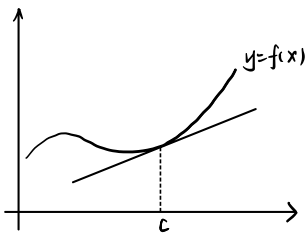
Lecture 16

§1 Differentiation

Recall: 在Calculus中, $f'(c) = \lim_{x\to c} \frac{f(x)-f(c)}{x-c}$ 表示

- slope of tangent line
- rate of change
- velocity



Q: What if f is vector-valued ($f\in\mathbb{R}^n$) and $x\in\mathbb{R}^n$?

Recall: 在Calculus中, 有另一种表示:

$$\frac{f(x)-f(c)}{x-c} = f'(c) + R(x), \quad R(x)\to 0 \text{ as } x\to c$$

Notation:

$o(1)$: any function which converges to 0 as $x\to c$

$O(1)$: any function which is bdd as $x\to c$

运算:

- $o(g(x)) = g(x)\, o(1)$
- $O(g(x)) = g(x)\, O(1)$
- $o(1)+o(1) = o(1)$
- $o(1)\cdot O(1) = o(1)$
- $\sin(o(1)) = o(1)$

将 $R(x)$ 替换为 $o(1)$, 有:

$$f(x)-f(c) = f'(c)(x-c) + \underbrace{(x-c)\cdot o(1)}_{} = o(x-c), \quad \text{as } x\to c$$

$$\underbrace{f(x) = f(c) + f'(c)(x-c)}_{L(x):\ \text{linearization of } f \text{ at } c} + \underbrace{o(x-c)}_{\text{error}}$$

$y = L(x)$ graph is tangent line

$f(x)\approx L(x)$, $x\approx c$

1. **Definition:** differentiability (可微)

Let $f: D\,(\text{open in } \mathbb{R}^n) \to \mathbb{R}^m$, $c\in D$, 则称 f is differentiable at c, 若

$\exists A_{m\times n}$, s.t. $\underbrace{f(\vec{x})}_{\in\mathbb{R}^m} = \underbrace{f(\vec{c})}_{\in\mathbb{R}^m} + A_{m\times n}\underbrace{(\vec{x}-\vec{c})}_{\in\mathbb{R}^n} + \underbrace{\vec{o}(\|\vec{x}-\vec{c}\|)}_{\in\mathbb{R}^m}$ as $\vec{x}\to\vec{c}$ or $\vec{x}\approx\vec{c}$

若上式成立, 则 A 被称为 total derivative (全导数) of f at c

Notation: $f'(c) = A_{m\times n} = Df(c)$

注: "h-notation": $x = c+h$

$$f(c+h) = f(c) + Ah + o(\|h\|), \quad \text{as } h\to 0$$

e.g. $f(x) = A_{m\times n}\vec{x}$, $\vec{x}\in\mathbb{R}^n$

Q: $\forall c\in\mathbb{R}^n$, f differentiable at c?

A: Yes, with $f'(c) = A$

$f(x) = Ax$

$f(c) + A(x-c) = Ac + A(x-c) = A(x)$

$\therefore$ f differentiable at c, $f'(c) = A$

§2 关于 Differentiation 的 facts

对于: $\vec{x} \in D \subset R^n$

$$\vec{f}(x) = \begin{bmatrix} f_1(x_1, x_2, \cdots, x_n) \\ f_2(x_1, x_2, \cdots, x_n) \\ \vdots \\ f_m(x_1, x_2, \cdots, x_n) \end{bmatrix} \in R^m$$

1、 Fact 1: 可微 $\Rightarrow$ 偏导存在，且全导数取决于偏导

differentiability $\Rightarrow$ partial differentiability

证明:

$\because$ f differentiable at c

$\therefore$ $f(x) = f(c) + A_{m\times n}(x-c) + o(|x-c|)$ as $x \to c$ / $x \approx c$ (*)

Take $x = \begin{bmatrix} x_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} \Rightarrow x - c = \begin{bmatrix} x_1 - c_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$

Then by (*):

$$\begin{bmatrix} f_1(x_1, c_2, \cdots, c_n) \\ f_2(x_1, c_2, \cdots, c_n) \\ \vdots \\ f_m(x_1, c_2, \cdots, c_n) \end{bmatrix} = \begin{bmatrix} f_1(c_1, c_2, \cdots, c_n) \\ f_2(c_1, c_2, \cdots, c_n) \\ \vdots \\ f_m(c_1, c_2, \cdots, c_n) \end{bmatrix} + \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 - c_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + o(|x_1 - c_1|)$$

