第11.3节

方向导数与梯度（gradient）

Definition

设 $\vec{l}\in\mathbb{R}^n$, 且 $\vec{l}\neq 0$. 过 X_0 引一条射线 L 使得其方向(包括指向)与 $\vec{l}$ 一致. 如果 $\lim_{L\ni X\to X_0}\frac{f(X)-f(X_0)}{|X-X_0|}$ 收敛, 则称其极限值为 $f(X)$ 在 X_0 沿 $\vec{l}$ 的方向导数, 记为 $\frac{\partial f}{\partial \vec{l}}(X_0)$.

不妨设 $\vec{l}=(\cos\alpha_1,\cdots,\cos\alpha_n)$, $\cos^2\alpha_1+\cdots+\cos^2\alpha_n=1$,则

$$
\begin{aligned}
\frac{\partial f}{\partial \vec{l}}(X_0) &= \lim_{t\to 0+}\frac{f(X_0+t\,\vec{l})-f(X_0)}{t}\\
&= \lim_{t\to 0+}\frac{f(x_{01}+t\,\cos\alpha_1,\cdots,x_{0n}+t\cos\alpha_n)-f(x_{01},\cdots,x_{0n})}{t}.
\end{aligned}
$$

$f(X)$在X_0沿x_i轴正向和负向的方向导数分别记为$\dfrac{\partial f}{\partial \vec{x}_i^{+}}(X_0)$和$\dfrac{\partial f}{\partial \vec{x}_i^{-}}(X_0)$.

$\dfrac{\partial f}{\partial x_i}(X_0)$存在 $\Leftrightarrow$ $\dfrac{\partial f}{\partial \vec{x}_i^{+}}(X_0)$和$\dfrac{\partial f}{\partial \vec{x}_i^{-}}(X_0)$都存在, 且

$$\frac{\partial f}{\partial \vec{x}_i^{+}}(X_0) = -\frac{\partial f}{\partial \vec{x}_i^{-}}(X_0).$$

若$\dfrac{\partial f}{\partial x_i}(X_0)$存在, 则

$$\frac{\partial f}{\partial x_i}(X_0) = \frac{\partial f}{\partial \vec{x}_i^{+}}(X_0).$$

Example

$f(X)=|X|=\sqrt{x_1^2+\cdots+x_n^2}$, 对于任一方向$\vec{l}$, $|\vec{l}|=1$, 则

$$\frac{\partial f}{\partial \vec{l}}(O)=\lim_{t\to 0^+}\frac{f(t\vec{l})-f(O)}{t}=1.$$

故$f(X)$在O点沿任意方向的方向导数均为1, 从而$\dfrac{\partial f}{\partial x_i}(O)$不存在.

Example

$$f(x,y)=\begin{cases}\dfrac{|x|^{\frac{1}{2}}|y|^{\frac{1}{2}}}{\sqrt{x^2+y^2}}, & x^2+y^2\neq 0,\\ 0, & x^2+y^2=0.\end{cases}$$

显然$\dfrac{\partial f}{\partial x}(0,0)=\dfrac{\partial f}{\partial y}(0,0)=0$, 但取$\vec{l}=\left(\dfrac{\sqrt{2}}{2},\dfrac{\sqrt{2}}{2}\right)$, 由于

$$\lim_{t\to 0^+}\frac{f\left(\frac{\sqrt{2}}{2}t,\frac{\sqrt{2}}{2}t\right)-f(0,0)}{t}=\lim_{t\to 0^+}\frac{\sqrt{2}}{2t}=+\infty,$$

从而$\dfrac{\partial f}{\partial \vec{l}}(0,0)$不存在.

