Section 2.5

Properties of the derivative.

Let $f: U \subset \mathbb{R}^n \to \mathbb{R}^m$ so that the matrix of derivatives Df is a $m \times n$ matrix. If $g: U \subset \mathbb{R}^n \to \mathbb{R}^m$, then it is easy to see that:

$$D(f+g) = Df + Dg.$$

Indeed:

$$f = (f_1, \dots, f_m), \quad g = (g_1, \dots, g_m).$$

$$\Rightarrow f+g = (f_1+g_1, \dots, f_m+g_m)$$

$$D(f+g) = \begin{pmatrix} \frac{\partial (f_1+g_1)}{\partial x_1} & \cdots\cdots & \frac{\partial}{\partial x_n}(f_1+g_1) \\ \frac{\partial}{\partial x_1}(f_2+g_2) & \cdots\cdots & \frac{\partial}{\partial x_n}(f_2+g_2) \\ \vdots & & \\ \frac{\partial}{\partial x_1}(f_m+g_m) & \cdots\cdots & \frac{\partial}{\partial x_n}(f_m+g_m) \end{pmatrix}_{m \times n}$$

$$= \begin{pmatrix} \frac{\partial f_1}{\partial x_1} + \frac{\partial g_1}{\partial x_1} & \cdots\cdots & \frac{\partial f_1}{\partial x_n} + \frac{\partial g_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} + \frac{\partial g_2}{\partial x_1} & \cdots\cdots & \frac{\partial f_2}{\partial x_n} + \frac{\partial g_2}{\partial x_n} \\ \vdots & & \\ \frac{\partial f_m}{\partial x_1} + \frac{\partial g_m}{\partial x_1} & \cdots\cdots & \frac{\partial f_m}{\partial x_n} + \frac{\partial g_m}{\partial x_n} \end{pmatrix}$$

$$
= \begin{array}{c} m \times n \\ \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & & \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix} \end{array} + \begin{array}{c} m \times n \\ \begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \frac{\partial g_2}{\partial x_1} & \cdots & \frac{\partial g_2}{\partial x_n} \\ \vdots & & \\ \frac{\partial g_m}{\partial x_1} & \cdots & \frac{\partial g_m}{\partial x_n} \end{pmatrix} \end{array}
$$

$$= Df + Dg.$$

Suppose now that $g : U \subset \mathbb{R}^n \to \mathbb{R}$. Then:

$$gf = (gf_1, \dots, gf_m)$$

and:

$$
D(gf) = \begin{pmatrix} \frac{\partial}{\partial x_1}(gf_1) & \cdots & \frac{\partial}{\partial x_n}(gf_1) \\ \vdots & & \\ \frac{\partial}{\partial x_1}(gf_m) & \cdots & \frac{\partial}{\partial x_n}(gf_m) \end{pmatrix}
$$

$$
\overset{\text{Product rule}}{=} \begin{pmatrix} f_1 \frac{\partial g}{\partial x_1} + g \frac{\partial f_1}{\partial x_1} & \cdots & f_1 \frac{\partial g}{\partial x_n} + g \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ f_m \frac{\partial g}{\partial x_1} + g \frac{\partial f_m}{\partial x_1} & \cdots & f_m \frac{\partial g}{\partial x_n} + g \frac{\partial f_m}{\partial x_n} \end{pmatrix}
$$

$$= \begin{array}{c} m\times n \\ \begin{pmatrix} f_1 \frac{\partial g}{\partial x_1} & \cdots & f_1 \frac{\partial g}{\partial x_n} \\ \vdots & & \\ f_m \frac{\partial g}{\partial x_1} & \cdots & f_m \frac{\partial g}{\partial x_n} \end{pmatrix} \end{array} + g\, Df$$

$$= \underset{m\times 1}{\begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_m \end{pmatrix}} \underset{1\times n}{\begin{pmatrix} \frac{\partial g}{\partial x_1} & \cdots & \frac{\partial g}{\partial x_n} \end{pmatrix}} + g\, Df$$

