多元函数微分法

January 19, 2018

1 基本概念

二元有序实数组 (x, y) 的全体，即 $\boldsymbol{R}^2 = \boldsymbol{R} \times \boldsymbol{R} = \{(x, y)|x, y \in \boldsymbol{R}\}$ 就表示坐标平面。坐标平面上具有某种性质 P 的点的集合，称为平面点集，记作

$$E = \{(x, y)|(x, y)\text{具有性质}P\} \tag{1}$$

设 $P_0(x_0, y_0)$ 是 xOy 平面上的一个点，δ 是某一正数。与点 $P_0(x_0, y_0)$ 距离小于 δ 的点 $P(x, y)$ 的全体，称为点 P_0 的δ 邻域，记作 $U(P_0, \delta)$，即

$$U(P_0, \delta) = \{P|\ |PP_0| < \delta\} \ ,$$

即

$$U(P_0, \delta) = \{(x, y)|\sqrt{(x - x_0)^2 + (y - y_0)^2} < \delta\}$$

点 P_0 的去心 δ 邻域，记作 $\mathring{U}(P_0, \delta)$，即

$$\mathring{U}(P_0, \delta) = \{P|\ 0 < |PP_0| < \delta\} \ .$$

任意一点 $P \in \boldsymbol{R}^2$ 与任意一个点集 $E \subset \boldsymbol{R}^2$ 之间的有三种关系:

内点：若存在点 P 的某个邻域 $U(P)$，使得 $U(P) \subset E$，则称 P 为 E 的内点;

外点：若存在点 P 的某个邻域 $U(P)$，使得 $U(P) \bigcap E = \emptyset$，则称 P 为 E 的外点;

边界点：若点 P 的任一邻域内既含有属于 E 点，又含有不属于 E 的点，则称 P 为 E 的边界点。E 的边界点的全体，称为 E 的边界，记作 ∂E。

E 的内点必属于 E，E 的外点必不属于 E，E 的边界点可能属于 E，也可能不属于 E。

聚点：若对于任意给定的 $\delta > 0$，点 P 的去心邻域 $\mathring{U}(P_0, \delta)$ 内总有 E 中的点，则称 P 是 E 的聚点。点集 E 的聚点 P 本身，可以属于 E，也可以不属于 E。

开集：若点集 E 的点都是 E 的内点，则称 E 为开集。

闭集：若点集 E 的边界 $\partial E \subset E$，则称 E 为闭集。

连通集：若点集 E 内任何两点，都可用折线联结起来，且该折线上的点都属于 E，则称 E 为连通集。

区域(开区域)：连通的开集称为区域或开区域。

闭区域：开区域连同它的边界一起构成的点集称为闭区域。

有界集：对于平面点集 E，若存在某一正数 r，使得

$$E \subset U(O, r) ,$$

其中 O 是坐标原点，则称 E 为有界集。

无界集：一个集合如果不是有界集，就称这集合为无界集。

设 n 为取定的一个正整数，用 $\boldsymbol{R}^n$ 表示 n 元有序实数组 $(x_1, x_2, \cdots, x_n)$ 的全体所构成的集合，即

$$\boldsymbol{R}^n = \boldsymbol{R} \times \boldsymbol{R} \times \cdots \times \boldsymbol{R} = \{(x_1, x_2, \cdots, x_n) | x_i \in \boldsymbol{R}, i = 1, 2, \cdots, n\} .$$

$\boldsymbol{R}^n$ 中的元素 $(x_1, x_2, \cdots, x_n)$ 也可以用单个字母 $\boldsymbol{x}$ 表示，即 $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)$。当所

有的 $x_i(i=1,2,\cdots,n)$ 都为零时，称这样的元素为 $\boldsymbol{R}^n$ 中的零元，记为 $\mathbf{0}$ 或 O。

设 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n),\boldsymbol{y}=(y_1,y_2,\cdots,y_n)$ 为 $\boldsymbol{R}^n$ 中任意两个元素，$\lambda\in\boldsymbol{R}$，规定

$$\begin{aligned}&\boldsymbol{x}+\boldsymbol{y}=(x_1+y_1,x_2+y_2,\cdots,x_n+y_n)\ ,\\&\lambda\boldsymbol{x}=(\lambda x_1,\lambda x_2,\cdots,\lambda x_n)\ .\end{aligned}$$

