Section 2.5: Properties of the Derivative

Thm: Properties of the Derivative:

Assume $f, g: U \subset \mathbb{R}^n \to \mathbb{R}^m$ diff at $\vec{x}_0$, let c be a constant.

- Const Rule: $h(\vec{x}) = c f(\vec{x})$ Then $Dh(\vec{x}_0) = c\,Df(\vec{x}_0)$
- Sum Rule: $h(\vec{x}) = f(\vec{x}) + g(\vec{x})$ Then $Dh(\vec{x}_0) = Df(\vec{x}_0) + Dg(\vec{x}_0)$
- Product Rule: $m=1$, $h(\vec{x}) = f(\vec{x}) g(\vec{x})$ Then $Dh(\vec{x}_0) = Df(\vec{x}_0)\, g(\vec{x}_0) + f(\vec{x}_0)\, Dg(\vec{x}_0)$
- Quotient Rule: $m=1$, $h(\vec{x}) = f(\vec{x})/g(\vec{x})$ Then $Dh(\vec{x}_0) = \dfrac{Df(\vec{x}_0) g(\vec{x}_0) - f(\vec{x}_0) Dg(\vec{x}_0)}{g(\vec{x}_0)^2}$

Example: $f(x,y,z) = x^2+y^2+z^2 \quad g(x) = x^3+y^3$

$$h(x,y,z) = f(x,y,z)\,g(x,y,z) = x^5 + x^3y^2 + x^3z^2 + x^2y^3 + y^5 + z^2y^3$$

$$Dh(x,y,z) = \left(\frac{\partial h}{\partial x}, \frac{\partial h}{\partial y}, \frac{\partial h}{\partial z}\right) = \big(5x^4 + 3x^2y^2 + 3x^2z^2 + 2xy^3,\ 2x^3y + 3x^2y^2 + 5y^4 + 3z^2y^2,\ 2x^3z + 2zy^3\big)$$

$$Df(x,y,z) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right) = (2x, 2y, 2z) \qquad f(x,y,z) = x^2+y^2+z^2$$

$$Dg(x,y,z) = \left(\frac{\partial g}{\partial x}, \frac{\partial g}{\partial y}, \frac{\partial g}{\partial z}\right) = (3x^2, 3y^2, 0)$$

$$Df(x,y,z)\,g(x,y,z) + f(x,y,z)\,Dg(x,y,z) = (2x, 2y, 2z) * (x^3+y^3) + (x^2+y^2+z^2) * (3x^2, 3y^2, 0)$$

$$\cancel{= 3x^4 + 2x^3y + 2x^3z}$$

$$= \big(5x^4 + 3x^2y^2 + 3x^2z^2 + 2xy^3,\ 5y^4 + 2x^3y + 3x^2y^2 + 3z^2y^2,\ 2x^3z + 2y^3z\big)$$

↳ Note agree. Faster is to note symmetry in $x \to y$ and $y \to x$

Proof of product rule

↳ Use powerful technique of adding zero!

$$\lim_{\vec{x}\to\vec{x}_0} \frac{\| f(\vec{x})g(\vec{x}) - f(\vec{x}_0)g(\vec{x}_0) - [Df(\vec{x}_0)g(\vec{x}_0) + f(\vec{x}_0)Dg(\vec{x}_0)](\vec{x}-\vec{x}_0)\|}{\|\vec{x}-\vec{x}_0\|}$$

↳ add $-f(\vec{x}_0)g(\vec{x}) + f(\vec{x}_0)g(\vec{x})$

use triangle inequality: $\|\vec{P}+\vec{Q}\| \le \|\vec{P}\| + \|\vec{Q}\|$

$$\le \lim_{\vec{x}\to\vec{x}_0} \frac{\|[f(\vec{x}) - f(\vec{x}_0) - Df(\vec{x}_0)(\vec{x}-\vec{x}_0)]\|\,|g(\vec{x})|}{\|\vec{x}-\vec{x}_0\|}$$

$$+ \lim_{\vec{x}\to\vec{x}_0} \frac{\| Df(\vec{x}_0)(g(\vec{x})-g(\vec{x}_0))(\vec{x}-\vec{x}_0)\|}{\|\vec{x}-\vec{x}_0\|}$$

$$+ \lim_{\vec{x}\to\vec{x}_0} \frac{\|f(\vec{x}_0)\|\,\|g(\vec{x}) - g(\vec{x}_0) - Dg(\vec{x}_0)(\vec{x}-\vec{x}_0)\|}{\|\vec{x}-\vec{x}_0\|}$$

and each piece tends to zero. ∎

SECTION 2.5 (CONT)

