若 $y=f(x)$ 在开集 $D\subseteq D(f)$ 的每一点处都关于 $x_j$ 存在偏导数，

则可定义函数 $\frac{\partial f}{\partial x_j}(x)$，$x\in D$，称之为

$f$ 在 $D$ 上关于 $x_j$ 的偏导函数（简称偏导数）

定义 12.1.3

若 $n$ 元函数 $f$ 在 $x^0\in D(f)$ 处关于任意自变量的偏导数都存在，

则称列向量 $\left(\frac{\partial f}{\partial x_1}(x^0),\cdots,\frac{\partial f}{\partial x_n}(x^0)\right)^T$ 为 $f$ 在 $x^0$ 处的梯度，

记作 $\nabla f(x^0)$ 或 $\operatorname{grad} f(x^0)$.

命题 12.1.1　可微的必要条件

若 $f$ 在 $x\in D(f)^\circ$ 处可微，则 $f$ 在 $x$ 处的梯度存在，且

$df(x)=f'(x)\Delta x=\langle\nabla f(x),\Delta x\rangle$，$f'(x)=\langle\nabla f(x),\cdot\rangle$

注：$df=\sum_{j=1}^{n}\frac{\partial f}{\partial x_j}dx_j$，$f'=(\nabla f)^T=\left(\frac{\partial f}{\partial x_1},\cdots,\frac{\partial f}{\partial x_n}\right)$

Date.　　Page.

回顾

$f$ 在 $x^0 \in D(f)^\circ$ 处可微：若 $\exists a \in \mathbb{R}^n$ 使

$$f(x^0+\Delta x) - f(x^0) = \langle a, \Delta x \rangle + o(\|\Delta x\|) \qquad (\|\Delta x\| \to 0),$$

则称 $f$ 在 $x^0$ 处可微或可导，并称 $\langle a, \Delta x\rangle$ 为 $f$ 在 $x^0$ 处的微分，记作 $df(x^0)$，即

$$df(x^0) = \langle a, \Delta x \rangle,$$

并称映射（线性映射）

$$\Delta x \in \mathbb{R}^n \longmapsto \langle a, \Delta x \rangle \in \mathbb{R}$$

为 $f$ 在 $x^0$ 处的导数，记作 $f'(x^0)$ 或 $Df(x^0)$，即

$$f'(x^0) = \langle a, \cdot \rangle$$

梯度（偏导数 $\frac{\partial f}{\partial x_j}(x^0) := \lim_{t\to 0} \frac{f(x^0+te^j) - f(x^0)}{t}$）：

$$\nabla f(x^0) := \left(\frac{\partial f}{\partial x_1}(x^0), \cdots, \frac{\partial f}{\partial x_n}(x^0)\right)^T, \quad \mathrm{grad}\, f(x^0)$$

梯度的性质

(a) 若 $f \equiv \text{const}$，则 $\nabla f \equiv 0$

(b) 若 $\alpha, \beta$ 是常数，则 $\nabla(\alpha f + \beta g) = \alpha \nabla f + \beta \nabla g$

(c) $\nabla(f \cdot g) = g \nabla f + f \nabla g$

(d) 若 $g \neq 0$，则 $\nabla\left(\frac{f}{g}\right) = \frac{g\nabla f - f \nabla g}{g^2}$

三、多元函数的可微条件

定理 12.1.2　可微的必要条件　重要习题　习题13

若 $f$ 在 $x^0 \in D(f)^\circ$ 处可微，则 $f$ 在 $x^0$ 处连续且梯度 $\nabla f(x^0)$ 存在，并有

$$df(x^0) = \langle \nabla f(x^0), \Delta x \rangle, \quad f'(x^0) = \langle \nabla f(x^0), \cdot \rangle$$

注：$df = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j} dx_j$

定理 12.1.3　可微的充分条件　重要习题　习题14

若 $f$ 在 $x^0$ 的某邻域 $U \subseteq D(f)$ 上存在所有偏导数 $\frac{\partial f}{\partial x_j}$ $(j=1,2,\cdots,n)$，且这些偏导数都在 $x^0$ 处连续，那么 $f$ 在 $x^0$ 处可微

证明：利用一元函数的 Lagrange 中值定理，有

Date. Page.

$$f(x^0+\Delta x)-f(x^0)=f(x^0+\sum_{j=1}^{n}\Delta x_j e^j)-f(x^0)$$
$$=\sum_{k=1}^{n}\left[f(x^0+\sum_{j=1}^{k}\Delta x_j e^j)-f(x^0+\sum_{j=1}^{k-1}\Delta x_j e^j)\right]$$
$$\underline{\underline{0<\theta_k<1}}\ \sum_{k=1}^{n}\frac{\partial f}{\partial x_k}(x^0+\sum_{j=1}^{k-1}\Delta x_j e^j+\theta_k\Delta x_k)\Delta x_k$$
$$=\sum_{k=1}^{n}\frac{\partial f}{\partial x_k}(x^0)\Delta x_k+\sum_{k=1}^{n}\alpha_k\Delta x_k\ ,\ 其中$$
$$\alpha_k=\frac{\partial f}{\partial x_k}(x^0+\sum_{j=1}^{k-1}\Delta x_j e^j+\theta_k\Delta x_k)-\frac{\partial f}{\partial x_k}(x^0)\longrightarrow 0\ (\|\Delta x\|\to 0)$$

四. 多元函数的方向导数

定义 12.1.5　　重要例子　例 12.1.5

设 $y=f(x)$ 对于给定的 $x^0\in D(f)$ 和方向 $e\in S^{n-1}$，存在 $\delta>0$ 使得

$\forall t\in(0,\delta)$，$x^0+te\in D(f)$　　n维，模长为1

如果极限 $\lim\limits_{t\to 0+}\dfrac{f(x^0+te)-f(x^0)}{t}$

收敛，则称之为 $f$ 在 $x^0$ 处沿方向 $e$ 的方向导数。

记作 $f'_e(x^0)$，$D_ef(x^0)$，$\dfrac{\partial f}{\partial e}(x^0)$ 或 $\dfrac{\partial f}{\partial e}\Big|_{x^0}$

注：对 $j=1,2,\cdots,n$，显然有

$\dfrac{\partial f}{\partial x_j}(x^0)$ 存在 $\Longleftrightarrow$ $f'_{e_j}(x^0)$ 和 $f'_{-e_j}(x^0)$ 均存在，且 $f'_{e_j}(x^0)+f'_{-e_j}(x^0)=0$

定理 12.1.4

若 $f$ 在 $x^0\in D(f)^\circ$ 处可微，那么 $\forall e\in S^{n-1}$，方向导数 $f'_e(x^0)$ 存在，

且 $f'_e(x^0)=f'(x^0)e=\langle Df(x^0),e\rangle$，

$f'_e(x^0)+f'_{-e}(x^0)=0$.

定理 12.1.5

$f$ 在 $x^0\in D(f)^\circ$ 处可微的充要条件是下列二条成立：

(a) $f$ 在 $x^0$ 处的梯度 $\nabla f(x^0)$ 以及沿 $\forall e\in S^{n-1}$ 的方向导数 $f'_e(x^0)$

存在，且 $f'_e(x^0)=\langle\nabla f(x^0),e\rangle$；

(b) $\forall e\in S^{n-1}$，$\dfrac{f(x^0+rh)-f(x^0+re)}{r}\to 0$，

$(r,h)\to(0,e)$　　$(0,+\infty)\times S^{n-1}\to(r,h)$

定理 12.1.6

设 $y=f(x)$ 在 $x^0\in D(f)^\circ$ 处连续，那么

(a) $f$ 在 $x^0$ 处可微当且仅当其图像 $G(f)\subseteq\mathbb{R}^n\times\mathbb{R}$ 在点 $P_0(x^0,f(x^0))$

$P_0(x^0, f(x^0)) \in G(f)$处存在不平行于$y$轴的切平面$\pi$，即，$\pi$适合

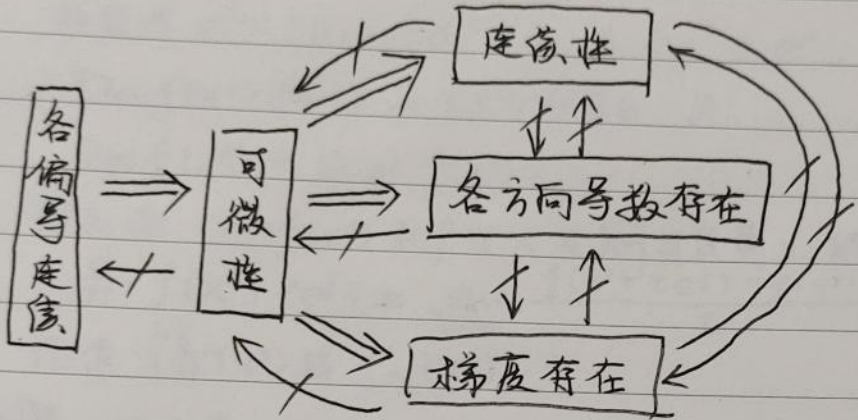
$$\lim_{P\in G(f)\to P_0} \frac{dist(P,\pi)}{dist(P,P_0)} = 0$$，其中 $\pi: y=f(x^0)+\langle a, x-x^0\rangle$

$a\in \mathbb{R}^n$ 是常向量

(b) $f$在$x^0$处可微时，$G(f)$在$P_0(x^0, f(x^0))$处的切平面方程为

$\pi: \ y = f(x^0) + \langle \operatorname{grad} f(x^0), x-x^0 \rangle$.

