# 微分形式导论

[美] M. Schreiber 著

白正国 译



人民教育出版社

# 微分形式导论

[美] M. Schreiber 著
白正国 译

人民教育出版社

本书力求有启发性地介绍微分形式．内容包括偏微分、微分形式、高维积分、外微分、在 $R^3$ 的向量运算、极值、积分几何等七章，还有三个附录，内容包括在一流形上的体积元素，形式的代数学以及关于 Curl[2] 的一个注意．

本书可供综合大学和师范院校数学系师生作教学参考用书．

**微分形式导论**

［美］M. Schreiber 著

白 正 国 译

*

人民教育出版社出版

新华书店北京发行所发行

北京印刷一厂印装

开本 850×1168 1/32　印张 3.5　字数 83,000

1980 年 8 月第 1 版　1981 年 2 月第 1 次印刷

印数 00,001—10,500

书号 13012·0510　定价 0.33 元

# 序

微分形式的一些实用知识深刻地影响着微积分学及其发展，应不迟延地把它列入课程中. 另一方面，关于微分形式的系统论证需要拓扑学和代数的工具，这对于低年级的大学生来说是困难的. 近年来有关高等微积分的一些教科书用到微分形式. 其中我们指出 Fleming[2], Nickerson-Spencer-Steenrod [3]和Spivak [6]的书作为代表①. 这些书虽然照顾到读者的程度，却不能减轻他们在这方面的负担，因为它们对此所提供的恰恰是全面地、一般性地、正规的论述. 因此，在这类书和传统的高等微积分之间有着一道鸿沟. 最近，我们借提供高等微积分的入门课程的机会，尝试用最少的工具和非常少的预备知识来表达微分形式的技巧. 这本书就是这个试验的结果.

我们要使该书的叙述成为有启发性的和具体化的. 概言之，我们把一个微分形式看作一个多维的被积式，如此它将自动地服从变数变换的运算规律. 它的积分区域(流形)是欧氏空间内明确给定的"曲面". 形式的微分（外微分）是函数的微分明显的推广，依赖这种运算完善了我们的论证. 为了避免在一般 Stokes 公式的正确证明中用到几何学方面的和一些不太简单的技巧起见，我们替代以简短的说得通的论证，希望它会具有吸引力和说服力. 这是我们为了把该书内容保持在一个初等水平所作的若干简略化中的一个项目.

该书的预备知识限于标准的微积分学初级教程稍稍加上一点其它知识. 这些知识并不是很专门的，例如 $k$ 维的欧氏空间，$k$-$k$

---

①括弧内的数字指的是书末的参考书目.

阶矩阵和它们相乘法则，关于 $k$-$k$ 阶行列式的简单知识等．在目前对正规的大学低年级学生，看来一般都具有这些知识。线性代数除了第六章一个地方外，一般是不需要的，在那里我们要把一个实的对称矩阵对角线化．关于这个定理和关于代数方面的其它若干知识（如上面所说的），我们指定 Schreier-Sperner 的教科书[5]作为参考书．本书在这些问题上的论述都不需要预备知识．在解析方面，我们指定 Courant[1]作为参考书．在设计本书时，我们力求除了这两本书外不需要参考其它的书，勤奋的读者在此基础上将可以顺利地自学．但本书较合适的用途是作为大学二年级学生或现代高等微积分的入门课程的一部分．

各章的内容从目录表便可一目了然，只有两个例外：在 6.2 我们给出关于几何平均和算术平均的定理和在 7.3 我们证明了等周不等式．我们用的记号都是标准的，列在记号页上．记号 $n.m(k)$ 是指 $n.m$ 节的公式 $(k)$．

我愉快地感谢：对 P. A. Griffiths，他鼓励了这个计划并建议本书包含积分几何的若干内容；对 Mary Ellen O'Brien，她给手稿打字；对学生们，他们自愿参加了这一试验．

