多元函数微分学

§1、偏导数与全微分

方向导数 $\vec{e}=(\cos\alpha,\sin\alpha)$ 则 $\dfrac{\partial f}{\partial \vec{e}}=\lim\limits_{t\to 0}\dfrac{f(x+t\cos\alpha,\ y+t\sin\alpha)-f(x,y)}{t}$ 称为f在点(x,y)沿$\vec{e}$的方向导数

偏导数 $\dfrac{\partial f}{\partial x}$ 即$\vec{e}=(1,0)$ $\dfrac{\partial f}{\partial y}$ 即$\vec{e}=(0,1)$

全微分 若∃$A,B\in\mathbb{R}$ 使$\sqrt{(\Delta x)^2+(\Delta y)^2}\to 0$时，$f(x+\Delta x,y+\Delta y)-f(x)=A\Delta x+B\Delta y+o\left(\sqrt{(\Delta x)^2+(\Delta y)^2}\right)$ 成立，则称**可微**

且 $df(x_0,y_0)=Adx+Bdy$ 称为f在点(x_0,y_0)的**全微分**

$\sqrt{h^2+k^2}$ 也可写作 $\alpha h+\beta k$，其中 $\lim\alpha=\lim\beta=0$

定理1 某点(x_0,y_0)可微 ⇒ (x_0,y_0)连续

pf: $\lim f(x,y)-f(x_0,y_0)=\lim A(x-x_0)+B(y-y_0)+o\left(\sqrt{(x-x_0)^2+(y-y_0)^2}\right)=0$ □

定理2 可微 ⇒ (x_0,y_0)方向导数存在

pf: 表达式中令$\Delta x=t\cos\theta$，$\Delta y=t\sin\theta$ 则 $\dfrac{\partial f}{\partial \vec{e}}=\lim\left(A\cos\theta+B\sin\theta+o(1)\right)=A\cos\theta+B\sin\theta=\dfrac{\partial f}{\partial x}\cos\theta+\dfrac{\partial f}{\partial y}\sin\theta$ □

推论: $\dfrac{\partial f}{\partial(\cos\theta,\sin\theta)}=\dfrac{\partial f}{\partial x}\cos\theta+\dfrac{\partial f}{\partial y}\sin\theta$，若可微

定理3: 若$U(x_0,y_0)$各偏导数存在且$\dfrac{\partial f}{\partial x}(x,y)$与$\dfrac{\partial f}{\partial y}(x,y)$在$(x_0,y_0)$连续 ⇒ 可微

pf: 记$h=x-x_0$，$k=y-y_0$，$f(x_0+h,y_0+k)-f(x_0,y_0)=f(x_0+h,y_0+k)-f(x_0,y_0+k)+f(x_0,y_0+k)-f(x_0,y_0)$

$\overset{\text{拉中}}{=\!=}f'_x(x_0+\theta h,y_0+k)h+f'_y(x_0,y_0+\omega k)k$ $0<\theta,\omega<1$

由于f'_x与f'_y在(x_0,y_0)邻域连续，故$f'_x(x_0+\theta h,y_0+k)=f'_x(x_0,y_0)+\alpha$，$\alpha\to 0$

$f'_y(x_0,y_0+\omega k)=f'_y(x_0,y_0)+\beta$，$\beta\to 0$ 即 $f(x_0+h,y_0+k)-f(x_0,y_0)=f'_x(x_0,y_0)h+f'_y(x_0,y_0)k+\alpha h+\beta k$

⇒ f在(x_0,y_0)可微 □

综上有关系 各偏导存在且连续 ⇒ 函数可微 ⇒ 函数连续 ⇑ ⇒ 各方向导数存在

m元函数情形 **全微分**: $\dfrac{\partial f}{\partial \vec{e}}(x_0)=\sum\limits_{i=1}^{m}A_ie^i$，其中$\vec{e}=(e^1,e^2,\cdots,e^m)$ $A_i=\dfrac{\partial f}{\partial e^i}$

§2 复合函数的偏导数与全微分

设m个一元函数$\varphi^1(t),\cdots,\varphi^m(t)$在$t_0$处可导，$\varphi^1(t_0)=x_0^1$，$\cdots$ $\varphi^m(t_0)=x_0^m$，$F(t)=f(\varphi^1(t),\cdots,\varphi^m(t))$

$F(t)-F(t_0)=f(\varphi^1(t),\cdots,\varphi^m(t))-f(\varphi^1(t_0),\cdots,\varphi^m(t_0))=\sum\limits_{i=1}^{m}\dfrac{\partial f}{\partial x^i}(x_0)\left(\varphi^i(t)-\varphi^i(t_0)\right)+o\left(\sqrt{\sum\limits_{i=1}^{m}(\varphi^i(t)-\varphi^i(t_0))^2}\right)$

又$\varphi^i(t)-\varphi^i(t_0)=\dfrac{d\varphi^i}{dt}(t_0)(t-t_0)+o(t-t_0)$，$i=1,2,\cdots,m$，即得$F(t)-F(t_0)=\sum\limits_{i=1}^{m}\dfrac{\partial f}{\partial x^i}(x_0)\dfrac{d\varphi^i}{dt}(t_0)(t-t_0)+o(\|\Delta x\|)$

这就证明了$F'(t_0)=\sum\limits_{i=1}^{m}\dfrac{\partial f}{\partial x^i}(x_0)\dfrac{\partial\varphi^i}{\partial t}(t_0)$，写作$\dfrac{\partial F}{\partial t}=\sum\limits_{i=1}^{m}\dfrac{\partial f}{\partial x^i}\dfrac{\partial x^i}{\partial t}$，当$t=(t_0^1,t_0^2,\cdots,t_0^k)$时，偏导$\dfrac{\partial F}{\partial t^j}=\sum\limits_{i=1}^{m}\dfrac{\partial f}{\partial x^i}\dfrac{\partial x^i}{\partial t^j}$

