第 14 章　多变量函数的微分学

§14.1　方向导数和偏导数

方向和方向余弦　称$\mathbb{R}^n$中的单位向量$\vec{u}$为一个方向；若$\theta_1,\theta_2,\cdots,\theta_n$分别是$\vec{u}$与坐标向量$\vec{e}_1,\vec{e}_2,\cdots,\vec{e}_n$的夹角，则$\vec{u}=(\cos\theta_1,\cos\theta_2,\cdots,\cos\theta_n)$，故方向又称为方向余弦.

定义 14.1　设f是非空开集$D\subset\mathbb{R}^n$上的函数，$a\in D$，$\vec{u}$是$\mathbb{R}^n$中的一个方向. 若$\lim\limits_{t\to 0}\dfrac{f(a+t\vec{u})-f(a)}{t}$存在且有限，则将该极限记成$\dfrac{\partial f}{\partial\vec{u}}(a)$或$\left.\dfrac{\partial f}{\partial\vec{u}}\right|_a$，称为$f$在$a$处沿方向$\vec{u}$的方向导数. 显然，方向导数$\dfrac{\partial f}{\partial\vec{u}}(a)$与方向$\vec{u}$有关. 特别的，$\dfrac{\partial f}{\partial(-\vec{u})}(a)=-\dfrac{\partial f}{\partial\vec{u}}(a)$.

定义 14.1′　通常将$\dfrac{\partial f}{\partial\vec{e}_k}(a)$记成$\dfrac{\partial f}{\partial x_k}(a)$或$\left.\dfrac{\partial f}{\partial x_k}\right|_a$或$D_k f(a)\ \ (k=1,2,\cdots,n)$，称为$f$在$a$处关于$x_k$的（1 阶）偏导数；若非空开集$D\subset\mathbb{R}^n$上的函数$f$在任意$x\in D$处都存在$\dfrac{\partial f}{\partial x_k}(x)$，则$\dfrac{\partial f}{\partial x_k}$也是$D$上的函数，称为$f$关于$x_k$的（1 阶）偏导函数.

命题　$\dfrac{\partial f}{\partial x_k}(a)$恰好是以$x_k$为自变量的函数$g(x_k)=f(a_1,\cdots,a_{k-1},x_k,a_{k+1},\cdots,a_n)$在$a_k$处的导数$g'(a_k)$.

证：$\dfrac{\partial f}{\partial x_k}(a)=\dfrac{\partial f}{\partial\vec{e}_k}(a)=\lim\limits_{t\to 0}\dfrac{f(a+t\vec{e}_k)-f(a)}{t}=\lim\limits_{t\to 0}\dfrac{g(a_k+t)-g(a_k)}{t}=g'(a_k)$. □

方向导数的物理意义　若$t(x,y,z)$是区域$D\subset\mathbb{R}^3$上的温度函数，$a\in D$，$\vec{u}$是$\mathbb{R}^3$中的一个方向，则$\dfrac{\partial t}{\partial\vec{u}}(a)$恰是温度沿方向$\vec{u}$的“变化率”.

偏导数的几何意义　设$\mathbb{R}^3$中的曲面$z=f(x,y)$与平面$y=y_0$相交成曲线γ_1，与平面$x=x_0$相交成曲线γ_2，那么$(1,0,\frac{\partial f}{\partial x}(x_0,y_0))$便是$\gamma_1$在

$(x_0, y_0, f(x_0,y_0))$ 处的一个切向量；$(0,1,\frac{\partial f}{\partial y}(x_0,y_0))$ 便是 γ_2 在 $(x_0, y_0, f(x_0,y_0))$ 处的一个切向量.

练习题 14.1（P_{110}） 2, 3, 4, 5(3), 6(1, 5, 8, 12).

§14.2　多变量函数的微分

定义 14.2(重要的概念)　设 f 是开集 $D\subset\mathbb{R}^n$ 上的函数, $a\in D$. 若存在常向量 $\lambda=(\lambda_1,\lambda_2,\cdots,\lambda_n)\in\mathbb{R}^n$ 使得

$$f(a+h)-f(a)=\sum_{k=1}^{n}\lambda_k h_k+o(\|h\|)\qquad(\|h\|\to 0),$$

则称 f 在 a 处可微分, 并称以 $\mathbb{R}^n$ 为定义域、以 $dx=(dx_1,dx_2,\cdots,dx_n)$ 为自变量的线性函数 $\sum_{k=1}^{n}\lambda_k dx_k$ 为 f 在 a 处的微分, 记作 $df(a)=\sum_{k=1}^{n}\lambda_k dx_k$；若 f 在 D 中的每个点处都可微分, 则称 f 是 D 上的可微函数, 此时 $df(x)=\sum_{k=1}^{n}\lambda_k(x)dx_k$ 是 $D\times\mathbb{R}^n$ 上以 (x,dx) 为自变量的 $2n$ 元函数.

定理 14.1　若多变量函数 f 在 a 处可微, 则它必在 a 处连续.

证: $f(x)-f(a)=\sum_{k=1}^{n}\lambda_k(x_k-a_k)+o(\|x-a\|)\qquad(\|x-a\|\to 0)$. □

命题 1 若 n 元函数 f 在 a 处可微, 则所有的 1 阶偏导数 $\frac{\partial f}{\partial x_k}(a)$ $(1\le k\le n)$ 都存在, 此时 $df(a)=\sum_{k=1}^{n}\frac{\partial f}{\partial x_k}(a)dx_k$； 反之, 结论可能不正确(见例 1).

证:
$$f(a+h)-f(a)=\sum_{k=1}^{n}\lambda_k h_k+o(\|h\|)\qquad(\|h\|\to 0),$$

$$\frac{\partial f}{\partial \vec{e}_k}(a)=\lim_{t\to 0}\frac{f(a+t\vec{e}_k)-f(a)}{t}=\lim_{t\to 0}\frac{\lambda_k t+o(\|t\vec{e}_k\|)}{t}=\lambda_k,$$

故
$$\frac{\partial f}{\partial x_k}(a)=\lambda_k. \ \square$$

例 1　设 $f(x,y)=\begin{cases}\dfrac{xy}{x^2+y^2}, (x,y)\ne(0,0);\\ 0, \qquad (x,y)=(0,0),\end{cases}$ 则 $\frac{\partial f}{\partial x}(0,0)=\frac{\partial f}{\partial y}(0,0)=0$,

显然 f 在 $(0,0)$ 处不可微(因为 f 在 $(0,0)$ 处不连续).

证: (略).

微分的几何意义 设 $z=f(x,y)$ 是 $\mathbb{R}^3$ 中的曲面. 记 $O'=(x_0,y_0,f(x_0,y_0))$, 将直角坐标系 $Oxyz$ 平移得到新坐标系 $O'(dx)(dy)(dz)$. 那么, 在新坐标

系$O'(dx)(dy)(dz)$下, “曲面$z=f(x,y)$”的图像在$O'=(x_0,y_0,f(x_0,y_0))$处的切平面方程恰为$dz=\frac{\partial f}{\partial x}(x_0,y_0)dx+\frac{\partial f}{\partial y}(x_0,y_0)dy$.

证: 切平面的法向量之一是

$$(1,0,\tfrac{\partial f}{\partial x}(x_0,y_0))\times(0,1,\tfrac{\partial f}{\partial y}(x_0,y_0))=\det\begin{pmatrix}\vec{e}_1, & \vec{e}_2, & \vec{e}_3\\ 1, & 0, & \frac{\partial f}{\partial x}(x_0,y_0)\\ 0, & 1, & \frac{\partial f}{\partial y}(x_0,y_0)\end{pmatrix}$$

$$=(-\tfrac{\partial f}{\partial x}(x_0,y_0),-\tfrac{\partial f}{\partial y}(x_0,y_0),1),$$

故切平面的方程为 $-\frac{\partial f}{\partial x}(x_0,y_0)dx-\frac{\partial f}{\partial y}(x_0,y_0)dy+dz=0$. □

多变量函数的Jacobi矩阵和梯度 若n元函数f在a处的1阶偏导数$\frac{\partial f}{\partial x_k}(a)$ $(1\le k\le n)$都存在, 则记$Jf(a)=(\frac{\partial f}{\partial x_1}(a),\frac{\partial f}{\partial x_2}(a),\cdots,\frac{\partial f}{\partial x_n}(a))$, 称为$f$在$a$处的Jacobi矩阵; 记$grad\,f(a)=\nabla f(a)=\frac{\partial f}{\partial x_1}(a)\vec{e}_1+\frac{\partial f}{\partial x_2}(a)\vec{e}_2+\cdots+\frac{\partial f}{\partial x_n}(a)\vec{e}_n$, 称为$f$在$a$处的梯度, 其中$\nabla$被称为Hamilton算子. 于是, 若$f$在$a$处可微, 则

$$df(a)=Jf(a)dx=<grad\,f(a),dx>.$$

命题2 若n元函数f在a处可微, 则f在a处沿$\mathbb{R}^n$中任意方向$\vec{u}$的方向导数都存在, 并且$\frac{\partial f}{\partial \vec{u}}(a)=<grad\,f(a),\vec{u}>$; 反之, 结论可能不成立(见例2).