M. Schreiber

1977 年 6 月 21 日

# 记　　号

$a \in A$　　$a$ 是集 $A$ 的一个元素

$\boldsymbol{x}$　　一个向量

$\|\boldsymbol{x}\|$　　$\boldsymbol{x}$ 的长度

$\boldsymbol{x} \cdot \boldsymbol{y}$　　$\boldsymbol{x}$ 和 $\boldsymbol{y}$ 的数量积

$\boldsymbol{x} \times \boldsymbol{y}$　　$\boldsymbol{x}$ 和 $\boldsymbol{y}$ 的向量积

$\binom{n}{k}$　　二项式系数 $\binom{n}{k} = \frac{n!}{(n-k)!k!}$

$R$　　实数集

$R^k$　　$k$ 维的欧氏空间

$f: R^n \to R^m$　　以 $n$ 向量为变量的 $m$ 向量值的函数

$Z(\phi)$　　$\phi: R^k \to R^1$ 的在 $R^k$ 中满足 $\phi = 0$ 的集

$f \circ g$　　函数 $f$ 和 $g$ 的复合 $f \circ g(x) = f(g(x))$.

$\omega \wedge \tau$　　微分形式 $\omega$ 和 $\tau$ 的楔积

$\Lambda^r$　　$r$ 次形式的空间

$W$　　向量场的集合

$S$　　数量场的集合

$|T|$　　矩阵 $T$ 的行列式

${}^tT$　　矩阵 $T$ 的转置

$\mathrm{tr}T$　　矩阵 $T$ 的迹

# 目　录

# 第一章　偏　微　分

## 1.1　偏导数

我们记实的有序 $k$-组 $\boldsymbol{x}=(x_1, x_2, \cdots, x_k)$ 为 $R^k$. 这样的一个 $k$-组叫做一个 $k$ 向量，数 $x_1$, $x_2$, $\cdots$, $x_k$ 是它的分量. $k$ 向量的相加和向量与数量的乘积运算是按逐个分量函数进行的，这事实在平面和三维空间已为大家所熟知. 在我们的记号下，平面和三维空间分别记为 $R^2$ 和 $R^3$.

$k$ 向量 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 的内积（数量积）$\boldsymbol{x}\cdot\boldsymbol{y}=\Sigma x_i y_i$ 决定了 $R^k$ 中的长度和交角，叙述如下. $\boldsymbol{x}$ 的长度 $\|\boldsymbol{x}\|$ 是 $\|\boldsymbol{x}\|=\sqrt{\boldsymbol{x}\cdot\boldsymbol{x}}$, $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 间的交角 $\theta$ 由式子（余弦定律）$\cos\theta=\boldsymbol{x}\cdot\mathbf{y}/\|\boldsymbol{x}\|\cdot\|\mathbf{y}\|$ 所定义. 这些是在 $R^2$ 和 $R^3$ 中的对应结构对于 $R^k$ 的恰当的类似. 特别，$\boldsymbol{x}\in R^k$ 当 $\|\boldsymbol{x}\|=1$ 时，称为单位向量，并且 $k$ 向量 $\boldsymbol{x}, \boldsymbol{y}$ 当 $\boldsymbol{x}\cdot\boldsymbol{y}=0$ 时是正交的.

一个 $n$ 变量的函数可以看做以一个 $n$ 向量作为变量的函数. 我们也将涉及取向量值的函数. 记号 $f$: $R^n\to R^m$ 表示 $f$ 是以一个 $n$ 向量作为变量而取 $m$ 向量值的函数. 由于 1 向量是一个数，函数 $f$: $R^n\to R^1$ 是 $n$ 变量的一个数量值的函数.

取在 $R^k$ 的单位向量 $\boldsymbol{e}_1=(1, 0, \cdots, 0)$, $\boldsymbol{e}_2=(0, 1, 0, \cdots, 0)$, $\cdots$, $\boldsymbol{e}_k=(0, 0, \cdots, 1)$ 作为正交参考标架. 这些向量所在的直线是 $R^k$ 中一个笛氏坐标系的轴. 每点 $\boldsymbol{x}\in R^k$ 关于这个参考标架有唯一的表达式 $\boldsymbol{x}=\Sigma x_i\boldsymbol{e}_i$, 它关于标架的笛氏坐标是 $(x_1, x_2, \cdots, x_k)$. 对每一坐标方向附上一个偏微分运算 $\dfrac{\partial}{\partial x_i}$, 它对一个向量变量的数量值函数 $f$: $R^k\to R^1$ 起这样的作用:

$$\left[\frac{\partial}{\partial x_i}f\right](\boldsymbol{x})=\lim_{\delta\to 0}\frac{f(\boldsymbol{x}+\delta\boldsymbol{e}_i)-f(\boldsymbol{x})}{\delta} \tag{1}$$