微分表示的不变性 无论$x=(x^1,x^2,\cdots,x^m)$是自变元，还是依赖于别的变元$t=(t^1,\cdots,t^k)$的函数，$f(x)=f(x^1,\cdots,x^m)$全微分表示不变.

pf: $df(x)=df(x^1,x^2,\cdots,x^m)=\sum\limits_{j=1}^{k}\dfrac{\partial}{\partial t^j}f(x^1,\cdots,x^m)dt^j=\sum\limits_{j=1}^{k}\sum\limits_{i=1}^{m}\dfrac{\partial f}{\partial x^i}\dfrac{\partial x^i}{\partial t^j}dt^j=\sum\limits_{i=1}^{m}\dfrac{\partial f}{\partial x^i}\sum\limits_{j=1}^{k}\dfrac{\partial x^i}{\partial t^j}dt^j=\sum\limits_{i=1}^{m}\dfrac{\partial f}{\partial x^i}d(x^i(t))$

$=\sum\limits_{i=1}^{m}\dfrac{\partial f}{\partial x^i}dx^i$ □

定理: 设u,v可微，$\lambda\in\mathbb{R}$，则 $du+dv=d(u+v)$，$d(\lambda u)=\lambda du$，$duv=(du)v+(dv)u$，$d\dfrac{u}{v}=\dfrac{du\cdot v-dv\cdot u}{v^2}$ $(v\neq 0)$

§3. 高阶偏导数

规定 $\dfrac{\partial}{\partial x^i}\left(\dfrac{\partial f}{\partial x^j}\right)=\dfrac{\partial^2 f}{\partial x^i\partial x^j}\triangleq f''_{x^ix^j}(x)$ 或 $f_{x^ix^j}$

定理1 若$f(x,y)$两个二阶偏导数$f_{xy}(x,y)$和$f_{yx}(x,y)$邻近存在且在(x_0,y_0)连续，则$f_{xy}(x_0,y_0)=f_{yx}(x_0,y_0)$

pf: 考虑 $\varphi(x)=f(x,y_0+k)-f(x,y_0)$ 与 $\psi(y)=f(x_0+h,y)-f(x_0,y)$，由拉格朗日中值定理，∃$0<\theta_1,\theta_2,\theta_3,\theta_4<1$

$\varphi(x_0+h)-\varphi(x_0)=\varphi'(x_0+\theta_1 h)h=[f_x(x_0+\theta_1 h,y_0+k)-f_x(x_0+\theta_1 h,y_0)]h=f_{xy}(x_0+\theta_1 h,y_0+\theta_2 k)hk$

$\psi(y_0+k)-\psi(y_0)=\psi'(y_0+\theta_3 k)k=[f_y(x_0+h,y_0+\theta_3 k)-f_y(x_0,y_0+\theta_3 k)]k=f_{yx}(x_0+\theta_4 h,y_0+\theta_3 k)hk$

容易验证 $\varphi(x_0+h)-\varphi(x_0)=\psi(y_0+k)-\psi(y_0)$ ，再让 $(h,k)\to(0,0)$ 取极限，利用 f_{xy} 与 f_{yx} 连续性即得 $f_{xy}(x_0,y_0)=f_{yx}(x_0,y_0)$ □

定义： 设 Ω 是 $\mathbb{R}^m$ 中的开集，我们约定 C^r 表示在 Ω 上 r 阶连续可微的函数组成的集合，C^0 表示连续函数构成集合 $(r\geq1)$

定理2： Ω 是 $\mathbb{R}^m$ 中的开集，$f\in C^r(\Omega)$，则函数 f 的 k 阶 $(2\leq k\leq r)$ 混合偏导数与求导顺序无关 □

例 $u=f(x(\xi,\eta),y(\xi,\eta))$，求 $\frac{\partial^2 u}{\partial\xi\partial\eta}$

解：$\frac{\partial^2 u}{\partial\xi\partial\eta}=\frac{\partial}{\partial\xi}\left(\frac{\partial u}{\partial\eta}\right)=\frac{\partial}{\partial\xi}\left(\frac{\partial f}{\partial x}\frac{\partial x}{\partial\eta}+\frac{\partial f}{\partial y}\frac{\partial y}{\partial\eta}\right)=\frac{\partial}{\partial\xi}\left(\frac{\partial f}{\partial x}\frac{\partial x}{\partial\eta}\right)+\frac{\partial}{\partial\xi}\left(\frac{\partial f}{\partial y}\frac{\partial y}{\partial\eta}\right)=\frac{\partial}{\partial\xi}\left(\frac{\partial f}{\partial x}\right)\frac{\partial x}{\partial\eta}+\frac{\partial^2 x}{\partial\xi\partial\eta}\frac{\partial f}{\partial x}+\frac{\partial}{\partial\xi}\left(\frac{\partial f}{\partial y}\right)\frac{\partial y}{\partial\eta}+\frac{\partial^2 y}{\partial\xi\partial\eta}\frac{\partial f}{\partial y}$

$=\left(\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right)\frac{\partial x}{\partial\xi}+\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)\frac{\partial y}{\partial\xi}\right)\frac{\partial x}{\partial\eta}+\frac{\partial^2 x}{\partial\xi\partial\eta}\frac{\partial f}{\partial x}+\left(\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)\frac{\partial x}{\partial\xi}+\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right)\frac{\partial y}{\partial\xi}\right)\frac{\partial y}{\partial\eta}+\frac{\partial^2 y}{\partial\xi\partial\eta}\frac{\partial f}{\partial y}$

$=\frac{\partial^2 f}{\partial x^2}\frac{\partial x}{\partial\xi}\frac{\partial x}{\partial\eta}+\frac{\partial^2 f}{\partial y^2}\frac{\partial y}{\partial\xi}\frac{\partial y}{\partial\eta}+\frac{\partial^2 f}{\partial x\partial y}\left(\frac{\partial y}{\partial\xi}\frac{\partial x}{\partial\eta}+\frac{\partial x}{\partial\xi}\frac{\partial y}{\partial\eta}\right)+\frac{\partial^2 x}{\partial\xi\partial\eta}\frac{\partial f}{\partial x}+\frac{\partial^2 y}{\partial\xi\partial\eta}\frac{\partial f}{\partial y}$ □

