Lecture 11: Chain rule

If f and g are functions of one variable t, the **single variable chain rule** tells us that $d/dt f(g(t)) = f'(g(t))g'(t)$. For example, $d/dt \sin(\log(t)) = \cos(\log(t))/t$.
It can be proven by linearizing the functions f and g and verifying the chain rule in the linear case. The **chain rule** is also useful:

For example, to find $\arccos'(x)$, we write $1 = d/dx \cos(\arccos(x)) = -\sin(\arccos(x)) \arccos'(x) = -\sqrt{1-\sin^2(\arccos(x))} \arccos'(x) = \sqrt{1-x^2} \arccos'(x)$ so that $\arccos'(x) = -1/\sqrt{1-x^2}$.

Define the **gradient** $\nabla f(x,y) = \langle f_x(x,y), f_y(x,y) \rangle$ or $\nabla f(x,y,z) = \langle f_x(x,y,z), f_y(x,y,z), f_z(x,y,z) \rangle$.

If $\vec{r}(t)$ is curve and f is a function of several variables we can build a function $t \mapsto f(\vec{r}(t))$ of one variable. Similarly, If $\vec{r}(t)$ is a parametrization of a curve in the plane and f is a function of two variables, then $t \mapsto f(\vec{r}(t))$ is a function of one variable.

The **multivariable chain rule** is $\frac{d}{dt} f(\vec{r}(t)) = \nabla f(\vec{r}(t)) \cdot \vec{r}\,'(t)$.

Proof. When written out in two dimensions, it is

$$\frac{d}{dt} f(x(t), y(t)) = f_x(x(t), y(t))x'(t) + f_y(x(t), y(t))y'(t) \; .$$

Now, the identity

$$\frac{f(x(t+h),y(t+h)) - f(x(t),y(t))}{h} = \frac{f(x(t+h),y(t+h)) - f(x(t),y(t+h))}{h} + \frac{f(x(t),y(t+h)) - f(x(t),y(t))}{h}$$

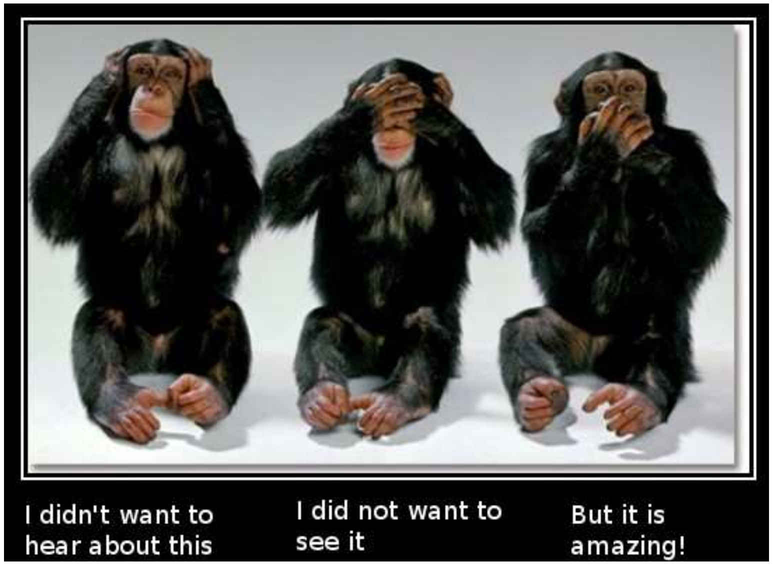
holds for every $h > 0$. The left hand side converges to $\frac{d}{dt} f(x(t), y(t))$ in the limit $h \to 0$ and the right hand side to $f_x(x(t), y(t))x'(t) + f_y(x(t), y(t))y'(t)$ using the single variable chain rule twice. Here is the proof of the later, when we differentiate f with respect to t and y is treated as a constant:

$$\frac{f(\text{x(t+h)}) - f(x(t))}{h} = \frac{[f(\text{x(t)} + \text{(x(t+h)-x(t))}) - f(x(t))]}{\text{[x(t+h)-x(t)]}} \cdot \frac{\text{[x(t+h)-x(t)]}}{h} \; .$$

Write $H(t) =$ x(t+h)-x(t) in the first part on the right hand side.

$$\frac{f(x(t+h)) - f(x(t))}{h} = \frac{[f(x(t)+H) - f(x(t))]}{H} \cdot \frac{x(t+h) - x(t)}{h} \; .$$

As $h \to 0$, we also have $H \to 0$ and the first part goes to $f'(x(t))$ and the second factor to $x'(t)$.