容易看到，对一函数 $f$: $R^k\to R^1$ 施行运算$\frac{\partial}{\partial x_i}$的结果是另一同类型的函数$\left[\frac{\partial}{\partial x_i}f\right]$: $R^k\to R^1$. 在不会混淆的情况下，我们把这新函数简记为 $f_i$. 即

$$f_i(\boldsymbol{x})=\left[\frac{\partial}{\partial x_i}f\right](\boldsymbol{x}) \tag{2}$$

$f_i$ 称为 $f$ 的第 $i$ 偏导数. 它的几何意义如下. 对已给 $f$: $R^k\to R^1$, 我们可以把方程

$$z=f(\boldsymbol{x}) \tag{3}$$

定义为 $R^{k+1}$ 中的一个 $k$ 维曲面. 如果对点 $\boldsymbol{x}\in R^k$, 除第 $i$ 坐标外固定其余所有坐标，而令第 $i$ 坐标 $x_i$ 任意变动，这些点画成 $R^k$ 的一直线，它通过 $\boldsymbol{x}$ 且和 $\boldsymbol{e}_i$ 平行. 这些点的坐标是$(x_1, x_2, \cdots, x_{i-1}, t, x_{i+1}, \cdots, x_k)$, $-\infty<t<+\infty$. 它们在 $f$ 下的象是在曲面 (3)上的 1 维曲线(具有一个自由度，即 $t$ 的变动). 作为 $t$ 的函数的这条曲线的斜率是 $f_i(x_1, x_2, \cdots, x_{i-1}, t, x_{i+1}, \cdots, x_k)$.

由于函数 $f$: $R^k\to R^1$ 的偏导数 $f_i$ 仍然是同一类型的函数，它们也可以偏微分. $f_i$ 关于第 $j$ 变量的微分结果可以记为 $f_{ij}$. 按照标准记号是

$$f_{ij}=\frac{\partial^2}{\partial x_j\partial x_i}f=\frac{\partial}{\partial x_j}\left[\frac{\partial}{\partial x_i}f\right] \tag{4}$$

必须注意下指标的次序是可以交换的. 由于 $f_{ij}$ 又是同一类型的函数，它们可以偏微分，记为 $f_{ijk}$. 按此进行，$f_i$ 称为一阶偏导数，$f_{ij}$ 为二阶偏导数，如此等等. 我们已经假定在讨论中所涉及的所有极限(它们都具有(1)的一般形状)都是存在的. 这就是说，在原则上有 $k$ 个一阶偏导数 $f_1, f_2, \cdots, f_k$; $k^2$ 个二阶偏导数 $f_{11}, f_{12}, \cdots,$

$f_{1k}, f_{21}, f_{22}, \cdots, f_{kk}$; $k^3$ 个三阶偏导数；如此等等. 另一方面，应用平均值定理不难指出，当所有的偏导数都是所含变量的连续函数时，则微分的次序是无关紧要的①. 例如，如果 $f_{12}$ 和 $f_{21}$ 都是连续函数，则 $f_{12}=f_{21}$；同样 $f_{112}=f_{121}=f_{211}$，如果三者都是连续的. 因此，如果高阶偏导数都是连续函数，则它们不相同的个数便显著地减少了. 混合高阶偏导数相等的规律可以叙述如次：如果两个高阶偏导数（都是连续函数）含有相同的指标且有相同的重复数，则二者是相等的.

## 1.2 可微分性，链的法则

由定义，一个实变量的实函数 $f$: $R^1 \to R^1$ 在 $x$ 是可微分的，如果

$$\frac{f(x+h)-f(x)}{h}=f'(x)+\varepsilon(h) \tag{1}$$

$$\varepsilon(h) \to 0 \quad \text{当 } h \to 0 \tag{2}$$

这就是说，当误差 $\eta(h)=h\cdot\varepsilon(h)$ 消失得比 $h$ 要快时，$\lim\limits_{h\to 0}\frac{\eta(h)}{h}=0$，增量

$$(\Delta f)(x, h)=f(x+h)-f(x)$$

可以随心所欲地趋近于微分

$$(df)(x, h)=f'(x)\cdot h \tag{3}$$

注意，微分是增量 $h$ 的一个线性函数.

