Diana Davis Math 34 February 7, 2018 Class #8 ①
§2.4

Mathematician spotlight: Ryan Hynd, University of Pennsylvania
- differential equations applied to inelastic collisions
- colloquium speaker here at Swarthmore yesterday.

Last time: A function is differentiable at a given point if its graph has a well-defined tangent plane there: no "sharp point" or "crease".

This time: Generalize, organize & compute the notion of a tangent plane, or best linear approximation, & differentiability, for $f: \mathbb{R}^m \to \mathbb{R}^n$ using "Jacobian matrix".

Recall: Our equation for the tangent plane (best linear approximation) to $z = f(x,y)$ at $(a,b,f(a,b))$ is $z = L(x,y) = f(a,b) + f_x(a,b)(x-a) + f_y(a,b)(y-b)$.

Changing notation, if $\vec{a} = (a,b)$ and $\vec{x} = (x,y)$: $= f(\vec{a}) + f_x(\vec{a})(x-a) + f_y(\vec{a})(y-b)$.

we can express this as a dot product: $= f(\vec{a}) + \begin{bmatrix} f_x(\vec{a}) \\ f_y(\vec{a}) \end{bmatrix} \bullet \begin{bmatrix} x-a \\ y-b \end{bmatrix}$

$$\boxed{\text{now define the "gradient of } f \text{ at } (a,b)\text{": } \nabla f = \begin{bmatrix} f_x \\ f_y \end{bmatrix} = \begin{bmatrix} f_x(a,b) \\ f_y(a,b) \end{bmatrix} = \begin{bmatrix} f_x(\vec{a}) \\ f_y(\vec{b}) \end{bmatrix}}$$

$$z = L(\vec{x}) = \underline{\underline{f(\vec{a}) + \nabla f(\vec{a}) \bullet (\vec{x} - \vec{a})}}.$$

Note: this is similar to point-slope form for the tangent line to $y = f(x)$ at $x=a$: $y = f(a) + f'(a)(x-a)$.

Also, we can now express the definition of differentiability more precisely:

$f(\vec{x})$ is differentiable at $\vec{a}$ if $\displaystyle\lim_{\vec{x} \to \vec{a}} \frac{f(\vec{x}) - \left[f(\vec{a}) + \nabla f(\vec{a}) \bullet (\vec{x} - \vec{a})\right]}{\|\vec{x} - \vec{a}\|} = 0.$
$(f: \mathbb{R}^n \to \mathbb{R})$

We can generalize this notion of "best linear approximation" to functions $f: \mathbb{R}^m \to \mathbb{R}^n$:

Now $f(\vec{x}) = f(x_1, \dots, x_m) = \begin{bmatrix} f_1(\vec{x}) \\ f_2(\vec{x}) \\ \vdots \\ f_n(\vec{x}) \end{bmatrix} = \begin{bmatrix} f_1(x_1, \dots, x_m) \\ f_2(x_1, \dots, x_m) \\ \vdots \\ f_n(x_1, \dots, x_m) \end{bmatrix}$. Instead of the column vector gradient ∇f, we now have the $n \times m$ Jacobian matrix Df:

$$Df = \begin{bmatrix} \partial f_1/\partial x_1 & \partial f_1/\partial x_2 & \cdots & \partial f_1/\partial x_m \\ \partial f_2/\partial x_1 & \partial f_2/\partial x_2 & \cdots & \partial f_2/\partial x_m \\ \vdots & \vdots & & \vdots \\ \partial f_n/\partial x_1 & \partial f_n/\partial x_2 & \cdots & \partial f_n/\partial x_m \end{bmatrix}.$$

For $f: \mathbb{R}^2 \to \mathbb{R}$, this reduces to the gradient: $Df = \left[\partial f_1/\partial x_1,\ \partial f_1/\partial x_2\right] = \left[f_x, f_y\right] = \nabla f$.

Now the best linear approximation of f at $\vec{a}$ is given by

$$L(\vec{x}) = f(\vec{a}) + Df(\vec{a})(\vec{x} - \vec{a}) = \underbrace{\begin{bmatrix} f_1(\vec{a}) \\ f_2(\vec{a}) \\ \vdots \\ f_n(\vec{a}) \end{bmatrix}}_{\text{numbers}} + \begin{bmatrix} Df(\vec{a}) \\ (n \times m) \end{bmatrix} \begin{bmatrix} \vec{x} - \vec{a} \\ (m \times 1) \end{bmatrix} = \overbrace{\begin{bmatrix} L_1(\vec{x}) \\ L_2(\vec{x}) \\ \vdots \\ L_n(\vec{x}) \end{bmatrix}}^{\text{all linear terms}},$$

a linear approximation in each coordinate.