关系示意图




五. 映射（向量值函数）的可微性

定义 12.1.5

设$f: D(f) \subseteq \mathbb{R}^n \longmapsto \mathbb{R}^m$，$x^0 \in D(f)^\circ$

若存在线性映射$T: \mathbb{R}^n \to \mathbb{R}^m$使得

$$\lim_{\|\Delta x\|\to 0} \frac{1}{\|\Delta x\|}\left[f(x^0+\Delta x) - f(x^0) - T\Delta x\right] = 0.$$

那么称$f$在$x^0$处可微或可导，此时，

$T$是唯一的，称作$f$在$x^0$处的导数，

记作$f'(x^0)$或$Df(x^0)$.

并称$T\Delta x$为$f$在$x^0$处的微分，记作$df(x^0)$.

如果对于开集$D \subseteq D(f)$，$f$在$\forall x \in D$处可微，则称$f$在$D$上是可微的

注：由于$f'(x^0): \mathbb{R}^n \to \mathbb{R}^m$是线性映射，所以对$\mathbb{R}^n$的标准基底和$\mathbb{R}^m$的标准基底，存在唯一的$m\times n$矩阵$A$，使得

$$f'(x^0)[e^1, \cdots, e^n] = [f'(x^0)e^1, \cdots, f'(x_0)e^n] = [\varepsilon_1, \cdots, \varepsilon_m]A,$$

显然，$df(x^0) = f'(x^0)\Delta x = A\Delta x$，$dx = \Delta x$，$\Delta x \in \mathbb{R}^n$，

定义 12.1.6

设 $f: D(f)\subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^m$，$x^0\in D(f)$，$h\in\mathbb{R}^n$，

并设存在 $\delta>0$ 使得 $\forall t\in(0,\delta)$，$x^0+th\in D(f)$，

如果极限 $\lim\limits_{t\to 0+}\dfrac{f(x^0+th)-f(x^0)}{t}$ 收敛，

则称它为 $f$ 在 $x^0$ 处沿向量 $h$ 的方向导数，

记作 $\dfrac{\partial f}{\partial h}(x^0)$，或 $D_h f(x^0)$.

如果对 $e^j=(0,\cdots,0,\underset{\text{第}j\text{位}}{1},0,\cdots,0)^T\in\mathbb{R}^n$，

$D_{e^j}f(x^0)$ 和 $D_{-e^j}f(x^0)$ 都存在，且

$D_{e^j}f(x^0)+D_{-e^j}f(x^0)=0$，

那么称 $D_{e^j}f(x^0)$ 为 $f$ 在 $x^0$ 处关于自变量 $x_j$ 的偏导数.

$$\frac{d}{dt}f(x^0+te^j):=\lim_{t\to 0}\frac{f(x^0+te^j)-f(x^0)}{t}=D_{e^j}f(x^0).$$

记作 $\dfrac{\partial f}{\partial x_j}(x^0)$ 或 $f'_{x_j}(x^0)$

命题 12.1.7

设 $f: D(f)\subseteq\mathbb{R}^n\longrightarrow\mathbb{R}^m$，$x^0\in D(f)^\circ$，$h\in\mathbb{R}^n$，$1\le j\le n$，那么

(a) $D_h f(x^0)$ 存在 $\Longleftrightarrow$ 所有的 $D_h f_i(x^0)$ $(i=1,2,\cdots,m)$ 都存在，此时

$$D_h f(x^0)=\begin{bmatrix} D_h f_1(x^0)\\ \vdots \\ D_h f_m(x^0)\end{bmatrix}$$

(b) $\dfrac{\partial f}{\partial x_j}(x^0)$ 存在 $\Longleftrightarrow$ 所有 $\dfrac{\partial f_i}{\partial x_j}(x^0)$ $(i=1,2,\cdots,m)$ 都存在；此时

$$\frac{\partial f}{\partial x_j}(x^0)=\begin{bmatrix}\dfrac{\partial f_1}{\partial x_j}(x^0)\\ \vdots \\ \dfrac{\partial f_m}{\partial x_j}(x^0)\end{bmatrix}$$

定理 12.1.8 可微的必要条件

设 $f: D(f)\subseteq\mathbb{R}^n\longrightarrow\mathbb{R}^m$ 在 $x^0\in D(f)^\circ$ 处可微，那么

(a) $f$ 在 $x^0$ 处连续，且在 $x^0$ 处存在唯一导数

(b) $f$ 在 $x^0$ 处沿任意~~方向导数~~向量 $h$ 的方向导数存在，且

$$\begin{cases} D_h f(x^0)=f'(x^0)h \\ D_h f(x^0)+D_{-h}f(x^0)=0\end{cases}$$

(c) $f$ 在 $x^0$ 处关于每个自变量 $x_j$ $(j=1,2,\cdots,n)$ 的偏导数存在，且导数矩阵

$$f'(x^0)=\left[\frac{\partial f}{\partial x_1}(x^0),\cdots,\frac{\partial f}{\partial x_n}(x^0)\right]$$

(d) $f$ 的各分量函数 $f_i\ (i=1,2,\cdots,m)$ 在 $x^0$ 处可微，且导数矩阵为

$$f'(x^0)=\begin{bmatrix} f_1'(x^0) \\ \vdots \\ f_m'(x^0) \end{bmatrix}=[\nabla f_1(x^0),\cdots,\nabla f_m(x^0)]^T$$

定理 12.1.9 可微的充分条件

设 $f: D(f)\subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^m$，$x^0\in D(f)^\circ$，那么下列各条等价：

(a) $f$ 在 $x^0$ 处可微

(b) $f$ 的各分量函数 $f_i\ (i=1,2,\cdots,m)$ 都在 $x^0$ 处可微

(c) $f$ 在 $x^0$ 处沿任意向量的方向导数存在，并且 $\forall h,k\in\mathbb{R}^n$，$\forall\lambda\in\mathbb{R}$ 有

$$D_{h+\lambda k}f(x^0)=D_hf(x^0)+\lambda D_kf(x^0)$$

还有，对 $\forall e\in S^{n-1}$，

$$\frac{1}{r}(f(x^0+rh)-f(x^0+re))\longrightarrow 0\ ((r,h)\in(0,+\infty)\times S^{n-1}\longrightarrow(0,e))$$

定理 12.1.10

设 $f: D(f)\subseteq\mathbb{R}^n\longrightarrow\mathbb{R}^m$，$x^0\in D(f)^\circ$，

如果在 $x^0$ 的某邻域 $U\subseteq D(f)$ 上存在所有偏导数 $\frac{\partial f_i}{\partial x_j}\ (i=1,\cdots,m;j=1,\cdots,n)$

并且它们都在 $x^0$ 处连续，

那么 $f$ 在 $x^0$ 处可微.

## §12.2 映射(函数)的求导法则

### 一、映射的求导法则

**定理 12.2.1** 线性运算求导法则和内积求导法则

设 $f: D(f) \subseteq \mathbb{R}^n \to \mathbb{R}^m$,

$g: D(g) \subseteq \mathbb{R}^n \to \mathbb{R}^m$,

$\lambda \in \mathbb{R}$,

如果 $f$ 和 $g$ 均在 $x^0$ 处可微，那么 $f+\lambda g$ 和 $\langle f, g\rangle$ 均在 $x^0$ 处可微，且

$$(f+\lambda g)'(x^0) = f'(x^0) + \lambda g'(x^0)$$

$$\langle f, g\rangle'(x^0) = g^T(x^0) f'(x^0) + f^T(x^0) g'(x^0)$$

引进记号: Jacobi矩阵

对于映射 $y = f(x)$, $f: D(f) \subseteq \mathbb{R}^n \to \mathbb{R}^m$，引进记号

$$\left.\frac{\partial(y_{i_1}, \cdots, y_{i_k})}{\partial(x_{j_1}, \cdots, x_{j_l})}\right|_{x^0} := \begin{bmatrix} \frac{\partial y_{i_1}}{\partial x_{j_1}} & \cdots & \frac{\partial y_{i_1}}{\partial x_{j_l}} \\ \vdots & & \vdots \\ \frac{\partial y_{i_k}}{\partial x_{j_1}} & \cdots & \frac{\partial y_{i_k}}{\partial x_{j_l}} \end{bmatrix}_{x^0} \equiv \begin{bmatrix} \frac{\partial f_{i_1}}{\partial x_{j_1}} & \cdots & \frac{\partial f_{i_1}}{\partial x_{j_l}} \\ \vdots & & \vdots \\ \frac{\partial f_{i_k}}{\partial x_{j_1}} & \cdots & \frac{\partial f_{i_k}}{\partial x_{j_l}} \end{bmatrix}_{x^0}$$

$$= \left.\frac{\partial(f_{i_1}, \cdots, f_{i_k})}{\partial(x_{j_1}, \cdots, x_{j_l})}\right|_{x^0}$$

其中 $1 \le i_1 < \cdots < i_k \le m$, $1 \le j_1 < \cdots < j_l \le n$, $x^0 \in D(f)$.