设对一已给函数 $f$: $R^1 \to R^1$ 和一已给 $x \in R^1$，存在一常数 $A$ 和一函数 $\alpha$: $R^1 \to R^1$ 使得

$$\begin{gathered}(\Delta f)(x, h)=Ah+\alpha(h) \\ \lim_{h\to 0}\frac{\alpha(h)}{h}=0\end{gathered} \tag{4}$$

①参看 Courant [1], vol. II, *pp.* 55—58.

这就是说，增量$(\Delta f)(x, h)$可以用$h$的一个线性函数逼近到所要的精确性. 函数$f$的这种性质蕴涵着$f$在$x$是可微分的，并且

$$f'(x)=A$$

把前面两段放在一起，便可看出，对单变量函数来说，可微分性和线性逼近是一样的. 对多变量函数可微分性的推广可由推广线性逼近的概念到多变量来做出，叙述如下.

一函数$f$: $R^k\to R^1$ 在 $\boldsymbol{x}\in R^k$ 称为是(定义)可微分的，如果存在常数$A_1, A_2, \cdots, A_k$和一函数$\alpha$: $R^k\to R^1$使得

$$(\Delta f)(\boldsymbol{x}, \boldsymbol{h})=\sum A_i h_i+\alpha(\boldsymbol{h}) \tag{5}$$

$$\lim_{\boldsymbol{h}\to\boldsymbol{0}}\frac{\alpha(\boldsymbol{h})}{\|\boldsymbol{h}\|}=0 \tag{6}$$

其中$\boldsymbol{h}=(h_1, h_2, \cdots, h_k)$是增量向量，而$(\Delta f)(\boldsymbol{x},\boldsymbol{h})=f(\boldsymbol{x}+\boldsymbol{h})-f(\boldsymbol{x})$是$f$的对应增量.

我们把(4)的线性函数$Ah$用增量向量$\boldsymbol{h}$的分量$h_1, \cdots, h_k$的一个线性函数来替代，而“$h\to0$”成为“$\boldsymbol{h}\to\boldsymbol{0}$”. 这是(4)的恰当的类似. 注意对任何$\boldsymbol{x}\in R^k$有

$$|x_j|\leqslant\|\boldsymbol{x}\|\leqslant\sqrt{k}\cdot\max\{|x_1|, \cdots, |x_k|\} \tag{7}$$

这是由于$x_j^2\leqslant\sum x_i^2\leqslant k\cdot\max\{x_1^2, \cdots, x_k^2\}$的缘故. 因此，$\boldsymbol{x}$的所有分量是小的当且只当$\|\boldsymbol{x}\|$是小的，并且$\boldsymbol{h}\to\boldsymbol{0}$当且只当$\|\boldsymbol{h}\|\to0$，当且只当，对每个$j$, $h_j\to0$.

设$f$在这意义下是可微分的，在(5)式置$\boldsymbol{h}=\delta\boldsymbol{e}_i$，使得

$$\frac{f(\boldsymbol{x}+\delta\boldsymbol{e}_i)-f(\boldsymbol{x})}{\delta}=A_i+\frac{\alpha(\boldsymbol{h})}{\delta}$$

由此当$\delta\to0$(即$\boldsymbol{h}\to\boldsymbol{0}$)极限是存在的，并且$A_i=f_i(\boldsymbol{x})$. 这就是说，上面所定义的可微分性蕴涵着一阶偏导数的存在性. 其逆，如$f$在一点$\boldsymbol{x}$的近傍有连续的一阶偏导数，我们现在即将证明，$f$在$\boldsymbol{x}$是可微分的. 这证明在记号上看来似乎是麻烦的，不过它的概念非常简单：我们把从$\boldsymbol{x}$到$\boldsymbol{x}+\boldsymbol{h}$的变化分成若干步骤而每一步只