THM: THE CHAIN RULE: f: open $V \subset \mathbb{R}^m \to \mathbb{R}^p$ and g: open $U \subset \mathbb{R}^n \to \mathbb{R}^m$ with $g(U) \subset V$. Suppose g is diff at $\vec{x}_0$ and f is diff at $\vec{y}_0 = g(\vec{x}_0)$. Then $h(\vec{x}) = (f \circ g)(\vec{x}) = f(g(\vec{x}))$ is diff at $\vec{x}_0$ and $Dh(\vec{x}_0) = Df(\cancel{\vec{x}_0})\, Dg(\vec{x}_0)$

Example: $g(x,y) = (\cos x, \sin x, y)$

$$f(u,v,w) = u^2 + v^2 + uvw$$

$$h(x,y) = f(g(x,y)) = f(\cos x, \sin x, y)$$

$$= \cos^2 x + \sin^2 x + \cos x \cdot \sin x \cdot y$$

$$(Dh)(x,y) = \left(\frac{\partial h}{\partial x}, \frac{\partial h}{\partial y}\right) = \left((-\sin^2 x + \cos^2 x)y,\ \cos x \cdot \sin x\right)$$

$$\cancel{Dg(x,y) = \left(\frac{\partial g}{\partial x}, \frac{\partial g}{\partial y}\right) = (-\sin x}$$

$$\cancel{Dg(x,y) = \begin{pmatrix} \frac{\partial g_1}{\partial x} & \frac{\partial g_2}{\partial x} & \frac{\partial g_3}{\partial x} \\ \frac{\partial g_1}{\partial y} & \frac{\partial g_2}{\partial y} & \frac{\partial g_3}{\partial y} \end{pmatrix}}$$

$$Dg(x,y) = \begin{pmatrix} \frac{\partial g_1}{\partial x} & \frac{\partial g_1}{\partial y} \\ \frac{\partial g_2}{\partial x} & \frac{\partial g_2}{\partial y} \\ \frac{\partial g_3}{\partial x} & \frac{\partial g_3}{\partial y} \end{pmatrix} = \begin{pmatrix} -\sin x, & 0 \\ \cos x, & 0 \\ 0, & 1 \end{pmatrix}$$

SECTION 2.5 (CONT)

$$Df(u,v,w) = \left(\frac{\partial f}{\partial u}, \frac{\partial f}{\partial v}, \frac{\partial f}{\partial w}\right) = \cancel{(2u, 2v)}$$

$$= (2u + vw,\ 2v + uw,\ uv)$$

$$Df(g(x,y)) = (2\cos x + y\sin x,\ 2\sin x + y\cos x,\ \cos x\sin x)$$

$$Df(g(x,y))\,Dg(x,y) = (2\cos x + y\sin x,\ 2\sin x + y\cos x,\ \cos x \sin x)\begin{pmatrix} -\sin x & 0 \\ \cos x & 0 \\ 0 & 1 \end{pmatrix}$$

$$= \left(-2\cos x\sin x - y\sin^2 x + 2\cos x\sin x + y\cos^2 x,\ \cos x\sin x\right)$$

CAVEATS

① Construct the right matrix of partials:

$g: \mathbb{R}^n \to \mathbb{R}^m$ so $g(x_1,\dots,x_n) = (g_1(x_1,\dots,x_n),\dots,g_m(x_1,\dots,x_n))$

Then
$$Dg(\vec{x}) = \begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial g_m}{\partial x_1} & \cdots & \frac{\partial g_m}{\partial x_n} \end{pmatrix} \begin{array}{l} \leftarrow \text{first row all } g_1 \\ \vdots \\ \leftarrow \text{last row all } g_m \end{array}$$