§4. 有限增量公式与泰勒公式

定理1. 设 D 是 $\mathbb{R}^m$ 一个开集，$a=(a_1,\cdots a_m)$ 与 $a+h=(a_1+h_1,\cdots,a_m+h_m)$ 是 $\bar{D}$ 中两点，且联结该两点的开线段 $(a,a+h)\subset D$

若 f 在 $\bar{D}$ 连续，D 可微，则 $\exists\,0<\theta<1$，$f(a+h)=f(a)+\sum_{i=1}^m\frac{\partial f}{\partial x_i}(a+\theta h)h_i$

pf：考察函数 $\varphi(t)=f(a+th)$，$\varphi'(t)=\sum_{i=1}^m\frac{\partial f}{\partial x_i}\frac{\partial x_i}{\partial t}=\sum_{i=1}^m\frac{\partial f}{\partial x_i}(a+th)h_i$，注意到 $\varphi(t)$ 在 $[0,1]$ 连续，$(0,1)$ 可微

故由一元函数有限增量公式，$\exists\,0<\theta<1$，$\varphi(1)=\varphi(0)+\varphi'(\theta)(1-0)$ 从而 $f(a+h)=f(a)+\sum_{i=1}^m\frac{\partial f}{\partial x_i}(a+\theta h)h_i$ □

定理2 设 $a,a+h$ 是 $\mathbb{R}^m$ 中两点，D 是包含 $[a,a+h]$ 的一个开集，若 $f(x)=f(x_1,x_2,\cdots x_m)$ 在 D 连续可微，

则 $f(a+h)=f(a)+\sum_{i=1}^m\left(\int_0^1\frac{\partial f}{\partial x_i}(a+th)dt\right)h_i$ （多元函数的有限增量公式）

pf：φ 同上，由牛顿-莱布尼茨，$\varphi(1)=\varphi(0)+\int_0^1\varphi'(t)dt$，又 $\varphi'(t)=\sum_{i=1}^m\frac{\partial f}{\partial x_i}(a+th)h_i$ □

定理3 设 f 在开区域 $\Omega\subset\mathbb{R}^m$ 可微，若 $\frac{\partial f}{\partial x_1}(x)=\cdots=\frac{\partial f}{\partial x_m}(x)=0$，$\forall x\in\Omega$，则 f 在 Ω 上恒常

pf：先证 Ω 是开球情形，由定理2易证，再由连通性，$\exists\gamma:[0,1]\to\Omega$，$\gamma(0)=a$，$\gamma(1)=b$，且 γ 连续，记 $\alpha=\sup\{t\mid t\in[0,1],f(\gamma(t))=f(a)\}$

由 f,α 连续性 $\Rightarrow f(\gamma(\alpha))=f(a)$，若 $\alpha<1$，由连续性与开集，$\exists U(\alpha,\delta)\subset\Omega$，利用定理2找到更大 α，矛盾！故 $\alpha=1$，即 $f(a)=f(b)$ □

定理4： 设 D 是 $\mathbb{R}^m$ 中开集，若 $f\in C^{n+1}(D)$，$[a,a+h]\subset D$，则 $f(a+h)=T_n+R_{n+1}$，$T_n=\sum_{p=0}^n\frac{1}{p!}\left(h_1\frac{\partial}{\partial x_1}+\cdots+h_m\frac{\partial}{\partial x_m}\right)^p f(a)$

余项 R_{n+1} 可以为拉格朗日形式：$\frac{1}{(n+1)!}\left(h_1\frac{\partial}{\partial x_1}+\cdots+h_m\frac{\partial}{\partial x_m}\right)^{n+1}f(a+\theta h)$ $(0<\theta<1)$ 或积分形式 $\frac{1}{n!}\int_0^1(1-t)^n\left(h_1\frac{\partial}{\partial x_1}+\cdots+h_m\frac{\partial}{\partial x_m}\right)^{n+1}f(a+th)dt$

或 peano 形式：$o(\|h\|^n)$ （Taylor 公式）

pf：利用 $\varphi(t)=f(a+th)$，对 $\varphi(t)$ 用一元函数情形，注意到 $\varphi^{(k)}(t)=\left(h_1\frac{\partial}{\partial x_1}+\cdots+h_m\frac{\partial}{\partial x_m}\right)^k f(a+th)$ □

§5. 隐函数定理

隐函数： $D\subset\mathbb{R}$，$E\subset\mathbb{R}$，$F(x,y)$ 定义在 $D\times E$ 上，若 $\forall x\in D$，$\exists$ 唯一 $y\in E$ 使 $F(x,y)=0$，则我们称 $F(x,y)=0$ 确定了由 D 到 E 的一个隐函数

定理1： 设 $F(x,y)$ 在包含 (x_0,y_0) 的一个开集 Ω 上连续可微，并满足条件 $F(x_0,y_0)=0$，$\frac{\partial F}{\partial y}(x_0,y_0)\neq0$，则存在以 (x_0,y_0) 为中心的开方块

$D\times E\subset\Omega$，$D=(x_0-\delta,x_0+\delta)$，$E=(y_0-\eta,y_0+\eta)$ 使得：(1) $\forall x\in D$，恰存在唯一 $y\in E$ 满足 $F(x,y)=0$

(2) $y=f(x)$ 在 D 连续可微，$\frac{dy}{dx}=-\frac{\frac{\partial F}{\partial x}(x,y)}{\frac{\partial F}{\partial y}(x,y)}$ 特殊的隐函数定理

pf：不妨 $\frac{\partial F}{\partial y}(x_0,y_0)>0$，则由 $\frac{\partial F}{\partial y}(x,y)$ 连续性，存在 (x_0,y_0) 附近区间 $[x_0-\gamma,x_0+\gamma]\times[y_0-\eta,y_0+\eta]\subset\Omega$，使 $\frac{\partial F}{\partial y}(x,y)$ 在上面恒 >0.

