x^2 + y^2 + z^2 - 1)}{\partial_z (x^2 + y^2 + z^2 - 1)} = -\frac{2x}{2z} = -\frac{x}{z} \\
\partial_y z &= -\frac{y}{z}.
\end{aligned}$$

Directional Derivatives

Consider the derivative of a function $\boldsymbol{f}$ along some direction $\hat{\boldsymbol{v}}$. The *directional derivative* can be thought of as a derivative of a function inside a "cross section":





The line along the direction of differentiation is parametrized by some

$$\boldsymbol{l}(t) = \boldsymbol{p} + t\hat{\boldsymbol{v}},$$

so that the function along the line

$$\boldsymbol{f}\big|_{\boldsymbol{l}(t)} = \boldsymbol{f}(\boldsymbol{p} + t\hat{\boldsymbol{v}}),$$

so that the directional derivative can be found by the chain rule

$$D_{\hat{\boldsymbol{v}}}\boldsymbol{f} = \mathrm{d}_t \boldsymbol{f}(\boldsymbol{p} + t\hat{\boldsymbol{v}}) = D\boldsymbol{f} \cdot \hat{\boldsymbol{v}}.$$