Theorem

设$f(X)$在$X_0 \in \mathbb{R}^n$点可微, 则对于任意$\vec{l} = (\cos\alpha_1, \cdots, \cos\alpha_n), \cos^2\alpha_1 + \cdots + \cos^2\alpha_n = 1$, $\dfrac{\partial f}{\partial \vec{l}}(X_0)$存在, 且

$$\frac{\partial f}{\partial \vec{l}}(X_0) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(X_0)\ \cos\alpha_i = \langle \nabla f(X_0), \vec{l}\,\rangle.$$

一般地, 对任意非零向量$\vec{l}$, 有

$$\frac{\partial f}{\partial \vec{l}}(X_0) = \langle \nabla f(X_0), \frac{1}{|\vec{l}\,|}\vec{l}\,\rangle = \frac{1}{|\vec{l}\,|}\langle \nabla f(X_0), \vec{l}\,\rangle.$$

证明 由$f(X)$在X_0点可微知

$$f(X_0+t\vec{l})-f(X_0)=t\sum_{i=1}^{n}\frac{\partial f}{\partial x_i}(X_0)\cos\alpha_i+o(|t|).$$

由此可得

$$\begin{aligned}\lim_{t\to0^+}\frac{f(X_0+t\vec{l})-f(X_0)}{t} &= \lim_{t\to0^+}\left[\sum_{i=1}^{n}\frac{\partial f}{\partial x_i}(X_0)\cos\alpha_i+\frac{o(|t|)}{t}\right]\\ &= \sum_{i=1}^{n}\frac{\partial f}{\partial x_i}(X_0)\cos\alpha_i,\end{aligned}$$

即$\dfrac{\partial f}{\partial\vec{l}}(X_0)$存在.

Example

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

可知$f'_x(0,0)=f'_y(0,0)=0$, 沿$\vec{l}=(1,1)$的方向导数为$\dfrac{\partial f}{\partial \vec{l}}(0,0)=\dfrac{1}{2}\neq\langle\nabla f,\vec{l}\rangle=0.$

Example

$$f(x,y)=\begin{cases}\dfrac{x^3y}{x^8+y^2}, & (x,y)\neq(0,0)\\ 0, & (x,y)=(0,0).\end{cases}$$

沿任何方向$\vec{l}$: $\dfrac{\partial f}{\partial \vec{l}}(0,0)=0$, 偏导数$f'_x(0,0)=f'_y(0,0)=0$.

$f(x,y)$在$(0,0)$点不连续, 进而$f(x,y)$在$(0,0)$点不可微. 事实上，沿曲线$\Gamma: y=x^4, x>0$有

$$\lim_{\Gamma\ni(x,y)\to(0,0)} f(x,y)=\lim_{x\to0^+}\frac{1}{2x}=+\infty.$$

设$f(X)$在$X_0 \in \mathbb{R}^n$可微, $|\vec{l}\,| = 1$, 由$\dfrac{\partial f}{\partial \vec{l}}(X_0) = \langle \nabla f(X_0), \vec{l}\,\rangle$
$\leqslant |\nabla f(X_0)|$.

(i) $\nabla f(X_0) = 0$, 则对任意方向$\vec{l}$均有$\frac{\partial f}{\partial \vec{l}}(X_0) = 0$.

(ii) $\nabla f(X_0) \neq 0$, 上式等号成立当且仅当向量$\vec{l}$与向量$\nabla f(X_0)$方向相同, 则存在唯一的向量$\vec{l}_0$(即$\nabla f(X_0)$)使得$\dfrac{\partial f}{\partial \vec{l}_0}(X_0) = |\nabla f(X_0)| = |\vec{l}_0|$, 这时$f(X)$在$X_0$沿该方向的方向导数取值最大, 从而沿这个方向函数增长最快. 因此当梯度非零时, 它是函数增长最快的方向.

设$H:\mathbb{R}^n\to\mathbb{R}$,考虑如下方程

$$\dot{\phi}(t)=-\nabla H(\phi(t)).$$

方程的解$\phi:\mathbb{R}\to\mathbb{R}^n$称为$H$的负梯度流.

函数H沿着负梯度流的值是**单调不增的**.

