17.4 多元微积分链式法则

摘要

1. （链式法则的应用其一）设$f:\mathbb{R}^n\to\mathbb{R}$和$g:\mathbb{R}^n\to\mathbb{R}$都是可微函数，我们把函数$h:\mathbb{R}^n\to\mathbb{R}^2$定义为$h(x):=(f(x),g(x))$，然后令$k:\mathbb{R}^2\to\mathbb{R}$表示乘法函数$k(a,b):=ab$，然后可以注意到：

$$Dh(x_0)=\begin{pmatrix}\nabla f(x_0)\\ \nabla g(x_0)\end{pmatrix}\qquad Dk(a,b)=(b,a)$$

   然后应用链式法则，我们有：

$$D(k\circ h)(x_0)=(f(x_0),g(x_0))\begin{pmatrix}\nabla f(x_0)\\ \nabla g(x_0)\end{pmatrix}=g(x_0)\nabla f(x_0)+f(x_0)\nabla g(x_0)$$

   然后又有$k\circ h=fg$，因此$D(fg)=\nabla(fg)$，从而这便给出了**乘积法则**：

$$\nabla(fg)=g\nabla f+f\nabla g$$

   类似地，我们可以给出**和法则**$\nabla(f+g)=\nabla f+\nabla g$与**差法则**$\nabla(f-g)=\nabla f-\nabla g$。

2. （链式法则的应用其二）设$T:\mathbb{R}^n\to\mathbb{R}^m$是一个线性变换，由于$T$在每一点$x$处都连续可微且$T'(x)=T$（见习题17.4.1）。因此，对于任意的可微函数$f:E\to\mathbb{R}^n$，$Tf:E\to\mathbb{R}^m$也是可微的，并且可以通过链式法则给出$(Tf)'(x)=T(f'(x_0))$。这事实上相当于一元微积分里面法则$(cf)'=c(f')$（其中$c$是常数）。

   此外还有一个链式法则的特殊情形，设$f:\mathbb{R}^n\to\mathbb{R}^m$是一个可微函数，并且对于每一个$j=1,\dots,n$，$x_j:\mathbb{R}\to\mathbb{R}$都是可微函数。那么有：

$$\frac{\mathrm{d}}{\mathrm{d}t}f(x_1(t),\dots,x_n(t))=\sum_{j=1}^{n}x_j'(t)\frac{\partial}{\partial t}f(x_1(t),\dots,x_n(t))$$

   （注：尝试使用链式法则对函数$f\circ\pi$求导，其中$\pi(t):=(x_1(t),\dots,x_n(t))$，从而我们有$D\pi=(x_i')^T_{1\le i\le n}$与$Df(x_0)=\left(\frac{\partial f}{\partial x_1}(x_0)^T,\frac{\partial f}{\partial x_2}(x_0)^T,\cdots,\frac{\partial f}{\partial x_n}(x_0)^T\right)$，做矩阵乘法即可得到上面的结论）

命题

1. （17.4.1 多元微积分链式法则）设E是$\mathbb{R}^n$的子集，F是$\mathbb{R}^m$的子集。设$f:E\to F$是一个函数，$g:F\to\mathbb{R}^p$是另一个函数，并设x_0是E的内点。如果f在x_0处可微，并且$f(x_0)$是F的内点，同时g在$f(x_0)$处也是可微的，那么$g\circ f:E\to\mathbb{R}^p$在$x_0$处可微，而且还有等式：

$$(g\circ f)'(x_0)=g'(f(x_0))f'(x_0)$$

   （注：应当将这个同一元微积分中的链式法则（定理10.1.15）做对比；作为链式法则和引理17.1.16的一个推论，我们有$D(g\circ f)(x_0)=Dg(f(x_0))Df(x_0)$，也就是说，我们可以用矩阵和矩阵乘法来描述链式法则）

课后习题

17.4.1 设$T:\mathbb{R}^n\to\mathbb{R}^m$是一个线性变换，证明：$T$在任意点$x$处都是连续可微的，并且有$T'(x)=T$。思考：$DT$是什么（连续可微的定义见定义17.5.1）