特别地

$$\left.\frac{\partial(y_1, \cdots, y_m)}{\partial(x_1, \cdots, x_n)}\right|_{x^0} = \left.\frac{\partial(f_1, \cdots, f_m)}{\partial(x_1, \cdots, x_n)}\right|_{x^0} = f'(x^0)$$

当Jacobi矩阵是方阵时，其行列式常称作Jacobi行列式，也常以 $\frac{\partial(y_1, \cdots, y_n)}{\partial(x_1, \cdots, x_n)}$ 表示 $\det \frac{\partial(y_1, \cdots, y_n)}{\partial(x_1, \cdots, x_n)}$

**定理 12.2.2** 链式法则

设 $f: D(f) \subseteq \mathbb{R}^n \to \mathbb{R}^m$ 和 $g: D(g) \subseteq \mathbb{R}^m \to \mathbb{R}^l$，且 $R(f) \subseteq D(g)$

若 $f$ 在 $x^0 \in D(f)^\circ$ 处可微，$g$ 在 $f(x^0)$ 处可微，

那么 $g \circ f$ 在 $x^0$ 处可微，且

$$(g\circ f)'(x^0) = g'(f(x^0)) f'(x^0), \quad d(g\circ f)(x^0) = g'(f(x^0))\, df(x^0)$$

$$\left.\frac{\partial(z_{i_1}, \cdots, z_{i_p})}{\partial(x_{j_1}, \cdots, x_{j_q})}\right|_{x^0} = \left.\frac{\partial(z_{i_1}, \cdots, z_{i_p})}{\partial(y_1, \cdots, y_m)}\right|_{y^0} \cdot \left.\frac{\partial(y_1, \cdots, y_m)}{\partial(x_{j_1}, \cdots, x_{j_q})}\right|_{x^0},$$

Date. Page.

其中 $y=f(x)$，$z=g(y)$，$1\le i_1<\cdots<i_p\le l$，$1\le j_1<\cdots<j_q\le n$，$y^0=f(x^0)$.

证要：由 $f$ 在 $x^0$ 处可微，$g$ 在 $f(x^0)$ 处可微，

据 T2 - 第1题，存在矩阵值映射

$$A: D(f)\longrightarrow \mathbb{R}^{m\times n} \text{ 和 } B: D(g)\longrightarrow \mathbb{R}^{l\times n}$$

使得它们分别在 $x^0$ 和 $f(x^0)$ 处连续，且

$$f(x)-f(x^0)=A(x)(x-x^0),\quad x\in D(f),$$

$$g(y)-g(f(x^0))=B(y)(y-y^0),\quad y\in D(g).$$

于是 $$g\circ f(x)-g\circ f(x^0)=B(f(x))(f(x)-f(x^0))$$
$$=B(f(x))A(x)(x-x^0),\quad x\in D(f)$$

其中 $x\in D(f)\longmapsto B(f(x))A(x)\in\mathbb{R}^{l\times n}$ 在 $x^0$ 处连续.

再由 T2 - 第1题，即得所证.

例：

设 $W=\begin{bmatrix} f(x,\psi(x,y,\varphi(x,y))) \\ g(y,\varphi(x,y))\end{bmatrix}$，求 $W'$

解：$W'(x,y)=\dfrac{\partial(f,g)}{\partial(x,y)}=\begin{bmatrix}\dfrac{\partial f}{\partial(x,y)}\\ \dfrac{\partial g}{\partial(x,y)}\end{bmatrix}$

$$=\begin{bmatrix}\dfrac{\partial f}{\partial(x,\psi)}\cdot\dfrac{\partial(x,\psi)}{\partial(x,y)}\\ \dfrac{\partial g}{\partial(y,\varphi)}\cdot\dfrac{\partial(y,\varphi)}{\partial(x,y)}\end{bmatrix}=\begin{bmatrix}\dfrac{\partial f}{\partial(x,\psi)}\begin{bmatrix}\dfrac{\partial x}{\partial(x,y)}\\ \dfrac{\partial \psi}{\partial(x,y)}\end{bmatrix}\\ \dfrac{\partial g}{\partial(y,\varphi)}\begin{bmatrix}\dfrac{\partial y}{\partial(x,y)}\\ \dfrac{\partial \varphi}{\partial(x,y)}\end{bmatrix}\end{bmatrix}$$

$$=\begin{bmatrix}\dfrac{\partial f}{\partial(x,\psi)}\begin{bmatrix}\dfrac{\partial x}{\partial(x,y)}\\ \dfrac{\partial \psi}{\partial(x,y,\varphi)}\begin{bmatrix}\dfrac{\partial(x,y)}{\partial(x,y)}\\ \dfrac{\partial\varphi}{\partial(x,y)}\end{bmatrix}\end{bmatrix}\\ \dfrac{\partial g}{\partial(y,\varphi)}\begin{bmatrix}\dfrac{\partial y}{\partial(x,y)}\\ \dfrac{\partial \varphi}{\partial(x,y)}\end{bmatrix}\end{bmatrix}$$

$$=\begin{bmatrix}(f'_x,f'_\psi)\begin{bmatrix}(1,0)\\ (\psi'_x,\psi'_y,\psi'_\varphi)\begin{bmatrix}\begin{bmatrix}1&0\\0&1\end{bmatrix}\\ (\varphi'_x,\varphi'_y)\end{bmatrix}\end{bmatrix}\\ (g'_y,g'_\varphi)\begin{bmatrix}(0,1)\\ (\varphi'_x,\varphi'_y)\end{bmatrix}\end{bmatrix}$$

$$= \begin{bmatrix} f'_x + f'_\psi(\psi'_x + \psi'_\varphi \varphi'_x) & f'_\psi(\psi'_y + \psi'_\varphi \varphi'_y) \\ g'_\varphi \varphi'_x & g'_y + g'_\varphi \varphi'_y \end{bmatrix}$$

$$= \begin{bmatrix} W'_{1x} & W'_{1y} \\ W'_{2x} & W'_{2y} \end{bmatrix}$$

$$\frac{\partial W_1}{\partial x} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial \psi}\cdot\frac{\partial \psi}{\partial x} + \frac{\partial f}{\partial \psi}\cdot\frac{\partial \psi}{\partial \varphi}\cdot\frac{\partial \varphi}{\partial x}$$

$$\frac{\partial W_1}{\partial y} = \frac{\partial f}{\partial \psi}\cdot\frac{\partial \psi}{\partial y} + \frac{\partial f}{\partial \psi}\cdot\frac{\partial \psi}{\partial \varphi}\cdot\frac{\partial \varphi}{\partial y}$$

$$\frac{\partial W_2}{\partial x} = \frac{\partial g}{\partial \varphi}\cdot\frac{\partial \varphi}{\partial x}$$

$$\frac{\partial W_2}{\partial y} = \frac{\partial g}{\partial y} + \frac{\partial g}{\partial \varphi}\cdot\frac{\partial \varphi}{\partial y}$$

回顾

设 $f: D(f) \subseteq \mathbb{R}^n \to \mathbb{R}^m$ 在 $x^0 \in D(f)^\circ$ 处可微，

$g: D(g) \subseteq \mathbb{R}^m \to \mathbb{R}^l$ 在 $f(x^0)$ 处可微，且 $R(f) \subseteq D(g)$.

那么 $g\circ f$ 在 $x^0$ 处可微，且

$(g\circ f)'(x^0) = g'(f(x^0))f'(x^0)$,

$d(g\circ f)(x^0) = g'(f(x^0))\,d f(x^0)$ —— 一阶微分形式不变性

$$\left.\frac{\partial(z_{i_1},\cdots,z_{i_p})}{\partial(x_{j_1},\cdots,x_{j_q})}\right|_{x^0} = \left.\frac{\partial(z_{i_1},\cdots,z_{i_p})}{\partial(y_1,\cdots,y_m)}\right|_{y^0}\cdot\left.\frac{\partial(y_1,\cdots,y_m)}{\partial(x_{j_1},\cdots,x_{j_q})}\right|_{x^0},$$

其中 $y=f(x)$，$z=g(y)$，$1\le i_1<\cdots<i_p\le l$，$1\le j_1<\cdots<j_q\le n$，$y^0=f(x^0)$

注：$g$ 在 $f(x^0)$ 处可微不能减弱为 $g$ 在 $f(x^0)$ 处存在偏导数

公式

$d(u+\lambda v) = du + \lambda dv$ （$\lambda$ 是常数）

$d(uv) = v du + u dv$

$d\left(\frac{u}{v}\right) = \frac{v du - u dv}{v^2}$ $(v\ne 0)$

$d(\ln|v|) = \frac{dv}{v}$ $(v\ne 0)$

Date. Page.

# §12.3 中值定理与Taylor公式

## 一、高阶偏导数与高阶微分

设有$n$元函数$y=f(x)$，$x=\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix}\in D(f)$，并设$D(f)$是$\mathbb{R}^n$中开集，

如果$f$在$D(f)$上存在偏导数$\frac{\partial f}{\partial x_i}$，而$\frac{\partial f}{\partial x_i}$在$D(f)$又存在偏导数$\frac{\partial}{\partial x_j}\frac{\partial f}{\partial x_i}$，则记

$$\frac{\partial^2 f}{\partial x_j\partial x_i}:=\frac{\partial}{\partial x_j}\frac{\partial f}{\partial x_i},$$

称作$f$具有二阶偏导数$\frac{\partial^2 f}{\partial x_j\partial x_i}$

当$i=j$时，特记上述二阶偏导数为$\frac{\partial^2 f}{\partial x_i^2}$.

