Unit 11: Chain rule

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

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

> **Definition:** Define the **gradient** $\nabla f(x, y) = [f_x(x, y), f_y(x, y)]$ or $\nabla f(x, y, z) = [f_x(x, y, z), f_y(x, y, z), f_z(x, y, z)]$. It is the analog of the derivative in higher dimensions.

11.3. If $\vec{r}(t)$ is curve and f is a function of several variables we get a function $t \mapsto f(\vec{r}(t))$ of one variable. Similarly, if $\vec{r}(t)$ is a parametrization of a planar curve f is a function of two variables, then $t \mapsto f(\vec{r}(t))$ is a function of one variable.

> **Theorem:** $\frac{d}{dt} f(\vec{r}(t)) = \nabla f(\vec{r}(t)) \cdot \vec{r}\,'(t)$.

Proof. When written out in two dimensions, it is

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

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:

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

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 $x'(t)$.

11.4. The chain rule is powerful because it implies **all other differentiation rules**: the addition, product and quotient rule: $f(x, y) = x + y, x = u(t), y = v(t), d/dt(x + y) = f_x u' + f_y v' = u' + v'$.
$f(x, y) = xy, x = u(t), y = v(t), d/dt(xy) = f_x u' + f_y v' = vu' + uv'$.
$f(x, y) = x/y, x = u(t), y = v(t), d/dt(x/y) = f_x u' + f_y v' = u'/y - v'u/v^2$.

11.5. We mentioned that the chain rule can be derived from linearization. Let us look at this case: if f is a linear function $f(x, y) = ax + by - c$ and if $\vec{r}(t) = [x_0 + tu, y_0 + tv]$ is 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 $\vec{r}\,'(t) = (u, v)$. Because the chain rule only refers to the derivatives of the functions and the linearlization too, the chain rule is also true for general functions.

<div align="center">EXAMPLES</div>

11.6. A ladybug moves on a circle $\vec{r}(t) = [\cos(t), \sin(t)]$ on a table with temperature distribution $f(x, y) = x^2 - y^3$. Find the rate of change of the temperature $\nabla f(x, y) = (2x, -3y^2)$, $\vec{r}(t) = (-\sin(t), \cos(t))$ $d/dt f(\vec{r}(t)) = \nabla T(\vec{r}(t)) \cdot \vec{r}(t) = (2\cos(t), -3\sin(t)^2) \cdot (-\sin(t), \cos(t)) = -2\cos(t)\sin(t) - 3\sin^2(t)\cos(t)$.
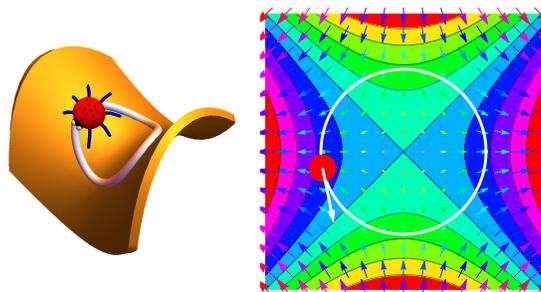
11.7. From $f(x, y) = 0$, one can express y as a function of x, at least near a point where f_y is not zero. From $\frac{d}{dx} f(x, y(x)) = \nabla f \cdot (1, y'(x)) = f_x + f_y y' = 0$, we obtain $y' = -f_x/f_y$. Even so, we do not know $y(x)$, we can compute its derivative! Implicit differentiation 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

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

so that $z_x(x, y) = -f_x(x, y, z)/f_z(x, y, z)$. This works at points where f_z is not zero.

11.8. 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: $z_x(x, y) = -f_x(2, 1, 1)/f_z(2, 1, 1) = -4/(2/9) = -18$.

HOMEWORK

This homework is due on Tuesday, 7/16/2024 in class.

Problem 11.1: You know that $d/dt f(\vec{r}(t)) = 19$ at $t = 7$ if $\vec{r}(t) = [t, t]$ and $d/dt f(\vec{r}(t)) = 11$ at $t = 7$. $\vec{r}(t) = [t, 14 - t]$. Find the gradient of f at $(7, 7)$.

Problem 11.2: 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 $\vec{r}(t) = [t, t, 1/t]$. Using the chain rule, compute the rate of change of the pressure the observer measures at time $t = 2$.

Problem 11.3: The chain rule is closely related to linearization. Lets get back to linearization a bit: A farm costs $f(x, y)$, where x is the number of cows and y is the number of ducks. There are 10 cows and 20 ducks and $f(10, 20) = 1000000$. We know that $f_x(x, y) = 2x$ and $f_y(x, y) = y^2$ for all x, y. Estimate $f(12, 19)$.

Here is a song out of this:
"Old MacDonald had a million dollar farm, E-I-E-I-O,
and on that farm he had $x = 10$ cows, E-I-E-I-O,
and on that farm he had $y = 20$ ducks, E-I-E-I-O,
with $f_x = 2x$ here and $f_y = y^2$ there,
and here two cows more, and there a duck less,
how much does the farm cost now, E-I-E-I-O?"

Problem 11.4: Find, using implicit differentiation the derivative $d/dx \, \text{arctanh}(x)$, 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 $\cosh^2(x) - \sinh^2(x) = 1$.

Problem 11.5: 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 $g(1.01, 0.999)$ using linear approximation.