证: 当$t\to 0$时, 有$f(a+t\vec{u})-f(a)=<grad\,f(a),t\vec{u}>+o(\|t\vec{u}\|)$, 故

$$\frac{\partial f}{\partial \vec{u}}(a)=\lim_{t\to 0}\frac{f(a+t\vec{u})-f(a)}{t}=<grad\,f(a),\vec{u}>. \ \square$$

例2 设 $f(x,y)=\begin{cases}\dfrac{x^2y}{x^4+y^2}, & (x,y)\ne(0,0);\\ 0, & (x,y)=(0,0),\end{cases}$ 则$\frac{\partial f}{\partial \vec{u}}(0,0)=\frac{u_1^2}{u_2},\ u_2\ne 0$; $\frac{\partial f}{\partial \vec{u}}(0,0)=0,\ u_2=0$. 显然$f$在$(0,0)$处不可微(因为$f$在$(0,0)$处不连续).

证: (略).

定理 14.2 n元函数f在a处可微$\Leftrightarrow$存在函数$\beta_k(h)$ $(k=1,2,\cdots,n)$, $\lim\limits_{h\to 0}\beta_k(h)=0$, 使得$f(a+h)-f(a)=Jf(a)h+\sum\limits_{k=1}^{n}\beta_k(h)h_k$.

证: "$\Rightarrow$". 令$\beta_k(h)=\begin{cases}\frac{f(a+h)-f(a)-Jf(a)h}{\|h\|^2}h_k, & h\neq 0;\\ 0, & h=0,\end{cases}$ 则

$$\sum_{k=1}^{n}\beta_k(h)h_k = f(a+h)-f(a)-Jf(a)h,$$

并且 $$\beta_k(h)=\tfrac{f(a+h)-f(a)-Jf(a)h}{\|h\|}\tfrac{h_k}{\|h\|}=\tfrac{o(\|h\|)}{\|h\|}\tfrac{h_k}{\|h\|}\to 0\ \ (\|h\|\to 0).$$

"$\Leftarrow$". $\left|\sum\limits_{k=1}^{n}\beta_k(h)h_k\right|\le\sum\limits_{k=1}^{n}|\beta_k(h)|\|h\|=o(\|h\|)$ $(\|h\|\to 0)$, 这说明f在a处可微. □

定理 14.3(可微的充分条件) 设f是开集$D\subset\mathbb{R}^n$上的函数, $a\in D$. 若$Jf(x)$在a的某个邻域上存在, 并且在a处连续, 则f必在a处可微.

证: 仅证$n=2$的情形.

$$\begin{aligned}f(a_1+h_1,a_2+h_2)-f(a_1,a_2)&=[f(a_1+h_1,a_2+h_2)-f(a_1,a_2+h_2)]\\&\quad+[f(a_1,a_2+h_2)-f(a_1,a_2]\\&=\frac{\partial f}{\partial x_1}(a_1+\theta_1h_1,a_2+h_2)h_1+\frac{\partial f}{\partial x_2}(a_1,a_2+\theta_2h_2)h_2,\end{aligned}$$ 其中

$$0<\theta_1=\theta_1(h)<1,\ 0<\theta_2=\theta_2(h)<1.$$

令 $$\beta_1(h)=\frac{\partial f}{\partial x_1}(a_1+\theta_1h_1,a_2+h_2)-\frac{\partial f}{\partial x_1}(a_1,a_2),$$

$$\beta_2(h)=\frac{\partial f}{\partial x_2}(a_1,a_2+h_2)-\frac{\partial f}{\partial x_2}(a_1,a_2),$$

则有 $$\lim_{h\to 0}\beta_k(h)=0\quad (k=1,2).$$

于是 $$f(a+h)-f(a)=\frac{\partial f}{\partial x_1}(a)h_1+\beta_1(h)h_1+\frac{\partial f}{\partial x_2}(a)h_2+\beta_2(h)h_2.$$

根据定理 14.2, f在a处可微. □

C^1**函数** 设f是开集$D\subset\mathbb{R}^n$上的函数. 若$Jf(x)$在D上处处存在, 并且处处连续, 则称f是D上的C^1函数; D上C^1函数的全体用$C^1(D)$表

示. 于是,

$\{C^1$ 函数$\}\subset\{$可微函数$\}\subset\{$连续函数$\}$；

$\{$可微函数$\}\subset\{$各方向导数存在的函数$\}$；

$\{$连续函数$\}$与$\{$各方向导数存在的函数$\}$互不包含.

微分用于近似计算　当$\|x-a\|$很小时，$f(x)\approx f(a)+Jf(a)(x-a)$.

练习题 14.2（P_{115}）　2, 3, 4(1, 3), 5(2, 4).

§14.3 映射的微分

定义 14.4 设 f 是从开集 $D\subset\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的映射, $a\in D$. 若存在常数矩阵 $A\in\mathbb{R}^{m\times n}$ 使得

$$f(a+h)-f(a)=Ah+o(\|h\|)\qquad(\|h\|\to 0),$$

则称 f 在 a 处可微分, 并称以 $\mathbb{R}^n$ 为定义域、以 $dx=(dx_1,dx_2,\cdots,dx_n)$ 为自变量的线性映射 Adx 为 f 在 a 处的微分, 记作 $df(a)=Adx$；若 f 在 D 中的每个点处都可微分, 则称 f 是从 D 到 $\mathbb{R}^m$ 的可微映射, 此时 $df(x)=A(x)dx$ 是从 $D\times\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的以 (x,dx) 为自变量的 $2n$ 元映射.

命题 n 元映射 $f=(f_1,f_2,\cdots,f_m)$ 在 a 处可微分当且仅当每个分量函数 $f_j\ (j=1,2,\cdots,m)$ 在 a 处可微分.

证:
$$\begin{pmatrix} f_1(a+h)\\ \vdots\\ f_m(a+h)\end{pmatrix}-\begin{pmatrix} f_1(a)\\ \vdots\\ f_m(a)\end{pmatrix}=\begin{pmatrix} a_{11},\cdots,a_{1n}\\ \vdots\ \ \ddots\ \ \vdots\\ a_{m1},\cdots,a_{mn}\end{pmatrix}\begin{pmatrix} h_1\\ \vdots\\ h_n\end{pmatrix}+o(\|h\|)\quad(\|h\|\to 0)$$

$$\Leftrightarrow f_j(a+h)-f_j(a)=\sum_{k=1}^{n}a_{jk}h_k+o(\|h\|)\ (\|h\|\to 0)\qquad(j=1,2,\cdots,m).\ \square$$

多变量映射的 Jacobi 矩阵 若 n 元映射 $f=(f_1,f_2,\cdots,f_m)$ 在 a 处的所有 1 阶偏导数 $\frac{\partial f_j}{\partial x_k}(a)\ (1\le j\le m,\ 1\le k\le n)$ 都存在, 则记

$$Jf(a)=\begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a),\frac{\partial f_1}{\partial x_2}(a),\cdots,\frac{\partial f_1}{\partial x_n}(a)\\ \frac{\partial f_2}{\partial x_1}(a),\frac{\partial f_2}{\partial x_2}(a),\cdots,\frac{\partial f_2}{\partial x_n}(a)\\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ \frac{\partial f_m}{\partial x_1}(a),\frac{\partial f_m}{\partial x_2}(a),\cdots,\frac{\partial f_m}{\partial x_n}(a)\end{pmatrix},$$

称为 f 在 a 处的 Jacobi 矩阵. 于是, 若 f 在 a 处可微分, 则

$$df(a)=Jf(a)dx.$$

定理 14.4（可微的充分条件） 设 f 是从开集 $D\subset\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的映射, $a\in D$. 若 $Jf(x)$ 在 a 的某个邻域上存在, 并且在 a 处连续, 则 f 必在 a 处可微.

证：　因为 $Jf_j(x)$ 在 a 的某个邻域上存在，并且在 a 处连续，故 f_j 在 a 处可微 $(j=1,2,\cdots,m)$，从而 f 在 a 处可微. □

练习题 14.3（P_{118}）　2, 3, 4, 6, 7, 8.

§ 14. 4　复合求导

引理　对于 $A\in\mathbb{R}^{m\times n}$, $h\subset\mathbb{R}^n$, 一定有 $\|Ah\|\le\|A\|\|h\|$.