$$= (f_1 \cdots f_m)^T Dg + g\, Df.$$

<u>Ex</u>: $f(x,y) = x^2 \tan y \qquad g(x,y) = y \ln(x^2+1)$

$Df = \left(2x \tan y \quad x^2 \sec^2 y\right)$

$Dg = \left(\frac{2xy}{x^2+1} \quad \ln(x^2+1)\right)$

$$D(fg) = x^2 \tan y \left(\frac{2xy}{x^2+1} \quad \ln(x^2+1)\right) + y\ln(x^2+1)\left(2x\tan y \quad x^2\sec^2 y\right)$$

Simplify to a 1×2 matrix, we can also think of these matrices as vectors and compute $\nabla(fg)$, which is a 1×2 vector with the same components as $D(fg)$

Since in the previous example,
$Df = \nabla f$, $Dg = \nabla g$, the product rule $D(fg)$
reduces to:

$$\nabla (fg) = f \nabla g + g \nabla f.$$

Indeed, we can check again for this case:

$$\nabla (fg) = \left(\frac{\partial}{\partial x_1}(fg), \ldots, \frac{\partial}{\partial x_n}(fg) \right)$$

$$= \left(f \frac{\partial g}{\partial x_1} + g \frac{\partial f}{\partial x_1}, \ldots, f \frac{\partial g}{\partial x_n} + g \frac{\partial f}{\partial x_n} \right)$$

$$= f \left(\frac{\partial g}{\partial x_1}, \ldots, \frac{\partial g}{\partial x_n} \right) + g \left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n} \right)$$

$$= f \nabla g + g \nabla f.$$

<u>Chain rule</u>: Recall, in elementary Calculus:

$$\frac{d}{dx} (f \circ g)(x) = f'(g(x)) g'(x),$$

this formula can also be written as follows:

$$y = f(u), \quad u = g(x)$$

$$\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx}$$

Ex: $f(u) = \sin u + e^u$ $\quad f'(u) = \cos u + e^u$

$g(x) = \ln(x^2+1)$ $\quad g'(x) = \frac{2x}{1+x^2}$

$$\frac{d}{dx}(f \circ g)(x) = f'(g(x)) \cdot g'(x)$$

$$= \left[\cos(\ln(1+x^2)) + e^{\ln(1+x^2)}\right] \cdot \frac{2x}{1+x^2}$$

$$= \left(\cos(\ln(1+x^2)) + 1 + x^2\right) \frac{2x}{1+x^2}$$

The most general formula for chain rule is:

theorem 3: Let $f: U \subset \mathbb{R}^m \to \mathbb{R}^p$, and $g: V \subset \mathbb{R}^n \to \mathbb{R}^m$. Then, if g is differentiable at $\vec{x}_0$ and f is differentiable at $g(\vec{x}_0)$, $f \circ g$ is differentiable at $\vec{x}_0$ and

$$D(f \circ g)(\vec{x}_0) = Df(g(\vec{x}_0)) \cdot Dg(\vec{x}_0).$$

Note that Df is a $p \times m$ matrix, Dg is a $m \times n$ matrix, so $D(f \circ g)$ is $p \times m$ matrix.

Ex: $f(x,y,z) = (xy^2, yz^2)$

$g(s,t) = (e^s, s\cos t, s+t)$

Compute $D(f\circ g)(s,t)$.

We first note that $f: \mathbb{R}^3 \to \mathbb{R}^2$ and $g: \mathbb{R}^2 \to \mathbb{R}^3$, so $f\circ g: \mathbb{R}^2 \to \mathbb{R}^2$, and hence $D(f\circ g)$ is a 2×2 matrix.

Each component of f and g is differentiable, which is easily seen from Theorem 2 in page 64.