设$f:\mathbb{R}^n\to\mathbb{R}$, $G:\mathbb{R}^l\to\mathbb{R}^n$, 令$U(X)=f(G(X))$,则

$$G(X)=\begin{pmatrix} y_1(x_1,\cdots,x_l)\\ \vdots \\ y_n(x_1,\cdots,x_l)\end{pmatrix},\quad (x_1,\cdots,x_l)\in\mathbb{R}^l.$$

$$\mathrm{d}U(X_0)=J_f(Y_0)dG(X_0)=J_f(Y_0)J_G(X_0)\Delta X^T.$$

$$\Rightarrow \frac{\partial U}{\partial x_i}(X_0)=\sum_{j=1}^{n}\frac{\partial f}{\partial y_j}(Y_0)\frac{\partial y_j}{\partial x_i}(X_0),\ i=1,\cdots,l,$$

Theorem (复合函数的链式法则：$\mathbb{R}^n \longrightarrow \mathbb{R}^m \longrightarrow \mathbb{R}$)

设$f(u_1,\cdots,u_m)$在$U_0=(u_{01},\cdots,u_{0m})\in\mathbb{R}^m$可微. 每个$u_i$在$X_0=(x_{01},\cdots,x_{0n})\in\mathbb{R}^n$可微, $u_i(X_0)=u_{0i}$, 记$g(X)=f(u_1(X),\cdots,u_m(X))$, 则$g$在$X_0$可微, 且

$$\nabla g(X_0)=\sum_{i=1}^{m}\frac{\partial f}{\partial u_i}(U_0)\nabla u_i(X_0),$$

$$\left(\tfrac{\partial g}{\partial x_1},\cdots,\tfrac{\partial g}{\partial x_n}\right)=\left(\tfrac{\partial f}{\partial u_1},\cdots,\tfrac{\partial f}{\partial u_m}\right)\begin{pmatrix}\dfrac{\partial u_1}{\partial x_1} & \cdots & \dfrac{\partial u_1}{\partial x_n}\\ \vdots & \vdots & \vdots\\ \dfrac{\partial u_m}{\partial x_1} & \cdots & \dfrac{\partial u_m}{\partial x_n}\end{pmatrix}$$

$$\nabla g(X_0)=\nabla f(U_0)J_U(X_0)=\nabla f(U_0)\frac{\partial(u_1,\cdots,u_m)}{\partial(x_1,\cdots,x_n)}(X_0).$$

Example

设函数$u(x,y)$有一阶连续偏导数，令$x=r\cos\theta$，$y=r\sin\theta$，证明

$$\left(\frac{\partial u}{\partial x}\right)^2+\left(\frac{\partial u}{\partial y}\right)^2=\left(\frac{\partial u}{\partial r}\right)^2+\frac{1}{r^2}\left(\frac{\partial u}{\partial \theta}\right)^2.$$

证明 由复合函数的链式法则得到

$$\begin{aligned}\frac{\partial u}{\partial r}&=\frac{\partial u}{\partial x}\frac{\partial x}{\partial r}+\frac{\partial u}{\partial y}\frac{\partial y}{\partial r}=\frac{\partial u}{\partial x}\cos\theta+\frac{\partial u}{\partial y}\sin\theta,\\\frac{\partial u}{\partial \theta}&=\frac{\partial u}{\partial x}\frac{\partial x}{\partial \theta}+\frac{\partial u}{\partial y}\frac{\partial y}{\partial \theta}=-r\frac{\partial u}{\partial x}\sin\theta+r\frac{\partial u}{\partial y}\cos\theta.\end{aligned}$$

Example

设函数$u(x,y)$有一阶连续偏导数，令$x=r\cos\theta$，$y=r\sin\theta$，证明

$$\left(\frac{\partial u}{\partial x}\right)^2+\left(\frac{\partial u}{\partial y}\right)^2=\left(\frac{\partial u}{\partial r}\right)^2+\frac{1}{r^2}\left(\frac{\partial u}{\partial \theta}\right)^2.$$

证明 由链式法则

$$\begin{aligned}\left(\frac{\partial u}{\partial r},\frac{\partial u}{\partial \theta}\right) &= \left(\frac{\partial u}{\partial x},\frac{\partial u}{\partial y}\right)\begin{pmatrix}\frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta}\\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta}\end{pmatrix}\\ &= \left(\frac{\partial u}{\partial x},\frac{\partial u}{\partial y}\right)\begin{pmatrix}\cos\theta & -r\sin\theta\\ \sin\theta & r\cos\theta\end{pmatrix}.\end{aligned}$$

$$\left(\frac{\partial u}{\partial r}, \frac{1}{r}\frac{\partial u}{\partial \theta}\right) = \left(\frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}\right) \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

