

理解李导数

Charles Daly

2021.09.16

摘要

这些笔记是我对李导数的一些思考。不久前，当我第一次看到李导数的定义，并且几乎不了解它时，William Goldman 告诉我一个让我印象深刻的事实。具体来说，向量场 X 和 Y 的李导数可以被视为它们各自流的对易子的二阶导数。这似乎是微分几何中众所周知的事实之一，但我尚未找到完整的证明资源，所以我想我会将它写下来，以防有人像我第一次看到李导数时一样，感到无助和迷茫。

1 动机

首先，除非另有说明，否则所有后续内容均假设为光滑的。设 M 和 N 分别为维度为 m 和 n 的光滑流形。给定一个光滑映射 $F : M \rightarrow N$ ，我们可以将一个丛同态 $dF : TM \rightarrow TN$ 与 F 关联起来。这个映射有时被称为 F 的全微分 (*global differential*)。它是一个光滑映射 $G : \mathbb{R}^m \rightarrow \mathbb{R}^n$ 的标准微分的流形类比，并因此可以被视为 F 的导数。全微分几乎在任何介绍性微分几何课程中都会出现。

当我第一次看到全微分时，感觉就像在视频游戏中获得了力量。突然间，我可以开始对任何我想要的映射进行导数运算，并对其有很好的理解。虽然这在某种程度上是正确的，但我记得我花了好几个小时试图解决以下问题。一个向量场的导数是什么？这个看似无害的问题引出了大量非常有趣的数学概念，包括在流形上的连通 (*connections on manifolds*) 和 Ehresmann 连通 (*Ehresmann connections*)。

当然，给定一个在 M 上的光滑向量场 Y ，它具有作为一个光滑映射 $Y : M \rightarrow TM$ 的性质，使得 $\pi \circ Y = \text{id}$ ，其中 $\pi : TM \rightarrow M$ 是标准的切丛投影。因此，向量场的导数具有完美的意义，其全微分由 $dY : TM \rightarrow T(TM)$ 给出。并且现在你也许可以看到这里存在一些技术困难。特别是，如果我们设 $p \in M$ ，则 $dY_p : T_p M \rightarrow T_{Y_p}(TM)$ 是从流形 (*manifold*) M 在 p 处的切空间到 M 的切丛 (*tangent bundle*) 在 Y_p 处的切空间的一个线性映射。虽然这在理论上都是正确的，但对于每个 $v \in T_p M$ ，我们希望 $dY_p(v)$ 是 $T_p M$ 的一个元素。这样，我们就可以在 v 和 $dY_p(v)$ 之间进行一些比较。

我认为，在简单地放弃和收工之前，探索几种潜在的‘解决方法’是很有价值的。例如，也许有人可以取向量 $dY_p(v) \in T_{Y_p}(TM)$ ，并将其与 $d\pi_{Y_p} : T_{Y_p}(TM) \rightarrow T_p M$ 下的象进行标识。毕竟，这是我们希望 $dY_p(v)$ 无论如何都在其中的空间。但请注意，根据链式法则，

$$d\pi_{Y_p}(dY_p(v)) = d(\pi \circ Y)_p(v) = d(\text{id})_p(v) = v. \quad (1)$$

我们可以看到，在试图通过投影映射将 $dY_p(v)$ 标识为切空间中的向量时，我们实际上忽略了关于 Y 本身的全部数据。不过，这不应该让人感到惊讶。如果你认为导数是在方向和曲线方面，例如，