包含着一个变量的变化，这样，定义1.1(1)和微分的基本性质(1)、(2)便可使用了．以下是证明．为了建立(5)、(6)，置 $\boldsymbol{h}=\sum_{i=1}^{k}h_i\boldsymbol{e}_i$，并且把$(\Delta f)(\boldsymbol{x},\ \boldsymbol{h})$分解为$(\Delta f)(\boldsymbol{x},\ \boldsymbol{h})=f(\boldsymbol{x}+\Sigma h_i\boldsymbol{e}_i)-f(\boldsymbol{x})=\sum_{j=1}^{k}\left\{f(\boldsymbol{x}+\sum_{i=j}^{k}h_i\boldsymbol{e}_i)-f(\boldsymbol{x}+\sum_{i=j+1}^{k}h_i\boldsymbol{e}_i)\right\}$，其中最后一项$(j=k)$应解释为$\{f(\boldsymbol{x}+h_k\boldsymbol{e}_k)-f(\boldsymbol{x})\}$．现在应有$f(\boldsymbol{x}+\sum_{i=j}^{k}h_i\boldsymbol{e}_i)-f(\boldsymbol{x}+\sum_{i=j+1}^{k}h_i\boldsymbol{e}_i)=h_jf_j(\boldsymbol{x}+\sum_{i=j+1}^{k}h_i\boldsymbol{e}_i)+\alpha_j$，而由1.1(1)和(1)，(2)当$h\to 0$时，$\alpha_j/h_j\to 0$．由一阶偏导数的连续性的假定，我们有$h_jf_j(\boldsymbol{x}+\sum_{i=j+1}^{k}h_i\boldsymbol{e}_i)=h_jf_j(\boldsymbol{x})+h_j\varepsilon_j$，而当$h_j\to 0$时$\varepsilon_j\to 0$．因此$(\Delta f)(\boldsymbol{x},\boldsymbol{h})=\sum_{j=1}^{k}h_jf_j(\boldsymbol{x})+\sum_{j=1}^{k}\{\alpha_j+h_j\varepsilon_j\}$．利用(7)的第一部分，我们看到误差项$\sum_{j=1}^{k}\{\alpha_j+h_j\varepsilon_j\}$满足(6)，证毕．

下面的记号是醒目的．置增量向量 $\boldsymbol{h}$ 为 $d\boldsymbol{x}=(dx_1,\ dx_2,\ \cdots,\ dx_k)$，且写(5)中的线性项为

$$(df)(\boldsymbol{x},d\boldsymbol{x})=\Sigma f_i(\boldsymbol{x})dx_i \tag{8}$$

我们要强调地指出，这个量是 $\boldsymbol{x}$ 和 $d\boldsymbol{x}$ 的函数，称为函数 $f$: $R^k\to R^1$ 的微分，并且注意到它是对一个单变量函数的相应公式(3)的类似式子．关于单变量或多变量函数的可微分性的定义，现在实质上可以用相同的格式给出：设单变量函数或多变量函数 $f$ 在一点近傍的增量 $\Delta f$ 分别在指定的精确度(4)或(6)的范围内逼近于对应的微分，则 $f$ 在这点是可微分的．对单变量函数来说，为达到这个精确度，导数的存在是充分条件，而对多变量函数我们已经证明

了一阶偏导数的存在和连续是充分条件.

设函数 $\boldsymbol{g}$: $R^1 \to R^k$ 具有分量 $g^i$: $R^1 \to R^1$, $i=1, 2, \cdots, k$, 即 $\boldsymbol{g}(t)=(g^1(t), g^2(t), \cdots, g^k(t))$, $t \in R^1$. 设每一 $g^i$ 是可微分的, $(\Delta g^i)(t, h)=h \cdot \dfrac{dg^i}{dt}+\alpha_i(h)$, 当 $h \to 0$ 时 $(\alpha_i(h)/h) \to 0$, 则 $(\Delta \boldsymbol{g})(t, h)=h \cdot \dfrac{d\boldsymbol{g}}{dt}+\boldsymbol{\alpha}(h)$, 式中 $\dfrac{d\boldsymbol{g}}{dt}=\left(\dfrac{dg^1}{dt}, \cdots, \dfrac{dg^k}{dt}\right)$, 且 $\boldsymbol{\alpha}(h)=(\alpha_1(h), \cdots, \alpha_k(h))$, 当 $h \to 0$ 显然有 $(\|\boldsymbol{\alpha}(h)\|/h) \to 0$. 在这情况下(即每一 $g^i$ 是可微分的)说 $\boldsymbol{g}$ 是可微分的, 这是这个术语一个自然的推广.