定义了线性运算的集合 $\boldsymbol{R}^n$ 称为n 维空间。

定义

设 D 是 $\boldsymbol{R}^2$ 的一个非空子集，称映射 $f:D\to\boldsymbol{R}$ 为定义在 D 上的二元函数，记为

$$z=f(x,y)\ ,(x,y)\in D$$

或

$$z=f(P)\ ,P\in D$$

其中点集 D 称为该函数的定义域，x,y 称为自变量，z 称为因变量。

定义

设二元函数 $f(P)=f(x,y)$ 的定义域为 D，$P_0(x_0,y_0)$ 是 D 的聚点。若存在常数 A，对于任意给定的正数 ε，总存在正数 δ，使得当点 $P(x,y)\in D\bigcap\mathring{U}(P_0,\delta)$ 时，都有

$$|f(P)-A|=|f(x,y)-A|<\varepsilon\ ,$$

成立，那么就称常数 A 为函数 $f(x,y)$ 当 $(x,y)\to(x_0,y_0)$ 时的极限，记作

$$\lim_{(x,y)\to(x_0,y_0)}f(x,y)=A\quad\text{或者}\quad f(x,y)\to A((x,y)\to(x_0,y_0))\ ,$$

也记作

$$\lim_{P\to P_0}f(P)=A\quad\text{或者}\quad f(P)\to A(P\to P_0)\ .$$

二元函数的极限也叫做二重极限。二重极限的存在，是指 $P(x,y)$ 以任何方式趋于 $P_0(x_0,y_0)$ 时，$f(x,y)$ 都无限接近于 A。若 $P(x,y)$ 以某一特殊方式，如沿着一条定直线或定曲线趋于 $P_0(x_0,y_0)$ 时，即使 $f(x,y)$ 无限接近于某一确定值，还是不能由此断定函数极限的存在。但反过来，若当 $P(x,y)$ 以不同方式趋于 $P_0(x_0,y_0)$ 时，$f(x,y)$ 趋于不同的值，可以断定函数的极限不存在。

2 偏导数

定义

设函数 $z=f(x,y)$ 在点 (x_0,y_0) 的某一领域内有定义，当 y 固定在 y_0 而 x 在 x_0 处有增量 δx 时，相应的函数有增量

$$f(x_0+\Delta x,y_0)-f(x_0,y_0)\ ,$$

若

$$\lim_{\Delta x\to 0}\frac{f(x_0+\Delta x,y_0)-f(x_0,y_0)}{\Delta x}\ ,$$

存在，则称此极限为函数 $z=f(x,y)$ 在点 (x_0,y_0) 处对 x 的偏导数，记作

$$\left.\frac{\partial z}{\partial x}\right|_{x=x_0,y=y_0},\left.\frac{\partial f}{\partial x}\right|_{x=x_0,y=y_0},\left.z_x\right|_{x=x_0,y=y_0},f_x(x=x_0,y=y_0)\ .$$

函数 $z=f(x,y)$ 在点 (x_0,y_0) 处对 y 的偏导数定义为

$$\lim_{\Delta y\to 0}\frac{f(x_0,y_0+\Delta y)-f(x_0,y_0)}{\Delta y}\ ,\tag{2}$$

记作

$$\left.\frac{\partial z}{\partial y}\right|_{x=x_0,y=y_0},\left.\frac{\partial f}{\partial y}\right|_{x=x_0,y=y_0},\left.z_y\right|_{x=x_0,y=y_0},f_y(x=x_0,y=y_0)\ .$$

定理

若函数 $z = f(x, y)$ 的两个二阶混合偏导数 $\dfrac{\partial^2 z}{\partial y \partial x}$ 以及 $\dfrac{\partial^2 z}{\partial x \partial y}$ 在区域 D 内连续，那么该区域内这两个二阶混合偏导数必相等。

