17.3 注释

张志聪

2025 年 5 月 12 日

1. 把 f 写成 $f=(f_1,f_2,\cdots,f_m)$，那么容易得出

$$\frac{\partial f}{\partial x_j}(x_0)=(\frac{\partial f_1}{\partial x_j}(x_0),...,\frac{\partial f_m}{\partial x_j}(x_0))$$

证明：

$$\begin{aligned}
\frac{\partial f}{\partial x_j}(x_0)&=\lim_{t\to 0;t\neq 0,x_0+te_j\in E}\frac{f(x_0+te_j)-f(x_0)}{t}\\
&=\lim_{t\to 0;t\neq 0,x_0+te_j\in E}\frac{(f_1(x_0+te_j),...,f_m(x_0+te_j))-(f_1(x_0),...,f_m(x_0))}{t}\\
&=\lim_{t\to 0;t\neq 0,x_0+te_j\in E}\frac{(f_1(x_0+te_j)-f_1(x_0),...,f_m(x_0+te_j)-f_m(x_0))}{t}\\
&=\lim_{t\to 0;t\neq 0,x_0+te_j\in E}(\frac{f_1(x_0+te_j)-f_1(x_0)}{t},...,\frac{f_m(x_0+te_j)-f_m(x_0)}{t})\\
&=(\frac{\partial f_1}{\partial x_j}(x_0),...,\frac{\partial f_m}{\partial x_j}(x_0))
\end{aligned}$$

2. 设 E 是 $\mathbb{R}^n$ 的子集，$f:E\to\mathbb{R}^m$ 是一个函数，F 是 E 的子集，并设 x_0 是 F 的内点。如果 f 在 F 上的全体偏导数 $\frac{\partial f}{\partial x_j}$ 都存在，并且它们在 x_0 处都是连续的。那么，把 f 写成 $f_1,f_2,\cdots,f_m$，（每一个 f_i 都是从 E 到 $\mathbb{R}$ 的函数）。对变量 x_1 使用平均值定理（推论 *10.2.9* 中值定理），存在一个介于 0 和 v_1 之间的 t_i，使得

$$f_i(x_0+v_1e_1)-f_i(x_0)=\frac{\partial f_i}{\partial x_1}(x_0+t_ie_1)v_1$$

我的问题是，按照平均值定理，等式应该是

$$f_i(x_0+v_1e_1)-f_i(x_0)=\frac{\partial f_i}{\partial x_1}(x_0+t_ie_1)v_1e_1$$

才对吧！因为 $(x_0+v_1e_1)-x_0=v_1e_1$。

证明：

我的理解是有问题的。

一方面，如果改为

$$f_i(x_0+v_1e_1)-f_i(x_0)=\frac{\partial f_i}{\partial x_1}(x_0+t_ie_1)v_1e_1$$

此时，等式左侧是标量，右侧是向量，存在矛盾。

另一方面，$f_i:\mathbb{R}^n\to\mathbb{R}$，对变量 x_1 求偏导数，即固定了 $x_2,x_3,\cdots,x_n$，仅让 x_1 变化，此时 f_i 可以看做单变量函数：

$$g(t)=f_i(x_0+te_1)\ t\in[0,v_1]$$

于是

$$g(v_1)-g(0)=g'(t_i)(v_1-0)$$

因为

$$\begin{aligned}g'(t_0)&=\lim_{t\to t_0}\frac{g(t)-g(t_0)}{t-t_0}\\&=\lim_{t\to t_0}\frac{f_i(x_0+te_1)-f_i(x_0+t_0e_1)}{t-t_0}\end{aligned}$$

令 $h=t-t_0$，则当 $t\to t_0$ 时，$h\to 0$。代入后：

$$g'(t_0)=\lim_{h\to 0}\frac{f_i(x_0+(t_0+h)e_1)-f_i(x_0+t_0e_1)}{h}$$

又因为

$$\begin{aligned}\frac{\partial f_i}{\partial x_1}(x_0+t_0e_1)&=\lim_{t\to 0;t\neq 0,x_0+te_1\in F}\frac{f_i(x_0+t_0e_1+te_1)-f_i(x_0+t_0e_1)}{t}\\&=\lim_{t\to 0;t\neq 0,x_0+te_1\in F}\frac{f_i(x_0+(t_0+t)e_1)-f_i(x_0+t_0e_1)}{t}\end{aligned}$$

由 t_0 的任意性可得

$$g'(t) = \frac{\partial f_i}{\partial x_1}(x_0 + te_1)$$

综上可得，

$$\begin{aligned} g(v_1) - g(0) &= f_i(x_0 + v_1 e_1) - f_i(x_0) \\ &= g'(t_i)v_1 \\ &= \frac{\partial f_i}{\partial x_1}(x_0 + t_i e_1)v_1 \end{aligned}$$

3. 特别地，如果 $E \subseteq \mathbb{R}^m$，$f : E \to \mathbb{R}$ 是一个实值函数，并且 f 在 x_0 处的梯度 $\nabla f(x_0)$ 被定义为 n 维行向量

$$\nabla f(x_0) := \left(\frac{\partial f}{\partial x_1}(x_0), \cdots, \frac{\partial f}{\partial x_n}(x_0)\right)$$

，那么只要 x_0 是某个梯度存在且连续的区域的内点，我们就有熟知的公式

$$D_v f(x_0) = v \cdot \nabla f(x_0)$$

说明：
这里是点乘，注意中间有个点。

4. 为什么导数矩阵是书中定义的格式（*P365*）$D_f(x_0) = (\frac{\partial f_i}{\partial x_j})_{1 \le i \le m; 1 \le j \le n}$。

按照可微性（定义 17.2.2）可知，$f'(x_0)$ 是一个线性变换，且 $f'(x_0) : \mathbb{R}^n \to \mathbb{R}^m$。于是按照引理 17.1.13 可知，存在一个 $m \times n$ 矩阵 A 使得 $f'(x_0) = L_A$。所以，$D_f(x_0) = A$ 应该是一个 $m \times n$ 矩阵，符合书中的定义。

不妨设

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

因为

$$
\begin{aligned}
\frac{\partial f}{\partial x_j}(x_0) &= f'(x_0)e_j \\
&= L_A e_j \\
&= \left(\sum_{k=1}^{n} a_{ik}x_k\right)_{1\leq i\leq m} \\
&= (a_{ij})_{1\leq i\leq m}
\end{aligned}
$$

这表示的是矩阵 A 的 j 列。

又因为

$$\frac{\partial f}{\partial x_j}(x_0) = (\frac{\partial f_1}{\partial x_j}(x_0), ..., \frac{\partial f_m}{\partial x_j}(x_0))$$

综上可得，

$$a_i j = \frac{\partial f_i}{\partial x_j}(x_0)$$

与书中定义一致。