> 显然根据可微的定义与线性变换的性质我们有：
>
> $$\lim_{y\to x}\frac{\|Ty-Tx-T(y-x)\|}{\|y-x\|}=\lim_{y\to x}\frac{0}{\|y-x\|}=0$$
>
> 因此我们有T在x处可微且全导数$T'(x)=T$。从而对任意的$1\le i\le n$有对变量x_i的偏导为$\frac{\partial T}{\partial x_i}(x)=T'(x)e_i=Te_i$是一个常值函数（因此显然是连续的），对应地导数矩阵有：

> $$DT(x)=\left(\left(\frac{\partial T}{\partial x_1}(x)\right)^T,\dots,\left(\frac{\partial T}{\partial x_n}(x)\right)^T\right)=((Te_1)^T,\dots,(Te_n)^T)$$

17.4.2 设E是$\mathbb{R}^n$的子集。证明：如果函数$f:E\to\mathbb{R}^m$在E的一个内点x_0处可微，那么f也在x_0处连续（提示：利用习题17.1.4）

> 设f在x_0处的全导数为L。
>
> 考虑任意的$\varepsilon>0$，由于f在x_0处可微，因此存在$\sigma>0$使得对任意的$x\in E$满足$0<\|x-x_0\|<\sigma$都有：
>
> $$\|[f(x)-f(x_0)]-L(x-x_0)\|<\varepsilon\|x-x_0\|$$
>
> 而根据习题17.1.4的结论，我们知道存在一个整数M使得$\|Ly\|\leq M\|y\|$对所有的$y\in\mathbb{R}^n$都成立，因此根据度量的三角不等式我们有：
>
> $$\|f(x)-f(x_0)\|\leq\|[f(x)-f(x_0)]-L(x-x_0)\|+\|L(x-x_0)\|<(\varepsilon+M)\|x-x_0\|$$
>
> 此时我们定义$\delta:=\min\left(\sigma,\dfrac{\varepsilon}{\varepsilon+M}\right)$，此时我们得到：
>
> > 对任意的$\varepsilon>0$，存在$\delta>0$使得对任意的$x\in E$满足$0<\|x-x_0\|<\delta$有：
> >
> > $$\|f(x)-f(x_0)\|<(\varepsilon+M)\|x-x_0\|\leq\frac{\varepsilon}{\varepsilon+M}(\varepsilon+M)=\varepsilon$$
>
> 即f在x_0处连续。
>
> 综上，结论得证。

17.4.3 证明定理17.4.1（提示：回顾一元微积分中链式法则（定理10.1.15）证明的全过程。一个可能有效的方式是利用由序列描述的极限定义（参见命题14.1.5(b)），同时利用习题17.1.4）

> 由于f在x_0处可微，因此有：
>
> $$\lim_{x\to x_0;x\in E\setminus\{x_0\}}\frac{\|f(x)-(f(x_0)+f'(x_0)(x-x_0))\|}{\|x-x_0\|}=0$$
>
> 也即对任意的$\varepsilon>0$，存在$\delta>0$使得对任意的$x\in E\setminus\{x_0\}$满足$\|x-x_0\|<\delta$都有：
>
> 同理，由于g在$f(x_0)$处可微，因此有：
>
> $$\lim_{y\to f(x_0);y\in F\setminus\{f(x_0)\}}\frac{\|g(y)-[g\circ f(x_0)+g'(f(x_0))(y-f(x_0))]\|}{\|y-f(x_0)\|}=0$$
>
> 也即对任意的$\varepsilon>0$，存在$\delta>0$使得对任意的$x\in F\setminus\{f(x_0)\}$满足$\|y-f(x_0)\|<\delta$都有$\|[g(y)-g\circ f(x_0)]-g'(f(x_0))(y-f(x_0))\|<\varepsilon\|y-f(x_0)\|$。
>
> 由于$f'(x_0)$与$g'(f(x_0))$都是线性变换，因此根据习题17.1.4分别存在正实数M,N满足：
>
> > - $\forall\ x\in\mathbb{R}^n$，$\|f'(x_0)x\|\leq M\|x\|$。
> > - $\forall\ y\in\mathbb{R}^m$，$\|g'(f(x_0))y\|\leq N\|y\|$。
>
> 于是对任意的$\varepsilon>0$，我们首先寻找三个大于零的常数：
>
> - δ_1:
>
> > 根据f可微的结论，我们知道存在$\delta_1>0$使得对任意的$x\in E\setminus\{x_0\}$满足$\|x-x_0\|<\delta_1$有：
> >
> > $$\|[f(x)-f(x_0)]-f'(x_0)(x-x_0)\|<\frac{\varepsilon}{1+N+M}\|x-x_0\|$$
>
> - δ_2:
>
> > 根据f可微的结论，我们知道存在$\delta_2>0$使得对任意的$x\in E\setminus\{x_0\}$满足$\|x-x_0\|<\delta_2$有：
> >
> > $$\|[f(x)-f(x_0)]-f'(x_0)(x-x_0)\|<\|x-x_0\|$$