证：记 $\vec{a}_j=(a_{j1},\cdots,a_{jn})$, 则 $\begin{pmatrix}a_{11},\cdots,a_{1n}\\ \vdots\ \ \ddots\ \ \vdots\\ a_{m1},\cdots,a_{mn}\end{pmatrix}\begin{pmatrix}h_1\\ \vdots\\ h_n\end{pmatrix}=\begin{pmatrix}<\vec{a}_1,h>\\ \vdots\\ <\vec{a}_m,h>\end{pmatrix}$. 故

$$\|Ah\|^2=\sum_{j=1}^{m}<\vec{a}_j,h>^2\le\sum_{j=1}^{m}\|\vec{a}_j\|^2\|h\|^2=\|A\|^2\|h\|^2.\ \square$$

定理 14. 5（复合映射求导的链式法则）　设 $D\subset\mathbb{R}^n$, $G\subset\mathbb{R}^m$ 是开集, $f:D\to G$, $g:G\to\mathbb{R}^l$ 是映射. 若 f 在 $x\in D$ 处可微分, g 在 $y=f(x)\in G$ 处可微分, 则复合映射 $g\circ f:D\to\mathbb{R}^l$ 在 x 处可微分, 并且

$$J(g\circ f)(x)=Jg(f(x))Jf(x).$$

证：记 $k=f(x+h)-f(x)=Jf(x)h+u(h)$, 其中 $u(h)=o(\|h\|)$ $(\|h\|\to 0)$,

则
$$g\circ f(x+h)-g\circ f(x)=g(y+k)-g(y)$$
$$=Jg(y)k+v(k),\ \text{其中}\ v(k)=o(\|k\|)\ (\|k\|\to 0).$$

于是
$$g\circ f(x+h)-g\circ f(x)=Jg(y)[Jf(x)h+u(h)]+v(k)$$
$$=Jg(y)Jf(x)h+Jg(y)u(h)+v(k).$$
$$\|Jg(y)u(h)+v(k)\|\le\|Jg(y)\|\|u(h)\|+\|v(k)\|$$
$$\le\begin{cases}\|Jg(y)\|\|u(h)\|+\dfrac{\|v(k)\|}{\|k\|}[\|Jf(x)\|\|h\|+\|u(h)\|], & \|k\|\ne 0;\\ \|Jg(y)\|\|u(h)\|, & \|k\|=0.\end{cases}$$

由 $\lim\limits_{\|h\|\to 0}k=\lim\limits_{\|h\|\to 0}[f(x+h)-f(x)]=0$, 便知 $\lim\limits_{\|h\|\to 0}\dfrac{\|Jg(y)u(h)+v(k)\|}{\|h\|}=0$.

故
$$g\circ f(x+h)-g\circ f(x)=Jg(y)Jf(x)h+o(\|h\|)\ (\|h\|\to 0).\ \square$$

推论 14. 5′（微分的形式不变性）　$dg(f(x))=Jg(f(x))df(x)$.

推论 14. 5″（复合函数求导的链式法则）　对于复合函数 $\varphi(x_1,\cdots,x_n)=g[f_1(x_1,\cdots,x_n),\cdots,f_m(x_1,\cdots,x_n)]$, 有

$$\frac{\partial \varphi}{\partial x_k}(x)=\sum_{j=1}^{m}\frac{\partial g}{\partial y_j}[f(x)]\frac{\partial f_j}{\partial x_k}(x),\ k=1,2,\cdots,n.$$

例 1（P_{124}，**第 10 题**）　设 $f(x,y,z)$ 是 3 元函数，$\vec{n}_1,\vec{n}_2,\vec{n}_3$ 是 $\mathbb{R}^3$ 中三个互相正交的方向．求证：

$$\left(\frac{\partial f}{\partial \vec{n}_1}\right)^2+\left(\frac{\partial f}{\partial \vec{n}_2}\right)^2+\left(\frac{\partial f}{\partial \vec{n}_3}\right)^2=\left(\frac{\partial f}{\partial x}\right)^2+\left(\frac{\partial f}{\partial y}\right)^2+\left(\frac{\partial f}{\partial z}\right)^2.$$

证：
$$\nabla f=<\nabla f,\vec{n}_1>\vec{n}_1+<\nabla f,\vec{n}_2>\vec{n}_2+<\nabla f,\vec{n}_3>\vec{n}_3$$
$$=\frac{\partial f}{\partial \vec{n}_1}\vec{n}_1+\frac{\partial f}{\partial \vec{n}_2}\vec{n}_2+\frac{\partial f}{\partial \vec{n}_3}\vec{n}_3,$$
两边同时求长度的平方，就得到结论．□

例 2（P_{124}，**第 8 题**）　设 $f(x,y,z)=F(u,v,w)$，其中 $x^2=vw,\ y^2=wu$，$z^2=uv$．求证：

$$x\frac{\partial f}{\partial x}+y\frac{\partial f}{\partial y}+z\frac{\partial f}{\partial z}=u\frac{\partial F}{\partial u}+v\frac{\partial F}{\partial v}+w\frac{\partial F}{\partial w}.$$

证：令 $\varphi(t)=f(tx,ty,tz)$，显然 $\varphi(t)=F(tu,tv,tw)$．于是有

$$\varphi'(1)=\frac{\partial f}{\partial x}(x,y,z)x+\frac{\partial f}{\partial y}(x,y,z)y+\frac{\partial f}{\partial z}(x,y,z)z;$$

又有
$$\varphi'(1)=\frac{\partial F}{\partial u}(u,v,w)u+\frac{\partial F}{\partial v}(u,v,w)v+\frac{\partial F}{\partial w}(u,v,w)w.$$ □

练习题 14.4（P_{123}）　2, 3, 5, 6, 7, 9.

§14.5 拟微分中值定理

定理 10.7(多变量函数的微分中值定理) 若 f 是凸区域 $D\subset\mathbb{R}^n$ 上的可微函数, 则 $\forall a,b\in D, a\neq b$, 必存在 $\xi\in(a,b)\subset D$, 使得

$$f(b)-f(a)=Jf(\xi)(b-a).$$

证: $\varphi(t)=f[(1-t)a+tb]$ 是 $[0,1]$ 上的单变量可微函数, 故 $\exists\theta\in(0,1)$ 使得 $\varphi'(\theta)=\varphi(1)-\varphi(0)=f(b)-f(a)$. 令 $\xi=(1-\theta)a+\theta b\in(a,b)$, 则有

$$\varphi'(\theta)=Jf(\xi)(b-a). \ \square$$

定理 10.8(单变量映射的拟微分中值定理) 若单变量连续映射 $f:[a,b]\to\mathbb{R}^m$ 在 (a,b) 上可微, 则必存在 $\xi\in(a,b)$, 使得

$$\|f(b)-f(a)\|\le\|Jf(\xi)\|(b-a).$$

证: 单变量函数 $\varphi(t)=<f(t),f(b)-f(a)>$ 在 $[a,b]$ 上连续, 在 (a,b) 上可微, 故存在 $\xi\in(a,b)$, 使得

$$\varphi(b)-\varphi(a)=\varphi'(\xi)(b-a),$$

即 $$\|f(b)-f(a)\|^2=<Jf(\xi),f(b)-f(a)>(b-a)$$

$$\le\|Jf(\xi)\|\|f(b)-f(a)\|(b-a). \ \square$$

注记 10. $8'$ 单变量映射不成立形如"$f(b)-f(a)=Jf(\xi)(b-a)$"微分中值定理.

证: $f(t)=(\cos t,\ \sin t)$ 在 $[0,2\pi]$ 上连续, 在 $(0,2\pi)$ 上可微. 但是,

$$f(2\pi)-f(0)=(0,0);$$

$$Jf(\xi)(2\pi-0)=(-\sin\xi,\cos\xi)2\pi\neq(0,0). \ \square$$

定理 10.9(多变量映射的拟微分中值定理) 若 f 是从凸区域 $D\subset\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的可微映射, 则 $\forall a,b\in D, a\neq b$, 必存在 $\xi\in(a,b)\subset D$, 使得

$$\|f(b)-f(a)\|\le\|Jf(\xi)\|\|b-a\|.$$

证: 单变量映射 $\varphi(t)=f[(1-t)a+tb]$ 在 $[0,1]$ 上可微, 故 $\exists\theta\in(0,1)$ 使得 $\|\varphi(1)-\varphi(0)\|\le\|J\varphi(\theta)\|$(定理 10.8). 令 $\xi=(1-\theta)a+\theta b\in(a,b)$, 则有

$J\varphi(\theta)=Jf(\xi)(b-a)$. 故 $\|f(b)-f(a)\|=\|\varphi(1)-\varphi(0)\|\le\|Jf(\xi)\|\|b-a\|$.

□

定理 10.10 设 f 是从区域(连通开集) $D\subset\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的可微映射. 若 $Jf=0$ 在 D 上成立, 则 f 在 D 上是常向量.