设 $\boldsymbol{g}$: $R^1 \to R^k$ 和 $f$: $R^k \to R^1$ 都是可微分的, 则其复合函数 $f \circ \boldsymbol{g}$: $R^1 \to R^1$ 是可微分的, 并且

$$\frac{d}{dt}(f \circ \boldsymbol{g})(t)=\Sigma f_i(\boldsymbol{g}(t)) \cdot \frac{d}{dt} g^i(t) \tag{9}$$

这公式可以从 (8) 猜出, 述之如次. 在(8)中置 $\boldsymbol{x}=\boldsymbol{g}(t)$, 得出 $df(\boldsymbol{g}(t), d\boldsymbol{g})=\Sigma f_i(\boldsymbol{g}(t)) dg^i$, 再除以 $dt$ 便仿佛得到(9)了. 现在给出(9)的证明. 由 $\boldsymbol{g}$ 的可微分性应有 $\boldsymbol{g}(t+h)=\boldsymbol{g}(t)+h \cdot \dfrac{d}{dt}\boldsymbol{g}(t)+\boldsymbol{\alpha}(h)$. 而 $(f \circ \boldsymbol{g})(t+h)=f(\boldsymbol{g}(t+h))$, 因此由 $f$ 的可微分性我们有 $(f \circ \boldsymbol{g})(t+h)=(f \circ \boldsymbol{g})(t)+\Sigma f_i(\boldsymbol{g}(t)) \cdot \left\{h \dfrac{d}{dt} g^i(t)+\alpha_i(h)\right\}+\beta(\Delta \boldsymbol{g}(t))=(f \circ \boldsymbol{g})(t)+h \Sigma f_i(\boldsymbol{g}(t)) \cdot \dfrac{d}{dt} g^i(t)+\Sigma f_i(\boldsymbol{g}(t)) \cdot \alpha_i(h)+\beta(\Delta \boldsymbol{g}(t))$. 从此看到, $\Delta(f \circ \boldsymbol{g})(t, h)$ 以误差 $\Sigma f_i(\boldsymbol{g}(t)) \alpha_i(h)+\beta(\Delta \boldsymbol{g}(t))$ 逼近于 $h\left\{\Sigma f_i(\boldsymbol{g}(t)) \cdot \dfrac{d}{dt} g^i(t)\right\}$. 但由于当 $h \to 0$ 时, $\dfrac{1}{h} \alpha_i(h) \to 0$, 因而当 $h \to 0$ 成立 $\dfrac{1}{h} \Sigma f_i(\boldsymbol{g}(t)) \cdot \alpha_i(h) \to 0$; 同时当 $h \to 0$ 有 $\dfrac{1}{h} \beta(\Delta \boldsymbol{g}(t))=\dfrac{\beta(\Delta \boldsymbol{g}(t))}{\|\Delta \boldsymbol{g}(t)\|} \dfrac{\|\Delta \boldsymbol{g}(t)\|}{h} \to 0$, 这是因为从 $f$ 的可微

分性得出第一因子应消失，而第二因子等于 $\left\|\frac{d}{dt}\boldsymbol{g}(t)+\frac{\boldsymbol{\alpha}(h)}{h}\right\|$，当 $h\to 0$ 它趋近于 $\left\|\frac{d}{dt}\boldsymbol{g}(x)\right\|$. 因此在要求的比率(4)下总误差消失了；这是说 $f\circ\boldsymbol{g}$ 是可微分的并且 (9) 成立. 这公式是链的法则对多变量的推广，我们在后面还要用到它.