1 We move on a circle $\vec{r}(t) = \langle \cos(t), \sin(t) \rangle$ on a table with temperature distribution $f(x,y) = x^2 - y^3$. Find the rate of change of the temperature $\nabla f(x,y) = (2x, -3y^2)$, $\vec{r}\,'(t) = (-\sin(t), \cos(t))$ $d/dt f(\vec{r}(t)) = \nabla T(\vec{r}(t)) \cdot \vec{r}\,'(t) = (2\cos(t), -3\sin(t)^2) \cdot (-\sin(t), \cos(t)) = -2\cos(t)\sin(t) - 3\sin^2(t)\cos(t)$.

From $f(x,y) = 0$ one can express y as a function of x. From $d/df(x, y(x)) = \nabla f \cdot (1, y'(x)) = f_x + f_y y' = 0$, we obtain $y' = -f_x/f_y$. Even so, we do not know $y(x)$, we can compute its derivative! Implicit differentiation works also in three variables. The equation $f(x,y,z) = c$ defines a surface. Near a point where f_z is not zero, the surface can be described as a graph $z = z(x,y)$. We can compute the derivative z_x without actually knowing the function $z(x,y)$. To do so, we consider y a fixed parameter and compute using the chain rule

$$f_x(x, y, z(x,y))1 + f_z(x,y) z_x(x,y) = 0$$

so that $z_x(x,y) = -f_x(x,y,z)/f_z(x,y,z)$.

2 The surface $f(x,y,z) = x^2 + y^2/4 + z^2/9 = 6$ is an ellipsoid. Compute $z_x(x,y)$ at the point $(x,y,z) = (2,1,1)$.
Solution: $z_x(x,y) = -f_x(2,1,1)/f_z(2,1,1) = -4/(2/9) = -18$.

The chain rule is powerful because it implies other differentation rules like the addition, product and quotient rule in one dimensions: $f(x,y) = x + y, x = u(t), y = v(t), d/dt(x+y) = f_x u' + f_y v' = u' + v'$.
$f(x,y) = xy, x = u(t), y = v(t), d/dt(xy) = f_x u' + f_y v' = vu' + uv'$.
$f(x,y) = x/y, x = u(t), y = v(t), d/dt(x/y) = f_x u' + f_y v' = u'/y - v'u/v^2$.
As in one dimensions, the chain rule follows from linearization. If f is a linear function $f(x,y) = ax + by - c$ and if the curve $\vec{r}(t) = \langle x_0 + tu, y_0 + tv \rangle$ parametrizes a line. Then $\frac{d}{dt} f(\vec{r}(t)) = \frac{d}{dt}(a(x_0 + tu) + b(y_0 + tv)) = au + bv$ and this is the dot product of $\nabla f = (a,b)$ with $\vec{r}\,'(t) = (u,v)$. Since the chain rule only refers to the derivatives of the functions which agree at the point, the chain rule is also true for general functions.

Homework

1 You know that $d/dtf(\vec{r}(t)) = 2$ if $\vec{r}(t) = \langle t, t\rangle$ and $d/dtf(\vec{r}(t)) = 3$ if $\vec{r}(t) = \langle t, -t\rangle$. Find the gradient of f at $(0, 0)$.

2 The pressure in the space at the position (x, y, z) is $p(x, y, z) = x^2+y^2-z^3$ and the trajectory of an observer is the curve $\vec{r}(t) = \langle t, t, 1/t\rangle$. Using the chain rule, compute the rate of change of the pressure the observer measures at time $t = 2$.

3 Mechanical systems are determined by the energy function $H(x, y)$, a function of two variables. The first, x is the position and the second y is the momentum. The equations of motion for the curve $\vec{r}(t) = \langle x(t), y(t)\rangle$ are

$$\begin{aligned} x'(t) &= H_y(x, y) \\ y'(t) &= -H_x(x, y) \end{aligned}$$

They are called called **Hamilton equations**. a) Using the chain rule, verify that in full generality, the energy of a Hamiltonian system is preserved: for every path $\vec{r}(t) = \langle x(t), y(t)\rangle$ solving the system, we have $H(x(t), y(t)) = const$.
b) Check this in the particular case of the **pendulum**, where $H(x, y) = y^2/2 - \sin(x)$.

4 Derive using implicit differentiation the derivative d/dx arctanh(x), where

$$\tanh(x) = \sinh(x)/\cosh(x) .$$

The **hyperbolic sine** and **hyperbolic cosine** are defined as are $\sinh(x) = (e^x - e^{-x})/2$ and $\cosh(x) = (e^x + e^{-x})/2$. We have $\sinh' = \cosh$ and $\cosh' = \sinh$ and $\cosh^2(x) - \sinh^2(x) = 1$.

5 The equation $f(x, y, z) = e^{xyz} + z = 1 + e$ implicitly defines z as a function $z = g(x, y)$ of x and y. Find formulas (in terms of x,y and z) for $g_x(x, y)$ and $g_y(x, y)$. Estimate $g(1.01, 0.99)$ using linear approximation.