② Matrix multiplication, and evaluate Df at $g(\vec{x}_0)$

Section 2.5 (cont)

Formula for the chain rule

$h : \mathbb{R}^m \to \mathbb{R}$, $f : \mathbb{R}^n \to \mathbb{R}$ $g : \mathbb{R}^m \to \mathbb{R}^n$

Say h is a function of $u_1, \ldots, u_m$, as is g

say f is a function of $x_1, \ldots, x_n$

Then $$\frac{\partial h}{\partial u_1} = \frac{\partial f}{\partial x_1}\frac{\partial x_1}{\partial u_1} + \cdots + \frac{\partial f}{\partial x_n}\frac{\partial x_n}{\partial u_1}$$

$$\vdots$$

$$\frac{\partial h}{\partial u_m} = \frac{\partial f}{\partial x_1}\frac{\partial x_1}{\partial u_m} + \cdots + \frac{\partial f}{\partial x_n}\frac{\partial x_n}{\partial u_m}$$

↳ where we have $x_1 = x_1(u_1, \ldots, u_m)$ thru $x_n = x_n(u_1, \ldots, u_m)$

<u>Special Case</u>: $c : \mathbb{R} \to \mathbb{R}^m$, $h(t) = f(c(t))$

Proof : $$Dh(t_0) = \lim_{t \to t_0} \frac{h(t) - h(t_0)}{t - t_0}$$ $c(t) = (x_1(t), \ldots, x_n(t))$

$$= \lim_{t \to t_0} \frac{f(x_1(t), \ldots, x_n(t)) - f(x_1(t_0), \ldots, x_n(t_0))}{t - t_0}$$

$$= \lim_{t \to t_0} \left(\frac{f(x_1(t), \ldots, x_n(t)) - f(x_1(t_0), x_2(t), \ldots, x_n(t))}{t - t_0}\right)$$

$$+ \cdots + \lim_{t \to t_0} \left(\frac{f(x_1(t_0), \ldots, x_{n-1}(t_0), x_n(t)) - f(x_1(t_0), \ldots, x_n(t_0))}{t - t_0}\right)$$

Math 350: The Chain Rule

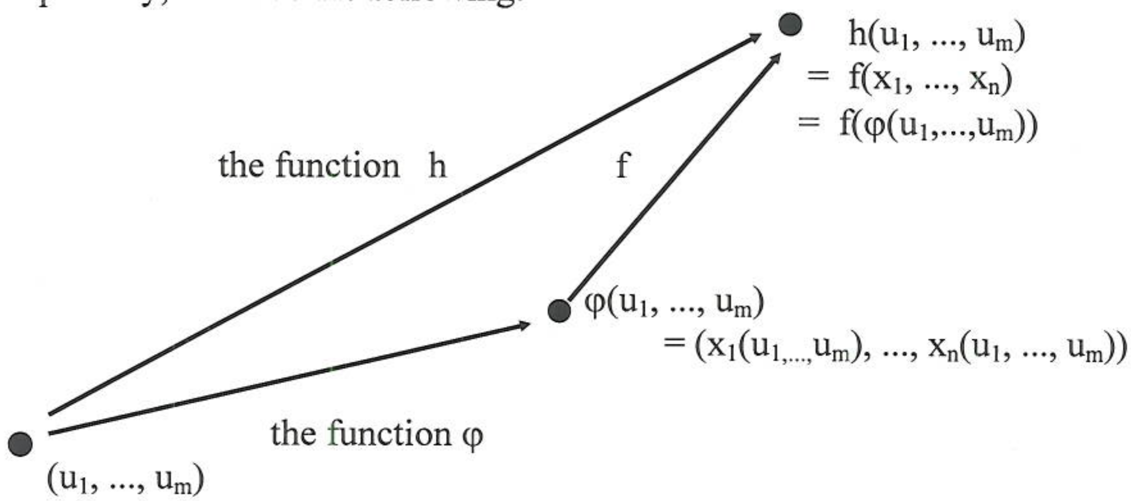
The Chain Rule is a very useful tool for analyzing the following: Say you have a function f of $(x_1, x_2, ..., x_n)$, and these variables are themselves functions of $(u_1, u_2, ..., u_m)$. How does our function f change as we vary u_1 thru u_m??? We'll state and explain the Chain Rule, and then give a DIFFERENT PROOF FROM THE BOOK, using *only* the definition of the derivative. This is a slight modification of notes I wrote years ago for a similar class at Princeton.