类似定义更高阶的偏导数.

连续可微符号 $C^k(D)$ ， $C^\infty(D)$

$C^0(D)=C(D)$ ， $C^1(D)$

如果$f$在开集$D(f)$上是$C^1$类的，那么

$$df=\langle\nabla f,\cancel{dx}\ dx\rangle=\sum_{j=1}^n\frac{\partial f}{\partial x_j}dx_j=\cancel{(\textstyle\sum_{j=1}^n \cdots)}\ \left(\sum_{j=1}^n dx_j\frac{\partial}{\partial x_j}\right)f=\langle dx,\nabla\rangle f$$

如果$f\in C^2(D)$，$D=D(f)$，那么$\nabla f$在$D$上是$C^1$类的，$\nabla f$在$D$上可微分

定义

$$d(df)=d\left[\left(\sum_{j=1}^n dx_j\frac{\partial}{\partial x_j}\right)f\right]=d[\langle dx,\nabla\rangle f]$$

$$d^2f=\langle d\nabla f,dx\rangle=\sum_{j=1}^n\left(d\frac{\partial f}{\partial x_j}\right)dx_j$$

依定义，特别地

$$d^2x_j=0\quad(j=1,2,\cdots,n)$$

一般地，如果$f\in C^{k+1}(D)$，那么定义

$$d^{k+1}f:=d(d^kf)=\sum_{j=1}^n d^k\left(\frac{\partial f}{\partial x_j}\right)\cdot dx_j\quad(k=1,2,\cdots)$$

称作$f$的高阶微分.

$$d^kf=\langle dx,\nabla\rangle^kf$$

定理 12.3.1 Clairaut定理

设有连续可微的 $n$ 元函数 $f$，$D(f)$ 是 $\mathbb{R}^n$ 中开集，对给定的下标 $i,j$ $(i\neq j)$ 和点 $x^0\in D(f)$，如果 $f$ 在 $D(f)$ 上存在二阶混合偏导数 $\dfrac{\partial^2 f}{\partial x_i\partial x_j}$ 和 $\dfrac{\partial^2 f}{\partial x_j\partial x_i}$，且它们都在 $x^0$ 处连续，那么

$$\frac{\partial^2 f}{\partial x_i\partial x_j}(x^0)=\frac{\partial^2 f}{\partial x_j\partial x_i}(x^0)$$

证要：令 $F(s,t):=f(x^0+se^i+te^j)-f(x^0+se^i)-f(x^0+te^j)$

$$+f(x^0),$$

$$\varphi(x):=f(x+se^i)-f(x),$$

$$\psi(x):=f(x+te^j)-f(x).$$

$e^i=\begin{pmatrix}0\\ \vdots\\ 0\\ 1\\ 0\\ \vdots\\ 0\end{pmatrix}\leftarrow$ 第 $i$ 位

那么，据Lagrange中值定理，

(据 $f,\varphi,\psi$ 定义)

$$F(s,t)=\varphi(x^0+te^j)-\varphi(x^0)=\psi(x^0+se^i)-\psi(x^0)$$

(中值定理)

$$\overset{0<\theta_1<1}{=\!=\!=}\ t\frac{\partial\varphi}{\partial x_j}(x^0+\theta_1 te^j)\overset{0<\theta_2<1}{=\!=\!=}\ s\frac{\partial\psi}{\partial x_i}(x^0+\theta_2 se^i)$$

注意到 $\dfrac{\partial\varphi}{\partial x_j}(x)=\dfrac{\partial f}{\partial x_j}(x+se^i)-\dfrac{\partial f}{\partial x_j}(x)$

$$\overset{0<\xi_1<1}{=\!=\!=}\ s\frac{\partial^2 f}{\partial x_i\partial x_j}(x+\xi_1 se^i),$$

$$\frac{\partial\psi}{\partial x_i}(x)\overset{0<\xi_2<1}{=\!=\!=}\ t\frac{\partial^2 f}{\partial x_j\partial x_i}(x+\xi_2 te^j)$$

我们得

$$F(s,t)\underset{0<\xi_1<1}{\overset{0<\theta_1<1}{=\!=\!=}}\ st\frac{\partial^2 f}{\partial x_i\partial x_j}(x^0+\theta_1 te^j+\xi_1 se^i)$$

$$\underset{0<\xi_2<1}{\overset{0<\theta_2<1}{=\!=\!=}}\ st\frac{\partial^2 f}{\partial x_j\partial x_i}(x^0+\theta_2 se^i+\xi_2 te^j)$$

从而 $\dfrac{\partial^2 f}{\partial x_i\partial x_j}(x^0)=\dfrac{\partial^2 f}{\partial x_j\partial x_i}(x^0)$

推论

如果 $f\in C^m(D)$，$D$ 是 $\mathbb{R}^n$ 中开集，那么对任意的自然数 $k\leq m$，$f$ 在 $D$ 上的任意 $k$ 阶偏导数与求导的次序无关.

二、多元函数的微分中值定理和Taylor公式

定理 12.3.2 中值定理

如果多元函数 $f$ 在开凸集 $D\subseteq D(f)$ 上可微，那么 $\forall a,b\in D$，$\exists\theta\in(0,1)$，使得

Date.　　Page.

$$f(b)-f(a)=f'((1-\theta)a+\theta b)(b-a)$$
$$=\langle \nabla f((1-\theta)a+\theta b),\ b-a\rangle$$

证要：令 $\phi(t):=f((1-t)a+tb)$，

则 $\phi(t)$ 在 $[0,1]$ 上连续，在 $(0,1)$ 内可导，并且

$$\phi'(t)=f'((1-t)a+tb)(b-a)$$
$$=\langle \nabla f((1-t)a+tb),\ b-a\rangle$$

据一元函数的中值定理，$\exists\,\theta\in(0,1)$，使

$$f(b)-f(a)=\phi(1)-\phi(0)$$
$$=\phi'(\theta)$$
$$=f'((1-t)a+tb)(b-a)$$

推论

若多元函数 $f$ 在开区域 $D$ 上各偏导数存在且 $\nabla f=0$，则 $f$ 在 $D$ 上是常值的.

定理 12.3.3　Taylor公式

设多元函数 $f\in C^{m+1}(D)$，其中 $D$ 是 $\mathbb{R}^n$ 中的开凸集，

那么 $\forall x^0,\ x^0+h\in D$，$\exists\,\theta\in(0,1)$，使得

$$f(x^0+h)=f(x^0)+\frac{1}{1!}\langle h,\nabla\rangle f(x^0)+\cdots+\frac{1}{m!}\langle h,\nabla\rangle^m f(x^0)$$
$$+\frac{1}{(m+1)!}\langle h,\nabla\rangle^{m+1} f(x^0+\theta h),$$

其中 $\langle h,\nabla\rangle^k=\left(h_1\frac{\partial}{\partial x_1}+\cdots+h_n\frac{\partial}{\partial x_n}\right)^k$，$h=(h_1,\cdots,h_n)^T$

证要：令 $\Phi(t):=f(x^0+th)$，则 $\Phi$ 在 $[0,1]$ 上有直到 $m+1$ 阶的连续导数，于是，据一元函数的Taylor定理，$\exists\,\theta\in(0,1)$ 使得

$$(0)\quad \Phi(1)=\Phi(0)+\frac{1}{1!}\Phi'(0)+\cdots+\frac{1}{m!}\Phi^{(m)}(0)+\frac{1}{(m+1)!}\Phi^{(m+1)}(\theta)$$

又 $\Phi'(t)=\langle h,\nabla\rangle f(x^0+th)$，

$$\Phi^{(k)}(t)=\langle h,\nabla\rangle^k f(x^0+th),\quad t\in[0,1]$$

代入 (0) 式得到 Taylor公式.

Date. Page.

注：对 $\forall \alpha=(\alpha_1,\cdots,\alpha_n)\in\mathbb{N}^n$，引入记号

$$|\alpha|:=\alpha_1+\cdots+\alpha_n$$

$$\alpha!:=\alpha_1!\cdot\cdots\cdot\alpha_n!$$

$$h^\alpha:=h_1^{\alpha_1}\cdot\cdots\cdot h_n^{\alpha_n}\quad \left(h=\begin{pmatrix}h_1\\ \vdots\\ h_n\end{pmatrix}\in\mathbb{R}^n\right)$$

$$\nabla^\alpha:=\frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1}\cdots\partial x_n^{\alpha_n}}$$

那么 $\langle h,\nabla\rangle^k=\left(h_1\frac{\partial}{\partial x_1}+\cdots+h_n\frac{\partial}{\partial x_n}\right)^k$

$$=\sum_{\alpha_1+\cdots+\alpha_n=k}\frac{k!}{\alpha_1!\cdot\cdots\cdot\alpha_n!}h_1^{\alpha_1}\cdot\cdots\cdot h_n^{\alpha_n}\frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}}\cdot\cdots\frac{\partial^{\alpha_n}}{\partial x_n^{\alpha_n}}$$

$$=\sum_{|\alpha|=k}\frac{k!}{\alpha!}h^\alpha\nabla^\alpha$$

因此，Taylor公式可表为

$$f(x^0+h)=\sum_{k=0}^{m}\sum_{|\alpha|=k}\frac{\nabla^\alpha f(x^0)}{\alpha!}h^\alpha+\sum_{|\alpha|=m+1}\frac{\nabla^\alpha f(x^0+\theta h)}{\alpha!}h^\alpha$$

一元情形：

$$f(x_0+h)=f(x_0)+\frac{1}{1!}f'(x_0)h+\cdots+\frac{1}{m!}f^{(m)}(x_0)h^m+\frac{1}{(m+1)!}f^{(m+1)}(x_0+\theta h)h^{m+1}$$

定理 12.3.4

设f如Th.12.3.3所设，那么

$$f(x^0+h)=\sum_{k=0}^{m}\sum_{|\alpha|=k}\frac{\nabla^\alpha f(x^0)}{\alpha!}h^\alpha+O(\|h\|^{m+1})\quad(\|h\|\to0)$$

$$=\sum_{k=0}^{m}\sum_{|\alpha|=k}\frac{\nabla^\alpha f(x^0)}{\alpha!}h^\alpha+o(\|h\|^{m+1})\quad(\|h\|\to0)$$

Hessian矩阵　　　　Hesse矩阵（PS 此老写法）.