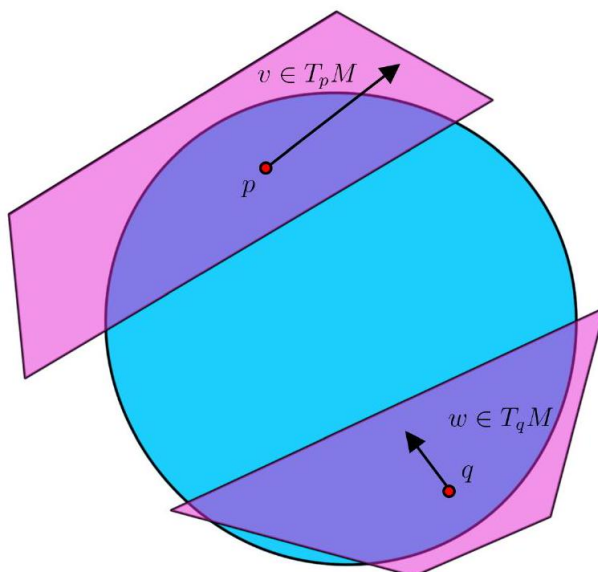


图 1: 注意这里, 在球面 $M = S^2$ 上有两个点 $p, q \in M$, 它们的切空间 $T_p M$ 和 $T_q M$ 用浅红色表示。向量 v 和 w 不能自然相减, 因为它们位于不同的切空间。这个困难的出现是因为在 $T_p M$ 和 $T_q M$ 之间没有自然的同构。

我们可能会取一个光滑路径 $\gamma: I \rightarrow M$ 以表示 v , 其中 $\gamma(0) = p$ 且 $\gamma'(0) = v$ 。这样做时, $dY_p(v)$ 由 $(Y \circ \gamma)'(0)$ 表示。注意, 这可以被视为取沿着 γ 的向量场的时变导数, 由 $Y_{\gamma(t)}$ 给出。我认为这种观点为我们提供了这里困难的一些洞察。具体来说, $Y_{\gamma(t)}$ 可以被视为 γ 到切丛的一个提升。根据交换图,

$$\begin{array}{ccc} & & TM \\ & \nearrow Y_\gamma & \downarrow \pi \\ I & \xrightarrow{\gamma} & M \end{array} \quad (2)$$

现在请注意, 随着 t 的变化, 对于每个 $t \in I$, $Y_{\gamma(t)} \in T_{\gamma(t)} M$ 。也就是, 如果我们把 TM 看做是在 M 之上的一个丛, 我们可以看到在微分上的困难。曲线 $Y_{\gamma(t)}$ 随着 t 的变化而改变 $\pi: TM \rightarrow M$ 的纤维, 这使得对于不同的 t , 它被传输位于 $Y_{\gamma(t)}$ 的不同的切空间中。理想情况下, 人们会试图定义

$$dY_p(v) := \lim_{t \rightarrow 0} \frac{1}{t} (Y_{\gamma(t)} - Y_p). \quad (3)$$

但请注意, $Y_{\gamma(t)}$ 和 Y_p 位于不同的切空间, 并因此将它们相减是没有意义的。下面的图 1 说明了这种困难。此外, 当我们将 $Y_{\gamma(t)}$ 与投影映射组合时, 我们只是得到 $\pi(Y_{\gamma(t)}) = \gamma(t)$, 从而消除了所有关于向量场的有关信息。向量场 Y 告诉我们为 t 的每个值, $Y_{\gamma(t)}$, 分配什么向量, 而对应于 v 的曲线 γ 告诉我们沿着流形对 Y 评估的方向。这种思维方式有助于我们得出一个概念, 即一个 Ehresmann 连通 (Ehresmann connection), 它本质上是将 $T(TM)$ 分解为两个丛的一种方法, 一个垂直丛对应于映射 $d\pi: T(TM) \rightarrow TM$ 的核, 另一个水平丛与垂直丛互补。垂直丛的良好定义与连通无关, 而水平丛是一个选择, 它本质上使我们能够将 $dY_p(v)$ 分解为垂直和水平分量。更多细节可以在 Timothy Goldberg 的一些优秀笔记中找到, *What is a Connection, and What is is Good For?* [1]

在结束本节之前, 我想指出, 在某些情况下, 这些困难不会出现。具体来说, 在 \mathbb{R}^n 的情况下,

我们有切丛自然地与 $\mathbb{R}^n \times \mathbb{R}^n$ 相标识。第一个因子可以被看作是流形分量，而第二个因子可以被看作是切空间分量。在这种情况下，在 \mathbb{R}^n 上的向量场 Y 是一个映射 $Y: \mathbb{R}^n \rightarrow T(\mathbb{R}^n) \simeq \mathbb{R}^n \times \mathbb{R}^n$ ，使得 $\pi_1 \circ Y = \text{id}$ ，其中 $\pi_1: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ 是盖射到第一个因子的投影。因此， Y 的第一个因子是标识，也就是，我们可以把 Y 标识为映射 $Y: \mathbb{R}^n \rightarrow \mathbb{R}^n$ ，对应于第二个因子。