## 1.3 泰乐定理

对已给 $f$: $R^k\to R^1$，定义 $F$: $R^1\to R^1$ 为

$$F(t)=f(\boldsymbol{x}+t\boldsymbol{h}) \tag{1}$$

式内 $\boldsymbol{x}$, $\boldsymbol{h}\in R^k$ 是任意的固定值. 把可微分性和收敛性暂置不论，$F$ 的泰乐级数是

$$F(t)=\Sigma\frac{t^n}{n!}F^{(n)}(0) \tag{2}$$

由 1.2(9)的链的法则有

$$F'(t)=\Sigma_i f_i(\boldsymbol{x}+t\boldsymbol{h})\cdot h_i$$

$$F''(t)=\Sigma_{i,j}f_{ij}(\boldsymbol{x}+t\boldsymbol{h})h_ih_j$$

一般

$$F^{(n)}(t)=\sum_{i_1,\cdots,i_n}f_{i_1,\cdots,i_n}(\boldsymbol{x}+t\boldsymbol{h})h_{i_1}h_{i_2}\cdots h_{i_n}$$

因此

$$F^{(n)}(0)=\sum_{i_1,\cdots,i_n}f_{i_1,\cdots,i_n}(\boldsymbol{x})h_{i_1}h_{i_2}\cdots h_{i_n} \tag{3}$$

把(3)代入(2)然后置 $t=1$ 便得

$$f(\boldsymbol{x}+\boldsymbol{h})=f(\boldsymbol{x})+\Sigma_i f_i(\boldsymbol{x})h_i+\Sigma_{i,j}f_{ij}(\boldsymbol{x})h_ih_j+\cdots \tag{4}$$

在这情形，链的法则指出若 $f$ 有直到 $n$ 阶的连续偏导数，则 $F$ 有直到 $n$ 阶的连续导数. 因此，若 $f$ 有直到 $n$ 阶的连续偏导数，则 $f$ 有直到 $n$ 阶的形(4)的展开式带一个误差项，写为 $R_n(\boldsymbol{x},\boldsymbol{h})$. 为完成(4)的讨论，我们需要对 $R_n(\boldsymbol{x},\boldsymbol{h})$ 当 $\boldsymbol{h}$ 接近于 0 时的估值. 对我们

的大部分工作而言我们只需对 $R_1$ 估值，即估计线性（或一阶）逼近的误差；至于对 $R_n$ 的估值方法可按照对 $R_1$ 的情况便不难明白，后者在记号上更为简单而已. 对余项用积分形式表示，$F$ 的第一阶泰乐展开为①

$$F(t)=F(0)+tF'(t)+\int_0^t(t-s)F''(s)ds \tag{5}$$

置 $t=1$，对 $f$ 展开式为

$$f(\boldsymbol{x}+\boldsymbol{h})=f(\boldsymbol{x})+\sum_i f_i(\boldsymbol{x})h_i+\int_0^1(1-s)F''(s)ds \tag{6}$$

因此 $R_1(\boldsymbol{x},\boldsymbol{h})=\int_0^1(1-s)F''(s)ds$. 因为 $F''(s)=\sum_{i,j}f_{ij}(\boldsymbol{x}+s\boldsymbol{h})\cdot h_ih_j$，我们有

$$|R_1(\boldsymbol{x},\boldsymbol{h})|\leqslant\{\max|f_{ij}|\}\cdot[\sum_{i,j}|h_ih_j|]\cdot\frac{1}{2} \tag{7}$$

式内因子 $\frac{1}{2}$ 是从 $(1-s)$ 的积分而来. 应用不等式 1.2(7) 到 $\boldsymbol{h}$ 得 $\sum_{i,j}|h_ih_j|\leqslant k^2\|\boldsymbol{h}\|^2$；再写 $M=\max|f_{ij}|$，便有

$$|R_1(\boldsymbol{x},\boldsymbol{h})|\leqslant\frac{1}{2}M\cdot k^2\cdot\|\boldsymbol{h}\|^2 \tag{8}$$

这蕴涵着所要求的

$$\frac{R_1(\boldsymbol{x},\ \boldsymbol{h})}{\|\boldsymbol{h}\|}\to 0 \qquad 当\ \boldsymbol{h}\to 0 \tag{9}$$