证:(反证法)假定 $\exists a,b\in D$ 使得 $f(a)\neq f(b)$. 令 $A=\{x\in D: f(x)=f(a)\}$, $B=\{x\in D: f(x)\neq f(a)\}$, 则 $D=A\bigcup B$ 是非空无交并分解. 对 f 应用拟微分中值定理易知 A 是开集；由 f 的连续性易知 B 也是开集. 这与 D 的连通性相矛盾(定理 13.15). □

§14.6 隐函数定理

隐函数问题 一、函数方程$F(x_1,\cdots,x_n,y)=0$在什么条件下能确保存在一个n元函数$y=f(x_1,\cdots,x_n)$是函数方程$F(x_1,\cdots,x_n,y)=0$的解, 即

$$F(x_1,\cdots,x_n,f(x_1,\cdots,x_n))\equiv 0\ ?$$

二、如何求出Jf, 即利用$JF=(J_xF,\frac{\partial F}{\partial y})$来表示$Jf$?

定理 14.12(隐函数定理) 若 $F(x,y)$ 是 $B_\delta(a,b)\subset\mathbb{R}^{n+1}$ 上的 C^k 函数(即 F 的所有 k 阶偏导函数都连续), $F(a,b)=0$, $\frac{\partial F}{\partial y}(a,b)\neq 0$, 则存在 $B_\varepsilon(a)\subset\mathbb{R}^n$ $(\varepsilon<\delta)$ 上唯一一个C^k函数$y=f(x)$, 使得

(1) $F(x,f(x))\equiv 0$;

(2) $b=f(a)$;

(3) 利用$F(x,f(x))\equiv 0$所确定的矩阵方程

$$J_xF(x,f(x))+\frac{\partial F}{\partial y}(x,f(x))Jf(x)=0,$$

能解出 $$Jf(x)=-[\frac{\partial F}{\partial y}(x,f(x))]^{-1}J_xF(x,f(x)).$$

证: 隐函数定理的证明主要利用“导数判别函数的增减性”和“连续函数的介值定理”, 很初等但也很罗嗦, 这里省略. □

例 1 设$z=z(x,y)$是由方程$x+y+z=e^{-(x+y+z)}$所确定的函数, 求出$\dfrac{\partial z}{\partial x},\dfrac{\partial z}{\partial y}$.

解: 在 $x+y+z(x,y)=e^{-(x+y+z(x,y))}$中分别对$x,y$求偏导数, 便得到

$$1+\frac{\partial z}{\partial x}(x,y)=-e^{-(x+y+z(x,y))}(1+\frac{\partial z}{\partial x}(x,y)),$$

$$1+\frac{\partial z}{\partial y}(x,y)=-e^{-(x+y+z(x,y))}(1+\frac{\partial z}{\partial y}(x,y)).$$

故 $\dfrac{\partial z}{\partial x}(x,y)+1\equiv 0,\ \dfrac{\partial z}{\partial y}(x,y)+1\equiv 0$, 从而$x+y+z(x,y)$是常数, 即

$z=z(x,y)$就是平面$x+y+z=c$, 其中常数c满足$ce^c=1$. □

例 2 研究由方程 $y-\varepsilon\sin y=x$ 所确定的隐函数 $y=y(x)$，其中 $\varepsilon\in(0,1)$ 是常数.

解: $f(y)=y-\varepsilon\sin y$ 是 $\mathbb{R}$ 上严格递增的 C^{∞} 函数, 值域为 $\mathbb{R}$, 故其反函数 $f^{-1}(x)$ 也是 $\mathbb{R}$ 上严格递增的 C^{∞} 函数,

$$(f^{-1})'(x)=\frac{1}{f'[f^{-1}(x)]}=\frac{1}{1+\varepsilon\cos f^{-1}(x)}.$$

这与隐函数定理的结论 $y'(x)+[\varepsilon\cos y(x)]y'(x)=1$ 相一致. □

例 3（P_{135}，第 3 题） **解**: $\dfrac{\partial x}{\partial y}$ 意为由 $F(x,y,z)=0$ 所确定的隐函数 $x=x(y,z)$ 关于 y 的偏导数, 其它类似.

由 $F(x(y,z),y,z)=0$，可得 $\dfrac{\partial F}{\partial x}\dfrac{\partial x}{\partial y}+\dfrac{\partial F}{\partial y}=0$；

由 $F(x,y(z,x),z)=0$，可得 $\dfrac{\partial F}{\partial y}\dfrac{\partial y}{\partial z}+\dfrac{\partial F}{\partial z}=0$；

由 $F(x,y,z(x,y))=0$，可得 $\dfrac{\partial F}{\partial x}+\dfrac{\partial F}{\partial z}\dfrac{\partial z}{\partial x}=0$.

乘起来便有 $\left(\dfrac{\partial F}{\partial x}\dfrac{\partial F}{\partial y}\dfrac{\partial F}{\partial z}\right)\left(\dfrac{\partial x}{\partial y}\dfrac{\partial y}{\partial z}\dfrac{\partial z}{\partial x}\right)=-\left(\dfrac{\partial F}{\partial x}\dfrac{\partial F}{\partial y}\dfrac{\partial F}{\partial z}\right)$. □

例 4 方程 $\sin x+\log y-xy^3=0$ 能在 $(0,1)$ 的附近确定唯一的隐函数 $y=y(x)$ 满足 $y(0)=1$，并且 $y'(0)=0$.

解: 记 $F(x,y)=\sin x+\log y-xy^3$，则 $F(0,1)=0$，$\dfrac{\partial F}{\partial y}(0,1)=1$，符合隐函数定理的条件. 故在 $(0,1)$ 的附近确定唯一的隐函数 $y=y(x)$ 满足 $y(0)=1$，并且 $y'(0)=-\left(\dfrac{\partial F}{\partial y}(0,1)\right)^{-1}\dfrac{\partial F}{\partial x}(0,1)=0$. □

练习题 14.6（P_{134}） 1(2), 2(1, 3), 4, 5, 6.

§14.7 隐映射定理

隐映射问题 一、函数方程组

$$F(x,y)=\begin{pmatrix} F_1(x_1,\cdots,x_n,y_1,\cdots,y_m) \\ \vdots \\ F_m(x_1,\cdots,x_n,y_1,\cdots,y_m) \end{pmatrix}=\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}=0$$

在什么条件下能确保存在一个n元映射

$$y=\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}=\begin{pmatrix} f_1(x_1,\cdots,x_n) \\ \vdots \\ f_m(x_1,\cdots,x_n) \end{pmatrix}=f(x)$$

是函数方程组$F(x,y)=0$的解, 即

$$F(x,f(x))\equiv 0\ ?$$

二、如何求出Jf, 即利用$JF=(J_xF,J_yF)$来表示Jf ？

定理 14.13(隐映射定理) 若$F(x,y)$是从$B_\delta(a,b)\subset\mathbb{R}^{n+m}$到$\mathbb{R}^m$的$C^k$映射(即每个$F_j$的所有$k$阶偏导函数都连续, $j=1,\cdots,m$), $F(a,b)=0$, $\det J_yF(a,b)\neq 0$, 则存在从$B_\varepsilon(a)\subset\mathbb{R}^n$ $(\varepsilon<\delta)$到$\mathbb{R}^m$的唯一一个C^k映射$y=f(x)$, 使得

(1) $F(x,f(x))\equiv 0$；

(2) $b=f(a)$；

(3) 利用$F(x,f(x))\equiv 0$所确定的矩阵方程

$$J_xF(x,f(x))+J_yF(x,f(x))Jf(x)=0,$$

能解出 $$Jf(x)=-[J_yF(x,f(x))]^{-1}J_xF(x,f(x)).$$

证： 不妨设$m=2$. 因为 $\det J_yF(a,b)=\det\begin{pmatrix} \frac{\partial F_1}{\partial y_1}(a,b),\frac{\partial F_1}{\partial y_2}(a,b) \\ \frac{\partial F_2}{\partial y_1}(a,b),\frac{\partial F_2}{\partial y_2}(a,b) \end{pmatrix}\neq 0$, 故 $\begin{pmatrix} \frac{\partial F_1}{\partial y_2}(a,b) \\ \frac{\partial F_2}{\partial y_2}(a,b) \end{pmatrix}\neq\begin{pmatrix} 0 \\ 0 \end{pmatrix}$. 可设$\frac{\partial F_2}{\partial y_2}(a,b)\neq 0$. 由隐函数定理, 存在$B_{\delta_1}(a,b_1)\subset\mathbb{R}^{n+1}$