- δ_3:

> 根据g可微的结论，我们知道存在$\sigma > 0$使得对任意的$y \in F \backslash \{f(x_0)\}$满足$\|y - f(x_0)\| < \sigma$有:
>
> $$\|[g(y) - g \circ f(x_0)] - g'(f(x_0))(y - f(x_0))\| < \frac{\varepsilon}{1+N+M}\|y - f(x_0)\|$$
>
> 而习题17.4.2可知f在x_0连续的，因此存在$\delta_3 > 0$对任意$x \in E \backslash \{x_0\}$满足$\|x - x_0\| < \delta_3$都有$\|f(x) - f(x_0)\| < \sigma$，从而上面的结论可以引申为:
>
> $$\|[g \circ f(x) - g \circ f(x_0)] - g'(f(x_0))(f(x) - f(x_0))\| < \frac{\varepsilon}{1+N+M}\|f(x) - f(x_0)\|$$

然后我们取$\delta := \min\{\delta_1, \delta_2, \delta_3\}$，于是对任意$x \in E \backslash \{x_0\}$满足$\|x - x_0\| < \delta$，通过多次应用三角不等式我们可以计算有:

$$\begin{aligned}
&\|[g \circ f(x) - g \circ f(x_0)] - g'(f(x_0))f'(x_0)(x - x_0)\| \\
\leq &\|[g \circ f(x) - g \circ f(x_0)] - g'(f(x_0))[f(x) - f(x_0)]\| + \|[g'(f(x_0))[f(x) - f(x_0)] - g'(f(x_0))f'(x_0)(x - x_0)\| \\
< &\frac{\varepsilon}{1+N+M}\|f(x) - f(x_0)\| + N\|[f(x) - f(x_0)] - f'(x_0)(x - x_0)\| \\
< &\frac{\varepsilon}{1+N+M}(\|[f(x) - f(x_0)] - f'(x_0)(x - x_0)\| + \|f'(x_0)(x - x_0)\|) + \frac{\varepsilon N}{1+N+M}\|x - x_0\| \\
< &\frac{\varepsilon}{1+N+M}(\|x - x_0\| + M\|x - x_0\|) + \frac{\varepsilon N}{1+N+M}\|x - x_0\| \\
= &\frac{\varepsilon(1+N+M)}{1+N+M}\|x - x_0\| = \varepsilon\|x - x_0\|
\end{aligned}$$

综上即有:

> 对任意的$\varepsilon > 0$，存在$\delta > 0$使得对任意的$x \in E \backslash \{x_0\}$满足$\|x - x_0\| < \delta$有:
>
> $$\frac{\|g \circ f(x) - (g \circ f(x_0) + g'(f(x_0))f'(x_0)(x - x_0))\|}{\|x - x_0\|} < \varepsilon$$

也即有$\lim\limits_{x \to x_0; x \in E \backslash \{x_0\}} \frac{\|g \circ f(x) - (g \circ f(x_0) + g'(f(x_0))f'(x_0)(x - x_0))\|}{\|x - x_0\|} = 0$，再考虑到由习题17.1.2 $g'(f(x_0))$与$f'(x_0)$的复合$g'(f(x_0))f'(x_0)$也是一个线性变换，于是根据全可微的定义我们有$g \circ f$是在x_0处可微的，并且导数为$g'(f(x_0))f'(x_0)$。

