17.2 多元微积分中的导数

摘录

1. (从$1\to1$到$n\to m$)在单变量微积分中，我们曾经给过一个函数$f:E\to\mathbb{R}$在$x_0\in E$处微分的定义，即：

$$f'(x_0):=\lim_{x\to x_0;x\in E\setminus\{x_0\}}\frac{f(x)-f(x_0)}{x-x_0}$$

基于我们以前的学习经历，我们也不妨从这个定义出发，尝试定义一个多变量函数$f:E\to\mathbb{R}^m$（其中$E\subseteq\mathbb{R}^n$）在$x_0\in E$处的导数为：

$$f'(x_0):=\lim_{x\to x_0;x\in E\setminus\{x_0\}}\frac{f(x)-f(x_0)}{x-x_0}$$

但是在这里我们遇到了困难，我们很难去定义一个m维向量除以一个n维向量到底要怎么解释，因此我们需要转变思路，我们将f在x_0处的可微性看为f在x_0附近的"**近似于线性性质**"的描述，基于这个前提展开本节的内容。

定义

1. (17.2.2 可微性) 设E是$\mathbb{R}^n$的子集，$f:E\to\mathbb{R}^m$是一个函数，$x_0\in E$是一个点，并设$L:\mathbb{R}^n\to\mathbb{R}^m$是一个线性变换。如果有：

$$\lim_{x\to x_0;x\in E\setminus\{x_0\}}\frac{\|f(x)-(f(x_0)+L(x-x_0))\|}{\|x-x_0\|}=0$$

其中$\|y\|$是y的长度（在l^2度量下），即：

$$\|(y_1,y_2,\dots,y_n)\|=\left(\sum_{i=1}^{n}y_i^2\right)^{1/2}$$

则我们称f**在x_0处是可微的**，并且**导数为L**。

（注：使用原始定义去计算导数太过麻烦（原书有个例子），之后会用更好的方式去计算；如同一元微积分一样，我们也可以将f在x_0处的导数记为$f'(x_0)$（这需要引理17.2.4去保证唯一性），于是$f'(x_0)$是满足$f(x)-f(x_0)\approx f'(x_0)(x-x_0)$的线性变换（这也被称为**牛顿近似**，与命题10.1.7比较）；有时候我们称f'为f的**全导数**（为了和后面的偏导数与方向导数区分），并且全导数f'同**导数矩阵**Df有密切的联系，这个是下一节的内容了）

命题

1. (17.2.2 可微性描述的改写？) 设E是$\mathbb{R}$的子集，$f:E\to\mathbb{R}$是一个函数，且$x_0\in E$且$L\in\mathbb{R}$。那么下面这两个命题是等价的：

> 1. f在x_0处是可微的，并且$f'(x_0)=L$。
> 2. $\displaystyle\lim_{x\to x_0;x\in E\setminus\{x_0\}}\frac{|f(x)-(f(x_0)+L(x-x_0))|}{|x-x_0|}=0$。

（注：这个等价关系相当于给出了近似式$f(x)-f(x_0)=L(x-x_0)$，这个定义看起来与微分差别不大，但是最注意到的是这个命题给出了不使用$x-x_0$作为除数，这样就规避了向量“除法”的问题。然后从这个等价定义出发，我们去思考找出如何有一个对应的L作用于n维向量$x-x_0$，使得m维向量$f(x)-f(x_0)$近似$L(x-x_0)$。从这个角度出发，我们不难想到L应该有一个线性变换的形式）

2. （17.2.4 导数的唯一性）设E是$\mathbb{R}^n$的子集，$f:E\to\mathbb{R}^m$是一个函数，$x_0\in E$是一个点，并设$L_1:\mathbb{R}^n\to\mathbb{R}^m$和$L_2:\mathbb{R}^n\to\mathbb{R}^m$都是线性变换。如果$f$在$x_0$处可微，并且导数为$L_1$的同时还有导数为$L_2$，那么$L_1=L_2$。