考虑 $\psi(y)=F(x_0,y)$，$\frac{d\psi}{dy}=\frac{\partial F}{\partial y}>0$ $\Rightarrow$ 关于 y 单增 $\Rightarrow F(x_0,y_0-\eta)<0<F(x_0,y_0+\eta)$，再考察 x 的函数 $F(x,y_0-\eta)$ 与

$F(x,y_0+\eta)$，由连续性 $\Rightarrow\exists\,\delta>0$，$\forall x\in(x_0-\delta,x_0+\delta)$，$F(x,y_0-\eta)<0<F(x,y_0+\eta)$，下验证 $D=(x_0-\delta,x_0+\delta)$ 与 $E=(y_0-\eta,y_0+\eta)$

为所求

① 唯一性：$\forall x_1\in D$，考察 $F(x_1,y)=\varphi(y)$，$\varphi(y_0-\eta)<0<\varphi(y_0+\eta)$ 且 $\varphi(y)\nearrow$，故 $\exists$ 唯一 y_1，$F(x_1,y_1)=0$ 零存+单调卡出唯一零点

② 连续性：$\forall x_1\in D$，$y_1=f(x_1)$，任意充分小 ε，使 $(y_1-\varepsilon,y_1+\varepsilon)\subset(y_0-\eta,y_0+\eta)$，类似可得 $F(x_1,y_1-\varepsilon)<0<F(x_1,y_1+\varepsilon)$，又由 $F(x,y_1\pm\varepsilon)$

连续性，存在 $\sigma>0$，$\forall x\in(x_1-\sigma,x_1+\sigma)$，上式依旧成立，则 $f(U(x_1,\sigma))\subset(y_1-\varepsilon,y_1+\varepsilon)$ $\Rightarrow$ f 连续 卡大小证连续

③ 可微性：考虑差商 $\frac{f(x+h)-f(x)}{h}$，令 $k=f(x+h)-f(x)$，$y=f(x)$，由二元函数有限增量公式，$0=F(x+h,y+k)-F(x,y)$ 增量公式刚推微分表达式

$$=\frac{\partial F}{\partial x}(x+\theta h,y+\theta k)h+\frac{\partial F}{\partial y}(x+\theta h,y+\theta k)k \Rightarrow \frac{k}{h}=-\frac{\frac{\partial F}{\partial x}(x+\theta h,y+\theta k)}{\frac{\partial F}{\partial y}(x+\theta h,y+\theta k)}$$，取 $h\to0$，则 $f'(x)=-\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial y}}(x,y)$ □

推论：若 $F(x,y)$ 在开集 Ω 上是 r 阶连续可微的，则 $y=f(x)$ 在 D 上也是 r 阶连续可微的 在证 C^r 微分同胚推论时用到

注记：1. 二元可平行推广至 m 元情形：$F(x_0^1,\cdots x_0^m,y_0)=0$ 且 $\frac{\partial F}{\partial y}(x_0^1,\cdots x_0^m,y)\neq0$，$\frac{\partial y}{\partial x^i}=-\frac{\frac{\partial F}{\partial x^i}}{\frac{\partial F}{\partial y}}(x^1,\cdots,x^m,y)$

2. 已知 f 可微后，$F(x,f(x))=0$，记 $\varphi(x)=F(x,f(x))=F(x,y)$，则 $0=\frac{d\varphi}{dx}=\frac{\partial F}{\partial x}\frac{\partial x}{\partial x}+\frac{\partial F}{\partial y}\frac{\partial y}{\partial x}=\frac{\partial F}{\partial x}+f'(x)\frac{\partial F}{\partial y}$，即证.

雅可比矩阵 $F^i:(x,y^1,\cdots,y^p)\to\mathbb{R}$，给定方程组 $F^i(x,y^1,\cdots y^p)=0$，$i=1,2,\cdots,p$，F^i 在开集 Ω 上连续可微

称 $\begin{bmatrix}\frac{\partial F^1}{\partial y^1} & \cdots & \frac{\partial F^1}{\partial y^p}\\ \vdots & & \vdots \\ \frac{\partial F^p}{\partial y^1} & \cdots & \frac{\partial F^p}{\partial y^p}\end{bmatrix}$ 为函数 $F^1,\cdots,F^p$ 对变元 $y^1,\cdots,y^p$ 的雅可比矩阵，其 det 记为 $\frac{\partial(F^1,\cdots F^p)}{\partial(y^1,\cdots,y^p)}$

定理2：设 $F^1(x,y^1,\cdots y^p),\cdots,F^p(x,y^1,\cdots y^p)$ 在包含 $(x_0,y_0^1,\cdots,y_0^p)$ 的一个开集上连续可微并满足条件

(1) $F^i(x_0,y_0^1,\cdots y_0^p)=0$，$(i=1,2,\cdots p)$ (2) $\frac{\partial(F_1,\cdots F_p)}{\partial(y^1,\cdots,y^p)}(x_0,y_0^1,\cdots,y_0^p)\neq0$

则存在以 $(x_0,y_0^1,\cdots y_0^p)$ 为中心的开方块 $D\times(E^1\times\cdots\times E^p)\subset\Omega$ 使得

(1) $\forall x\in D$，恰存在一组 $(y^1,y^2,\cdots y^p)\in E^1\times\cdots\times E^p$ 使 $F^i(x,y^1,\cdots,y^p)=0$，$i=1,2\cdots p$

(2) 考虑该方程组确定的由 D 分别到 $E^1,\cdots E^p$ 的函数 $y^i=f^i(x)$，$i=1,\cdots p$，则 $f^i(x)$ 连续可微

