Math 2210-1
Monday 10/18

§ 13.6-13.8 (continued Wednesday!)
affine approximation and differentiability; directional derivatives.

Recall on Friday we studied differential approximation $df_{\vec{x}} = \frac{\partial f}{\partial x_1}\Delta x_1 + \dots + \frac{\partial f}{\partial x_n}\Delta x_n$
as a way of approximating $\Delta f = f(\vec{x}+\Delta\vec{x}) - f(\vec{x})$.
And at the end of class we even considered.

$$\vec{F}(\vec{x}) = \begin{bmatrix} f_1(\vec{x}) \\ \vdots \\ f_m(\vec{x}) \end{bmatrix} \quad ; \quad \Delta\vec{F} = \vec{F}(\vec{x}+\Delta\vec{x}) - F(\vec{x}) \approx d\vec{F} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(\vec{x}), \frac{\partial f_1}{\partial x_2}, \dots \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} \dots \\ \vdots \\ \frac{\partial f_m}{\partial x_1} \dots \frac{\partial f_m(\vec{x})}{\partial x_n} \end{bmatrix} \begin{bmatrix} \Delta x_1 \\ \Delta x_2 \\ \vdots \\ \Delta x_n \end{bmatrix}.$$

called this $m \times n$ matrix $[F'(x)]$.

We never actually worried about when differential approximation works
~ which is exactly the condition we define as "<u>differentiable</u>":

$\vec{F}: D \to \mathbb{R}^m$ is differentiable at $\vec{a} \in D$ ($D \subset \mathbb{R}^n$)
iff the approximation formula

$$* \boxed{\vec{F}(\vec{a}+\vec{h}) - \vec{F}(\vec{a}) = C\vec{h} + \text{error}}$$

holds, where $C_{m\times n}$ is a matrix, and $\boxed{\frac{|\text{error}|}{|h|} \to 0 \text{ as } \vec{h} \to \vec{0}.}$

<u>Remarks</u>: The texts solves for "error" and writes
$$\frac{|\vec{F}(a+h) - \vec{F}(a) - Ch|}{|h|} \to 0 \text{ instead.}$$

Also, if $*$ holds (with the error estimate), then it is "easy" to deduce that C actually equals the derivative matrix $[F'(a)]$. $\longrightarrow \boxed{C = [F'(a)]}$
(However, the affine approx formula is not implied by just knowing that all the partial derivs of $\vec{F}$ exist at $\vec{a}$ — in general you need a little more, see below.)

e.g. $f: \mathbb{R}^n \to \mathbb{R}$

if $f(\vec{a}+\vec{h}) - f(\vec{a}) = [c_1, c_2, \dots c_n]\begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{bmatrix}$ + error

then
$$f(\vec{a}+h\vec{e}_j) - f(\vec{a}) = [c_1 \dots c_n]\begin{bmatrix} 0 \\ \vdots \\ h \\ \vdots \\ 0 \end{bmatrix} \leftarrow \text{entry}_j \quad + \text{error}$$
$$= hc_j + \text{error}$$

$\Rightarrow \frac{f(\vec{a}+h\vec{e}_j) - f(\vec{a})}{h} = c_j + \frac{\text{error}}{h}$. Now let $h \to 0$ (deduce $\frac{\partial f}{\partial x_j}(\vec{a}) = c_j$)

Examples

1 $\vec{r} : I \to \mathbb{R}^n$, $I \subset \mathbb{R}$

we said $\vec{r}$ diff'ble at t meant

old def: $\boxed{\lim_{h \to 0} \frac{\vec{r}(t+h) - \vec{r}(t)}{h} := \vec{r}\,'(t) \text{ exists.}}$

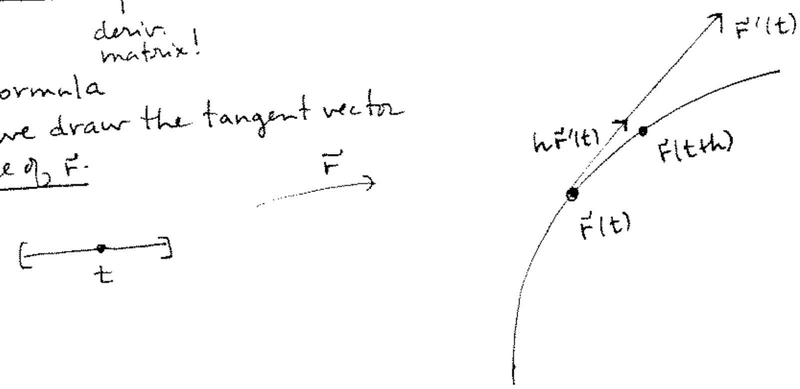
This is same as

$$\frac{\vec{r}(t+h) - \vec{r}(t)}{h} = \vec{r}\,'(t) + \varepsilon \quad , \quad \varepsilon \to 0 \text{ as } h \to 0$$