Hence, g is differentiable at any point (s,t) and f is differentiable at any point $g(s,t)$. Hence Theorem 3 yields that $f\circ g$ is differentiable at any point (s,t), and:

$$D(f\circ g)(s,t) = Df(g(s,t)) \cdot Dg(s,t)$$

We compute:

$$Df = \begin{pmatrix} y^2 & 2xy & 0 \\ 0 & z^2 & 2yz \end{pmatrix}$$

$\Rightarrow$

$$Df\,(g(s,t)) = \begin{pmatrix} s^2\cos^2 t & 2se^s\cos t & 0 \\ 0 & (s+t)^2 & 2s(s+t)\cos t \end{pmatrix}$$

$$Dg = \begin{pmatrix} e^s & 0 \\ \cos t & -s\sin t \\ 1 & 1 \end{pmatrix}$$

$\Rightarrow$

$$Df\,(g(s,t)) \cdot Dg(s,t) = \begin{pmatrix} s^2\cos^2 t & 2se^s\cos t & 0 \\ 0 & (s+t)^2 & 2s(s+t)\cos t \end{pmatrix} \begin{pmatrix} e^s & 0 \\ \cos t & -s\sin t \\ 1 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} e^s s^2\cos^2 t + 2se^s\cos^2 t & -2s^2e^s\sin t\,\cos t \\ (s+t)^2\cos t + 2s(s+t)\cos t & -s(s+t)^2\sin t + 2s(s+t)\cos t \end{pmatrix}$$

In particular, for example, to complute $D(f\circ g)(1,0)$ we plug:

$$D(f\circ g)(1,0) = \begin{pmatrix} e^1 + 2e^1 & 0 \\ 1+2 & 2 \end{pmatrix} = \begin{pmatrix} 3e & 0 \\ 3 & 2 \end{pmatrix}$$

Important special cases:

(1) $f: \mathbb{R}^3 \to \mathbb{R}$ $\quad g: \mathbb{R}^2 \to \mathbb{R}^3$

$f(x, y, z)$ $\quad g(u, v) = (x(u,v), y(u,v), z(u,v))$

Let $h = f\circ g : \mathbb{R}^2 \to \mathbb{R}$, Dh is a 1×2 matrix

$$D(f\circ g) = Df(g(u,v)) \cdot Dg(u,v)$$

$$= \underbrace{\left(\frac{\partial f}{\partial x}(g(u,v)) \quad \frac{\partial f}{\partial y}(g(u,v)) \quad \frac{\partial f}{\partial z}(g(u,v))\right)}_{1\times 3} \cdot \underbrace{\begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} \end{pmatrix}}_{3\times 2}$$

$$= \left(\frac{\partial f}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial f}{\partial y}\cdot\frac{\partial y}{\partial u} + \frac{\partial f}{\partial z}\cdot\frac{\partial z}{\partial u} \quad \frac{\partial f}{\partial x}\cdot\frac{\partial x}{\partial v} + \frac{\partial f}{\partial y}\cdot\frac{\partial y}{\partial v} + \frac{\partial f}{\partial z}\cdot\frac{\partial z}{\partial v}\right)$$

$$= \left(\frac{\partial h}{\partial u} \quad \frac{\partial h}{\partial v}\right)$$

Since $h = f\circ g : \mathbb{R}^2 \to \mathbb{R}$ is a real valued function, we can consider $D(f\circ g)$ as a vector $\nabla(f\circ g) = \left(\frac{\partial h}{\partial u}, \frac{\partial h}{\partial v}\right)$. Clearly:

$$\boxed{\begin{aligned} \frac{\partial h}{\partial u} &= \frac{\partial f}{\partial x}\cdot\frac{\partial x}{\partial u} + \frac{\partial f}{\partial y}\cdot\frac{\partial y}{\partial u} + \frac{\partial f}{\partial z}\cdot\frac{\partial z}{\partial u} \\ \frac{\partial h}{\partial v} &= \frac{\partial f}{\partial x}\cdot\frac{\partial x}{\partial v} + \frac{\partial f}{\partial y}\cdot\frac{\partial y}{\partial v} + \frac{\partial f}{\partial z}\cdot\frac{\partial z}{\partial v} \end{aligned}}$$