现在，对于每个切向量 $v \in T_p \mathbb{R}^n \simeq \mathbb{R}^n$ ，我们可以取 Y 在 v 方向上的导数。受方程 (3) 的启发，让我们通过曲线 $\gamma(t) = (\gamma^1(t), \dots, \gamma^n(t))$ 表示 v ，使得 $\gamma(0) = p$ 和 $\gamma'(0) = v$ 。同样，让我们通过 $Y(q) = (Y^1(q), \dots, Y^n(q))$ 在坐标中表达 Y 。回到方程 (3)，我们有

$$\begin{aligned} dY_p(v) &= \lim_{t \rightarrow 0} \frac{1}{t} (Y_{\gamma(t)} - Y_p) = \lim_{t \rightarrow 0} \frac{1}{t} (Y^i(\gamma(t)) - Y^i(p)) e_i = \frac{\partial Y^i}{\partial x^j}(\gamma(0)) \frac{d\gamma^j}{dt}(0) e_i \\ &= \left(\frac{\partial Y^1}{\partial x^j}(p) \frac{d\gamma^j}{dt}(0), \dots, \frac{\partial Y^n}{\partial x^j}(p) \frac{d\gamma^j}{dt}(0) \right), \end{aligned} \quad (4)$$

其中，第一行的第三个方程是通过应用链式法则以组合 $Y \circ \gamma$ 得到的。注意，方程 (4) 为我们提供了向量场在 \mathbb{R}^n 中导数的一个概念，而且，这正是任何人都能猜到的，即只是对分量进行微分。从表面上看，这真的只是坐标扩展的练习。事实上，值得尝试看看这在流形的通用情况下如何转换。

对于一个任意流形，我们不再享有切丛的全局平凡化能力，即 TM 不一定用 $M \times \mathbb{R}^n$ 标识。尽管如此，我们仍然享有切丛的局部平凡化。因此，对于一个点 $p \in M$ ，选取一个关于它的坐标图表 (U, ϕ) ，其中 $\phi: U \rightarrow \mathbb{R}^n$ 是一个微分同胚，盖射到 \mathbb{R}^n 的某些开子集上。这会诱导出切丛的一个局部帧，由 $\{\partial/\partial x^i\}$ 给出，因此我们可以表达一个向量场，对于所有的 $q \in U$ ，至少在 p 的局部，为 $Y|_q = Y^i(q) \partial/\partial x^i|_q$ 。此外，如果我们选取一个 $v \in T_p M$ ，我们可以根据该帧将其表达为 $v = v^j \partial/\partial x^j|_p$ 。也就是说，我们可以尝试定义

$$dY_p(v) := v^j \frac{\partial Y^i}{\partial x^j}(p) \frac{\partial}{\partial x^i} \Big|_p, \quad (5)$$

正如我们在方程 (4) 中所做的那样，其中每个 v^j 对应于 $d\gamma^j/dt(0)$ 。与在流形上的大多数构造一样，我们最好确保这个量独立于坐标贴片是良好定义的。为此，如果 (V, ψ) 是另一个坐标贴片，具有坐标 \tilde{x}^k ，我们注意到有方程 $Y|_q = Y^i(q) \partial/\partial x^i|_q = \tilde{Y}^k(q) \partial/\partial \tilde{x}^k|_q$ 和 $v = v^j \partial/\partial x^j|_p = \tilde{v}^l \partial/\partial \tilde{x}^l|_p$ 。他们的坐标关联为

$$Y^i(q) = \tilde{Y}^k(q) \frac{\partial x^i}{\partial \tilde{x}^k}(q) \text{ and } v^j = \tilde{v}^l \frac{\partial x^j}{\partial \tilde{x}^l}(p). \quad (6)$$