or

new def: $\boxed{\vec{r}(t+h) - \vec{r}(t) = \underbrace{\left[\vec{r}\,'(t)\right]}_{\text{deriv. matrix!}} h + \underbrace{\varepsilon h}_{\text{"error"}}} \quad , \quad \frac{|\text{error}|}{h} \to 0 \text{ as } h \to 0.$

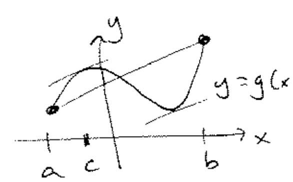
geometric rep. of deriv. matrix & approx. formula
$\to$ in this case we draw the tangent vector in the range of $\vec{r}$.



2 $f : D \to \mathbb{R}$, $D \subset \mathbb{R}^n$

Theorem: If all partial derivs $\frac{\partial f}{\partial x_j}$ exist and are continuous near $\vec{a} \in D$, then f is differentiable at $\vec{a}$ (so our differential approx's on Friday were close to Δf).

proof: we need to recall mean value thm (MVT) from Calc,
for $g : [a,b] \to \mathbb{R}$ differentiable
$\exists\, c \in [a,b]$ s.t. $g'(c) = \frac{g(b) - g(a)}{b-a}$
(or $g(b) - g(a) = (b-a)g'(c)$)

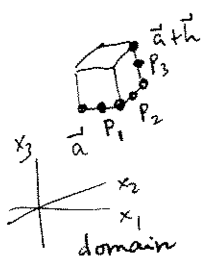


Now, (for $n=3$, e.g.)

$$f(\vec{a}+\vec{h}) - f(\vec{a}) = f\begin{bmatrix} a_1+h_1 \\ a_2+h_2 \\ a_3+h_3 \end{bmatrix} - f\begin{bmatrix} a_1+h_1 \\ a_2+h_2 \\ a_3 \end{bmatrix} \overset{\text{MVT}}{=} \frac{\partial f}{\partial x_3}(P_3)\, h_3 \qquad P_3 = \begin{bmatrix} a_1+h_1 \\ a_2+h_2 \\ c_3 \end{bmatrix}.$$

$$+ f\begin{bmatrix} a_1+h_1 \\ a_2+h_2 \\ a_3 \end{bmatrix} - f\begin{bmatrix} a_1+h_1 \\ a_2 \\ a_3 \end{bmatrix} = \frac{\partial f}{\partial x_2}(P_2)\, h_2 \qquad P_2 = \begin{bmatrix} a_1+h_1 \\ c_2 \\ a_3 \end{bmatrix}$$

$$+ f\begin{bmatrix} a_1+h_1 \\ a_2 \\ a_3 \end{bmatrix} - f\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \frac{\partial f}{\partial x_1}(P_1)\, h_1 \qquad P_1 = \begin{bmatrix} c_1 \\ a_2 \\ a_3 \end{bmatrix}$$



Write $[m(h)] = [\frac{\partial f}{\partial x_1}(P_1),\ \frac{\partial f}{\partial x_2}(P_2),\ \frac{\partial f}{\partial x_3}(P_3)]$ (for "matrix depending on h")

$[C] = [f'(a)] = [\frac{\partial f}{\partial x_1}(\vec{a}),\ \frac{\partial f}{\partial x_2}(\vec{a}),\ \frac{\partial f}{\partial x_3}(\vec{a})]$

Then

$$f(\vec{a}+\vec{h}) - f(\vec{a}) = [m(h)]\left[\vec{h}\right]$$

$$= [C]\left[\vec{h}\right] + \underbrace{[m(h)-c]\left[\vec{h}\right]}_{\text{"error"}},$$

$(= |m(h)-c| \cos\theta)$

$\frac{|\text{error}|}{|h|} \le |m(h)-c| \to 0$ as $h \to 0$ since partials are continuous near $\vec{a}$. ∎