A particular example of this first special case is:

$T(x,y,z)$ is temperature at the point $(x,y,z) \in \mathbb{R}^3$

The sphere $x^2+y^2+z^2=1$ can be parametrized as:
$g(\theta,\varphi) = (\cos\theta \sin\varphi, \sin\theta \sin\varphi, \cos\varphi)$

Hence, the composition function $h(\theta,\varphi) = T(g(\theta,\varphi))$ gives the temperature of the sphere at the point $g(\theta,\varphi)$ and $\frac{\partial h}{\partial \theta}, \frac{\partial h}{\partial \varphi}$ give the rate of change of the temperature with respect to the variables θ and φ.

(2) $f : \mathbb{R}^2 \to \mathbb{R}$ $\qquad g : \mathbb{R}^3 \to \mathbb{R}^2$

$f(x,y)$ $\qquad g(u,v,w) = (x(u,v,w), y(u,v,w))$

$h = f \circ g : \mathbb{R}^3 \to \mathbb{R}$, so Dh is a 1×3 matrix:

$$Dh = \left(\frac{\partial h}{\partial u} \quad \frac{\partial h}{\partial v} \quad \frac{\partial h}{\partial w} \right)$$

Proceeding as in (1), we compute:

$$\frac{\partial h}{\partial v} = \frac{\partial f}{\partial x} \cdot \frac{\partial x}{\partial v} + \frac{\partial f}{\partial y} \cdot \frac{\partial y}{\partial v},$$

similarly for $\frac{\partial h}{\partial u}$ and $\frac{\partial h}{\partial w}$.

3 Consider the path:

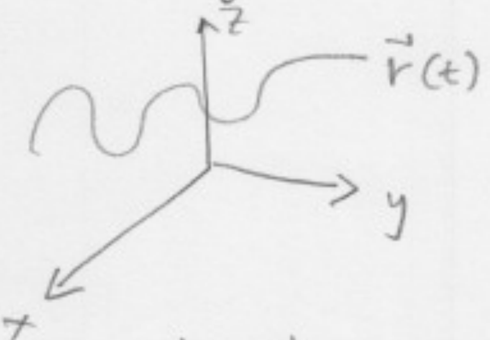
$$\vec{r}(t) = (x(t), y(t), z(t)) : \mathbb{R} \to \mathbb{R}^3,$$

and a function $f(x,y,z) : \mathbb{R}^3 \to \mathbb{R}$. You can think of $\vec{r}(t)$ as a wire in $\mathbb{R}^3$ and $f(x,y,z)$ a function giving density at (x,y,z).

The composition:

$$h(t) = (f \circ \vec{r})(t) = f(\vec{r}(t)) = f(x(t), y(t), z(t)),$$

gives the density of the wire at $(x(t), y(t), z(t))$



How is the density changing along the wire? We compute $h'(t)$:

$$h'(t) = \overset{1\times 3}{Df(\vec{r}(t))}\, \overset{3\times 1}{D(\vec{r}(t))}$$

$$= \left(\frac{\partial f}{\partial x}(\vec{r}(t)) \quad \frac{\partial f}{\partial y}(\vec{r}(t)) \quad \frac{\partial f}{\partial z}(\vec{r}(t)) \right) \begin{pmatrix} x'(t) \\ y'(t) \\ z'(t) \end{pmatrix}$$

$$= \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} + \frac{\partial f}{\partial z}\frac{dz}{dt}$$

In practice, sometimes we do not think too much about matrices and we just compute mechanically:

$$h = f(\vec{r}(t))$$
$$= f(x(t), y(t), z(t)).$$

$$\frac{\partial h}{\partial t} = \frac{\partial f}{\partial x} \cdot \frac{dx}{dt} + \frac{\partial f}{\partial y} \cdot \frac{dy}{dt} + \frac{\partial f}{\partial z} \cdot \frac{dz}{dt}$$

$$= \nabla f \, (\vec{r}(t)) \cdot \vec{r}\,'(t).$$

Implicit function.