它们的导数可通过解线性方程组求得：$\frac{\partial F^i}{\partial x}+\sum_{j=1}^{p}\frac{\partial F^i}{\partial y^j}\frac{df^j(x)}{dx}=0$，$i=1,2,\cdots,p$.

pf：推广即可，注意到 $\frac{\partial(F^1,\cdots F^p)}{\partial(y^1,\cdots y^p)}(x_0,y_0^1\cdots y_0^p)\neq0\Rightarrow(x_0,y_0^1,\cdots y_0^p)$ 附近 $\frac{\partial(F^1\cdots F^p)}{\partial(y^1\cdots y^p)}\neq0\Rightarrow$ 微分方程组系数矩阵满秩 $\Rightarrow$ 可解. □

§6. 线性映射

矩阵的范数 定义 $L(\mathbb{R}^m,\mathbb{R}^p)$ 表示全体 $p\times m$ 矩阵构成的集合. $L=\bigsqcup_{m,p\in\mathbb{N}}L(\mathbb{R}^m,\mathbb{R}^p)$ 为全体矩阵的集合

考虑实值函数 $N:L\to\mathbb{R}$，称为线性映射的范数(或矩阵的范数)，如果 (i) $\forall A\in L$，$N(A)\geq0$ 且 $N(A)=0$ 蕴含 $A=0$

(ii) $\forall A\in L$，$\lambda\in\mathbb{R}$，$N(\lambda A)=|\lambda|N(A)$ (iii) $\forall A,B\in L$，A,B 可加，$N(A+B)\leq N(A)+N(B)$

(iv) $\forall A,B\in L$，A,B 可乘. $N(AB)\leq N(A)N(B)$

例子 对 $A=\{a_{ij}\}_{i=1\cdots m,j=1\cdots p}$，$|A|=\max_{1\leq j\leq p}\{\sum_{i=1}^{m}|a_{ij}|\}$，即取各行元素绝对值之和最大值，$\|A\|=\sqrt{\sum_{i=1}^{m}\sum_{j=1}^{p}a_{ij}^2}$

设 $x\in\mathbb{R}^m$，$A\in L(\mathbb{R}^m,\mathbb{R}^p)$，则 $|Ax|\leq|A||x|$，$\|Ax\|\leq\|A\|\|x\|$

§7. 向量值函数的微分

微分 设 G 是 $\mathbb{R}^m$ 的一个开集，$f:G\to\mathbb{R}^p$ 是一个向量值函数，$x_0\in G$，若存在一个线性映射 $A\in L(\mathbb{R}^m,\mathbb{R}^p)$ 使 $\lim_{h\to0}\frac{|f(x_0+h)-f(x_0)-Ah|}{|h|}=0$

那么我们说向量值函数 f 在点 x_0 处可微分 若将 $|\cdot|$ 范数替为 $\|\cdot\|$ 或其它范数均可，因为不同范数在 $\mathbb{R}^m$ 上等价.

唯一性：$\forall\varepsilon>0$，$|(B-A)h|=\frac{1}{\varepsilon}|(B-A)(\varepsilon h)|\leq|h|\left(\frac{|f(x_0+\varepsilon h)-f(x_0)-A(\varepsilon h)|}{|\varepsilon h|}+\frac{|f(x_0+\varepsilon h)-f(x_0)-B(\varepsilon h)|}{|\varepsilon h|}\right)$，令 $\varepsilon\to0\Rightarrow(B-A)h=0$，$\forall h\Rightarrow B=A$

从而我们将唯一满足条件的线性映射 A 叫作向量值函数 f 在点 x_0 的微分，记为 $Df(x_0)=A$

定理1：G 是 $\mathbb{R}^m$ 中一个开集，$f:G\to\mathbb{R}^p$ 是一个映射，$x_0\in G$，若 f 在点 x_0 可微则存在点 x_0 的邻域 U 和正实数 γ，使

$|f(x)-f(x_0)|\leq\gamma|x-x_0|$，$\forall x\in U$，由此得 f 在 x_0 可微 $\Rightarrow$ f 在 x_0 连续.

pf: $|f(x)-f(x_0)| \le |f(x)-f(x_0)-A(x-x_0)| + |A(x-x_0)| \le \varepsilon|x-x_0| + |A||x-x_0| = (\varepsilon+|A|)|x-x_0|$ □

定理2: G是$\mathbb{R}^m$中一个开集，$f: G\to\mathbb{R}^p$与$g: G\to\mathbb{R}^p$，$\lambda\in\mathbb{R}$，f与g在x_0处均可微，则$f+g$可微，λf可微，

且 $D(f+g)(x_0) = Df(x_0) + Dg(x_0)$，$D(\lambda f)(x_0) = \lambda Df(x_0)$ □ （微分的线性）

定理3: 设G是$\mathbb{R}^m$的开集，H是$\mathbb{R}^p$开集，$f: G\to\mathbb{R}^p$，$g: H\to\mathbb{R}^q$，$f(G)\subset H$。若f在点x_0可微，g在点$f(x_0)\in H$可微

则 $g\circ f: G\to\mathbb{R}^q$ 在x_0可微，且 $D(g\circ f)(x_0) = Dg(f(x_0))\,Df(x_0)$ □ （链式法则）

定理4: 向量值函数$f(x) = (f_1(x),\cdots,f_p(x))$在点$x_0$可微的充要条件是：它的每个分量$f_i(x)$都在点$x_0$处可微，当这条件满足时

便有 $Df(x_0) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_p} \\ \vdots & & \vdots \\ \frac{\partial f_p}{\partial x_1} & \cdots & \frac{\partial f_p}{\partial x_p}\end{bmatrix}$ □ （Jacobi矩阵的显式表示）

定理5 设G是$\mathbb{R}^m$中的开集，$f: G\to\mathbb{R}^p$是在G上可微的向量值函数。又设a,b是G中2个点，满足$[a,b]\subset G$，则存在$c\in(a,b)$