二阶混合偏导数在连续的条件下与求导的次序无关。

3 全微分

4 多元复合函数的求导

5 隐函数的求导

6 多元函数微分学的几何应用

7 方向导数与梯度

偏导数反映的是函数沿坐标轴方向的变化率。

设 l 是 xOy 平面上以 $P_0(x_0, y_0)$ 为始点的一条射线，$\boldsymbol{e}_l = (\cos\alpha, \cos\beta)$ 是与 l 同方向的单位向量。射线 l 的参数方程为

$$x = x_0 + t\cos\alpha\ ,(t \geqslant 0)$$

$$y = y_0 + t\cos\beta\ ,$$

设函数 $z = f(x, y)$ 在点 $P_0(x_0, y_0)$ 的某个邻域 $U(P_0)$ 内有定义，$P(x_0+t\cos\alpha, y_0+t\cos\beta)$ 为 l 上另一点，且 $P \in U(P_0)$。若函数增量$f(x_0 + t\cos\alpha, y_0 + t\cos\beta) - f(x_0, y_0)$ 与P 到

P_0 的距离$|PP_0|=t$ 的比值

$$\frac{f(x_0+t\cos\alpha, y_0+t\cos\beta)-f(x_0,y_0)}{t}$$

当 P 沿着 l 趋于 P_0(即 $t\to 0^+$) 时的极限存在，称此极限为函数 $f(x,y)$ 在点 P_0 沿方向 l 的方向导数，记作 $\left.\frac{\partial f}{\partial l}\right|_{(x_0,y_0)}$，即

$$\left.\frac{\partial f}{\partial l}\right|_{x_0,y_0}=\lim_{t\to 0^+}\frac{f(x_0+t\cos\alpha, y_0+t\cos\beta)-f(x_0,y_0)}{t} \tag{3}$$

方向导数 $\left.\frac{\partial f}{\partial l}\right|_{(x_0,y_0)}$ 就是函数 $f(x,y)$ 在点 $P_0(x_0,y_0)$ 处沿方向 l 的变化率。若函数 $f(x,y)$ 在点 $P_0(x_0,y_0)$ 的偏导数存在，$\boldsymbol{e}_l=\boldsymbol{i}=(1,0)$，则

$$\left.\frac{\partial f}{\partial l}\right|_{(x_0,y_0)}=\lim_{t\to 0^+}\frac{f(x_0+t,y_0)-f(x_0,y_0)}{t}=f_x(x_0,y_0)\ ,$$

若 $\boldsymbol{e}_l=\boldsymbol{j}=(0,1)$，则

$$\left.\frac{\partial f}{\partial l}\right|_{(x_0,y_0)}=\lim_{t\to 0^+}\frac{f(x_0,y_0+t)-f(x_0,y_0)}{t}=f_y(x_0,y_0)\ ,$$

反之，若 $\boldsymbol{e}_l=\boldsymbol{i}$，$\left.\frac{\partial z}{\partial l}\right|_{(x_0,y_0)}$ 存在，则 $\left.\frac{\partial z}{\partial x}\right|_{(x_0,y_0)}$ 未必存在。

定理

如果函数 $f(x,y)$ 在点 $P_0(x_0,y_0)$ 可微分，那么函数在该点沿任一方向 l 的方向导数存在，且有

$$\left.\frac{\partial f}{\partial l}\right|_{(x_0,y_0)}=f_x(x_0,y_0)\cos\alpha+f_y(x_0,y_0)\cos\beta\ , \tag{4}$$

其中 $\cos\alpha,\cos\beta$ 是方向 l 的方向余弦。

设函数 $f(x,y)$ 在平面区域 D 内具有一阶连续偏导数，则对于每一点 $P_0(x_0,y_0)\in D$，都可定出一个向量

$$f_x(x_0,y_0)\boldsymbol{i}+f_y(x_0,y_0)\boldsymbol{j}\ , \tag{5}$$

这向量称为函数 $f(x,y)$ 在点 $P_0(x_0,y_0)$ 的梯度，记作 $\mathbf{grad}\,f(x_0,y_0)$ 或 $\nabla f(x_0,y_0)$，即

$$\mathbf{grad}\,f(x_0,y_0)=\nabla f(x_0,y_0)=f_x(x_0,y_0)\boldsymbol{i}+f_y(x_0,y_0)\boldsymbol{j}\ . \tag{6}$$