两边转置并相乘得

$$\begin{aligned} & \left(\frac{\partial u}{\partial r}\right)^2 + \frac{1}{r^2}\left(\frac{\partial u}{\partial \theta}\right)^2 \\ = & \left(\frac{\partial u}{\partial r}, \frac{1}{r}\frac{\partial u}{\partial \theta}\right)\left(\frac{\partial u}{\partial r}, \frac{1}{r}\frac{\partial u}{\partial \theta}\right)^T \\ = & \left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2. \end{aligned}$$

Example

设二阶偏导数连续的函数$z=z(x,y)$满足$z''_{xx}+z''_{yy}=0$.
令$x=r\cos\theta,\ y=r\sin\theta,\ w(r,\theta)=z(r\cos\theta,r\sin\theta)$,
求$w=w(r,\theta)$所满足的方程.

解

$$
\begin{aligned}
\frac{\partial w}{\partial r} &= \frac{\partial z}{\partial x}\cos\theta+\frac{\partial z}{\partial y}\sin\theta,\\
\frac{\partial w}{\partial \theta} &= -\frac{\partial z}{\partial x}r\sin\theta+\frac{\partial z}{\partial y}r\cos\theta.
\end{aligned}
$$

从上式两等式中解得

$$
\begin{aligned}
\frac{\partial z}{\partial x} &= \frac{\partial w}{\partial r}\cos\theta-\frac{\partial w}{\partial \theta}\frac{1}{r}\sin\theta,\\
\frac{\partial z}{\partial y} &= \frac{\partial w}{\partial r}\sin\theta+\frac{\partial w}{\partial \theta}\frac{1}{r}\cos\theta.
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial^2 z}{\partial x^2} &= \frac{\partial}{\partial x}\frac{\partial z}{\partial x} = \left(\cos\theta\frac{\partial}{\partial r} - \frac{1}{r}\sin\theta\frac{\partial}{\partial\theta}\right)\left(\frac{\partial w}{\partial r}\cos\theta - \frac{\partial w}{\partial\theta}\frac{1}{r}\sin\theta\right) \\
&= \frac{\partial^2 w}{\partial r^2}\cos^2\theta - \frac{\partial^2 w}{\partial r\partial\theta}\frac{1}{r}\sin\theta\cos\theta + \frac{\partial w}{\partial\theta}\frac{1}{r^2}\sin\theta\cos\theta \\
&\quad - \frac{\partial^2 w}{\partial\theta\partial r}\frac{1}{r}\sin\theta\cos\theta + \frac{\partial w}{\partial r}\frac{1}{r}\sin^2\theta + \frac{\partial^2 w}{\partial\theta^2}\frac{1}{r^2}\sin^2\theta \\
&\quad + \frac{\partial w}{\partial\theta}\frac{1}{r^2}\sin\theta\cos\theta.
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial^2 z}{\partial y^2} &= \frac{\partial^2 w}{\partial r^2}\sin^2\theta + \frac{\partial^2 w}{\partial r\partial\theta}\frac{1}{r}\sin\theta\cos\theta - \frac{\partial w}{\partial\theta}\frac{1}{r^2}\sin\theta\cos\theta \\
&\quad + \frac{\partial^2 w}{\partial\theta\partial r}\frac{1}{r}\sin\theta\cos\theta + \frac{\partial w}{\partial r}\frac{1}{r}\cos^2\theta + \frac{\partial^2 w}{\partial\theta^2}\frac{1}{r^2}\cos^2\theta \\
&\quad - \frac{\partial w}{\partial\theta}\frac{1}{r^2}\sin\theta\cos\theta.
\end{aligned}
$$

$$
\frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2} = \frac{\partial^2 w}{\partial r^2} + \frac{1}{r}\frac{\partial w}{\partial r} + \frac{1}{r^2}\frac{\partial^2 w}{\partial\theta^2} = 0.
$$