If $H(x,y,z)=0$ is an equation and a function $z=f(x,y)$ is such that

$$H(x,y,f(x,y))=0,$$

then we say that $f(x,y)$ is defined implicitly by $H(x,y,f(x,y))=0$.

<u>Ex</u>: The function $z=f(x,y)=\sin x+\cos y$ is defined implicitly by the equation

$$z\sin x - z^2 - \sin^2 y + \sin x\cos y + 1 = 0$$

because

$$(\sin x+\cos x)\sin x - (\sin x+\cos y)^2 - \sin^2 y + \sin x\cos y + 1$$

$$= \cancel{\sin^2 x} + \cancel{\sin x\cos x} \cancel{-\sin^2 x} \cancel{- 2\sin x\cos y} \cancel{-\cos^2 y}$$
$$\cancel{-\sin^2 y} + \cancel{\sin x\cos y} \cancel{+1} = 0$$

Even though it is often difficult to compute the implicit function, it is easy to find their derivatives.

Ex: $z = f(x,y)$ is defined implicitly by $H(x, y, f(x,y)) = 0$. Compute $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$.

The equation $H(x, y, f(x,y))$ can be seen as a composition of the function $H: \mathbb{R}^3 \to \mathbb{R}$ and $g: \mathbb{R}^2 \to \mathbb{R}^3$, with $g(x, y) = (x, y, f(x,y))$. Taking derivatives on both sides of the equation $H(x, y, f(x,y)) = 0$ we get:

$$\frac{\partial}{\partial x} H(g(x,y)) = 0.$$

Using the chain rule:

$$\frac{\partial H}{\partial x} \cdot \cancel{\frac{dx}{dx}}^{1} + \frac{\partial H}{\partial y} \cancel{\frac{dy}{\partial x}}^{0} + \frac{\partial H}{\partial z} \cdot \frac{\partial z}{\partial x} = 0$$

$$\therefore \quad \frac{\partial z}{\partial x} = -\frac{\partial H/\partial x}{\partial H/\partial z}, \quad \text{if } \frac{\partial H}{\partial z} \neq 0.$$

Similarly:

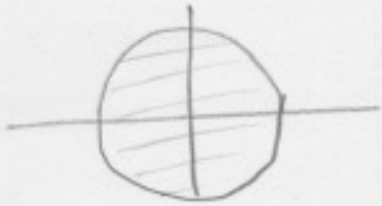
$$\frac{\partial z}{\partial y} = -\frac{\partial H/\partial y}{\partial H/\partial z}.$$

Another example of the use of chain rule: converting to polar coordinates.

Motivation: Let $u(x,y): \mathbb{R}^2 \to \mathbb{R}$, the Laplace operator, denoted as Δ, applied to u is defined as:

$$\Delta u := \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}$$
$$= \frac{\partial}{\partial x}\left(\partial u/\partial x\right) + \frac{\partial}{\partial y}\left(\frac{\partial u}{\partial y}\right)$$

Consider the disk $x^2+y^2 \leq 1$:



Suppose we want to find a function $u(x,y)$ such that the following is true:

$$(*) \begin{cases} \Delta u(x,y) = 0, & \text{if } (x,y) \in \{x^2+y^2<1\} \\ u(x,y) = g(x,y), & \text{if } (x,y) \text{ is such that } x^2+y^2=1 \end{cases}$$

The function g is given to us and we need to find u. For example, suppose:

$g(x,y) \equiv 1$, for all (x,y) in the boundary of the disk.