Definition. The Jacobian matrix of a function $f:\mathbb{R}^m \to \mathbb{R}^n$ is the $n \times m$ matrix Df of partial derivatives of f. If $f:\mathbb{R}^m \to \mathbb{R}$, the Jacobian matrix is a row vector, and is called the gradient of f, denoted by ∇f. Given a point $\vec{a}$, the function $L(\vec{x}) = f(\vec{a}) + Df(\vec{a})(\vec{x}-\vec{a})$ provides the best linear approximation of f in the region near the point $\vec{a}$.

Example. Let $\underbrace{f:\mathbb{R}^3 \to \mathbb{R}^2}$ be the function $f(x,y,z) = \begin{bmatrix} xy^2z + z\cdot e^y \\ z + \sin(xyz) \end{bmatrix} = \begin{bmatrix} f_1(x,y,z) \\ f_2(x,y,z) \end{bmatrix}$

think of it as the wind direction vector for any point in 3-space.

Let's find the best linear approximation of f at the point $\vec{a} = (1,1,1)$:

(1) Find the Jacobian matrix of f:

$$Df = \begin{bmatrix} \partial f_1/\partial x & \partial f_1/\partial y & \partial f_1/\partial z \\ \partial f_2/\partial x & \partial f_2/\partial y & \partial f_2/\partial z \end{bmatrix} = \begin{bmatrix} y^2z \underline{\qquad} & \underline{\qquad} & \underline{\qquad} \\ \underline{\qquad} & \underline{\qquad} & \underline{\qquad} \end{bmatrix} \Rightarrow Df(1,1,1) = \begin{bmatrix} 1 & 2+e & 1+e \\ \cos 1 & \cos 1 & 1+\cos 1 \end{bmatrix}$$

(2) Find the linear approximation of f at $(1,1,1)$:

all linear terms! No polynomials, exponents, trig... just numbers & x, y, z.

$$L(\vec{x}) = f(\vec{a}) + Df(\vec{a})(\vec{x}-\vec{a}) = f(1,1,1) + Df(1,1,1)\begin{bmatrix} x-1 \\ y-1 \\ z-1 \end{bmatrix}$$

$$= \begin{bmatrix} 1+e \\ 1+\sin 1 \end{bmatrix} + \begin{bmatrix} 1 & 2+e & 1+e \\ \cos 1 & \cos 1 & 1+\cos 1 \end{bmatrix}\begin{bmatrix} x-1 \\ y-1 \\ z-1 \end{bmatrix} = \begin{bmatrix} -3-e+x+(2+e)y+(1+e)z \\ \sin 1 - 3\cos 1 + \cos 1(x+y) + (1+\cos 1)z \end{bmatrix}.$$

We generalized derivatives to higher dimensions. Now we'll do higher-order derivatives:

Example. Let $f(x,y) = xy^2 + e^{xy}$.

Then $f_x(x,y) = \dfrac{\partial f}{\partial x} = \underline{\qquad\qquad\qquad}$

$f_y(x,y) = \dfrac{\partial f}{\partial y} = \underline{\qquad\qquad\qquad}$

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = f_{xx}(x,y) = y^2 e^{xy}.$$

$$\frac{\partial^2 f}{\partial y\,\partial x} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = f_{xy}(x,y) = 2y + e^{xy} + xy\,e^{xy}.$$

$$\frac{\partial^2 f}{\partial x\,\partial y} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = f_{yx}(x,y) = 2y + e^{xy} + xy\,e^{xy}.$$

$$\frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) = f_{yy}(x,y) = 2x + x^2 e^{xy}.$$

Notice: $f_{xy} = f_{yx}$. This is always true:

Clairaut's Theorem: If $f(x_1,\ldots,x_m)$ has continuous 1^{st} and 2^{nd} partial derivatives, then the order of differentiation does not matter: $f_{x_i x_j} = f_{x_j x_i}$ for all i,j.

In fact, if f has continuous $1^{st}, 2^{nd}, \ldots, k^{th}$ partial derivatives, you can take them in any order: $f_{xyzxyz} = f_{xxyyzz} = f_{zxyxyz} = \ldots$ (any reordering of 2 x's, 2 y's, 2 z's).

So, use the order that is easiest and makes things (terms) disappear early!