对于n元函数f，记

$$Hf(x^0):=\begin{pmatrix}\frac{\partial^2 f}{\partial x_1^2}(x^0) & \cdots & \frac{\partial^2 f}{\partial x_1\partial x_n}(x^0)\\ \vdots & & \vdots\\ \frac{\partial^2 f}{\partial x_n\partial x_1}(x^0) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(x^0)\end{pmatrix}$$

称为f的Hessian矩阵.

当 $f\in C^2$ 时，Hesse矩阵是一个n阶对称矩阵，且

$$\sum_{i,j=1}^{n}\frac{\partial^2 f}{\partial x_i\partial x_j}(x^0)h_ih_j=h^T Hf(x^0)h$$

从而有Taylor公式如下：

$$f(x^0+h)=f(x^0)+\langle h,\nabla\rangle f(x^0)+\frac{1}{2}h^T Hf(x^0)h+o(\|h\|^2)$$

Date. Page.

三、映射的微分中值不等式与Taylor公式

定理 12.3.5 微分中值不等式

设 $f: D(f) \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^m$ 在开凸集 $D \subseteq D(f)$ 上可微，

那么 $\forall a, b \in D$，存在 $\xi = a + \theta(b-a)$，$\theta \in (0,1)$，使得

$$\| f(b) - f(a) \| \leq \| f'(\xi) \| \cdot \| b-a \|$$

推论 1.

设 $f: D(f) \subseteq \mathbb{R}^n \to \mathbb{R}^m$ 在开凸集 $D \subseteq D(f)$ 上可微，

那么 $\forall a, b, c \in D$，存在 $\xi = a + \theta(b-a)$，$\theta \in (0,1)$，使得

$$\| f(b) - f(a) - f'(c)(b-a) \| \leq \| f'(\xi) - f'(c) \| \cdot \| b-a \|.$$

推论 2.

≠开凸集

若 $f: D(f) \subseteq \mathbb{R}^n \to \mathbb{R}^m$ 在开区域 $D \subseteq D(f)$ 上满足 $f' = 0$，

那么 $f$ 在 $D$ 上是常向量.

定理 12.3.6 Taylor公式

设 $f \in C^{r+1}(D, \mathbb{R}^m)$，$D$ 是 $\mathbb{R}^n$ 中的开凸集，那么

对任意的 $x$，$x+h \in D$，有

"$D_h$" 方向导数算子

$$f(x+h) = f(x) + \frac{1}{1!} D_h f(x) + \cdots + \frac{1}{r!} D_h^r f(x)$$
$$+ \frac{1}{r!} \int_0^1 (1-t)^r D_h^{r+1} f(x+th)\, dt$$

其中 $D_h^k := \underbrace{D_h D_h \cdots D_h}_{k个}.$

注：对于 $r$ 次 $n$ 元多项式 $P(x) = \sum_{|\alpha| \leq r} a_\alpha x^\alpha \ (x \in \mathbb{R}^n)$，规定

$$P(\nabla) = \sum_{|\alpha| \leq r} a_\alpha \nabla^\alpha$$

$$P(\nabla) f = \begin{bmatrix} P(\nabla) f_1 \\ \vdots \\ P(\nabla) f_m \end{bmatrix} \quad \left( f = \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix} \in C^r(D, \mathbb{R}^m) \right)$$

$$D_h^k = \langle h, \nabla \rangle^k = \sum_{|\alpha| = k} \frac{k!}{\alpha!} h^\alpha \nabla^\alpha$$

记号：$|\alpha| = \alpha_1 + \cdots + \alpha_n$ $\quad \alpha! = \alpha_1! \cdot \cdots \cdot \alpha_n!$

$$x^\alpha = x_1^{\alpha_1} \cdot \cdots \cdot x_n^{\alpha_n}$$

$$\nabla^\alpha = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}}$$

# §12.4 隐函数定理

## 一、隐函数

对于 $n+1$ 元函数 $z=F(x,y)$，$(x,y)\in D(F)\subseteq \mathbb{R}^n\times\mathbb{R}$，其中 $D(F)$ 构成了 $\mathbb{R}^n\times\mathbb{R}$ 中的开区域，考虑方程

$F(x,y)=0$ $\quad F^{-1}(0):=\{(x,y)\in D(F)\mid F(x,y)=0\}$

的解集 $F^{-1}(0)$ 的性态，我们感兴趣 $F^{-1}(0)$ 能否构成 $y=f(x)$ 这种形式的函数

$$\underbrace{F(x,y)=0}_{f}:\ D\longrightarrow E\subseteq\mathbb{R}$$

$$D(f)\longrightarrow R(f)$$

例0. $x^2+y^2-1=0$：$[-1,1]\to[0,+\infty)$ 隐函数.

$x\longmapsto y=\sqrt{1-x^2}$

例1. $x^2+y^2-1=0$：$[-1,1]\longrightarrow\mathbb{R}$ 不构成函数.

如果 $n$ 元函数 $y=f(x)$，$x\in D(f)\subseteq\mathbb{R}^n$，满足

$G(f)\subseteq D(F)$，$F(G(f))=\{0\}$，

并且

$\forall x\in D(f)$，$\exists!\, y\in R(f)$，使得 $(x,y)\in D(F)$ 且 $F(x,y)=0$.

则称 $y=f(x)$ 是方程 $F(x,y)=0$ 的一个隐函数.

## 二、隐函数定理

定理 12.4.1 隐函数存在性定理

设有 $n+1$ 元函数 $z=F(x,y)$，$(x,y)\in D(F)\subseteq\mathbb{R}^n\times\mathbb{R}$，其中 $D(F)$ 是 $\mathbb{R}^n\times\mathbb{R}$ 中的开区域.

如果 (a) $F$ 处处连续，$\frac{\partial F}{\partial y}$ 处处存在且连续.

(b) 在 $P_0(x^0,y^0)\in D(F)$ 处，$F(x^0,y^0)=0$，$\frac{\partial F}{\partial y}(x^0,y^0)\neq 0$.

那么 (i) 存在 $x^0$ 的邻域 $B(x^0,\varepsilon)\subseteq\mathbb{R}^n$ 和 $y^0$ 的邻域 $(y^0-\delta,y^0+\delta)\subseteq\mathbb{R}$，

使得 $F(x,y)=0$：$x\in B(x^0,\varepsilon)\longmapsto y\in(y^0-\delta,y^0+\delta)$

确定出（唯一的）一个隐函数 $y=f(x)$，$x\in D(f):=B(x^0,\varepsilon)$

$R(f)\subseteq(y^0-\delta,y^0+\delta)$，$F(x,f(x))\equiv 0$，$x\in D(f)$

Date. Page.

(ii) 隐函数 $f$ 满足：$f \in C(D(f))$，$y^0 = f(x^0)$

证：由 (a)(b)，不妨设 $\frac{\partial F}{\partial y}(x^0, y^0) > 0$，并且存在 $r>0$ 和 $\delta>0$ 使得

$\overline{B(x^0, r)} \times [y^0-\delta, y^0+\delta] \subseteq D(F)$，　　$\overline{B}$：闭包

且 $\forall (x,y) \in \overline{B(x^0, r)} \times [y^0-\delta, y^0+\delta]$，$\frac{\partial F}{\partial y}(x,y) > 0$.

因此，任意固定 $x \in B(x^0, r)$，

$F(x,y)$ 作为 $y \in [y^0-\delta, y^0+\delta]$ 的函数，

连续且严格增.

