17.4 注释

张志聪

2025 年 5 月 12 日

说明 1. *作为链式法则和引理 17.1.16（以及引理 17.1.13）的一个推论，我们得到*

$$D(g\circ f)(x_0) = Dg(f(x_0))Df(x_0)$$

证明：

由链式法则，我们有

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

$(g\circ f)'(x_0)$ 是一个线性变换，由 P364 中对导数矩阵的构造方式可知

$$(g\circ f)'(x_0) = L_{D(g\circ f)(x_0)}$$

同理可得

$$g'(f(x_0)) = L_{Dg(f(x_0))}$$
$$f'(x_0) = L_{Df(x_0)}$$

综上，

$$L_{D(g\circ f)(x_0)} = L_{Dg(f(x_0))}L_{Df(x_0)}$$

利用引理 17.1.16，我们得到

$$L_{D(g\circ f)(x_0)} = L_{Dg(f(x_0))Df(x_0)}$$

由引理 17.1.13 可知线性变换存在唯一的对应矩阵，于是我们有

$$D(g\circ f)(x_0) = Dg(f(x_0))Df(x_0)$$

说明 2. $D(fg) = \nabla(fg)$

证明：

$k \circ h = fg : \mathbb{R}^n \to \mathbb{R}$，按照书中关于梯度的定义（P364），我们有

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

而 $D(fg)$ 也将是 $1 \times n$ 的矩阵。而且

$$\begin{aligned} D(fg) &= \left(\frac{\partial fg}{\partial x_1}, \cdots, \frac{\partial fg}{\partial x_n}\right) \\ &= \nabla(fg) \end{aligned}$$