$(\delta_1 < \delta)$ 上唯一一个 C^k 函数 $y_2 = g(x, y_1)$ 使得 $F_2(x, y_1, g(x, y_1)) \equiv 0,\ b_2 = g(a, b_1)$，$\dfrac{\partial g}{\partial y_1}(a, b_1) = -[\frac{\partial F_2}{\partial y_2}(a,b)]^{-1}\frac{\partial F_2}{\partial y_1}(a,b)$. 再考虑方程 $F_1(x, y_1, g(x, y_1)) = 0$. 令 $G(x, y_1) = F_1(x, y_1, g(x, y_1))$，则 $G(x, y_1)$ 是 $B_{\delta_1}(a, b_1) \subset \mathbb{R}^{n+1}$ 上的 C^k 函数，$G(a, b_1) = F_1(a, b) = 0$, $\frac{\partial G}{\partial y_1}(a, b_1) = \frac{\partial F_1}{\partial y_1}(a, b) + \frac{\partial F_1}{\partial y_2}(a, b)\frac{\partial g}{\partial y_1}(a, b_1) = \frac{\partial F_1}{\partial y_1}(a, b) - \frac{\partial F_1}{\partial y_2}(a, b)[\frac{\partial F_2}{\partial y_2}(a, b)]^{-1}\frac{\partial F_2}{\partial y_1}(a, b) = [\frac{\partial F_2}{\partial y_2}(a, b)]^{-1}\det\begin{pmatrix}\frac{\partial F_1}{\partial y_1}(a,b), \frac{\partial F_1}{\partial y_2}(a,b) \\ \frac{\partial F_2}{\partial y_1}(a,b), \frac{\partial F_2}{\partial y_2}(a,b)\end{pmatrix} \neq 0$. 再由隐函数定理, 存在 $B_\varepsilon(a) \subset \mathbb{R}^n$ $(\varepsilon < \delta_1)$ 上唯一一个 C^k 函数 $y_1 = f_1(x)$ 使得 $G(x, f_1(x)) = F_1(x, f_1(x), g(x, f_1(x))) \equiv 0, b_1 = f_1(a)$. 令 $y_2 = f_2(x) = g(x, f_1(x))$，则 $\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} f_1(x) \\ f_2(x) \end{pmatrix}$ 是 $B_\varepsilon(a)$ 上的 C^k 映射, 满足

$$\begin{pmatrix} F_1(x, f_1(x), f_2(x)) \\ F_2(x, f_1(x), f_2(x)) \end{pmatrix} \equiv \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} f_1(a) \\ f_2(a) \end{pmatrix}. \ \square$$

例（P_{143}，第 5 题）　解：将 $\begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} u(x,y) \\ v(x,y) \\ w(x,y) \end{pmatrix}$ 代入方程组, 并对方程组两边关于 x, y 求 Jacobi 矩阵, 便得

$$\begin{pmatrix} 0, & 0 \\ 0, & -\frac{\pi}{2} \\ 0, & 0 \end{pmatrix} + \begin{pmatrix} 2, & -1, & 0 \\ 2, & 2, & 0 \\ 1, & 1, & 1 \end{pmatrix}\begin{pmatrix} \frac{\partial u}{\partial x}, \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x}, \frac{\partial v}{\partial y} \\ \frac{\partial w}{\partial x}, \frac{\partial w}{\partial y} \end{pmatrix} = \begin{pmatrix} 0, & 0 \\ 0, & 0 \\ 0, & 0 \end{pmatrix},$$

$$\begin{pmatrix} 2, & -1, & 0 \\ 2, & 2, & 0 \\ 1, & 1, & 1 \end{pmatrix}\begin{pmatrix} \frac{\partial u}{\partial x}, \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x}, \frac{\partial v}{\partial y} \\ \frac{\partial w}{\partial x}, \frac{\partial w}{\partial y} \end{pmatrix} = \begin{pmatrix} 0, & 0 \\ 0, & \frac{\pi}{2} \\ 0, & 0 \end{pmatrix},$$

$$\begin{pmatrix} \frac{\partial u}{\partial x}, \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x}, \frac{\partial v}{\partial y} \\ \frac{\partial w}{\partial x}, \frac{\partial w}{\partial y} \end{pmatrix} = \begin{pmatrix} 0, & \frac{\pi}{12} \\ 0, & \frac{\pi}{6} \\ 0, & -\frac{\pi}{4} \end{pmatrix}. \ \square$$

练习题 14.7（P_{142}）　2, 3, 4, 6.

§14.8　逆映射定理

同胚映射　设$E\subset\mathbb{R}^n$是非空点集，$f: E\to\mathbb{R}^m$是单射. 若f在E上连续，f^{-1}在$f(E)$上连续, 则称f是从E到$f(E)$上的同胚映射, 并称E与$f(E)$同胚.

注记　若$m\neq n$, 则$\mathbb{R}^n$中的区域与$\mathbb{R}^m$中的区域一定不同胚.

邻域　设$a\in\mathbb{R}^n$. 若开集$U\subset\mathbb{R}^n$含有a, 则称U是a的一个邻域.

定理 14.14（局部逆映射定理）　设f是从开集$D\subset\mathbb{R}^n$到$\mathbb{R}^n$的C^k映射, $a\in D$. 若$\det Jf(a)\neq 0$, 则存在a的邻域$U\subset D$和$b=f(a)$的邻域$V\subset f(D)$, 使得

（1）　$f|_U$是从U到V上的C^k同胚映射（即$f|_U$和$(f|_U)^{-1}$都是C^k映射）；

（2）　$J(f|_U)^{-1}(y)=[J(f|_U)(x)]^{-1}$, $\forall y=f(x)\in V$.

证：对方程组$\begin{pmatrix} f_1(x)-y_1 \\ \vdots \\ f_n(x)-y_n \end{pmatrix}=\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$应用隐映射定理即得到结论（细节省略）. □

推论 14.14′　设f是从开集$D\subset\mathbb{R}^n$到$\mathbb{R}^n$的C^k映射. 若$\forall x\in D$都有$\det Jf(x)\neq 0$, 则$G=f(D)$是$\mathbb{R}^n$中的开集.

定理 14.15（整体逆映射定理）　设f是从开集$D\subset\mathbb{R}^n$到$\mathbb{R}^n$的C^k单射, 并且$\forall x\in D$都有$\det Jf(x)\neq 0$, 则

（1）　f是从D到$f(D)$上的C^k同胚映射；

（2）　$Jf^{-1}(y)=[Jf(x)]^{-1}$, $\forall y=f(x)\in f(D)$.

证：由局部逆映射定理及其推论. □

练习题 14.8（P_{147}）　1, 2.

§14.9 高阶偏导数

高阶偏导数 对 n 元函数 f 可定义 n 个 1 阶偏导函数 $\dfrac{\partial f}{\partial x_i}$ $(1\le i\le n)$；n^2 个 2 阶偏导函数 $\dfrac{\partial^2 f}{\partial x_i\partial x_j}=\dfrac{\partial}{\partial x_i}(\dfrac{\partial f}{\partial x_j})(1\le i,j\le n,i\ne j)$ 和 $\dfrac{\partial^2 f}{\partial x_i^2}=\dfrac{\partial}{\partial x_i}(\dfrac{\partial f}{\partial x_i})$ $(1\le i\le n)$；类似地, 可定义 n^k 个 k 阶偏导函数.

定理 14.16 设 f 是开集 $D\subset\mathbb{R}^n$ 上的函数, $a\in D$. 若 $\dfrac{\partial^2 f}{\partial x_i\partial x_j}$ 和 $\dfrac{\partial^2 f}{\partial x_j\partial x_i}$ 都在 a 的某个邻域上存在, 并且在 a 处连续, 则 $\dfrac{\partial^2 f}{\partial x_i\partial x_j}(a)=\dfrac{\partial^2 f}{\partial x_j\partial x_i}(a)$.

证: 不妨设 $n=2$. 对充分小的常数 $h_1,h_2>0$, 令

$$\varphi(x_2)=f(a_1+h_1,x_2)-f(a_1,x_2),$$

$$\psi(x_1)=f(x_1,a_2+h_2)-f(x_1,a_2).$$

则

$$\varphi(a_2+h_2)-\varphi(a_2)=f(a_1+h_1,a_2+h_2)-f(a_1,a_2+h_2)$$

$$-f(a_1+h_1,a_2)+f(a_1,a_2)=\psi(a_1+h_1)-\psi(a_1).$$

一方面,

$$\varphi(a_2+h_2)-\varphi(a_2)=\varphi'(a_2+\theta_2h_2)h_2$$

$$=[\frac{\partial f}{\partial x_2}(a_1+h_1,a_2+\theta_2h_2)-\frac{\partial f}{\partial x_2}(a_1,a_2+\theta_2h_2)]h_2$$

$$=\frac{\partial^2 f}{\partial x_1\partial x_2}(a_1+\theta_1h_1,a_2+\theta_2h_2)h_1h_2,\quad 0<\theta_1,\theta_2<1.$$

另一方面,

$$\psi(a_1+h_1)-\psi(a_1)=\psi'(a_1+\eta_1h_1)h_1$$

$$=[\frac{\partial f}{\partial x_1}(a_1+\eta_1h_1,a_2+h_2)-\frac{\partial f}{\partial x_1}(a_1+\eta_1h_1,a_2)]h_1$$

$$=\frac{\partial^2 f}{\partial x_2\partial x_1}(a_1+\eta_1h_1,a_2+\eta_2h_2)h_1h_2,\quad 0<\eta_1,\eta_2<1.$$

于是,

$$\frac{\partial^2 f}{\partial x_1\partial x_2}(a_1+\theta_1h_1,a_2+\theta_2h_2)=\frac{\partial^2 f}{\partial x_2\partial x_1}(a_1+\eta_1h_1,a_2+\eta_2h_2).$$

令 $(h_1,h_2)\to(0,0)$, 就得到 $\dfrac{\partial^2 f}{\partial x_1\partial x_2}(a_1,a_2)=\dfrac{\partial^2 f}{\partial x_2\partial x_1}(a_1,a_2)$. □

定理 14.16′ 若 n 元函数 f 的两个 k 阶偏导函数都在 a 处连续, 仅仅是求导次序不同, 则这两个 k 阶偏导函数在 a 处相等.