(I). Statement:

We'll state the Chain Rule. First, some notation:

Let $h: \mathbb{R}^m \to \mathbb{R}$	say h is a function of $(u_1, u_2, ..., u_m)$
$f: \mathbb{R}^n \to \mathbb{R}$	say f is a function of $(x_1, x_2, ..., x_n)$
$\varphi: \mathbb{R}^m \to \mathbb{R}^n$	say φ is a function of $(u_1, u_2, ..., u_m)$

Graphically, we have the following:



Our function h lives on $\mathbb{R}^m$. So, you give it an m-tuple, $(u_1, ..., u_m)$, and it will give you a real number back. The function f lives on $\mathbb{R}^n$. If you give it an n-tuple, $(x_1, ..., x_n)$, it will give you back a number. And what of the variables x_1 thru x_n? Well, they can be thought of as functions on $\mathbb{R}^m$: you give them an m-tuple, $(u_1, ..., u_m)$, and they'll return a number.

We cannot look at $f(x_1(u_1, ..., u_m))$, for f composed with x_1 doesn't make sense: x_1 gives us just ONE number; f needs n numbers.

What do we do? Remember, we're trying to understand the beast:

$$h(u_1, ..., u_m) = f(x_1(u_1, ..., u_m), ..., x_n(u_1, ..., u_m))$$

We define an auxiliary function, φ, to help us. What will $\varphi(u_1, \ldots, u_m)$ be? Whatever we want. We now look for something useful. Look at the Right Hand Side above—wouldn't it be nice if we could choose a φ that would give us this? We can! Just let:

$$\varphi(u_1, \ldots, u_m) = (x_1(u_1, \ldots, u_m), x_2(u_1, \ldots, u_m), \ldots, x_n(u_1, \ldots, u_m))$$

Now we can write $h = f \circ \varphi$, f composed with φ. The advantage of this is that we know that often compositions of nice functions are nice: if we compose two continuous functions, we get a continuous function. In one dimension, we have the 1-dimensional chain rule for compositions. We hope to be able to do something similar here. Anyway, here is the long awaited statement of:

The Chain Rule:

$$\begin{aligned}(Dh)(u_1, \ldots, u_m) &= (Df)(\varphi\,(u_1, \ldots, u_m))\,(D\varphi)(u_1, \ldots, u_m)\\ &= (Df)(x_1, \ldots, x_n)\,(D\varphi)(u_1, \ldots, u_m)\end{aligned}$$

Let's write out what this is: for the sake of space, I will not explicitly write WHERE the functions are being evaluated—we always evaluate h at $(u_1, \ldots, u_m)$, f at $\varphi(u_1, \ldots, u_m) = (x_1(u_1, \ldots, u_m), \ldots, x_n(u_1, \ldots, u_m))$, and φ at $(u_1, \ldots, u_m)$.

The Chain Rule:

$$\mathbf{Dh} = \left(\frac{\partial h}{\partial u_1}, \frac{\partial h}{\partial u_2}, \ldots, \frac{\partial h}{\partial u_m}\right) \quad \mathbf{Df} = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n}\right)$$

$D\varphi$ is more complicated: Unlike Df and Dh, which are vectors, $D\varphi$ is a matrix quantity. This is because φ is really a collection of m functions,

$$\begin{aligned}\varphi(u_1, \ldots, u_m) &= (\varphi_1(u_1, \ldots, u_m), \ldots., \varphi_n(u_1, \ldots, u_m))\\ &= (x_1(u_1, \ldots, u_m), \ldots., x_n(u_1, \ldots, u_m))\end{aligned}$$

We obtain:

$$
(D\varphi) = \begin{pmatrix} \frac{\partial x_1}{\partial u_1}, & \frac{\partial x_1}{\partial u_2}, & \ldots, & \frac{\partial x_1}{\partial u_m} \\ \frac{\partial x_2}{\partial u_1}, & \frac{\partial x_2}{\partial u_2}, & \ldots, & \frac{\partial x_2}{\partial u_m} \\ & & & \\ \frac{\partial x_n}{\partial u_1}, & \frac{\partial x_n}{\partial u_2}, & \ldots, & \frac{\partial x_n}{\partial u_m} \end{pmatrix}
$$