③ $\vec{F}: D \to \mathbb{R}^m$, $D \subset \mathbb{R}^n$

$\vec{F}(\vec{x}) = \begin{bmatrix} f_1(\vec{x}) \\ f_2(\vec{x}) \\ \vdots \\ f_m(\vec{x}) \end{bmatrix}$. If all partials of all $f_i(\vec{x})$ exist and are continuous near $\vec{x} = \vec{a}$ then $\vec{F}$ is differentiable at $\vec{a}$. $[\vec{F}'(a)]_{m \times n}$!

proof

$$\begin{bmatrix} f_1(\vec{a}+\vec{h}) \\ f_2(\vec{a}+\vec{h}) \\ \vdots \\ f_m(\vec{a}+\vec{h}) \end{bmatrix} - \begin{bmatrix} f_1(\vec{a}) \\ f_2(\vec{a}) \\ \vdots \\ f_m(\vec{a}) \end{bmatrix} = \begin{bmatrix} [f_1'(a)] \\ [f_2'(a)] \\ \vdots \\ [f_m'(a)] \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{bmatrix} + \begin{bmatrix} \text{error}_1 \\ \text{error}_2 \\ \vdots \\ \text{error}_m \end{bmatrix}$$

by applying ② to each component fcn f_i. ∎

Return to case ②.

In this case we call the derivative matrix $[\frac{\partial f}{\partial x_1}(\vec{a}), \ldots \frac{\partial f}{\partial x_n}(\vec{a})]$ the gradient of f (at $\vec{a}$), and write $\nabla f(\vec{a})$.

It is interesting to plot the gradient in the domain of f, as a vector field.

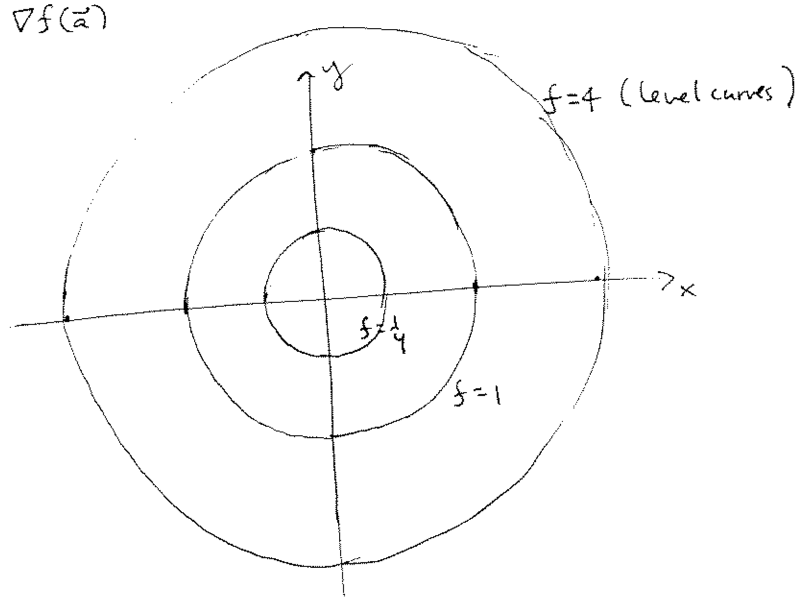
→ at each $\vec{a}$, plot a vector displacement $\nabla f(\vec{a})$

example: $f(x,y) = x^2+y^2$

$\nabla f(x,y) = \langle 2x, 2y \rangle$. ← may rescale these.

Do you notice anything?

(x,y)	$\nabla f(x,y)$
$(0,0)$	
$(1,0)$	
$(0,1)$	
(x,y)	