$$= \begin{bmatrix} f_1(c_1, c_2, \cdots, c_n) \\ f_2(c_1, c_2, \cdots, c_n) \\ \vdots \\ f_m(c_1, c_2, \cdots, c_n) \end{bmatrix} + \begin{bmatrix} a_{11}(x_1 - c_1) \\ a_{21}(x_1 - c_1) \\ \vdots \\ a_{m1}(x_1 - c_1) \end{bmatrix} + o(|x_1 - c_1|)$$

$$\Rightarrow \begin{cases} f_1(x_1, c_2, \cdots, c_n) = f_1(c_1, c_2, \cdots, c_n) + a_{11}(x_1 - c_1) + o(|x_1 - c_1|) & \frac{d f_1(x_1, c_2, \cdots, c_n)}{d x_1}\Big|_{x_1 = c_1} = \frac{\partial f_1}{\partial x_1}(c) \\ \vdots \\ f_m(x_1, c_2, \cdots, c_n) = f_m(c_1, c_2, \cdots, c_n) + a_{m1}(x_1 - c_1) + o(|x_1 - c_1|) & \frac{d f_m(x_1, c_2, \cdots, c_n)}{d x_1}\Big|_{x_1 = c_1} = \frac{\partial f_m}{\partial x_1}(c) \end{cases}$$

Then $a_{n1} = \frac{\partial f_n}{\partial x_1}(c)$

Moral of the story:

If f differentiable at c, then all $f_i(x_1, \cdots, x_n)$ have partial derivatives at c

Moreover, $a_{ij} = \frac{\partial f_i}{\partial x_j}(c)$, $i = 1, \cdots, m$, $j = 1, \cdots, n$

$$A = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(c) & \frac{\partial f_1}{\partial x_2}(c) & \cdots & \frac{\partial f_1}{\partial x_n}(c) \\ \frac{\partial f_2}{\partial x_1}(c) & \frac{\partial f_2}{\partial x_2}(c) & \cdots & \frac{\partial f_2}{\partial x_n}(c) \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(c) & \frac{\partial f_m}{\partial x_2}(c) & \cdots & \frac{\partial f_m}{\partial x_n}(c) \end{bmatrix}_{m\times n}$$
$$\begin{array}{cccc} \downarrow & \downarrow & & \downarrow \\ \frac{\partial \vec{f}}{\partial x_1} & \frac{\partial \vec{f}}{\partial x_2} & \cdots & \frac{\partial \vec{f}}{\partial x_n} \end{array}$$

is called Jacobian matrix of $\vec{f}$ at c

$$A = \begin{array}{c} \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(c) & \frac{\partial f_1}{\partial x_2}(c) & \cdots & \frac{\partial f_1}{\partial x_n}(c) \\ \frac{\partial f_2}{\partial x_1}(c) & \frac{\partial f_2}{\partial x_2}(c) & \cdots & \frac{\partial f_2}{\partial x_n}(c) \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(c) & \frac{\partial f_m}{\partial x_2}(c) & \cdots & \frac{\partial f_m}{\partial x_n}(c) \end{bmatrix}_{m\times n} \end{array} \begin{array}{l} \to \nabla f_1 \\ \to \nabla f_2 \\ \\ \to \nabla f_m \end{array}$$

注：可微 $\Rightarrow$ 偏导存在．

但通常情况下，偏导存在 $\not\Rightarrow$ 可微．除非偏导连续

2. **Fact 2:** 可微 $\Rightarrow$ 连续

(total) differentiability $\Rightarrow$ continuity

证明：

$\because$ f is differentiable at c

$\therefore$ $f(x) = f(c) + A_{m\times n}(x-c) + o(\|x-c\|)$ as $x \to c$

$\therefore$ $\lim\limits_{x\to c} f(x) = f(c) + 0 + 0 = f(c)$

$\therefore$ f is continuous at c

e.g. $f(x,y) = \begin{cases} \frac{xy}{x^2+y^2}, & (x,y)\neq(0,0) \\ 0, & (x,y)=(0,0) \end{cases}$

$f_x(0,0) = \frac{d f(x,0)}{dx}\Big|_{x=0} = 0$

$f_y(0,0) = \frac{d f(0,y)}{dy}\Big|_{y=0} = 0$

Claim: $\lim\limits_{(x,y)\to(0,0)} f(x,y)$ D.N.E. $\Rightarrow$ f not continuous at $(0,0)$ $\Rightarrow$ f not differentiable at $(0,0)$

$\lim\limits_{(x,y)\to(0,0)} f(x,y)\big|_{y=kx} = \lim\limits_{x\to 0} \frac{kx^2}{x^2+k^2x^2} = \frac{k}{1+k^2}$

$\therefore$ f not differentiable at $(0,0)$

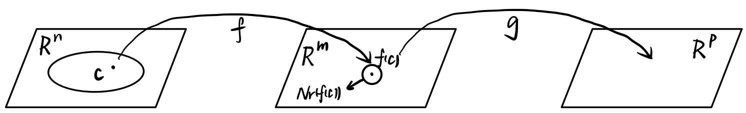
3. **Fact 3:** Chain rule (复合函数的可微性)