Combining the above expressions for Dh, Df, and $D\varphi$ yields:

Chain Rule:

$$
\frac{\partial h}{\partial u_1} = \frac{\partial f}{\partial x_1}\frac{\partial x_1}{\partial u_1} + \frac{\partial f}{\partial x_2}\frac{\partial x_2}{\partial u_1} + \ldots + \frac{\partial f}{\partial x_n}\frac{\partial x_n}{\partial u_1}
$$

$$
\frac{\partial h}{\partial u_2} = \frac{\partial f}{\partial x_1}\frac{\partial x_1}{\partial u_2} + \frac{\partial f}{\partial x_2}\frac{\partial x_2}{\partial u_2} + \ldots + \frac{\partial f}{\partial x_n}\frac{\partial x_n}{\partial u_2}
$$

and so on till

$$
\frac{\partial h}{\partial u_m} = \frac{\partial f}{\partial x_1}\frac{\partial x_1}{\partial u_m} + \frac{\partial f}{\partial x_2}\frac{\partial x_2}{\partial u_m} + \ldots + \frac{\partial f}{\partial x_n}\frac{\partial x_n}{\partial u_m}
$$

<u>(II). One Dimensional Case:</u>

OK. We now have the above formula, but WHERE DID IT COME FROM? Let's go back to one-dimension, and take a look at what is happening:

Translating from our language to what we spoke in High School:

$$h(u) = f(\varphi(u)) \rightarrow h'(u) = f'(\varphi(u))\,\varphi'(u)$$

How do we go about proving this? Always go back to what you know: here we're trying to find the derivative. Okay, so, let's recall the definition of the derivative. We know that. The derivative is defined by:

$$\begin{aligned} h'(u) &= \lim_{y\to u} \{h(y) - h(u)\} / \{y-u\} \\ &= \lim_{y\to u} \{f(\varphi(y)) - f(\varphi(u))\} / \{y-u\} \\ &= \lim_{y\to u} \frac{f(\varphi(y)) - f(\varphi(u))}{\varphi(y) - \varphi(u)} * \frac{\varphi(y) - \varphi(u)}{y - u} \end{aligned}$$

All we did was multiply by 1 in a very clever way. Why did we do this? Our function f is a function of one variable. The second term looks like $\varphi'(u)$ in the limit, and the first term looks like f' evaluated at $\varphi(u)$. As the two limits exist, the limit of the product is the product of the limits, so we can conclude:

$$h'(u) = f'(\varphi(u))\,\varphi'(u)$$

Why isn't this proof rigorous? The definition of $f'(z)$ is the following:

$$f'(z) = \lim_{w\to z} \{f(w) - f(z)\} / \{w - z\}$$

We cheated in the above: this limit has to hold *FOR ALL* paths where w heads to z. We didn't consider *all* paths, only a special path. But maybe this isn't too bad: if the limit exists, then it doesn't matter *WHICH* path we take. In better words: look, I know $f'(z)$ exists, and I know the value is *INDEPENDENT* of the path I take. So why don't I just make life easy on myself and take this nice path? What a great idea! We leave for the interested, rigorous reader what to do if $\varphi(y)$ equals $\varphi(u)$ infinitely often (this cannot happen if $\varphi'(u) \neq 0$). Hint: go back to the definition of $\partial h/\partial u$ and calculate it directly, going along points where $\varphi(y) = \varphi(u)$.