Theorem

设$D \subseteq \mathbb{R}^n$是凸区域, $f: D \to \mathbb{R}$是可微函数, 则对任意$X_1, X_2 \in D$, 在以X_1, X_2为端点的线段上存在一点ξ, 使得

$$f(X_2) - f(X_1) = \langle \nabla f(\xi), X_2 - X_1 \rangle.$$

注: 该结论对向量值函数不成立. 例如$F: [0, 2\pi] \to \mathbb{R}^2$,

$$F(t) = \begin{pmatrix} \sin t \\ \cos t \end{pmatrix}, \ t \in [0, 2\pi].$$

(反证) 若存在$\xi \in [0, 2\pi]$使得

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} = F(2\pi) - F(0) = J_F(\xi)(2\pi - 0) = 2\pi \begin{pmatrix} \cos \xi \\ -\sin \xi \end{pmatrix}, \text{ 矛盾.}$$

以X_1, X_2为端点的线段上的任意一点可表示为

$$X = X_1 + t(X_2 - X_1),\ 0 \leqslant t \leqslant 1.$$

考虑一元函数$g(t) = f(X_1 + t(X_2 - X_1))$,
有$g'(t) = \langle \nabla f(X_1 + t(X_2 - X_1)), X_2 - X_1 \rangle$. 由一元函数的微分中值定理, 存在$t_0 \in (0,1)$使得

$$\begin{aligned} f(X_2) - f(X_1) &= g(1) - g(0) = g'(t_0) \\ &= \langle \nabla f(X_1 + t_0(X_2 - X_1)), X_2 - X_1 \rangle. \end{aligned}$$

因此取$\xi = X_1 + t_0(X_2 - X_1)$即可.

设D是$\mathbb{R}^n$中的凸区域, $F: D \to \mathbb{R}^m$可微, 对于$X_1, X_2 \in D$, 求证: 存在$\theta \in (0,1)$, 使得

$$|F(X_2) - F(X_1)| \leqslant |J_F(X_1 + \theta(X_2 - X_1))||X_2 - X_1|.$$

Proof. 可考虑

$$G(t) = \langle F(X_1 + t(X_2 - X_1)) - F(X_1), F(X_2) - F(X_1) \rangle, \ t \in [0,1].$$

Corollary

若定义在区域$D \subseteq \mathbb{R}^n$内的函数$f(X)$的偏导数恒为零,则f为常值函数.

证明 D是道路连通的,只须证$f(X_1) = f(X_2), \forall\ X_1, X_2 \in D$. 取连续道路$\gamma : [0,1] \to D$满足$\gamma(0) = X_1$, $\gamma(1) = X_2$. $[0,1]$紧$\Rightarrow$ $\gamma([0,1])$紧. 对于任意$X \in \gamma([0,1])$, 存在$r = r(X) > 0$, 使得$B_r(X) \subseteq D$, 则$\gamma([0,1]) \subseteq \{B_r(X) |\ X \in \gamma([0,1])\}$开覆盖, 故存在有限子覆盖:

$$\gamma([0,1]) \subseteq \bigcup_{i=1}^{k} B_{r_i}(X_i) \subseteq D.$$

而每个$B_{r_i}(X_i)$均为凸区域, 故$f(X)$在每个$B_{r_i}(X_i)$均为常数, 特别地, $f(X)$沿$\gamma([0,1])$取常值, 故有$f(X_1) = f(X_2)$.

任取一个点$X_0 \in D$, 令

$$U = \{X \in D |\ f(X) = f(X_0)\}.$$

只须证U是D的非空相对开且相对闭子集, 由D的连通性, 即得$U = D$.

(1) 非空. 由于$X_0 \in U$.

(2) 相对闭. 若$X_m \in U$且$X_m \to X^* \in D$,

则$f(X^*) = \lim\limits_{m\to\infty} f(X_m) = f(X_0)$, 即$X^* \in U$, 从而$U$是$D$相对闭.