   （注：需要注意的是这里强调了x_0是一个内点，这个结论在边界点上面并不一定，书里给了个很极端的单点集的例子）

课后习题

17.2.1 证明引理17.2.1

> 我们知道$\lim\limits_{x\to x_0;x\in E\backslash\{x_0\}}\dfrac{|f(x)-(f(x_0)+L(x-x_0))|}{|x-x_0|}=0$即有：对任意的$\varepsilon>0$，存在$\delta>0$使得对任意的$x\in E\backslash\{x_0\}$满足$|x-x_0|\leq\delta$都有：
>
> $$\frac{|f(x)-(f(x_0)+L(x-x_0))|}{|x-x_0|}\leq\varepsilon\iff|f(x)-(f(x_0)+L(x-x_0))|\leq\varepsilon|x-x_0|$$
>
> 而我们又注意到$x=x_0$处$|f(x)-(f(x_0)+L(x-x_0))|\leq\varepsilon|x-x_0|$是显然成立的，与$f$和$\varepsilon$无关。因此我们可以得到等价关系有：$\lim\limits_{x\to x_0;x\in E\backslash\{x_0\}}\dfrac{|f(x)-(f(x_0)+L(x-x_0))|}{|x-x_0|}=0$当且仅当对任意的$\varepsilon>0$，存在$\delta>0$使得对任意的$x\in E$满足$|x-x_0|\leq\delta$都有$|f(x)-(f(x_0)+L(x-x_0))|\leq\varepsilon|x-x_0|$。然后利用牛顿逼近法（命题10.1.7）我们就可以进一步得到这等价于f在x_0处的导数为L。

17.2.2 证明引理17.2.4（提示：利用反证法。如果$L_1\neq L_2$，那么存在一个向量v使得$L_1v\neq L_2v$，并且这个向量一定不是零向量。（为什么？）然后再利用导数的定义，专门考察$x=x_0+tv$（其中t时一个标量）时的情景来导出矛盾）

> 使用反证法，若$L_1\neq L_2$，那么至少存在一个$v\in\mathbb{R}^n$使得$L_1v\neq L_2v$。特别地，我们可以假定这个v不是零向量，这是因为：
>
> > 由于L_1,L_2都是线性变换，若v是零向量，且对任意其它的$v'\in\mathbb{R}^n$都有$L_1v'=L_2v'$。则我们任取$v_0\in\mathbb{R}^n$，根据线性变换的性质有：
> >
> > $$L_1v_0=L_1(v_0+v)=L_1v_0+L_1v\neq L_2v_0+L_2v=L_2(v_0+v)=L_2v_0$$
> >
> > 导出了矛盾，因此除了零向量以外至少存在一个$v\in\mathbb{R}^n$满足$L_1v\neq L_2v$。
>
> 于是回到可微的定义。根据命题14.1.5，我们特别考虑收敛于x_0的点列$x_n:=x_0+\dfrac{c}{n}v$（其中$n\geq1$，c是一个足够小的正常数满足点列x_n包含于E，由于x_0是内点因此c是存在的），则由于L_1,L_2都是导数应当有极限：
>
> $$\lim_{n\to\infty}\frac{\|f(x_n)-(f(x_0)+L_1(x_n-x_0))\|}{\|x_n-x_0\|}=\lim_{n\to\infty}\frac{\|f(x_n)-(f(x_0)+L_2(x_n-x_0))\|}{\|x_n-x_0\|}=0$$
>
> 也即：
>
> $$\lim_{n\to\infty}\frac{\|f(x_0+\frac{c}{n}v)-f(x_0)-\frac{c}{n}L_1v\|}{\|\frac{c}{n}v\|}=\lim_{n\to\infty}\frac{\|f(x_0+\frac{c}{n}v)-f(x_0)-\frac{c}{n}L_2v\|}{\|\frac{c}{n}v\|}=0$$

> 再结合到v是一个确定的向量，c是一个常数，运用极限定律，我们可以化简为：
>
> $$\lim_{n\to\infty}\left\|nf\left(x_0+\frac{c}{n}v\right)-nf(x_0)-cL_1v\right\|=\lim_{n\to\infty}\left\|nf\left(x_0+\frac{c}{n}v\right)-nf(x_0)-cL_2v\right\|=0$$
>
> 从而利用三角不等式，我们有：
>
> $$\begin{aligned}&\|cL_1v-cL_2v\|\le\left\|cL_1v-nf\left(x_0+\frac{c}{n}v\right)+nf(x_0)\right\|+\left\|nf\left(x_0+\frac{c}{n}v\right)-nf(x_0)-cL_2v\right\|\\\Longrightarrow&\|L_1v-L_2v\|=0\end{aligned}$$
>
> 这与$L_1v\neq L_2v$导出了矛盾，于是反证结束，反证假设不成立，只能有$L_1=L_2$。

本节相关跳转

[实分析 10.1 基本定义]

[实分析 17.3 偏导数和方向导数]