注记 14.16″ 若 n 元函数 f 的所有 k 阶偏导函数都在 a 处连续, 则 f 在 a 处的 k 阶偏导数的全体便是 $\dfrac{\partial^k f}{\partial x_1^{i_1}\partial x_2^{i_2}\cdots\partial x_n^{i_n}}(a)$ $(0\le i_1, i_2,\cdots,i_n\le k$, $i_1+i_2+\cdots+i_n=k)$, 一共有 C_{k+n-1}^{n-1} 个.

证: f 在 a 处的所有 k 阶偏导数的个数 $=(x_1+\cdots+x_n)^k$ 所含单项式的个数. 下面对自变量的个数 n 应用数学归纳法. 1个自变量时, x_1^k 所含单项式的个数为 1, 结论成立; 假定 $n-1$ 个自变量时, $(x_1+\cdots+x_{n-1})^k$ 所含单项式的个数为 C_{k+n-2}^{n-2} ; 于是,

$$
\begin{aligned}
&(x_1+\cdots+x_n)^k \text{所含单项式的个数}\\
&=\sum_{i=0}^{k} C_k^i x_n^{k-i}(x_1+\cdots+x_{n-1})^i \text{所含单项式的个数}\\
&=\sum_{i=0}^{k} C_{i+n-2}^{n-2}.
\end{aligned}
$$

因为
$$
\begin{aligned}
\sum_{k=0}^{\infty} C_{k+n-1}^{n-1}x^k &= (1-x)^{-n}=(1-x)^{-(n-1)}(1-x)^{-1}\\
&=\left(\sum_{i=0}^{\infty} C_{i+n-2}^{n-2}x^i\right)\left(\sum_{j=0}^{\infty} x^j\right)=\sum_{k=0}^{\infty}\left(\sum_{i=0}^{k} C_{i+n-2}^{n-2}\right)x^k,
\end{aligned}
$$

故
$$\sum_{i=0}^{k} C_{i+n-2}^{n-2}=C_{k+n-1}^{n-1}.$$

这说明 n 个自变量时, 结论也成立. □

例 1（P_{154}, **第 6 题**） **解**:

(1)
$$\frac{\partial^2 u}{\partial x^2}=0 \Rightarrow \frac{\partial u}{\partial x}=f(y,z);$$
$$\frac{\partial}{\partial x}[u-f(y,z)x]=0\Rightarrow u-f(y,z)x=g(y,z).$$

故
$$u(x,y,z)=f(y,z)x+g(y,z).$$

(2)
$$\frac{\partial^2 u}{\partial x\partial y}=0 \Rightarrow \frac{\partial u}{\partial y}=f_1(y,z);$$

$$\frac{\partial}{\partial y}[u-\int f_1(y,z)dy]=0 \Rightarrow u-\int f_1(y,z)dy=g(z,x).$$

故 $\quad u(x,y,z)=f(y,z)+g(z,x)$.

(3) $\quad \frac{\partial^3 u}{\partial x\partial y\partial z}=0 \Rightarrow \frac{\partial u}{\partial z}=f_1(y,z)+g_1(z,x)$;

$$\frac{\partial}{\partial z}[u-\int f_1(y,z)dz-\int g_1(z,x)dz]=0 \Rightarrow$$

$$u-\int f_1(y,z)dz-\int g_1(z,x)dz=h(x,y).$$

故 $\quad u(x,y,z)=f(y,z)+g(z,x)+h(x,y)$. □

例 2 设 $z=f(x,y),\ x=\varphi(u,v),\ y=\psi(u,v),\ z(u,v)=f[\varphi(u,v),\psi(u,v)]$,

求 $\frac{\partial^2 z}{\partial u^2}(u,v),\ \frac{\partial^2 z}{\partial u\partial v}(u,v),\ \frac{\partial^2 z}{\partial v^2}(u,v)$.

解: 仅求出 $\frac{\partial^2 z}{\partial u^2}(u,v)$.

$$\frac{\partial z}{\partial u}=\frac{\partial f}{\partial x}\frac{\partial \varphi}{\partial u}+\frac{\partial f}{\partial y}\frac{\partial \psi}{\partial u};$$

$$\frac{\partial^2 z}{\partial u^2}=[\frac{\partial^2 f}{\partial x^2}\frac{\partial \varphi}{\partial u}+\frac{\partial^2 f}{\partial x\partial y}\frac{\partial \psi}{\partial u}]\frac{\partial \varphi}{\partial u}+\frac{\partial f}{\partial x}\frac{\partial^2 \varphi}{\partial u^2}$$

$$+[\frac{\partial^2 f}{\partial x\partial y}\frac{\partial \varphi}{\partial u}+\frac{\partial^2 f}{\partial y^2}\frac{\partial \psi}{\partial u}]\frac{\partial \psi}{\partial u}+\frac{\partial f}{\partial y}\frac{\partial^2 \psi}{\partial u^2}$$

$$=\frac{\partial^2 f}{\partial x^2}\left(\frac{\partial \varphi}{\partial u}\right)^2+2\frac{\partial^2 f}{\partial x\partial y}\frac{\partial \varphi}{\partial u}\frac{\partial \psi}{\partial u}+\frac{\partial^2 f}{\partial y^2}\left(\frac{\partial \psi}{\partial u}\right)^2+\frac{\partial f}{\partial x}\frac{\partial^2 \varphi}{\partial u^2}+\frac{\partial f}{\partial y}\frac{\partial^2 \psi}{\partial u^2}.$$ □

练习题 14.9（P_{153}） 1(3, 10), 2, 4, 5.

§14.10 Taylor 公式

本节内容可用于函数估计、极值判断、近似计算等，具有较高的理论价值.

引理 14.1(推广的二项式定理) 对于$k\in\mathbb{N}$, $\alpha=(\alpha_1,\cdots,\alpha_n)\in\mathbb{N}^n$ 和 $x=(x_1,\cdots,x_n)\in\mathbb{R}^n$, 有如下的多项式展开

$$(x_1+\cdots+x_n)^k=\sum_{\alpha_1+\cdots+\alpha_n=k}\frac{k!}{\alpha_1!\cdots\alpha_n!}x_1^{\alpha_1}\cdots x_n^{\alpha_n}=\sum_{|\alpha|=k}\frac{k!}{\alpha!}x^\alpha .$$

证： 对 n 应用数学归纳法. □

定理 14.17(带 Lagrange 余项的 Taylor 公式) 设$D\subset\mathbb{R}^n$是凸区域. 若 f 是D上的C^{m+1}函数, 则$\forall a,x\in D, x\neq a$, 必$\exists\xi\in(a,x)\subset D$, 使得

$$f(x)=\sum_{k=0}^{m}\frac{1}{k!}\left(\sum_{j=1}^{n}(x_j-a_j)\frac{\partial}{\partial x_j}\right)^k f(a)+\frac{1}{(m+1)!}\left(\sum_{j=1}^{n}(x_j-a_j)\frac{\partial}{\partial x_j}\right)^{m+1} f(\xi)$$

$$=\sum_{k=0}^{m}\sum_{|\alpha|=k}\frac{1}{\alpha!}\frac{\partial^k f}{\partial x^\alpha}(a)(x-a)^\alpha+\sum_{|\alpha|=m+1}\frac{1}{\alpha!}\frac{\partial^{m+1} f}{\partial x^\alpha}(\xi)(x-a)^\alpha .$$

证： $\varphi(t)=f[(1-t)a+tx]$是$[0,1]$上的C^{m+1}函数, 故存在$\theta\in(0,1)$使得

$$\varphi(1)=\sum_{k=0}^{m}\frac{1}{k!}\varphi^{(k)}(0)+\frac{1}{(m+1)!}\varphi^{(m+1)}(\theta).$$

由
$$\frac{d}{dt}\varphi(t)=\sum_{j=1}^{n}\frac{\partial f}{\partial x_j}[(1-t)a+tx](x_j-a_j)$$

$$=\left(\sum_{j=1}^{n}(x_j-a_j)\frac{\partial}{\partial x_j}\right)f[(1-t)a+tx],$$