注意到 $F(x^0, y^0) = 0$，就有

$$F(x^0, y^0-\delta) < 0 < F(x^0, y^0+\delta)$$

再注意到 $F(x, y^0 \pm \delta)$ 作为 $x \in B(x^0, r)$ 的函数是连续的，

因而具有局部保号性，于是存在 $B(x^0, \varepsilon) \subseteq B(x^0, r)$ 使得

$$\forall x \in B(x^0, \varepsilon),\quad F(x, y^0-\delta) < 0 < F(x, y^0+\delta)$$

现利用一元连续函数的介值定理，得

对任意固定的 $x \in B(x^0, \varepsilon)$，存在唯一的 $f(x) \in (y^0-\delta, y^0+\delta)$

使得 $F(x, f(x)) = 0$，即，确定一个隐函数

$$F(x,y) = 0 : x \in B(x^0, \varepsilon) \longmapsto y = f(x) \in (y^0-\delta, y^0+\delta)$$

由此证明了 (i)

显然，$y^0 = f(x^0)$，再证 $f$ 是连续的：

对任意固定 $\tilde{x} \in B(x^0, \varepsilon)$，$f(\tilde{x}) \in (y^0-\delta, y^0+\delta)$

对任意的 $0 < \eta \leq \min\{f(\tilde{x}) - y^0 + \delta, y^0 + \delta - f(\tilde{x})\}$，有

$$y^0 - \delta \leq f(\tilde{x}) - \eta < f(\tilde{x}) + \eta \leq y^0 + \delta.$$

$$F(\tilde{x}, f(\tilde{x}) - \eta) < F(\tilde{x}, f(\tilde{x})) = 0 < F(\tilde{x}, f(\tilde{x}) + \eta)$$

由于 $F(x, f(\tilde{x}) \pm \eta)$ 作为 $x \in B(x^0, \varepsilon)$ 的连续函数，

存在 $B(\tilde{x}, \rho) \subseteq B(x^0, \varepsilon)$ 使

$$\forall x \in B(\tilde{x}, \rho),\quad F(x, f(\tilde{x}) - \eta) < 0 < F(x, f(\tilde{x}) + \eta),$$

$F(x, f(x)) = 0$，因而

$f(\tilde{x}) - \eta < f(x) < f(\tilde{x}) + \eta$，此证明了 $f$ 在 $\tilde{x}$ 处连续.

定理 12.4.2　　隐函数的连续可微性

设有 $n+1$ 元函数 $z=F(x,y)$，$(x,y)\in D(F)\subseteq \mathbb{R}^n\times\mathbb{R}$，

其中 $D(F)$ 是 $\mathbb{R}^n\times\mathbb{R}$ 中的开区域

如果 (a) $F\in C^k(D(F))$

(b) 在 $P_0(x^0,y^0)\in D(F)$ 处，$F(x^0,y^0)=0$，$\frac{\partial F}{\partial y}(x^0,y^0)\neq 0$

那么 $x^0$ 的邻域 $B(x^0,\varepsilon)\subseteq\mathbb{R}^n$ 和 $y^0$ 的邻域 $(y^0-\delta,y^0+\delta)\subseteq\mathbb{R}$，

使得　$F(x,y)=0: x\in B(x^0,\varepsilon)\longmapsto y\in(y^0-\delta,y^0+\delta)$

确立的隐函数 $y=f(x)$，$x\in D(f):=B(x^0,\varepsilon)$

适合 $f\in C^k(D(f))$，且

$$\frac{\partial f}{\partial x_j}=-\frac{\partial F}{\partial x_j}\Big/\frac{\partial F}{\partial y},\quad j=1,2,\cdots,n$$

$$df(x)=-\left\langle \frac{\nabla_x F(x,f(x))}{\frac{\partial F}{\partial y}(x,f(x))}, dx\right\rangle$$

其中 $\nabla_x := \left(\frac{\partial}{\partial x_1},\cdots,\frac{\partial}{\partial x_n}\right)^T$

Date.　　Page.

回顾

设 $F\in C^1(D(F))$，$D(F)$ 是 $\mathbb{R}^n\times\mathbb{R}$ 中开区域，如果在 $P_0(x^0,y^0)\in D(F)$ 处有

$$F(x^0,y^0)=0,\quad \frac{\partial F}{\partial y}(x^0,y^0)\neq 0,$$

那么存在 $x^0$ 的邻域 $B(x^0,\varepsilon)\subseteq\mathbb{R}^n$ 和 $y^0$ 的邻域 $(y^0-\delta,y^0+\delta)$ 使得

$$F(x,y)=0:\ x\in B(x^0,\varepsilon)\longmapsto y\in(y^0-\delta,y^0+\delta)$$

确立出一个隐函数 $y=f(x)$，$x\in B(x^0,\varepsilon)$，并且 $f\in C^1(B(x^0,\varepsilon))$，

$$\frac{\partial f}{\partial x_j}=-\frac{\partial F}{\partial x_j}\Big/\frac{\partial F}{\partial y},\quad j=1,2,\cdots,n$$

$$df(x)=-\left\langle\frac{\nabla_x F(x,f(x))}{\frac{\partial F}{\partial y}(x,f(x))},\ dx\right\rangle,\ 其中\ \nabla_x:=\begin{pmatrix}\frac{\partial}{\partial x_1}\\ \vdots\\ \frac{\partial}{\partial x_n}\end{pmatrix}$$

三、隐函数定理的解释

满足方程

$F(x,y)=0$，其中 $F\in C^1(D)$，$D$ 是 $\mathbb{R}^2$ 中的开区域，

是 $\mathbb{R}^3$ 中的曲面 $z=F(x,y)$ 与坐标平面 $Oxy$（即 $z=0$）的交集，

要想此交集构成某个函数

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

的图像，自然要求：

$\exists\, P_0(x_0,y_0)\in D$，适合

(i) $F(x_0,y_0)=0$，此保证 $F^{-1}(0)\neq\emptyset$，

(ii) $z=F(x,y)$ 在 $P_0(x_0,y_0)$ 处的切平面（$\left(\frac{\partial F}{\partial x}\Big|_{P_0},\frac{\partial F}{\partial y}\Big|_{P_0},-1\right)$ 法矢量）

$$\pi:\ \frac{\partial F}{\partial x}\Big|_{P_0}(x-x_0)+\frac{\partial F}{\partial y}\Big|_{P_0}(y-y_0)-z=0$$

不与 $Oxy$ 平行，即避免 $\frac{\partial F}{\partial x}\Big|_{P_0}=\frac{\partial F}{\partial y}\Big|_{P_0}=0$（此保证 $F^{-1}(0)\neq D$）

(iii) $\frac{\partial F}{\partial y}\Big|_{P_0}\neq 0$ 保证了 $\pi\cap Oxy$ 可构成 $y=kx+b$ 的形式：

$$y=-\frac{F_x'(P_0)}{F_y'(P_0)}(x-x_0)+y_0$$

二维情形

$$\begin{cases}F(x,y,u,v)=0\\ G(x,y,u,v)=0\end{cases}$$

其中 $\begin{pmatrix}F\\G\end{pmatrix}\in C^1$

符号统一：$\dfrac{\partial(F,G)}{\partial(u,v)}$ 视为矩阵

$$
\begin{cases}
\text{(i)} & \begin{cases} F(x_0, y_0, u_0, v_0) = 0 \\ G(x_0, y_0, u_0, v_0) = 0 \end{cases} \\
\text{(ii)} & \begin{pmatrix} F \\ G \end{pmatrix} \in C^1(D) \\
\text{(iii)} & \det \dfrac{\partial(F,G)}{\partial(u,v)}\Big|_{(x_0,y_0,u_0,v_0)} \neq 0
\end{cases}
$$

$$
\Longrightarrow \begin{pmatrix} F \\ G \end{pmatrix} = 0 : \begin{cases} u = u(x,y) \\ v = v(x,y) \end{cases} \quad (x,y) \in O((x_0,y_0), \rho)
$$

$$
\begin{pmatrix} x \\ y \end{pmatrix} \in O((x_0,y_0), \rho) \longmapsto \begin{pmatrix} u \\ v \end{pmatrix} \in O((u_0,v_0), \sigma)
$$

$$
\begin{pmatrix} u \\ v \end{pmatrix} = \Phi\begin{pmatrix} x \\ y \end{pmatrix} \in C^1
$$

且

$$
\frac{\partial(u,v)}{\partial(x,y)} = -\begin{pmatrix} F'_u & F'_v \\ G'_u & G'_v \end{pmatrix}^{-1} \begin{pmatrix} F_x & F_y \\ G_x & G_y \end{pmatrix}
$$

$$
= -\left[\frac{\partial(F,G)}{\partial(u,v)}\right]^{-1} \frac{\partial(F,G)}{\partial(x,y)}
$$

Date.　　Page.

# §12.5 隐映射定理与逆映射定理

## 一、隐映射定理

设 $F: D(F) \subseteq \mathbb{R}^n \times \mathbb{R}^m \longrightarrow \mathbb{R}^l$，$X \subseteq \mathbb{R}^n$，$Y \subseteq \mathbb{R}^m$，如果

$\forall x \in X$，$\exists!\, y \in Y$，使得

$(x, y) \in D(F)$ 且 $F(x, y) = 0$.

那么称

$F(x, y) = 0 : x \in X \longmapsto y \in Y$

确立了一个隐映射.

考虑上述 $F: D(F) \subseteq \mathbb{R}^n \times \mathbb{R}^m \longrightarrow \mathbb{R}^l$，$(x, y) \longmapsto F(x, y)$，

其中 $D(F)$ 是 $\mathbb{R}^n \times \mathbb{R}^m$ 中的开集，

对于 $(x^0, y^0) \in D(F)$，

如果线性映射 $A: \mathbb{R}^n \longrightarrow \mathbb{R}^l$ （$B: \mathbb{R}^m \longrightarrow \mathbb{R}^l$）使得

$$\lim_{\|\Delta x\| \to 0} \frac{F(x^0 + \Delta x, y^0) - F(x^0, y^0) - A\Delta x}{\|\Delta x\|} = 0$$

$$\left( \lim_{\|\Delta y\| \to 0} \frac{F(x^0, y^0 + \Delta y) - F(x^0, y^0) - B\Delta y}{\|\Delta y\|} = 0 \right)$$

那么称 $F$ 在 $(x^0, y^0)$ 处关于 $x$（关于 $y$）可导或可微，记作

$\partial_x F(x^0, y^0) = A$，$\frac{\partial F}{\partial x}(x^0, y^0) = A$，或 $F_x'(x^0, y^0) = A$

（$\partial_y F(x^0, y^0) = B$，$\frac{\partial F}{\partial y}(x^0, y^0) = B$，或 $F_y'(x^0, y^0) = B$

Date.　　Page.