使 $|f(b)-f(a)| \le |Df(c)||b-a|$ （有限增量估计）

pf: 考察向量值函数$f(x)$的各个分量$f_1(x),\cdots,f_p(x)$，不妨设$|f_k(b)-f_k(a)| = \max\limits_{1\le i\le p}|f_i(b)-f_i(a)| = |f(a)-f(b)|$

对$f_k(x)$用有限增量定理，存在$c\in(a,b)$，使$f_k(b)-f_k(a) = \sum\limits_{i=1}^m \frac{\partial f_k}{\partial x}(c)(b_i-a_i)$，于是 $|f_k(b)-f_k(a)| \le$

$\sum\limits_{i=1}^m \left|\frac{\partial f_k}{\partial x}(c)\right||b_i-a_i| \le \sum\limits_{i=1}^m \left|\frac{\partial f_k}{\partial x}(c)\right|\cdot\max\limits_{1\le i\le m}\{|b_i-a_i|\} \le \max\limits_{1\le j\le p}\left\{\sum\limits_{i=1}^m\left|\frac{\partial f_j}{\partial x_i}(c)\right|\right\}\cdot|b-a| = |Df(c)||b-a|$ □

note: 将|·|换为‖·‖后也成立：考察函数 $g(x) = \sum\limits_{j=1}^p (f_j(b)-f_j(a))f_j(x)$，由 $g(b)-g(a) = \sum\limits_{i=1}^m \frac{\partial g}{\partial x_i}(c)(b_i-a_i)$，$c\in(a,b)$

即 $\sum\limits_{j=1}^p (f_j(b)-f_j(a))^2 = \sum\limits_{j=1}^p\left((f_j(b)-f_j(a))\cdot\sum\limits_{i=1}^m\frac{\partial f_j}{\partial x_i}(c)(b_i-a_i)\right) \overset{\text{Cauchy}}{\le} \sqrt{\sum\limits_{j=1}^p(f_j(b)-f_j(a))^2}\cdot\sqrt{\sum\limits_{j=1}^p\left(\sum\limits_{i=1}^m\frac{\partial f_j}{\partial x_i}(c)(b_i-a_i)\right)}$

$\overset{\text{Cauchy}}{\le} \|f(b)-f(a)\|\cdot\sqrt{\sum\limits_{j=1}^p\sum\limits_{i=1}^m\left(\frac{\partial f_j}{\partial x_i}(c)\right)^2}\sqrt{\sum\limits_{i=1}^m(b_i-a_i)^2} = \|f(b)-f(a)\|\cdot\|Df(c)\|\cdot\|b-a\|$ 整理即证 □

定理6: 设Ω是$\mathbb{R}^{m+p} = \mathbb{R}^m\times\mathbb{R}^p$中的一个开集，$F: \Omega\to\mathbb{R}^q$是一个映射，$(x_0,y_0)\in\Omega$。若$F(x,y)$在点$(x_0,y_0)$（作为依赖$m+p$个变元的向量值函）

是可微分的，那么它在这点对变元x和y的偏微分均存在，并且

$D_xF(x_0,y_0)h = DF(x_0,y_0)(h,0)$，$\forall h\in\mathbb{R}^m$，$D_yF(x_0,y_0)k = DF(x_0,y_0)(0,k)$，$\forall k\in\mathbb{R}^p$

pf: 定义A_1与A_2：$A_1h = A(h,0)$，$\forall h\in\mathbb{R}^m$；$A_2k = A(0,k)$，$\forall k\in\mathbb{R}^p$，易验证$A_1 = D_xF$，$A_2 = D_yF$。□

注: $DF(x_0,y_0) = \left[\frac{\partial}{\partial x}\text{部分} \quad \frac{\partial}{\partial y}\text{部分}\right] = \left[D_xF(x_0,y_0) \quad D_yF(x_0,y_0)\right]$ （分块矩阵）

§8. 逆映射定理

微分同胚 设G，H是$\mathbb{R}^m$中的两个开集，$\varphi: G\to H$是双射，且φ与φ^{-1}均为C^r映射，那么称φ是由G到H的一个C^r同胚

局部微分同胚 设Ω是$\mathbb{R}^m$一个开集，$f: \Omega\to\mathbb{R}^m$是一个C^r映射，$a\in\Omega$，若存在包含点a的开集U和包含点$b=f(a)$的开集V

使f限制在U上是从U到V的一个C^r同胚，那么我们说f在a点是局部C^r同胚。

逆映射定理 设Ω是$\mathbb{R}^m$中的一个开集，$f: \Omega\to\mathbb{R}^m$是一个连续可微映射，$a\in\Omega$，如果$\det Df(a)\ne 0$，则f在a点上是局部微分同胚。

pf: 考虑 $F: \Omega\times\mathbb{R}^m\to\mathbb{R}^m$，$F(x,y) = f(x)-y$，$\forall (x,y)\in\Omega\times\mathbb{R}^m$ 则显F连续可微且$F(a,b) = f(a)-b = 0$ $\det D_xF(a,b) = \det Df(a)\ne 0$

对$F(x,y)=0$应用一般隐函数定理，则$\exists$以a为中心的开方块$W\subset\Omega$与以点b为中心的开方块$V\subset\mathbb{R}^m$使 (1) $\forall y\in V$，$\exists$唯一$x\in W$

满足$F(x,y) = f(x)-y=0$ 因而方程$f(x)-y=0$定义了$V\to W$：$x = g(y)$ (2) $V\to W\subset\mathbb{R}^m$连续可微，且$Dg(y) = -(D_xF)^{-1}D_yF$

$=(Df(x))^{-1}$。考虑$U = \{x\in W \mid f(x)\in V\}$，因为$W$与$V$均为开集，$f$是连续映射，所以$\forall x_0\in U$，$\exists\delta>0$使$U(x_0,\delta)\subset W$

且$f(U(x_0,\delta))\subset U(f(x_0),\varepsilon)\subset V$ $\Rightarrow U(x_0,\delta)\subset U$ $\Rightarrow U$为开集。故f在a点为局部微分同胚 □