对第一个方程为 $Y^i(q)$ 求微分，并抑制在 p 处的求值，产生

$$\frac{\partial Y^i}{\partial x^j} = \frac{\partial}{\partial x^j} \left(\tilde{Y}^k \frac{\partial x^i}{\partial \tilde{x}^k} \right) = \frac{\partial \tilde{x}^m}{\partial x^j} \frac{\partial}{\partial \tilde{x}^m} \left(\tilde{Y}^k \frac{\partial x^i}{\partial \tilde{x}^k} \right) = \frac{\partial \tilde{x}^m}{\partial x^j} \frac{\partial \tilde{Y}^k}{\partial \tilde{x}^m} \frac{\partial x^i}{\partial \tilde{x}^k} + \tilde{Y}^k \frac{\partial \tilde{x}^m}{\partial x^j} \frac{\partial^2 x^i}{\partial \tilde{x}^m \partial \tilde{x}^k}. \quad (7)$$

将方程 (7) 代回方程 (5)，并在方程 (6) 中改变 v^j 的坐标，产生

$$\begin{aligned} dY_p(v) &= \tilde{v}^l \frac{\partial x^j}{\partial \tilde{x}^l} \left(\frac{\partial \tilde{x}^m}{\partial x^j} \frac{\partial \tilde{Y}^k}{\partial \tilde{x}^m} \frac{\partial x^i}{\partial \tilde{x}^k} + \tilde{Y}^k \frac{\partial \tilde{x}^m}{\partial x^j} \frac{\partial^2 x^i}{\partial \tilde{x}^m \partial \tilde{x}^k} \right) \frac{\partial}{\partial x^i} \\ &= \tilde{v}^l \left(\frac{\partial x^j}{\partial \tilde{x}^l} \frac{\partial \tilde{x}^m}{\partial x^j} \right) \frac{\partial \tilde{Y}^k}{\partial \tilde{x}^m} \left(\frac{\partial x^i}{\partial \tilde{x}^k} \frac{\partial}{\partial x^i} \right) + \tilde{v}^l \tilde{Y}^k \left(\frac{\partial x^j}{\partial \tilde{x}^l} \frac{\partial \tilde{x}^m}{\partial x^j} \right) \frac{\partial^2 x^i}{\partial \tilde{x}^m \partial \tilde{x}^k} \frac{\partial}{\partial x^i} \\ &= \tilde{v}^l \delta_l^m \frac{\partial \tilde{Y}^k}{\partial \tilde{x}^m} \frac{\partial}{\partial \tilde{x}^k} + \tilde{v}^l \tilde{Y}^k \delta_l^m \frac{\partial^2 x^i}{\partial \tilde{x}^m \partial \tilde{x}^k} \frac{\partial}{\partial x^i} \\ &= \tilde{v}^l \frac{\partial \tilde{Y}^k}{\partial \tilde{x}^l} \frac{\partial}{\partial \tilde{x}^k} + \tilde{v}^l \tilde{Y}^k \frac{\partial^2 x^i}{\partial \tilde{x}^l \partial \tilde{x}^k} \frac{\partial}{\partial x^i}. \end{aligned} \quad (8)$$

虽然这些计算非常冗长，但它们确实提供了一些在我们的定义中出了什么问题的洞察。如果我们将方程 (8) 与方程 (5) 进行比较，我们可以清楚地看到出了什么问题。特别是，方程 (8) 中出现的第二项阻碍了该量独立于坐标图表的良好定义。

也就是说，我们应该花时间考虑为什么这在 \mathbb{R}^n 的情况下有效，而不是在一般流形中。注意在方程 (8) 中，如果我们假设我们的流形 M 被一个坐标图集 $\{(U_\alpha, \phi_\alpha)\}$ 覆盖，使得坐标的变化具有消失 (为零) 的 Hessian 矩阵，则方程 (5) 是良好定义的，并且这个导数依赖于坐标图集。事实上，一个非常自然的流形类是仿射流形 (affine manifolds)。仿射流形是一类其坐标变化是局部仿射映射的流形，即来自 $\mathbb{R}^n \rightarrow \mathbb{R}^n$ 的映射形式为 $F(x) = Ax + b$ ，其中 $A \in \text{GL}(n, \mathbb{R})$ 和 $b \in \mathbb{R}^n$ 。对于那些熟悉连通的人来说，值得验证这个导数对应于在 \mathbb{R}^n 上的标准连通的拉回连通。