回顾

考虑 $F: D(F) \subseteq \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^\ell$，$(x,y) \in D(F) \longmapsto F(x,y) \in \mathbb{R}^\ell$.

已定义的偏导~~数~~算子 $\frac{\partial F}{\partial x}(x^0,y^0)$（简记为 $\partial_x F(x^0,y^0)$，$\frac{\partial F}{\partial y}(x^0,y^0)$（简记为 $\partial_y F(x^0,y^0)$.

可视作其导数矩阵，即

$$\frac{\partial F}{\partial x}(x^0,y^0) = \frac{\partial(F_1,\cdots,F_\ell)}{\partial(x_1,\cdots,x_n)}\Big|_{(x^0,y^0)},$$

$$\frac{\partial F}{\partial y}(x^0,y^0) = \frac{\partial(F_1,\cdots,F_\ell)}{\partial(y_1,\cdots,y_m)}\Big|_{(x^0,y^0)}$$

$F = \begin{pmatrix} F_1 \\ \vdots \\ F_\ell \end{pmatrix}$

定理 12.5.1 隐映射定理

设 $F: D(F) \subseteq \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^\ell$，其中 $D(F)$ 是 $\mathbb{R}^n \times \mathbb{R}^m$ 中开集

如果 $F \in C^r$ $(r \geq 1)$，在 $(x^0,y^0) \in D(F)$ 处满足

$$F(x^0,y^0) = 0, \quad \det\left(\frac{\partial F}{\partial y}(x^0,y^0)\right) \neq 0$$

那么

(i) 存在 $x^0$ 的邻域 $U \subseteq \mathbb{R}^n$ 和 $y^0$ 的邻域 $V \subseteq \mathbb{R}^m$ 使得

$$F(x,y) = 0: \ x \in U \longmapsto y \in V$$

确定出一个隐映射 $y = f(x)$，$x \in U$；

即 $\begin{cases} F(x,f(x)) \equiv 0,\ x \in U \\ y^0 = f(x^0) \end{cases}$

(ii) $f \in C^r(U, \mathbb{R}^m)$ 且

$$f'(x) = -(\partial_y F)^{-1}\,\partial_x F\big|_{(x,f(x))}, \quad x \in U$$

二、逆映射定理

定理 12.5.2 逆映射定理

设 $f \in C^r(D(f), \mathbb{R}^n)$ $(r \geq 1)$，其中 $D(f)$ 在 $\mathbb{R}^n$ 中的开集.

如果存在 $x^0 \in D(f)$ 使得 $\det f'(x^0) \neq 0$，

那么存在 $x^0$ 的邻域 $U \subseteq D(f)$ 满足

(a) $f|_U : U \to \mathbb{R}^n$ 是单射

(b) $f(U)$ 在 $\mathbb{R}^n$ 中开

(c) $g := (f|_U)^{-1} \in C^r(f(U), \mathbb{R}^n)$，

且 $g'(y) = [f'(g(y))]^{-1}$，$y \in f(U)$

推论

设 $f \in C^1(D(f), \mathbb{R}^n)$，其中 $D(f)$ 是 $\mathbb{R}^n$ 中的开集，如果

$\forall x \in D(f)$，$\det f'(x) \neq 0$，

那么 $f$ 是开映射，即

若 $U \subseteq D(f)$ 在 $\mathbb{R}^n$ 中开，则 $f(U)$ 也在 $\mathbb{R}^n$ 中开；

特别地，$R(f)$ 在 $\mathbb{R}^n$ 中开.

定义　正则映射

设 $f: D(f) \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^n$，如果 $D \subseteq D(f)$ 是 $\mathbb{R}^n$ 中开集，使得 $f \in C^1(D, \mathbb{R}^n)$，且

(a) $\forall x \in D$，$\det f'(x) \neq 0$

(b) $f|_D: D \longrightarrow \mathbb{R}^n$ 是单射

那么称 $f$ 是 $D$ 上的正则映射.

三、隐函映射定理的几何意义　（欧氏空间中的微分流形）

定义1

设 $M \subseteq \mathbb{R}^n$ 非空，$1 \leq k < n$，$r \geq 1$，如果 $\forall p \in M$，

存在 $p$ 在 $\mathbb{R}^n$ 中的开集 $U$ 和映射 $\Phi \in C^r(U, \mathbb{R}^{n-k})$，使得

(a) $M \cap U = \{x \in U \mid \Phi(x) = 0\}$

(b) $\forall x \in U$，$\operatorname{rank} \Phi'(x) = n-k$

那么称 $M$ 是 $\mathbb{R}^n$ 中的一个 $k$ 维 $C^r$ 类流形，并将配对 $(U|\Phi)$ 称作 $M$ 在 $p$ 附近的局部隐式表示.

$C^1$ 类流形也称作微分流形或光滑流形，

1维微分流形称作光滑曲线，2维微分流形称作光滑曲面

定义2

设 $M$ 是 $\mathbb{R}^n$ 中的一个 $k$ 维微分流形，$p \in M$，若存在 $\gamma \in C^1((-\delta, \delta), M)$ $(\delta > 0)$，使得 $\gamma(0) = p$，则称

$v := \gamma'(0)$ $(\in \mathbb{R}^n)$ 为 $M$ 在 $p$ 处的一个切向量.

$M$ 在 $p$ 处的全体切向量组成的集合称作 $M$ 在 $p$ 处的切空间，记作 $T_p(M)$

Date. Page.

定理 1

设 $M$ 是 $\mathbb{R}^n$ 中的 $k$ 维微分流形，$(U|\Phi)$ 是 $M$ 在 $p$ 处的局部隐式表示，那么 $T_p(M)$ 是 $\mathbb{R}^n$ 中的 $k$ 维子空间，且 $T_p(M)=\ker(\Phi'(p))$，即

$$T_pM=\{v\in\mathbb{R}^n \mid \Phi'(p)v=0\}.$$

定义 3

设 $M$ 是 $\mathbb{R}^n$ 中的 $k$ 维微分流形，$p\in M$. 若 $n\in\mathbb{R}^n$ 使得

$$\forall v\in T_pM,\quad \langle n,v\rangle=0,$$

则称 $n$ 是 $M$ 在 $p$ 处的一个法向量.

全体法向量所组成的集合称作 $M$ 在 $p$ 处的法空间. 记作 $T_p^{\perp}M$.

定理 2

设 $M$ 是 $\mathbb{R}^n$ 中的 $k$ 维微分流形，$(U|\Phi)$ 是 $M$ 在 $p\in M$ 处的局部隐式表示，那么法空间 $T_p^{\perp}M$ 是 $\mathbb{R}^n$ 的 $n-k$ 维子空间，且

$$T_p^{\perp}M=\mathrm{span}\{\nabla\Phi_1(p),\cdots,\nabla\Phi_{n-k}(p)\},$$

其中 $\nabla\Phi_1(p),\cdots,\nabla\Phi_{n-k}(p)$ 恰是导数矩阵 $\Phi'(p)$ 的 $n-k$ 个行向量

## §12.6 极值问题

### 一、无条件极值

定义 12.6.1

设有 $n$ 元函数 $f$，如果对 $x^0 \in D(f)^\circ$，存在 $B(x^0,\delta) \subseteq D(f)$，使得
$\forall x \in B(x^0,\delta) \setminus \{x^0\}$，$f(x) \geq f(x^0)$（$f(x) > f(x^0)$），
那么称 $x^0$ 是 $f$ 的极小值点（严格极小值点），并称 $f(x^0)$ 是 $f$ 的极小值（严格极小值）

如果 $x^0$ 是 $-f$ 的极小值点（严格极小值点），那么称 $x^0$ 是 $f$ 的极大值点（严格极大值点），并称 $f(x^0)$ 是 $f$ 的极大值（严格极大值）

定理 12.6.1 极值的必要条件

设 $n$ 元函数 $f$ 在 $x^0 \in D(f)^\circ$ 处取得极值

(a) 若 $\nabla f(x^0)$ 存在，则必有 $\nabla f(x^0) = 0$.

(b) 若 $f$ 在 $x^0$ 的某邻域上二阶连续可微，

那么当 $x^0$ 是极小值点时，

Hesse 矩阵 $Hf(x^0)$ 正定或半正定.

当 $x^0$ 是极大值点时，

$Hf(x^0)$ 负定或半负定.

证明：(a) 对每个 $j = 1,2,\cdots,n$，考虑 $\varphi_j(t) := f(x^0 + te_j)$，

则 $\varphi_j'(0)$ 存在，且 $\varphi_j$ 在 $t=0$ 处取得极值

因而，$\varphi_j'(0) = 0$，即

$\frac{\partial f}{\partial x_j}(x^0) = 0$，从而 $\nabla f(x^0) = 0$.