推论 将定理1中的连续可微替为C^r连续可微，可推出f是C^r微分同胚

定理2 设Ω是$\mathbb{R}^m$中的开集，$f: \Omega\to\mathbb{R}^m$是一个连续可微映射，若$\det Df(x)\ne 0$，$\forall x\in\Omega$

则f把Ω的任何开子集仍映为$\mathbb{R}^m$的开集

pf: $\forall H\subset\Omega$, H为开集. 由逆映射定理 $\forall a\in H$, $\exists a$的邻域U, 使f在a为局部微分同胚, 且$f(U)=V$为开集.
而$b\in V$, 故$f(H)$中任一元素均为$f(H)$内点 $\Rightarrow f(H)$为开集. □

定理3 设G是$\mathbb{R}^m$开集, $f: G\to\mathbb{R}^m$是r阶连续可微的映射 $(r\geq 1)$. 如果(1) f单 (2) $\det Df\neq 0$, 则f是C^r微分同胚

pf: 由(1)知f是$G\to f(G)$的双射, 因此$\exists f^{-1}: f(G)\to G$. 又由逆映射定理及推论, 在每点$b\in f(G)$邻近, f^{-1}都是
r阶可微的. 从而我们说f^{-1}在$f(G)$是C^r的. □

§9. 多元函数的极值

定理1 数值函数$f(x)=f(x_1,\cdots,x_m)$在点$a=(a_1,\cdots,a_m)$附近有定义, 若f在a处取得极值, 必有$\frac{\partial f}{\partial x_i}(a)=0$, $i=1,2,\cdots m$.

pf: 设$A_i=\frac{\partial f}{\partial x_i}(a)$, 则$f(a+h)-f(a)=\sum_{i=1}^m A_ih_i+o(\|h\|)$, 取$h_i=A_i\varepsilon$, $\varepsilon>0$充分小, 则$f(a+h)-f(a)=\left(\sum_{i=1}^m A_i^2\right)\varepsilon+o(\|h\|)$
若$\sum_{i=1}^m A_i^2\neq 0$, 则当$\varepsilon\to 0$, $f(a+h)-f(a)=\left(\sum_{i=1}^m A_i^2\right)\varepsilon+o(\varepsilon)$ 符号由其主部$\pm\left(\sum_{i=1}^m A_i^2\right)\varepsilon$决定 $\Rightarrow f(a+h)>f(a)>f(a-h)$ 矛盾! □

临界点: 设f在a处可微, 若f在a点的任一偏导数均为0, 则称a为f的一个临界点.

引理2 设$\xi\in\mathbb{R}^m$, $\xi=(\xi_1,\cdots,\xi_m)$ 且二次型$Q(\xi)=\sum_{i,j=1}^m c_{ij}\xi_i\xi_j$ 是正定的 $(c_{ij}=c_{ji})$, 则$\exists\sigma>0$, $Q(\xi)\geq\sigma\|\xi\|^2$, $\forall\xi\in\mathbb{R}^m$

pf: 考虑$S=\{\xi\in\mathbb{R}^m \mid \|\xi\|=1\}$, S是有界闭集, 则$Q(\xi)$在S上取得最小值σ. 又Q正定, 故$\sigma>0$. 易知σ为所求 □

黑塞方阵 设函数f在点$a=(a_1,a_2,\cdots,a_m)$至少是二阶连续可微的, 我们称$H_f(a)=\left[\frac{\partial^2 f}{\partial x_i\partial x_j}(a)\right]_{m\times m}=\begin{bmatrix}\frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_2\partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_m\partial x_1}\\ \frac{\partial^2 f}{\partial x_1\partial x_2} & & & \\ & & \ddots & \\ \frac{\partial^2 f}{\partial x_1\partial x_m} & & & \frac{\partial^2 f}{\partial x_m\partial x_m}\end{bmatrix}$
为黑塞(Hasse)方阵

定理2 设$f(x)=f(x_1,x_2,\cdots,x_m)$在点$a=(a_1,a_2,\cdots,a_m)$邻近至少是$C^2$, a是f的一个临界点, 若f在点a的黑塞方阵$H_f(a)$
是正定(负定)的, 那么f在点a处取得严格极小值(极大值)

pf: 记$A_{ij}=\frac{\partial^2 f}{\partial x_i\partial x_j}(a)$, $i,j=1,2\cdots m$, 由引理, $\exists\sigma>0$, 使$\sum_{i,j=1}^m A_{ij}h_ih_j\geq\sigma\|h\|^2$, $\forall h\in\mathbb{R}^m$. 利用带peano余项的Taylor公式
展开2项, 我们有$f(a+h)-f(a)=(\text{一阶项}=0)+\frac{1}{2}\sum_{i,j=1}^m A_{ij}h_ih_j+o(\|h\|^2)\geq\left(\frac{1}{2}\sigma+o(1)\right)\|h\|^2\geq\frac{1}{4}\sigma\|h\|^2>0$ $(0<\|h\|<\delta)$ □
类似可证若$H_f(a)$负定, 那么f在点a处取严格极大值. □

注: 若$Q(\xi)=\sum_{i,j=1}^m A_{ij}\xi_i\xi_j$是不定的, 即存在$\xi',\xi''$, $Q(\xi')>0>Q(\xi'')$, 则$f(a+\varepsilon\xi'')-f(a)<0$, $f(a+\varepsilon\xi')-f(a)>0$, 则$f$在$a$处不取极值