Let $f: D(\text{open in } R^n) \to R^m$ be differentiable at $c \in D$

Let $g: N_r(f(c)) \subset R^m \to R^p$ be differentiable at $f(c)$

Then $g\circ f$ is differentiable at c

Moreover, $\underbrace{D(g\circ f)(c)}_{p\times n} = \underbrace{Dg(f(c))}_{p\times m} \cdot \underbrace{Df(c)}_{m\times n}$



证明：

$\because$ f differentiable at c

$\therefore$ $f(c+h) = f(c) + Df(c)h + o(\|h\|)$, as $h \in R^n \to 0$ (#)

$\because$ g differentiable at $f(c)$

$\therefore$ $g(f(c)+l) = g(f(c)) + Dg(f(c))\cdot l + o(\|l\|)$, as $l \in R^m \to 0$ (*)

Take $l = f(c+h) - f(c)$

$\because$ f continuous at c

$\therefore$ $l \to 0$ as $h \to 0$

$\therefore$ $g(f(c+h)) = g(f(c)+l)$

$= g(f(c)) + Dg(f(c)) \cdot l + o(|l|)$ by $(*)$

$= g(f(c)) + Dg(f(c)) \cdot [f(c+h) - f(c)] + o([f(c+h) - f(c)])$

$= g(f(c)) + Dg(f(c)) \cdot [Df(c) \cdot h + o(|h|)] + o([Df(c) \cdot h + o(|h|)])$ as $h \to 0$ by $(\#)$

$= g(f(c)) + Dg(f(c)) \cdot Df(c) \cdot h + o(|h|) + o(1) \cdot [Df(c) \cdot h + o(|h|)]$

$= g(f(c)) + Dg(f(c))\, Df(c) \cdot h + o(|h|) + o(1) \cdot O(|h|)$

$= g(f(c)) + Dg(f(c))\, Df(c) \cdot h + o(|h|)$ as $h \to 0$

Thus $g(f(x))$ is total differential at c & $D(g \circ f)(c) = Dg(f(c)) \cdot Df(c)$

e.g. $z = g(x, y)$, $(x, y) \in \mathbb{R}^2$

$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} f_1(u,v) \\ f_2(u,v) \end{pmatrix} = f(u,v)$, $(u,v) \in \mathbb{R}^2$

Then $D(g \circ f)(u,v) \overset{\text{Fact}}{\underset{3}{=}} Dg(f(u,v)) \cdot Df(u,v)$

$$= \left[\frac{\partial g}{\partial x}(f(u,v)),\ \frac{\partial g}{\partial y}(f(u,v)) \right]_{1\times 2} \begin{bmatrix} \frac{\partial f_1}{\partial u}(u,v) & \frac{\partial f_1}{\partial v}(u,v) \\ \frac{\partial f_2}{\partial u}(u,v) & \frac{\partial f_2}{\partial v}(u,v) \end{bmatrix}_{2\times 2}$$

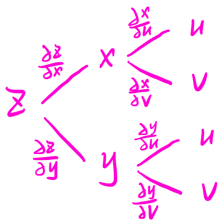
$$= \left[\frac{\partial g}{\partial x} \cdot \frac{\partial f_1}{\partial u} + \frac{\partial g}{\partial y} \cdot \frac{\partial f_2}{\partial u},\ \frac{\partial g}{\partial x} \frac{\partial f_1}{\partial v} + \frac{\partial g}{\partial y} \cdot \frac{\partial f_2}{\partial v} \right]$$

$$= \left[\frac{\partial (g \circ f)}{\partial u},\ \frac{\partial (g \circ f)}{\partial v} \right]$$

注: $\frac{\partial (g \circ f)}{\partial u} = \frac{\partial g}{\partial x} \cdot \frac{\partial f_1}{\partial u} + \frac{\partial g}{\partial y} \cdot \frac{\partial f_2}{\partial u}$

$\frac{\partial (g \circ f)}{\partial v} = \frac{\partial g}{\partial x} \frac{\partial f_1}{\partial v} + \frac{\partial g}{\partial y} \cdot \frac{\partial f_2}{\partial v}$

与 Calculus II 的联系:



$\frac{\partial z}{\partial u} = \frac{\partial z}{\partial x} \cdot \frac{\partial x}{\partial u} + \frac{\partial z}{\partial y} \cdot \frac{\partial y}{\partial v}$

$\frac{\partial z}{\partial v} = \frac{\partial z}{\partial x} \cdot \frac{\partial x}{\partial v} + \frac{\partial z}{\partial y} \cdot \frac{\partial y}{\partial v}$