(III). Higher Dimensions:

We now argue as in above, but in higher dimensions. To make things easier to view, let's just look at $n = 3$, $m = 2$, so we have (x_1, x_2, x_3), which we denote by (x, y, z) for convenience, and (u_1, u_2), which we denote by (u, w).

$$h(u,w) = f(x(u,w), y(u,w), z(u,w))$$

We calculate $\partial h/\partial u$, at the point (u,w), and compare with $\partial h/\partial u_1$ from page 3.

$$\partial h/\partial u = \lim_{v \to u} \{ h(v, w) - h(u, w) \} / \{ v - u \}$$

$$= \lim_{v \to u} \frac{f(x(v,w), y(v,w), z(v,w)) - f(x(u,w), y(u,w), z(u,w))}{v - u}$$

So, we start at the point (x(u,w), y(u,w), z(u,w)) and we finish at the point (x(v,w), y(v,w), z(v,w)). We cannot directly mimic the 1-dimensional case, but what if our starting point were (x(u,w), y(v,w), z(v,w))? Then all we would've done is change the x-coordinate of the 3-tuple, and we could multiply and divide by x(v,w) - x(u,w). We would then have:

$$\partial f/\partial x \quad \partial x/\partial u$$

Sadly, life isn't quite that simple: we don't have that as our starting point. But, what if we added and subtracted f(x(u,w), y(v,w), z(v,w)) in the numerator? Then we would get:

$$\frac{\partial h}{\partial u} = \lim_{v \to u} \frac{f(x(v,w),y(v,w),z(v,w)) - f(x(u,w),y(v,w),z(v,w))}{v - u} +$$

$$\lim_{v \to u} \frac{f(x(u,w),y(v,w),z(v,w)) - f(x(u,w),y(u,w),z(u,w))}{v - u}$$

We now multiply the first term by 1:

$$\frac{\partial h}{\partial u} = \lim_{v \to u} \frac{f(x(v,w),y(v,w),z(v,w)) - f(x(u,w),y(v,w),z(v,w))}{x(v,w) - x(u,w)} * \frac{x(v,w) - x(u,w)}{v - u}$$

$$+ \lim_{v \to u} \frac{f(x(u,w),y(v,w),z(v,w)) - f(x(u,w),y(u,w),z(u,w))}{v - u}$$

$$\frac{\partial h}{\partial u} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial u} + \lim_{v \to u} \frac{f(x(u,w),y(v,w),z(v,w)) - f(x(u,w),y(u,w),z(u,w))}{v - u}$$

Now we just repeat what we did before! We've got two points, start at (x(u,w),y(u,w),z(u,w)), end at (x(u,w),y(v,w),z(v,w)). Again, what if our first point were (x(u,w),y(u,w),z(v,w))? Then all we would've done is change the y-coordinate of the 3-tuple, and we could multiply and divide by y(v,w) - y(u,w). We would then (in the limit) get $\partial f/\partial y\ \partial y/\partial u$, plus another term, the difference of the point we added and our *true* first point. Let's do it!

$$\frac{\partial h}{\partial u} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial u} + \lim_{v \to u} \frac{f(x(u,w),y(v,w),z(v,w)) - f(x(u,w),y(u,w),z(v,w))}{v - u}$$
$$+ \lim_{v \to u} \frac{f(x(u,w),y(u,w),z(v,w)) - f(x(u,w),y(u,w),z(u,w))}{v - u}$$

Multiplying the first limit by {y(v,w) - y(u,w)} / {y(v,w) - y(u,w)} we get:

$$\frac{\partial h}{\partial u} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial u} + \lim_{v \to u} \frac{f(x(u,w),y(u,w),z(v,w))-f(x(u,w),y(u,w),z(u,w))}{v - u}$$

Multiplying the last term by {z(v,w) - z(v,w)} / {z(v,w) - z(v,w)}, we get that this term, in the limit, is just $\partial f/\partial z\ \partial z/\partial u$.

Hence we get:

$$\boxed{\frac{\partial h}{\partial u} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial u} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial u} \quad \textbf{which is The Chain Rule!}}$$

Consider $f(x(v,w), y(v,w), z(v,w)) - f(x(u,w), y(v,w), z(v,w))$

(MORE CHAIN RULE)