How do we find u?, We notice that, for example, if $u(x,y)=x^2-y^2$, we have:

$$\frac{\partial^2 u}{\partial x^2}=\frac{\partial}{\partial x}\left(\frac{\partial u}{\partial x}\right)=\frac{\partial}{\partial x}(2x)=2$$

$$\frac{\partial^2 u}{\partial y^2}=\frac{\partial}{\partial y}\left(\frac{\partial u}{\partial y}\right)=\frac{\partial}{\partial y}(-2y)=-2$$

$$\therefore \quad \frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}=0 \; ; \quad \text{that is, } \Delta u=0 .$$

<u>Def</u>: If u solves the Laplace equation $\Delta u=0$, then u is called a harmonic function

However if $g\equiv 1$, then the boundary condition $u(x,y)=g(x,y)$ is not satisfied.

Thus, in order to solve $(*)$, it is better if we work in polar coordinates and we look for a function $u(r,\theta)$ instead of $u(x,y)$.

Hence, we need to write the Laplace operator in polar coordinates. We have:

$$u(r,\theta), \quad r(x,y)=\sqrt{x^2+y^2}$$

$$\theta(x,y)=\tan^{-1}\frac{y}{x}, \quad \text{since} \quad \begin{array}{l} x=r\cos\theta \\ y=r\sin\theta \end{array}$$

We first write $\frac{\partial u}{\partial x}$, $\frac{\partial u}{\partial y}$ in terms of r and θ.

$$\frac{\partial u}{\partial x} = \frac{\partial u}{\partial r}\cdot\frac{\partial r}{\partial x} + \frac{\partial u}{\partial \theta}\cdot\frac{\partial \theta}{\partial x}$$

$$\frac{\partial u}{\partial y} = \frac{\partial u}{\partial r}\cdot\frac{\partial r}{\partial y} + \frac{\partial u}{\partial \theta}\cdot\frac{\partial \theta}{\partial y}$$

$$\Rightarrow r(x,y) = (x^2+y^2)^{1/2} \Rightarrow \frac{\partial r}{\partial x} = \frac{1}{2}(x^2+y^2)^{-1/2}(2x) = \frac{x}{\sqrt{x^2+y^2}} = \frac{x}{r} = \cos\theta$$

$$\frac{\partial r}{\partial y} = \frac{1}{2}(x^2+y^2)^{-1/2}(2y) = \frac{y}{\sqrt{x^2+y^2}} = \frac{y}{r} = \sin\theta$$

$$\Rightarrow \tan\theta = \frac{y}{x} \Rightarrow \sec^2\theta\cdot\frac{\partial\theta}{\partial x} = \frac{-y}{x^2} \Rightarrow \frac{\partial\theta}{\partial x} = \frac{-y}{x^2\sec^2\theta}$$

$$= \frac{-r\sin\theta}{r^2}$$

$$= \frac{-\sin\theta}{r}$$

$$\sec^2\theta\,\frac{\partial\theta}{\partial y} = \frac{1}{x} \Rightarrow \frac{\partial\theta}{\partial y} = \frac{1}{x\sec^2\theta}$$

$$= \frac{\cos^2\theta}{r\cos\theta}$$

$$= \frac{\cos\theta}{r}$$

$$\therefore \boxed{\begin{aligned} \frac{\partial u}{\partial x} &= \frac{\partial u}{\partial r} \cdot \cos\theta - \frac{\partial u}{\partial \theta} \frac{\sin\theta}{r} \\ \frac{\partial u}{\partial y} &= \frac{\partial u}{\partial r} \sin\theta + \frac{\partial u}{\partial \theta} \frac{\cos\theta}{r} \end{aligned}}$$

We now need to take one more derivative and compute $\frac{\partial}{\partial x}(\partial u/\partial x)$, $\frac{\partial}{\partial y}(\partial u/\partial y)$ in terms of r and θ. We will do that in section 3.1.