其中 $\nabla=\dfrac{\partial}{\partial x}\boldsymbol{i}+\dfrac{\partial}{\partial y}\boldsymbol{j}$ 称为 (二维的) 向量微分算子或 Nabla 算子，$\nabla f=\dfrac{\partial f}{\partial x}\boldsymbol{i}+\dfrac{\partial f}{\partial y}\boldsymbol{j}$。

若函数 $f(x,y)$ 在点 $P_0(x_0,y_0)$ 可微分，$\boldsymbol{e}_l=(\cos\alpha,\cos\beta)$ 是与方向 l 同向的单位向量，则

$$\begin{aligned}\left.\frac{\partial f}{\partial l}\right|_{(x_0,y_0)}&=f_x(x_0,y_0)\cos\alpha+f_y(x_0,y_0)\cos\beta\\&=\mathbf{grad}\,f(x_0,y_0)\cdot\boldsymbol{e}_l=|\mathbf{grad}\,f(x_0,y_0)|\cos\theta\ ,\end{aligned} \tag{7}$$

其中 $\theta=(\mathbf{grad}\,f(x_0,y_0),\boldsymbol{e}_l)$。

当 $\theta=0$，即方向 $\boldsymbol{e}_l$ 与梯度 $\mathbf{grad}\,f(x_0,y_0)$ 的方向相同时，函数 $f(x,y)$ 增加最快。函数在这个方向的方向导数达到最大值，即为梯度 $\mathbf{grad}\,f(x_0,y_0)$ 的模，即

$$\left.\frac{\partial f}{\partial l}\right|_{(x_0,y_0)}=|\mathbf{grad}\,f(x_0,y_0)|\ . \tag{8}$$

当 $\theta=\pi$，即方向 $\boldsymbol{e}_l$ 与梯度 $\mathbf{grad}\,f(x_0,y_0)$ 的方向相反时，函数 $f(x,y)$ 减少最快。函数在这个方向的方向导数达到最小值，即

$$\left.\frac{\partial f}{\partial l}\right|_{(x_0,y_0)}=-|\mathbf{grad}\,f(x_0,y_0)|\ . \tag{9}$$

当 $\theta=\dfrac{\pi}{2}$，即方向 $\boldsymbol{e}_l$ 与梯度 $\mathbf{grad}\,f(x_0,y_0)$ 的方向正交时，函数的变化率为 0，即

$$\left.\frac{\partial f}{\partial l}\right|_{(x_0,y_0)}=0\ . \tag{10}$$

7.1 Directional Derivatives and the Gradient Vector

[1] If $z = f(x, y)$, then the partial derivatives f_x and f_y are defined as

$$f_x(x_0, y_0) = \lim_{h\to 0} \frac{f(x_0 + h, y_0) - f(x_0, y_0)}{h} \tag{11}$$

$$f_y(x_0, y_0) = \lim_{h\to 0} \frac{f(x_0, y_0 + h) - f(x_0, y_0)}{h} \tag{12}$$

and represent the rates of change of z in the x- and y-directions, that is, in the directions of the unit vectors $\boldsymbol{i}$ and $\boldsymbol{j}$.

Suppose that we now wish to find the rate of change of z at $(x_0, y_0$ in the direction of an arbitrary unit vector $\boldsymbol{u} = \langle a, b\rangle$. Consider the surface S with the equation $z = f(x, y)$ (the graph of f) and we let $z_0 = f(x_0, y_0)$. Then the point $P(x_0, y_0, z_0$ lies on S. The vertical plane that passes through P in the direction of $\boldsymbol{u}$ intersects S in a curve C. The slope of the tangent line T to C at the point P is the rate of change of z in the direction of $\boldsymbol{u}$.