定理3 考察$f(x)=f(x_1,\cdots,x_{m+p})$, 变元满足约束$g_i(x)=g_i(x_1,\cdots,x_{m+p})=0$, $i=1,2\cdots p$ 且 $\operatorname{rank}\begin{bmatrix}\frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_p}{\partial x_1}\\ \frac{\partial g_p}{\partial x_1} & \cdots & \frac{\partial g_p}{\partial x_p}\end{bmatrix}=p$ ($f,g_i\in C^N$, N充大)
并在$a=(a_1,\cdots,a_{m+p})$处取得极值, 则存在$(\lambda_1,\cdots,\lambda_p)\in\mathbb{R}^p$ 使(a,λ)是$F(x,\lambda)=f(x)+\sum_{r=1}^p\lambda_rg_r(x)$的临界点.
换言之, a与λ满足方程组 $\frac{\partial f}{\partial x_k}(a)+\sum_{r=1}^p\lambda_r\frac{\partial g_r}{\partial x_k}(a)=0$, $k=1,2\cdots m+p$ 与 $g_r(a)=0$, $r=1,2\cdots p$. 拉乘

pf: 由rank$=p$. 故可取到向量的p个线性无关向量, 不妨取$m+1\cdots m+p$列, $\frac{\partial(g_1,\cdots,g_p)}{\partial(x_{m+1},\cdots,x_{m+p})}\neq 0$, 于是由隐函数
定理, 可解出$x_{m+i}=\varphi_i(x_1,\cdots,x_m)$, $i=1,2,\cdots p$, 只需求$\varphi(x_1,\cdots,x_m)=f(x_1,\cdots,x_m,\varphi_1(x_1,\cdots,x_m),\cdots,\varphi_p(x_1,\cdots,x_m))$
无条件极值. 因此在$a=(a_1,\cdots,a_m)$处, $\frac{\partial\varphi}{\partial x_i}=0$, $i=1,\cdots,m$ 即 $\frac{\partial f}{\partial x_i}+\sum_{j=1}^p\frac{\partial f}{\partial x_{m+j}}\frac{\partial\varphi_j}{\partial x_i}=0$, $i=1,2\cdots m$ --($*_1$)
下面设法消去$\frac{\partial\varphi_i}{\partial x_i}$, 为此利用条件, 求偏导: $\frac{\partial g_r}{\partial x_i}+\sum_{j=1}^p\frac{\partial g_r}{\partial x_{m+j}}\frac{\partial\varphi_j}{\partial x_i}=0$, $r=1,2\cdots p$ ---($*_2$)
将($*_2$)乘上对应$\lambda_1,\cdots,\lambda_p$, 加到($*_1$)上去, 得 $\frac{\partial f}{\partial x_i}+\sum_{r=1}^p\lambda_r\frac{\partial g_r}{\partial x_i}+\sum_{j=1}^p\left(\frac{\partial f}{\partial x_{m+j}}+\sum_{r=1}^p\lambda_r\frac{\partial g_r}{\partial x_{m+j}}\right)\frac{\partial\varphi_j}{\partial x_i}=0$ $(i=1,2,\cdots m)$
由rank$\frac{\partial(g_1,\cdots,g_p)}{\partial(x_{m+1},\cdots,x_{m+p})}(a)=p$. 故存在$\lambda_1,\cdots,\lambda_p$使$\frac{\partial f}{\partial x_{m+j}}(a)+\sum_{r=1}^p\lambda_r\frac{\partial g_r}{\partial x_{m+j}}(a)=0$, $j=1,2\cdots p$ $\Rightarrow\frac{\partial f}{\partial x_i}(a)+\sum_{r=1}^p\lambda_r\frac{\partial g_r}{\partial x_i}(a)=0$
$(i=1,2\cdots m)$ 再加条件 □

定理4 以上函数至少是二阶连续可微的, 设a与λ同上, 并记$F(x,\lambda)=f(x)+\sum_{r=1}^p\lambda_rg_r(x)$, 则若方阵$\left[\frac{\partial^2 F}{\partial x_k\partial x_l}(a,\lambda)\right]_{k,l=1}^{m+p}$
是正定的(负定的), 则f在g_r约束下在a点取到严格极小值(极大值). 取等充分条件

pf：由2阶Taylor公式，$f(a+h)-f(a)=\sum\limits_{k=1}^{m+p}\frac{\partial f}{\partial x_k}(a)h_k+\frac{1}{2}\sum\limits_{k,l=1}^{m+p}\frac{\partial^2 f}{\partial x_k\partial x_l}(a)h_kh_l+o(\|h\|^2)$ --(*1)，$0=g_r(a+h)-g_r(a)=\sum\limits_{k=1}^{m+p}\frac{\partial g_r}{\partial x_k}(a)h_k+\frac{1}{2}\sum\limits_{k,l=1}^{m+p}\frac{\partial^2 g_r}{\partial x_k\partial x_l}(a)h_kh_l+o(\|h\|^2)$，$r=1,2\cdots p$ (*2)　将(*2)各式乘以λ_r加到(*1)上可得

$f(a+h)-f(a)=\frac{1}{2}\sum\limits_{k,l=1}^{m+p}\frac{\partial^2 F}{\partial x_k\partial x_l}(a,\lambda)h_kh_l+o(\|h\|^2)$，再考虑二次型 □.

例　设有对称方阵 $\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & & & \vdots \\ a_{n1} & \cdots & \cdots & a_{nn} \end{bmatrix}=A$　$a_{ji}=a_{ij}$，$i,j=1,2\cdots n$ 我们来求二次型 $f(x)=\sum\limits_{i,j=1}^{n}a_{ij}x_ix_j$

在约束 $\sum\limits_{i=1}^{n}x_i^2=1$ 时最大值与最小值，为此引入辅助函数 $F(x,\lambda)=f(x)-\lambda(\sum\limits_{i=1}^{n}x_i^2-1)$，由拉格朗日乘数法，

$\frac{\partial F}{\partial x_i}=\sum\limits_{j=1}^{n}a_{ij}x_j-\lambda x_i=0$ --(*1)　且 $\sum\limits_{i=1}^{n}x_i^2=1$ --(*2)　即 $AX=\lambda X$，即λ为A的特征值．将(*1)分别乘上x_i并相加

可知在$(x_1,\cdots,x_n)$为特征值λ的特征向量时取极值λ．故最大值与最小值为A最大特征值与最小特征值 □