By the defn of the deriv

$$f(\vec{r}_1) - f(\vec{r}_0) - (Df)(\vec{r}_0)(\vec{r}_1 - \vec{r}_0) = \text{small}$$

$$f(\vec{r}_2) - f(\vec{r}_0) - (Df)(\vec{r}_0)(\vec{r}_2 - \vec{r}_0) = \text{small}$$

where $\vec{r}_0 = (x(u,w), y(u,w), z(u,w))$ all u's

$\vec{r}_1 = (x(u,w), y(v,w), z(v,w))$ one u, two v's

$\vec{r}_2 = (x(v,w), y(v,w), z(v,w))$ all v's

Thus $f(\vec{r}_2) - f(\vec{r}_1) = [f(\vec{r}_2) - f(\vec{r}_0)] - [f(\vec{r}_1) - f(\vec{r}_0)]$

$$= -(Df)(\vec{r}_0)(\vec{r}_2 - \vec{r}_0) + (Df)(\vec{r}_0)(\vec{r}_1 - \vec{r}_0) + \text{small}$$

$$= -(Df)(\vec{r}_0)\vec{r}_2 + \underbrace{(Df)(\vec{r}_0)\vec{r}_0 + (Df)(\vec{r}_0)\vec{r}_1 - (Df)(\vec{r}_0)\vec{r}_0}_{\text{cancel}} + \text{small}$$

$$= -(Df)(\vec{r}_0)(\vec{r}_2 - \vec{r}_1) + \text{small}$$

$$= -(Df)(\vec{r}_0)\cdot(x(v,w) - x(u,w), 0, 0)$$

This will lead to $\frac{\partial f}{\partial x}\frac{\partial x}{\partial u}$, as we divide by $v-u$ and then $\frac{x(v,w) - x(u,w)}{v-u} \to \frac{\partial x}{\partial u}$

SECTION 2.6: GRADIENTS AND DIRECTIONAL DERIVATIVES

- **Gradient**: $\text{grad}(f) = \nabla f = \left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}\right)$

  ↳ This is derivative Df written as a vector

- **Directional Deriv**: Directional derivative of f at $\vec{x}$ along vector $\vec{v}$ (usually unit length) is $\frac{d}{dt} f(\vec{x} + t\vec{v})$. If $\|\vec{v}\| = 1$ say the directional derivative in the direction of $\vec{v}$.

  ↳ if $\|\vec{v}\| \neq 1$, changing scale

  ↳ Equivalent to $\lim\limits_{h \to 0} \frac{f(\vec{x} + h\vec{v}) - f(\vec{x})}{h}$

> THM: $f: \mathbb{R}^n \to \mathbb{R}$ diff. Then all derivs exist and the dir deriv in the dir of $\vec{v}$ is $Df(\vec{x})\vec{v} = \nabla f(\vec{x}) \cdot \vec{v} = \frac{\partial f}{\partial x_1}(\vec{x}) v_1 + \cdots + \frac{\partial f}{\partial x_n}(\vec{x}) v_n$

Proof: Let $c(t) = \vec{x} + t\vec{v}$ and use chain rule

Note $c'(0) = \vec{v}$, $c(0) = \vec{x}$

Thus $\frac{d}{dt} f(c(t)) = Df(c(0))\, c'(t) = \nabla f(c(0)) \cdot \vec{v}$ ∎

> THM: GEOMETRIC INTERPRETATION:
>
> If $\nabla f(\vec{x}_0) \neq \vec{0}$ then $\nabla f(\vec{x}_0)$ points in dir of fastest increase of f

Proof: Rate of change of f in unit dir $\vec{n}$ is $\nabla f(\vec{x}_0) \cdot \vec{n}$

Have magnitude $\|\nabla f(\vec{x}_0)\| \cdot \|\vec{n}\| \cdot |\cos\theta|$, largest when $\theta = 0, \pi$

so parallel ($\theta = 0$ gives max, $\theta = \pi$ gives min)

Section 2.6 (Cont)

THM: Gradient is normal to level surfaces. Specifically, $f: \mathbb{R}^n \to \mathbb{R}$ a C^1 map, $\vec{x}_0$ in level surface S defined by $f(\vec{x}) = k$. If curve $c(t)$ in S with $c(0) = \vec{x}_0$ and $\vec{v} = c'(t)$ is the tangent vector at $t=0$, then $\nabla f(\vec{x}_0) \cdot \vec{v} = \vec{0}$

Proof: Chain rule again!