因此，对一具有直到 2 阶连续偏导数的多变量函数 $f$，在线性逼近后的误差项，在(9)的意义下，当 $\boldsymbol{h}\to 0$ 时，比增量向量 $\boldsymbol{h}$ 的长度 $\|\boldsymbol{h}\|$ 消失得更快. 一般，当 $f$ 有直到 $n+1$ 阶的连续偏导数时，误差 $R_n(\boldsymbol{x},\boldsymbol{h})$ 在下面意义下比 $\|\boldsymbol{h}\|^n$ 消失得更快

$$\frac{R_n(\boldsymbol{x},\ \boldsymbol{h})}{\|\boldsymbol{h}\|^n}\to 0 \qquad 当\ \boldsymbol{h}\to 0 \tag{10}$$

①参看 Courant [1], vol. I, $p$. 323.

这事实可由导出(9)的同样方法得到. 带有满足(10)的误差项 $R_n$ 的直到 $n$ 阶的展开式(4)的存在构成一个 $k$ 变量函数 $f$: $R^k \to R^1$ 的泰乐定理.

现在考察一个 $k$ 向量变量的 $k$ 向量值函数 $\boldsymbol{f}$: $R^k \to R^k$. 设 $f^\nu$: $R^k \to R^1$, $\nu=1, 2, \cdots, k$ 是 $\boldsymbol{f}$ 的分量. 即

$$\boldsymbol{f}(\boldsymbol{x})=(f^1(\boldsymbol{x}), f^2(\boldsymbol{x}), \cdots, f^k(\boldsymbol{x}))$$

对 $\boldsymbol{f}$ 的泰乐展开可由对分量 $f^\nu$ 的泰乐展开式集合组成. 高阶的项是复杂的, 但是我们只需要一阶的项, 它是很简洁的. 我们对每一 $f^\nu$ 写出(4), 得知 $\boldsymbol{f}(\boldsymbol{x}+\boldsymbol{h})$ 的一阶逼近是一向量它的第 $\nu$ 分量是 $\sum_i f_i^\nu(\boldsymbol{x}) h_i$. 这事提醒我们做一个矩阵, 它的元素是向量函数 $f^\nu$ 的一阶偏导数. 记这矩阵为 $d\boldsymbol{f}(\boldsymbol{x})$, 且如所知, 它是 $\boldsymbol{f}$ 的**微分**. 按定义有

$$d\boldsymbol{f}(\boldsymbol{x})=((f_i^\nu(\boldsymbol{x}))) \tag{11}$$

式内上(分量)指标是行(横行)指标, 而下(导数)指标是列指标. 因此

$$\boldsymbol{f}(\boldsymbol{x}+\boldsymbol{h})=\boldsymbol{f}(\boldsymbol{x})+d\boldsymbol{f}(\boldsymbol{x})\boldsymbol{h}+\cdots \tag{12}$$

式中第二项是应用矩阵 $d\boldsymbol{f}(\boldsymbol{x})$ 到向量 $\boldsymbol{h}$ 的结果. 这样, 对函数 $\boldsymbol{f}$: $R^k \to R^k$ 有了直到第一阶的通常形状的单变量线性逼近公式.

下面的名词是标准的. 一个函数 $f$: $R^k \to R^1$ 称为一个**数量场**, 并且可以把它设想为对 $R^k$ 的每点 $\boldsymbol{x}$ 附上一个数量 $f(\boldsymbol{x})$. 一个函数 $\boldsymbol{f}$: $R^k \to R^k$ 称为一个**向量场**, 并且可以把它设想为对 $R^k$ 的每点 $\boldsymbol{x}$ 附上一个向量 $\boldsymbol{f}(\boldsymbol{x})$. 利用这些名词, (4)是一个数量场 $f$ 的泰乐展开, 而(12), 我们已经看到是一个向量场 $\boldsymbol{f}$ 到第一阶的展开.

关于向量场微分(11)有一个重要事实, 即两个向量场的复合的微分是它们各自微分的矩阵乘积. 详言之这结论就是说, 设向量场 $\boldsymbol{f}$ 和 $\boldsymbol{g}$ 都是可微分的, 则 $\boldsymbol{f}\circ\boldsymbol{g}$ 也是可微分的并且