可得
$$\frac{d^k}{dt^k}\varphi(t)=\left(\sum_{j=1}^{n}(x_j-a_j)\frac{\partial}{\partial x_j}\right)^k f[(1-t)a+tx].$$

故
$$\varphi^{(k)}(0)=\left(\sum_{j=1}^{n}(x_j-a_j)\frac{\partial}{\partial x_j}\right)^k f(a),\quad k=0,1,\cdots,m,$$

$$\varphi^{(m+1)}(\theta)=\left(\sum_{j=1}^{n}(x_j-a_j)\frac{\partial}{\partial x_j}\right)^{m+1} f(\xi),\quad \xi=(1-\theta)a+\theta x .$$

带入前式即得到结论. □

注记 14.17′ $f(x_1,\cdots,x_n)$在$a=(a_1,\cdots,a_n)$处的 Taylor 展开式的前三项为

$$f(x)=f(a)+Jf(a)\begin{pmatrix}x_1-a_1\\ \vdots\\ x_n-a_n\end{pmatrix}+\frac{1}{2!}(x_1-a_1,\cdots,x_n-a_n)Hf(a)\begin{pmatrix}x_1-a_1\\ \vdots\\ x_n-a_n\end{pmatrix}+\cdots,$$

其中$Hf(a)=\left(\dfrac{\partial^2 f}{\partial x_i\partial x_j}(a)\right)$是$n$阶对称方阵, 称为$f$在$a$处的 Hessian 方阵.

定理 14.18(带 Peano 余项的 Taylor 公式) 若$f(x_1,\cdots,x_n)$是a的邻域上的C^m函数, 则

$$\begin{aligned}f(x)&=\sum_{k=0}^{m}\frac{1}{k!}\left(\sum_{j=1}^{n}(x_j-a_j)\frac{\partial}{\partial x_j}\right)^k f(a)+o(\|x-a\|^m)\\&=\sum_{k=0}^{m}\sum_{|\alpha|=k}\frac{1}{\alpha!}\frac{\partial^k f}{\partial x^\alpha}(a)(x-a)^\alpha+o(\|x-a\|^m)\quad(x\to a).\end{aligned}$$

证: $\varphi(t)=f[a+t\dfrac{x-a}{\|x-a\|}]$在$t=0$处$m$阶可导, 故

$$\varphi(t)=\sum_{k=0}^{m}\frac{1}{k!}\varphi^{(k)}(0)t^k+0(t^m)\quad(t\to 0).$$

由

$$\begin{aligned}\frac{d}{dt}\varphi(t)&=\sum_{j=1}^{n}\frac{\partial f}{\partial x_j}[a+t\frac{x-a}{\|x-a\|}]\frac{(x_j-a_j)}{\|x-a\|}\\&=\frac{1}{\|x-a\|}\left(\sum_{j=1}^{n}(x_j-a_j)\frac{\partial}{\partial x_j}\right)f[a+t\frac{x-a}{\|x-a\|}],\end{aligned}$$

可得

$$\frac{d^k}{dt^k}\varphi(t)=\frac{1}{\|x-a\|^k}\left(\sum_{j=1}^{n}(x_j-a_j)\frac{\partial}{\partial x_j}\right)^k f[a+t\frac{x-a}{\|x-a\|}],$$

故

$$\varphi^{(k)}(0)=\frac{1}{\|x-a\|^k}\left(\sum_{j=1}^{n}(x_j-a_j)\frac{\partial}{\partial x_j}\right)^k f(a),\quad k=0,1,\cdots,m.$$

将$t=\|x-a\|$带入前式即得到结论. □

Taylor 公式用于近似计算 当$\|x-a\|$很小时,

$$f(x)\approx\sum_{k=0}^{m}\frac{1}{k!}\left(\sum_{j=1}^{n}(x_j-a_j)\frac{\partial}{\partial x_j}\right)^k f(a).$$

§14.11 极值

定义 14.7 设 n 元函数 f 在 a 的邻域上有定义. 若 $\exists\delta>0$, 使得 $\forall x\in B_\delta(a)\subset\mathbb{R}^n$, 成立 $f(x)\le f(a)$, 则称 $f(a)$ 是 f 的一个极大值, a 是 f 的一个极大值点；类似地, 可以定义 f 的极小值和极小值点； f 的极大值和极小值统称为 f 的极值, f 的极大值点和极小值点统称为 f 的极值点.

定理 14.19（极值点的必要条件） 若 a 是 n 元函数 f 的极值点, 并且 $Jf(a)$ 存在, 则 a 是 f 的一个驻点, 即 $Jf(a)=0$.

证: 单变量函数 $\varphi(t)=f(a_1,\cdots,a_{k-1},t,a_{k+1},\cdots,a_n)$ 在 a_k 处取得极值, 故 $\varphi'(a_k)=\dfrac{\partial f}{\partial x_k}(a)=0$. □

注记 14.19′ 驻点可能不是极值点. 例如, $f(x,y)=xy$ 以 $(0,0)$ 为驻点, 但 $(0,0)$ 并非是 f 的极值点.

例 1（P_{135}, 第 6 题） 解: “8”的交点是方程组的一组解, 否则与隐函数定理中解的唯一性相矛盾；函数 F 在“8”的每个“圆盘”中都至少有一个极值点, 这两个极值点便是方程组的两组解. 故方程组至少有三组解. □

定理 14.23（极值点的充分条件） 若 f 是 a 的邻域上的 C^2 函数, $Jf(a)=0$, 则

(1) 若 $Hf(a)>0$（即 $Hf(a)$ 正定）, 则 $f(a)$ 是 f 的严格极小值；

(2) 若 $Hf(a)<0$（即 $-Hf(a)$ 正定）, 则 $f(a)$ 是 f 的严格极大值；

(3) 其它情形时, 各种可能性都有.

证: (1) 单位球面 $\partial B_1(0)$ 是紧集 $\Rightarrow \min\limits_{\xi\in\partial B_1(0)}\xi' Hf(a)\xi=M>0$.

由 $$f(a+h)=f(a)+Jf(a)h+\frac{1}{2!}h'Hf(a)h+o(\|h\|^2)\quad(\|h\|\to0)$$

可得 $$\frac{f(a+h)-f(a)}{\|h\|^2}\geq\frac{1}{2}M+\frac{o(\|h\|^2)}{\|h\|^2}\quad(\|h\|\to 0).$$

故$\exists\delta>0$, 使得$\forall h\in B_\delta(0)\setminus\{0\}$, 成立$\dfrac{f(a+h)-f(a)}{\|h\|^2}\geq\dfrac{1}{3}M>0$. 这说明$f(a)$是$f$的严格极小值.

(2) 证明与(1)类似.

(3) $Hf(a)>0$和$Hf(a)<0$以外的情形, 易举例说明各种可能都有. □

例 2（最小二乘法） 设平面上n个彼此不同的点$\{(x_i,y_i)\in\mathbb{R}^2:1\leq i\leq n\}$不同时位于一条平行于$y$轴的直线上. 求一条直线$y=a_0x+b_0$使得误差$\sum\limits_{i=1}^n(a_0x_i+b_0-y_i)^2$最小.（在实际问题中, 就是求线性函数与测量数据误差最小）

解: 问题为求 2 元函数$F(a,b)=\sum\limits_{i=1}^n(ax_i+b-y_i)^2$的最小值.

令 $$\frac{\partial F}{\partial a}(a,b)=\sum_{i=1}^n 2(ax_i+b-y_i)x_i=0,$$
$$\frac{\partial F}{\partial b}(a,b)=\sum_{i=1}^n 2(ax_i+b-y_i)=0,$$

可得二元一次方程组

$$\begin{cases}\left(\sum\limits_{i=1}^n x_i^2\right)a+\left(\sum\limits_{i=1}^n x_i\right)b=\sum\limits_{i=1}^n x_iy_i\\ \left(\sum\limits_{i=1}^n x_i\right)a+nb=\sum\limits_{i=1}^n y_i\ .\end{cases}$$

因为n个点$\{(x_i,y_i)\in\mathbb{R}^2:\ 1\leq i\leq n\}$不同时位于一条平行于$y$轴的直线上, 故向量$\vec{x}=(x_1,x_2,\cdots,x_n)$与$\vec{1}=(1,1,\cdots,1)$不平行, 从而

$$\|\vec{1}\|^2\|\vec{x}\|^2-(<\vec{1},\vec{x}>)^2=n\left(\sum_{i=1}^n x_i^2\right)-\left(\sum_{i=1}^n x_i\right)^2>0.$$

这说明上述二元一次方程组有唯一的解(a_0,b_0), 它就是函数F的驻点.

此外, 向量$(a_0,-1,b_0)$与$(\sum\limits_{i=1}^n x_i,\sum\limits_{i=1}^n y_i,n)$和$(\sum\limits_{i=1}^n x_i^2,\sum\limits_{i=1}^n x_iy_i,\sum\limits_{i=1}^n x_i)$都正交.