Apply to $h(t) = f(c(t)) = k$ ∎

Note: These results will be <u>VERY</u> useful for max/min problems

<u>Defn: Tangent Plane</u>: S be surface $f(\vec{x}) = k$. The tangent plane at $\vec{x}_0$ is defined by $\{\vec{x} : \nabla f(\vec{x}_0) \cdot (\vec{x} - \vec{x}_0) = 0\}$

Often call ∇f the gradient vector field; means at point $\vec{P}$ draw vector $\nabla f(\vec{P})$.

↳ Example: Gravity: $\vec{F}(x,y,z) = -G\frac{m_1 m_2}{r^2}\vec{n} = \nabla\left(\frac{G m_1 m_2}{r}\right)$

where $\vec{n} = \vec{r}/r$, $\vec{r} = (x,y,z)$

Homework: #2ab, #4a, #6a, #16 (Ralph), #18

Suggested: #5a, #12, #17, #21, #23

Review Problems

HW: Pg 176: #23 homog

#47 chemistry

Suggested: Pg 176: #26, #41

#42

Additional Comments: Section 2.6: Gradients and Dir Derivs

Start with $\frac{\partial f}{\partial x_i} = \lim_{h \to 0} \frac{f(x_1, \ldots, x_{i-1}, x_i + h, x_{i+1}, \ldots, x_n) - f(x_1, \ldots, x_n)}{h}$

$$= \lim_{h \to 0} \frac{f(\vec{x} + h\vec{e}_i) - f(\vec{x})}{h}$$

Generalizes to $\lim_{h \to 0} \frac{f(\vec{x} + h\vec{v}) - f(\vec{x})}{h}$, recover when $\vec{v} = \vec{e}_i$

↳ limit formula gives directional deriv, but computationally bad

↳ ex: consider $f(x,y) = x^3 y + x\cos y$, $\vec{v} = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$

↳ gives $\lim_{h \to 0} \frac{f(x + h/\sqrt{2},\ y + h/\sqrt{2}) - f(x,y)}{h}$

↳ don't want to compute limit! is there a better way?

<u>Chain Rule</u>: $c(t) = \vec{x} + t\vec{v}$, so $c(0) = \vec{x}$, $c(h) = \vec{x} + t\vec{v}$

↳ Note $A(t) = f(c(t)) \Rightarrow \frac{dA}{dt}(0) = (Df)(c(0)) \cdot c'(0)$

And $\frac{dA}{dt}(0) = \lim_{h \to 0} \frac{f(\vec{x} + h\vec{v}) - f(\vec{x})}{h}$ equal

↳ Thus can compute dir derivs easily

Do Example: #2C from book: $f(x,y) = e^x \cos \pi y$ at $(0,-1)$ in dir $\left(-\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}\right)$

Move onto geometric interpretation

↳ emphasize geometry, dividends from dotproduct formula

Level sets: see how useful, give example

Additional Comments: Sec 2.6: CONTINUED

Thm: Gradient is normal to Surface S where $f(\vec{x}) = k$ constant

↳ Example: $f(x,y,z) = x^2+y^2+z^2$

$$c(t) = (\cos t, \sin t, 0) \qquad c(0) = (1,0,0) = \vec{x_0}$$

$$c'(t) = (-\sin t, \cos t, 0) \qquad c'(0) = (0,1,0) = \vec{v}$$

$$(\nabla f)(\vec{x}) = (2x, 2y, 2z) \quad (\nabla f)(1,0,0) = (2,0,0)$$

$$\text{So } (\nabla f)(1,0,0)\cdot(0,1,0) = (2,0,0)\cdot(0,1,0) = 0$$

Do example of computing tangent plane:

$$f(x,y,z) = xyz = 1, \quad \text{point } (1,1,1)$$

Other Remarks:

In $A(t) = f(c(t))$, get $\frac{dA}{dt} = (Df)(c(0))\ c'(0)$

becomes $(\nabla f)(\vec{x})\cdot\vec{v}$; note $c'(0)$ column vector.

Why? $c(t) = (x_1+tv_1, \dots, x_n+tv_n)$ has n components

$$(Dc)(t) = \begin{pmatrix} \nabla c_1 \\ \vdots \\ \nabla c_n \end{pmatrix} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$