17.4.4 叙述并证明多元函数（即形如$f : E \to \mathbb{R}$的函数，其中E是$\mathbb{R}^n$的子集）的商法则（即叙述一个法则，使得该法则能够给出一个有关商函数f/g的公式）。将你给出的答案同[定理10.1.13(h)]对比，注意务必要明晰你的假设前提都是什么

> 我们可以给出下面的商法则：设E是$\mathbb{R}^n$的子集，x_0是E的内点，并设$f : E \to \mathbb{R}$与$g : E \to \mathbb{R}$是函数。若有f, g均在x_0处可微且$g(x_0) \neq 0$，则f/g也是在x_0处可微的，并且有:
>
> $$\nabla(f/g)(x_0) = \frac{g(x_0)\nabla f(x_0) - f(x_0)\nabla g(x_0)}{g(x_0)^2}$$
>
> 下面我们证明这个结论。
>
> ---
>
> 类似本节摘录1中的内容，我们定义函数$h : \mathbb{R}^n \to \mathbb{R}^2$为$h(x) := (f(x), g(x))$与$k : \mathbb{R}^2 \to \mathbb{R}$为$k(a, b) := a/b$，然后计算导数矩阵有:
>
> $$Dh(x_0) = \begin{pmatrix} \nabla f(x_0) \\ \nabla g(x_0) \end{pmatrix} \qquad Dk(a, b) = \begin{pmatrix} \frac{1}{b}, -\frac{a}{b^2} \end{pmatrix}$$
>
> 再注意到$k \circ h = f/g$，于是根据链式法则有:
>
> $$D(f/g)(x_0) = Dk(f(x_0), g(x_0))Dh(x_0) = \begin{pmatrix} \frac{1}{g(x_0)}, -\frac{f(x_0)}{g(x_0)^2} \end{pmatrix} \begin{pmatrix} \nabla f(x_0) \\ \nabla g(x_0) \end{pmatrix} = \frac{g(x_0)\nabla f(x_0) - f(x_0)\nabla g(x_0)}{g(x_0)^2}$$
>
> 于是结论得证。

17.4.5 设$\vec{x}:\mathbb{R}\to\mathbb{R}^3$是一个可微函数，并设$r:\mathbb{R}\to\mathbb{R}$是函数$r(t):=\|\vec{x}(t)\|$，其中$\|\vec{x}\|$表示$\vec{x}$在通常的$l^2$度量下的长度。设$t_0$是一个实数，证明：如果$r(t_0)\neq 0$，那么$r$在$t_0$就是可微的，并且有

$$r'(t_0):=\frac{\vec{x}'(t_0)\vec{x}(t_0)}{r(t_0)}$$

（提示：利用定理17.4.1）

> 我们设$g:\mathbb{R}^3\to\mathbb{R}$为$g(x_i)_{1\leq i\leq 3}:=\|(x_i)_{1\leq i\leq 3}\|=\sqrt{x_1^2+x_2^2+x_3^2}$，于是显然可以求得：
>
> $$\forall\, 1\leq i\leq 3,\ \frac{\partial g}{\partial x_i}(x_i)_{1\leq i\leq 3}=\frac{x_i}{\sqrt{x_1^2+x_2^2+x_3^2}}$$
>
> 并且显然有只要$x=(x_i)_{1\leq i\leq 3}\neq 0$就有$\frac{\partial g}{\partial x_i}$在$x$处连续，因此根据命题17.3.8我们有$f$在$x$处可微且对任意的$v=(v_i)_{1\leq i\leq 3}$有：
>
> $$f'(x)(v)=\sum_{i=1}^{3}\frac{x_iv_i}{\sqrt{x_1^2+x_2^2+x_3^2}}=\frac{xv}{\|x\|}$$
>
> 再注意到$r=g\circ\vec{x}$，从而我们有：
>
> $$r'(t_0)=g'(\vec{x}(t_0))x'(t_0)=\frac{\vec{x}(t_0)x'(t_0)}{\|\vec{x}(t_0)\|}=\frac{\vec{x}'(t_0)\vec{x}(t_0)}{r(t_0)}$$
>
> 于是结论得证。

本节相关跳转

实分析 10.1 基本定义

实分析 14.1 函数的极限值

实分析 17.1 线性变换

实分析 17.5 二阶导数和克莱尔定理