易算出$HF(a_0,b_0)=\begin{pmatrix} 2\sum_{i=1}^{n}x_i^2, & 2\sum_{i=1}^{n}x_i \\ 2\sum_{i=1}^{n}x_i, & 2n \end{pmatrix}>0$, 故$F(a_0,b_0)$是函数$F$的最小值.

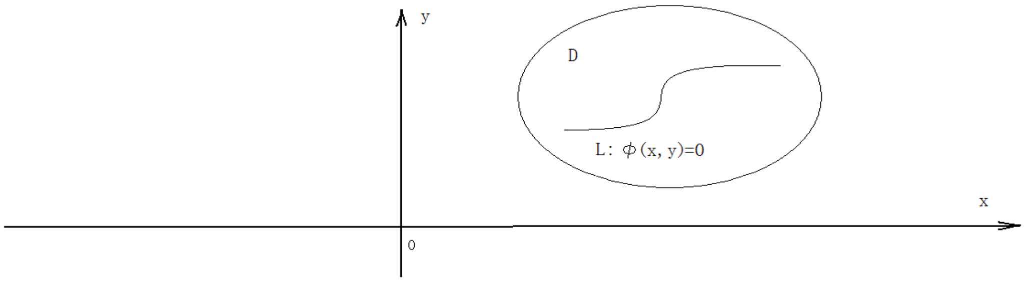
下面求出直线$y=a_0x+b_0$. 易看出, 点(x,y)位于直线$y=a_0x+b_0$上$\Leftrightarrow$ $(x,y,1)$与$(a_0,-1,b_0)$正交$\Leftrightarrow$ $(x,y,1)$, $(\sum_{i=1}^{n}x_i,\sum_{i=1}^{n}y_i,n)$, $(\sum_{i=1}^{n}x_i^2,\sum_{i=1}^{n}x_iy_i,\sum_{i=1}^{n}x_i)$共面$\Leftrightarrow$ $\det\begin{pmatrix} x, & y, & 1 \\ \sum_{i=1}^{n}x_i, & \sum_{i=1}^{n}y_i, & n \\ \sum_{i=1}^{n}x_i^2, & \sum_{i=1}^{n}x_iy_i, & \sum_{i=1}^{n}x_i \end{pmatrix}=0$. 这就是所求直线的方程. □

练习题 14.11（P_{167}） 2, 4.

§14.12 条件极值

本节内容可用于解决实际问题和得到一些有用的不等式.

条件极值问题 设 $D\subset\mathbb{R}^n\times\mathbb{R}^m$ 是开集，$f(x_1,\cdots,x_n,y_1,\cdots,y_m)$ 是 D 上的函数，$\varphi(x_1,\cdots,x_n,y_1,\cdots,y_m)$ 是从 D 到 $\mathbb{R}^m$ 的映射，$L=\{(x,y)\in D:\varphi(x,y)=0\}$. "求 $f\big|_L$ 的极值"这件事称为"函数 $f(x,y)$ 在条件 $\varphi(x,y)=0$ 下的条件极值问题".



显然, 若 $y=g(x)$ 是由 $\varphi(x,y)=0$ 所确定的隐映射并能解出来, 则条件极值问题便化为求 $f(x,g(x))$ 的极值(此路不通).

定理 14.25(Lagrange 乘数法) 设 $D\subset\mathbb{R}^n\times\mathbb{R}^m$ 是开集, $f(x,y)$ 是 D 上的 C^1 函数, $\varphi(x,y)$ 是从 D 到 $\mathbb{R}^m$ 的 C^1 映射, $(x_0,y_0)\in L=\{(x,y)\in D:\varphi(x,y)=0\}$, $\det J_y\varphi(x_0,y_0)\neq 0$, $\lambda_0=-J_yf(x_0,y_0)[J_y\varphi(x_0,y_0)]^{-1}$. 若 (x_0,y_0) 是 $f\big|_L$ 的极值点, 则 (x_0,y_0) 是辅助函数 $F(x,y)=f(x,y)+\lambda_0\varphi(x,y)$ 的驻点.

注记 14.25′(解条件极值问题的实际步骤) 求出方程组

$$\begin{cases} Jf(x,y)+\lambda J\varphi(x,y)=0 \\ \varphi(x,y)=0 \end{cases}$$

的解 (λ_0,x_0,y_0), 其中的 (x_0,y_0) 便可能是条件极值点(即 Lagrange 乘数法给出了条件极值点的必要条件).

证: 由隐映射定理, 存在从 $B_\varepsilon(x_0)\subset\mathbb{R}^n$ 到 $\mathbb{R}^m$ 的唯一的 C^1 映射 $y=g(x)$ 使得 $\varphi(x,g(x))\equiv 0$, $y_0=g(x_0)$. 故 x_0 是 $f(x,g(x))$ 的极值点, 从而

$$J_xf(x_0,y_0)+J_yf(x_0,y_0)Jg(x_0)=0.$$

因为

$$Jg(x_0)=-[J_y\varphi(x_0,y_0)]^{-1}J_x\varphi(x_0,y_0),$$

代入前式便得到

$$J_x f(x_0,y_0)-J_y f(x_0,y_0)[J_y\varphi(x_0,y_0)]^{-1}J_x\varphi(x_0,y_0)=0.$$

于是
$$\begin{cases}J_x f(x_0,y_0)+\lambda_0 J_x\varphi(x_0,y_0)=0\\ J_y f(x_0,y_0)+\lambda_0 J_y\varphi(x_0,y_0)=0,\end{cases}$$

此即 $JF(x_0,y_0)=Jf(x_0,y_0)+\lambda_0 J\varphi(x_0,y_0)=0$. □

定理 14.26(条件极值点的充分条件) 设 $D\subset\mathbb{R}^n\times\mathbb{R}^m$ 是开集, $f(x,y)$ 是 D 上的 C^2 函数, $\varphi(x,y)$ 是从 D 到 $\mathbb{R}^m$ 的 C^2 映射, (λ_0,x_0,y_0) 是方程组

$$\begin{cases}Jf(x,y)+\lambda J\varphi(x,y)=0\\ \qquad\varphi(x,y)=0\end{cases}$$

的解, $F(x,y)=f(x,y)+\lambda_0\varphi(x,y)$, 那么有如下结论

(1) 若 $HF(x_0,y_0)>0$ (即 $HF(x_0,y_0)$ 正定), 则 f 在 (x_0,y_0) 处取得严格条件极小值;

(2) 若 $HF(x_0,y_0)<0$ (即 $-HF(x_0,y_0)$ 正定), 则 f 在 (x_0,y_0) 处取得严格条件极大值;

(3) 其它情形时, 各种可能性都有.

证: 记 $z=(x,y)$. 对于 $z_0+h\in L=\{z\in D:\varphi(z)=0\}$, 若注意到 z_0 是辅助函数 $F(z)=f(z)+\lambda_0\varphi(z)$ 的驻点, 便得到

$$f(z_0+h)-f(z_0)=F(z_0+h)-F(z_0)=h'HF(z_0)h+o(\|h\|^2)\ \ (\|h\|\to 0).$$

由此即证明了(1)和(2). 通过举例可证明(3). □

例 将正数 M 分解成 n 个正数之和, 使其乘积最大.

解: 问题就是求 $f(x_1,x_2\cdots,x_n)=x_1x_2\cdots x_n\ (x_1>0,x_2>0,\cdots,x_n>0)$ 在条件 $x_1+x_2+\cdots+x_n=M$ 下的最大值.

令 $\varphi(x_1,x_2\cdots,x_n)=x_1+x_2+\cdots+x_n-M$, 解方程组

$$\begin{cases}Jf(x)+\lambda J\varphi(x)=0\\ \qquad\varphi(x)=0,\end{cases}$$

即 $$\begin{cases} f(x)\left(\frac{1}{x_1},\frac{1}{x_2},\cdots,\frac{1}{x_n}\right)+\lambda(1,1,\cdots,1)=(0,0,\cdots,0) \\ x_1+x_2+\cdots+x_n=M. \end{cases}$$

由第一个方程得 $x_1=x_2=\cdots=x_n$；由第二个方程得 $x_1=x_2=\cdots=x_n=\dfrac{M}{n}$. 根据题意，$f(x)$ 在条件 $\varphi(x)=0$ 下的最大值是存在的，因而这个最大值就是 $\left(\dfrac{M}{n}\right)^n$. 于是，当 $x_1>0, x_2>0,\cdots, x_n>0, x_1+x_2+\cdots+x_n=M$ 时，成立

$$x_1x_2\cdots x_n \le \left(\frac{M}{n}\right)^n = \left(\frac{x_1+x_2+\cdots+x_n}{n}\right)^n,$$

即 $$\sqrt[n]{x_1x_2\cdots x_n} \le \frac{x_1+x_2+\cdots+x_n}{n}. \square$$

练习题 14.12（P_{175}）　1(3), 2, 3.