If $Q(x, y, z)$ is another point on C and P', Q' are the projections of P, Q onto the xy-plane, then the vector $\overrightharpoon{P'Q'}$ is parallel to $\boldsymbol{u}$ and so

$$\overrightharpoon{P'Q'} = h\boldsymbol{u} = \langle ha, hb\rangle$$

for some scalar h. Therefore $x - x_0 = ha$, $y - y_0 = hb$, so $x = x_0 + ha$, $y = y_0 + hb$, and

$$\frac{\Delta z}{h} = \frac{z - z_0}{h} = \frac{f(x_0 + ha, y_0 + hb) - f(x_0, y_0)}{h}$$

If we take the limit as $h \to 0$, we obtain the rate of change of z (with respect to distance) in the direction of $\boldsymbol{u}$, which is called the directional derivative of f in the direction of $\boldsymbol{u}$.

Definition

The directional derivative of f at (x_0, y_0) in the direction of a unit vector $\boldsymbol{u} = \langle a, b\rangle$ is

$$D_{\boldsymbol{u}} f(x_0, y_0) = \lim_{h\to 0} \frac{f(x_0 + ha, y_0 + hb) - f(x_0, y_0)}{h}$$

if this limit exists.

Theorem

If f is a differentiable function of x and y, then f has a directional derivative in the direction of any unit vector $\boldsymbol{u} = \langle a, b\rangle$ and

$$D_{\boldsymbol{u}} f(x, y) = f_x(x, y)a + f_y(x, y)b$$

if this limit exists.

If the unit vector $\boldsymbol{u}$ makes an angle θ with the positive x-axis, then we can write $\boldsymbol{u} = \langle \cos\theta, \sin\theta$ and it becomes

$$D_{\boldsymbol{u}} f(x, y) = f_x(x, y)\cos\theta + f_y(x, y)\sin\theta\ . \tag{13}$$

The directional derivative of a differentiable function can be written as the dot product of two vectors:

$$\begin{aligned} D_{\boldsymbol{u}} f(x, y) &= f_x(x, y)a + f_y(x, y)b \\ &= \langle f_x(x, y), f_y(x, y)\rangle \cdot \langle a, b\rangle \\ &= \langle f_x(x, y), f_y(x, y)\rangle \cdot \boldsymbol{u}\ . \end{aligned}$$

Definition

If f is a function of two variables x and y, then the gradient of f is the vector function ∇f defined by

$$\nabla f = \langle f_x(x,y), f_y(x,y)\rangle = \frac{\partial f}{\partial x}\boldsymbol{i} + \frac{\partial f}{\partial y}\boldsymbol{j}\ .$$

$$D_{\boldsymbol{u}} f(x,y) = \nabla f(x,y) \cdot \boldsymbol{u}\ . \tag{14}$$

This expresses the directional derivative in the direction of a unit vector $\boldsymbol{u}$ as the scalar projection of the gradient vector onto $\boldsymbol{u}$.

Definition

The directional derivative of f at (x_0, y_0, z_0) in the direction of a unit vector $\boldsymbol{u} = \langle a, b, c\rangle$ is

$$D_{\boldsymbol{u}} f(x_0, y_0, z_0) = \lim_{h\to 0} \frac{f(x_0 + ha, y_0 + hb, z_0 + hc) - f(x_0, y_0, z_0)}{h}\ .$$

if this limit exists.

$$D_{\boldsymbol{u}} f(\boldsymbol{x}_0) = \lim_{h\to 0} \frac{f(\boldsymbol{x}_0 + h\boldsymbol{u}) - f(\boldsymbol{x}_0)}{h}\ . \tag{15}$$

Theorem

Suppose f is a differentiable function of two or three variables. The maximum value of the directional derivative $D_{\boldsymbol{u}} f(\boldsymbol{x})$ is $|\nabla f(\boldsymbol{x})|$ and it occurs when $\boldsymbol{u}$ has the same direction as the gradient vector $\nabla f(\boldsymbol{x})$.

8　多元函数的极值

8.1　多元函数的极值及最大值、最小值

8.2　条件极值　　拉格朗日乘数法

对于函数的自变量，除了限制在函数的定义域内外，无其他条件，称为无条件极值。

对自变量有附加条件的极值称为条件极值。

9　二元函数的泰勒公式

10　最小二乘法

References

[1] J. Stewart. *Calculus*. Cengage Learning, 2015.