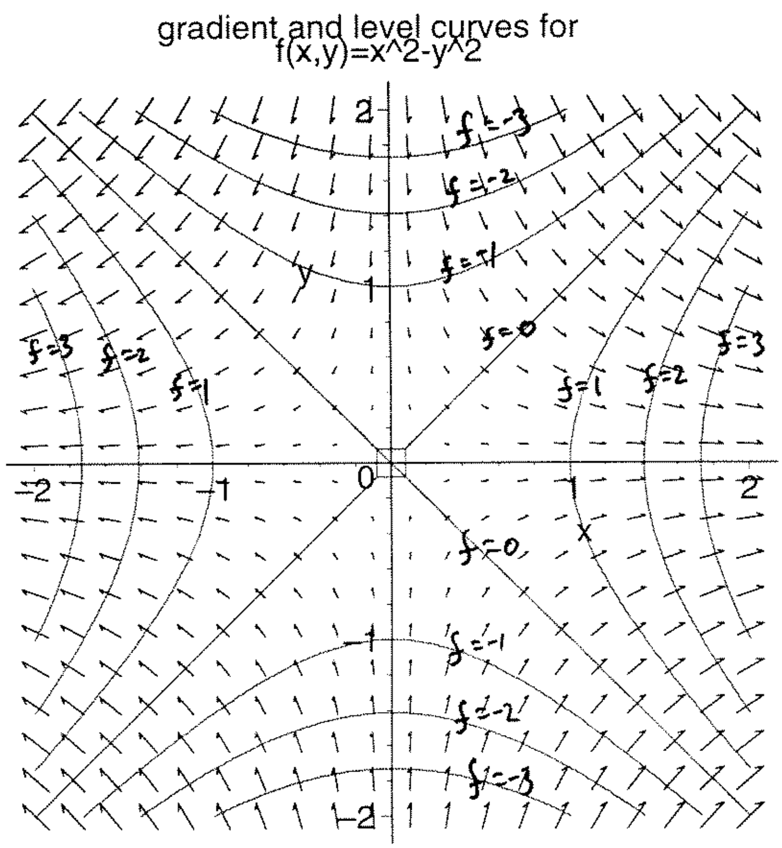


Geometric meaning of the gradient
Math 2210-1
Monday October 18, 2005

```
> with(plots):
Warning, the name changecoords has been redefined
> f:=(x,y)->x^2-y^2;
```

$$f := (x, y) \rightarrow x^2 - y^2$$

```
> plot0:=implicitplot(f(x,y)=0,x=-2..2,y=-2..2):
  plot1:=implicitplot(f(x,y)=1,x=-2..2,y=-2..2):
  plot2:=implicitplot(f(x,y)=-1,x=-2..2,y=-2..2):
  plot3:=implicitplot(f(x,y)=2,x=-2..2,y=-2..2):
  plot4:=implicitplot(f(x,y)=-2,x=-2..2,y=-2..2):
  plot5:=implicitplot(f(x,y)=3,x=-2..2,y=-2..2):
  plot6:=implicitplot(f(x,y)=-3,x=-2..2,y=-2..2):
  plot7:=gradplot(f(x,y),x=-2..2,y=-2..2):
  display({plot0,plot1,plot2,plot3,plot4,plot5,
  plot6,plot7},title=`gradient and level curves for
  f(x,y)=x^2-y^2`);
```



Do you notice anything?

$f : D \to \mathbb{R}^m$, $D \subset \mathbb{R}^n$

$\vec{u}$ a unit vector

$$\boxed{D_{\vec{u}} f(\vec{x}) := \lim_{h \to 0} \frac{f(\vec{x}+h\vec{u}) - f(\vec{x})}{h}}$$

(directional derivative of f in direction of $\vec{u}$
→ generalizes partial derivs, for which you only move in the coord directions).

If f is differentiable then

$$\frac{f(\vec{x}+h\vec{u}) - f(\vec{x})}{h} = \frac{\left[\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots \frac{\partial f}{\partial x_n}\right] \begin{bmatrix} hu_1 \\ hu_2 \\ \vdots \\ hu_n \end{bmatrix} + \text{error}}{h}$$

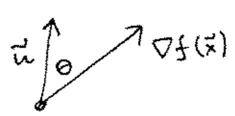
$$= \nabla f(\vec{x}) \cdot \vec{u} + \frac{\text{error}}{h}$$

take $\lim_{h \to 0}$:

since $|h\vec{u}| = |h|$, as $h \to 0$ $\frac{\text{error}}{h} \to 0$

so
$$\boxed{\begin{aligned} D_{\vec{u}} f(\vec{x}) &= \nabla f(\vec{x}) \cdot \vec{u} \\ &= |\nabla f(\vec{x})| \cdot 1 \cdot \cos\theta \end{aligned}}$$

Does this explain observations on page 3,4?



<u>example</u>: for $f(x,y) = x^2 - y^2$
in which direction at $P = (1,1)$
is f increasing most rapidly?
How fast is f increasing in that direction?