(3) 相对开. 若$Y \in U$, 由于D开, 从而存在$\delta > 0$,

使$B_\delta(Y) \subseteq D$. 由于$B_\delta(Y)$凸, 对任意$X \in B_\delta(Y)$, 有

$$f(X) - f(Y) = \langle \nabla f(\xi), X - Y \rangle = 0.$$

故$B_\delta(Y) \subseteq U$,则U是D的相对开集.

注: 由$g(t)$是一元凸函数,则

$$\varphi(t)=\frac{g(t)-g(t_0)}{t-t_0},\ t\neq t_0$$

是增函数, 由单调有界原理, $g'_+(t_0)$与$g'_-(t_0)$都存在, 并且有

$$g'_+(t_0)\geq g'_-(t_0).$$

(1) 设$f(X)$是$\mathbb{R}^n$上的凸函数，任取$X_0 \in \mathbb{R}^n$和单位方向向量$\vec{l}$，则$\dfrac{\partial f}{\partial \vec{l}}(X_0)$存在. 记$\dfrac{\partial f}{\partial \vec{l}^-}(X_0)$表示$f$在$X_0$沿$-\vec{l}$的方向导数, 则

$$\frac{\partial f}{\partial \vec{l}}(X_0) \geqslant -\frac{\partial f}{\partial \vec{l}^-}(X_0).$$

证明

令$g(t) = f(X_0 + t\vec{l}), t \in \mathbb{R}$. 则$g(t)$是$\mathbb{R}$上的凸函数.

且有$\dfrac{\partial f}{\partial \vec{l}}(X_0) = g'_+(0), \dfrac{\partial f}{\partial \vec{l}^-}(X_0) = -g'_-(0).$

(2) 设$f(X)$是$\mathbb{R}^n$上的凸函数, 若在X_0点所有偏导数存在, 则$f(X)$ 在X_0可微.

证明 不妨设$X_0 = O$. 函数$f(tX)$关于t是凸的，从而

$$\varphi(t) = \frac{f(tX) - f(O)}{t}, \ t \neq 0$$

是增函数.

$\forall X=(x_1,\cdots,x_n)\in\mathbb{R}^n$, 取$\mathbb{R}^n$的标准正交基$\{e_i\}$,

$$\begin{aligned}
f(X)-f(O) &= \varphi(1)\geqslant\lim_{t\to0^+}\varphi(t)\geqslant\lim_{t\to0^-}\varphi(t)\\
&= \lim_{t\to0^-}\frac{f(O)-f(tX)}{-t}\\
\left(tX=\frac{\sum_{i=1}^n ntx_ie_i}{n}\text{和}f\text{凸}\right) &\geq \lim_{t\to0^-}\frac{\displaystyle\sum_{i=1}^n(f(O)-f(ntx_ie_i))}{-nt}\\
&= \sum_{i=1}^n x_i\frac{\partial f}{\partial x_i}(O),
\end{aligned}$$

$$f(X)-f(O)\geq\sum_{i=1}^n\frac{\partial f}{\partial x_i}(O)x_i.$$

注: $f(X)-f(O)\geq\langle\nabla f(O),X\rangle$.

另一方面, 由f的凸性可得

$$\begin{aligned} f(X)-f(O) &\leqslant \frac{1}{n}\sum_{i=1}^{n}\left[f(nx_i\vec{e_i})-f(O)\right] \\ &= \sum_{i=1}^{n}\frac{1}{n}\left[f(0,\cdots,nx_i,\cdots,0)-f(0,\cdots,0)\right] \\ &= \sum_{i=1}^{n}\frac{\partial f}{\partial x_i}(O)x_i+\frac{1}{n}\sum_{i=1}^{n}\rho_i|x_i|, \end{aligned}$$

由柯西不等式可得

$$\left|\sum_{i=1}^{n}\rho_i|x_i|\right| \leqslant \sqrt{\rho_1^2+\cdots+\rho_n^2}|X| = o(|X|)$$

综上可知

$$f(X)-f(O)=\sum_{i=1}^{n}\frac{\partial f}{\partial x_i}(O)x_i+o(|X|).$$