(b) 由 $f$ 在 $x^0$ 的某邻域上是 $C^2$ 的，据 Taylor 公式，

并注意 $\nabla f(x^0) = 0$，得

$$
\begin{aligned}
f(x^0+h) - f(x^0) &= \frac{1}{2}h^T Hf(x^0+\theta h)h \\
&= \frac{1}{2}h^T Hf(x^0)h + \frac{1}{2}h^T(Hf(x^0+\theta h) - Hf(x^0))h \\
&= \frac{1}{2}\|h\|^2\left[\left(\frac{h}{\|h\|}\right)^T Hf(x^0)\frac{h}{\|h\|} + \omega(h)\right],
\end{aligned}
$$

其中 $\omega(h) := \left(\frac{h}{\|h\|}\right)^T (Hf(x^0+\theta h) - Hf(x^0))\frac{h}{\|h\|} \xrightarrow{\|h\|\to 0} 0$

现对任意给定的 $e \in \mathbb{R}^n \setminus \{0\}$，取 $h = te$ 代入上式得

$0 \le f(x^0 + te) - f(x^0) = \frac{1}{2}t^2\left[e^T Hf(x^0)e + \|e\|^2\omega(te)\right]$，当 $0<|t|<<1$ 时.

$e^T Hf(x^0) e \ge 0 \quad (\forall e \ne 0)$

定理 12.6.2 极值充分条件

设 $n$ 元函数 $f$ 在 $x^0 \in D(f)^\circ$ 的某邻域上二阶连续可微，

且 $x^0$ 是 $f$ 的驻点 $(\nabla f(x^0) = 0)$.

(a) 若 Hesse 矩阵 $Hf(x^0)$ 正定（负定），则

$x^0$ 是 $f$ 的一个严格极小值点（严格极大值点）

(b) 若 $Hf(x^0)$ 不定，则 $x^0$ 不是 $f$ 的极值点

定理 12.6.3 条件极值必要条件

设 $f \in C^1(D(f), \mathbb{R})$，其中 $D(f)$ 是 $\mathbb{R}^n$ 中开集，

又设 $M$ 是 $\mathbb{R}^n$ 中的 $k$ 维微分流形，适合 $M \subseteq D(f)$.

如果 $f|_M$ 在 $p \in M$ 处达到极大或极小，$(U, \Phi)$ 是 $M$ 在 $p$ 附近的局部隐式表示，那么存在唯一的 $\lambda \in \mathbb{R}^{n-k}$ 使得

$$F(x) := f(x) - \langle \lambda, \Phi(x) \rangle, \quad x \in U$$

以 $p$ 为驻点，即

$$F'(p) = f'(p) - \lambda^T \Phi'(p) = 0^T.$$

证：$\forall v \in T_p(M)$，$\exists \gamma \in C^1((-\delta, \delta), M)$ 使得

$\gamma(0) = p$，$\gamma'(0) = v$，

且 $\varphi(t) := f(\gamma(t))$ 在 $t=0$ 处达到极值.

从而 $f'(p)v = \varphi'(0) = 0$

这证明了 $f'(p) \in T_p^\perp M$. 据法空间的定理（定理2）

$$T_p^\perp M = \mathrm{span}\{\nabla \Phi_1(p), \cdots, \nabla \Phi_{n-k}(p)\}$$

故 $\exists \lambda = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_{n-k} \end{pmatrix} \in \mathbb{R}^{n-k}$，使得 ~ 略

Lagrange 乘数法

设 $f \in C^1(D(f), \mathbb{R})$，其中 $D(f)$ 是 $\mathbb{R}^n$ 中开集，又设 $M(\Phi(x)=0)$ 是 $\mathbb{R}^n$ 中的 k 维微分流形，适合 $M \subseteq D(f) \cap D(\Phi)$，

为求出 $f|_M$ 的极值点，令

$$L(x,\lambda) := f(x) - \langle \lambda, \Phi(x) \rangle,$$

$$(x,\lambda) \in (D(f) \cap D(\Phi)) \times \mathbb{R}^{n-k},$$

称之为 Lagrang 函数，求 $L(x,\lambda)$ 的驻点：

$$\begin{cases} \dfrac{\partial L}{\partial x} = f'(x) - \lambda^T \Phi'(x) = 0^T \\ \dfrac{\partial L}{\partial \lambda} = -(\Phi(x))^T = 0^T \end{cases} \quad L(x,\lambda)\text{的驻点方程}$$

据条件极值的必要条件，$f|_M$ 的极值点 p 连带某个 $\lambda \in \mathbb{R}^{n-k}$ 适合上面的驻点方程。

定理　条件极值的充分条件

设 $f \in C^2(D(f), \mathbb{R})$，$D(f)$ 是 $\mathbb{R}^n$ 中开集，

又设 $M(\Phi(x)=0)$ 是 $\mathbb{R}^n$ 中的 k 维 $C^2$ 类流形，适合 $M \subseteq U := D(f) \cap D(\Phi)$

对于 Lagrange 函数 $L(x,\lambda)$，如果 $(x^0,\lambda^0) \in U \times \mathbb{R}^{n-k}$ 为其驻点，那么：

(a) 当 $\tilde{H}_L(x^0,\lambda^0)$ 正定时，$f|_M$ 在 $x^0$ 处取到严格极小值

(b) 当 $\tilde{H}_L(x^0,\lambda^0)$ 负定时，$f|_M$ 在 $x^0$ 处取到严格极大值

其中 $\tilde{H}_L(x^0,\lambda^0)$ 定义如下：

$$\tilde{H}_L(x^0,\lambda^0) := \begin{bmatrix} \dfrac{\partial^2 L}{\partial x_1^2} & \cdots & \dfrac{\partial^2 L}{\partial x_1 \partial x_n} \\ \vdots & & \vdots \\ \dfrac{\partial^2 L}{\partial x_n \partial x_1} & \cdots & \dfrac{\partial^2 L}{\partial x_n^2} \end{bmatrix}_{(x^0,\lambda^0)}$$

偏 Hesse 矩阵

注：当 $\tilde{H}_L(x^0,\lambda^0)$ 不定时，$f|_M$ 在 $x^0$ 处仍有可能取到极值（有例子）

数学分析(3)

# 第十二章 多元微分学

## §12.1 多元函数与映射的可微性

符号：$A^\circ$ 内部，$A'$ 聚点集，$\partial A$ 边界，$\bar{A}$ 闭包

$x^\circ = \begin{pmatrix} x_1^\circ \\ \vdots \\ x_n^\circ \end{pmatrix}$

### 一、多元函数的全微分与导数

定义 12.1.1

设有 $n$ 元函数 $y=f(x)$，$x\in D(f)\subseteq \mathbb{R}^n$，

并设 $x^\circ \in D(f)^\circ$。

如果存在 $a\in\mathbb{R}^n$ 使得

$$f(x^\circ+\Delta x)-f(x^\circ)=\langle a,\Delta x\rangle+o(\|\Delta x\|)\quad(\|\Delta x\|\to 0),$$

那么称 $f$ 在 $x^\circ$ 处可微或可导，称 $\langle a,\Delta x\rangle$ 为 $f$ 在 $x^\circ$ 处的全微分（简称微分），记作 $dy|_{x^\circ}=\langle a,\Delta x\rangle$ 或 $df(x^\circ)=\langle a,\Delta x\rangle$

并称线性函数

$$\Delta x\in\mathbb{R}^n \xmapsto{a} \langle a,\Delta x\rangle\in\mathbb{R}$$

为 $f$ 在 $x^\circ$ 处的导数，记作 $f'(x^\circ)$ 或 $Df(x^\circ)$，即

$$f'(x^\circ)\equiv Df(x^\circ):=\langle a,\cdot\rangle\equiv a^T$$（行向量，而 $a$ 作列向量看）

如果 $f$ 在开集 $D\subseteq D(f)$ 的每一点处都可微，则称 $f$ 是 $D$ 上的可微函数或可导函数

命题 12.1.0

若 $f$ 在 $x^\circ\in D(f)^\circ$ 处可微，则 $f$ 在 $x^\circ$ 处的导数 $f'(x^\circ)$ 是唯一的，且 $df(x^\circ)=f'(x^\circ)\Delta x$.

### 二、多元函数的偏导数与梯度

$e^i = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \leftarrow i$

定义 12.1.2

设有 $n$ 元函数 $y=f(x)$，$x\in D(f)\subseteq\mathbb{R}^n$，并设

对 $x^\circ\in D(f)$ 和自然数 $j$ $(1\le j\le n)$，存在 $\delta>0$，使得

$\forall t\in(-\delta,\delta)$，$x^\circ+te^j\in D(f)$，其中 $e^j=(0,0,\cdots,0,\underset{\text{第}j\text{位}}{1},0,\cdots,0)$

如果极限 $\lim\limits_{t\to0}\dfrac{f(x^\circ+te^j)-f(x^\circ)}{t}$

收敛，则称它为 $f$ 在 $x^\circ$ 处关于第 $j$ 个坐标 $x^j$ 的偏导数，

记作 $\dfrac{\partial f}{\partial x_j}(x^\circ)$，$\dfrac{\partial f}{\partial x_j}\Big|_{x^\circ}$，$f_{x_j}'(x^\circ)$，$f_j'(x^\circ)$，或 $\dfrac{\partial y}{\partial x_j}(x^\circ)$，$y_{x_j}'(x^\circ)$ 等