在另一种情况下，我们可以在不必借助连通机制的情况下定义一个向量场的导数，这种情况发生在 M 是可平行的时候，也就是，在 TM 和 $M \times \mathbb{R}^n$ 之间存在一个丛同构。因此，在 M 上的一个向量场作为一个光滑映射 $Y: M \rightarrow TM$ 与具有一个光滑映射 $Y: M \rightarrow M \times \mathbb{R}^n$ 的情况相同，其中 Y 在第一个分量上是标识。因此，一个向量场的数据实际上只是一个映射 $Y: M \rightarrow \mathbb{R}^n$ 。我们现在可以按照方程 (3) 定义 $dY_p(v)$ ，其中由于切丛的全局平凡化，现在 $Y_{\gamma(t)} - Y_p$ 的差分是有意义的。当然，我们总是可以局部地平凡化，就像我们在方程 (8) 中尝试的那样。但是，让这个差商依赖于局部标识似乎是不合理的，而让它依赖于一个单一固定的标识则更为合理。话虽如此，可能值得再次强调的是，这样的定义实际上仍然依赖于选择，特别是在这里我们依赖于 TM 和 $M \times \mathbb{R}^n$ 之间的标识，而这个标识通常是非正规的。

2 李导数

考虑到上述讨论，定义流形上向量场的有意义导数似乎没有希望。主要问题之一是无法比较切空间。这是微分几何的一个基本困难，并且已经被广泛研究。实际上，人们经常会发现关于滚动而不滑动 (rolling without slipping) 的微积分问题，这与通过称为平行传输 (parallel transport) 的构造来标识在不同点处的切空间的概念密切相关。这个想法是，平行传输为我们提供了一种通过在流形上的曲线来标识在不同点处的切空间的方法。当然，人们沿着流形行进的方式，通常会影响这些标识。例如，球面上的两条曲线和下面的图 2。

受这种看待切空间之间标识的想法的启发，人们可能会考虑流形上的微分同胚群。如果我们想理解量 $Y_{\gamma(t)} - Y_p$ ，我们需要一些方法来标识切空间。微分同胚当然提供了这种方法。特别是，如果 F 是流形 M 的一个微分同胚，将一个点 p 映射到另一个点 q ，则微分 $dF_p: T_p M \rightarrow T_q M$ 是向量空间之间的一个同构。现在，假设我们有微分同胚的一条路径 F_t ，使得 $F_t(p) = \gamma(t)$ 。这为我们提供了通过 $d(F_t)_p$ 对 $T_p M$ 和 $T_{\gamma(t)} M$ 进行标识的方法。在这种情况下，我们不一定能理解 $Y_{\gamma(t)} - Y_p$ ，但似乎很自然地考虑了量

$$\left(d(F_t)_p\right)^{-1} (Y_{\gamma(t)} - Y_p). \quad (9)$$

在继续之前，让我们确定这是一个合理的量。当然， $Y_{\gamma(t)} \in T_{\gamma(t)} M$ 。另一方面，由于 $d(F_t)_p: T_p M \rightarrow T_{\gamma(t)} M$ ，我们知道其逆映射，由于 F_t 是一个微分同胚，将 $T_{\gamma(t)} M$ 映射到 $T_p M$ ，因此定义良好。因此，在方程 (9) 中的量确实有意义。此外，我们还可以考虑极限

$$\lim_{t \rightarrow 0} \frac{1}{t} \left(\left(d(F_t)_p\right)^{-1} (Y_{\gamma(t)} - Y_p) \right). \quad (10)$$

对于微分同胚的一条通用路径，人们不应该期望上述结果收敛，因为我们无法控制例如 $d(F_t)_p$ 在

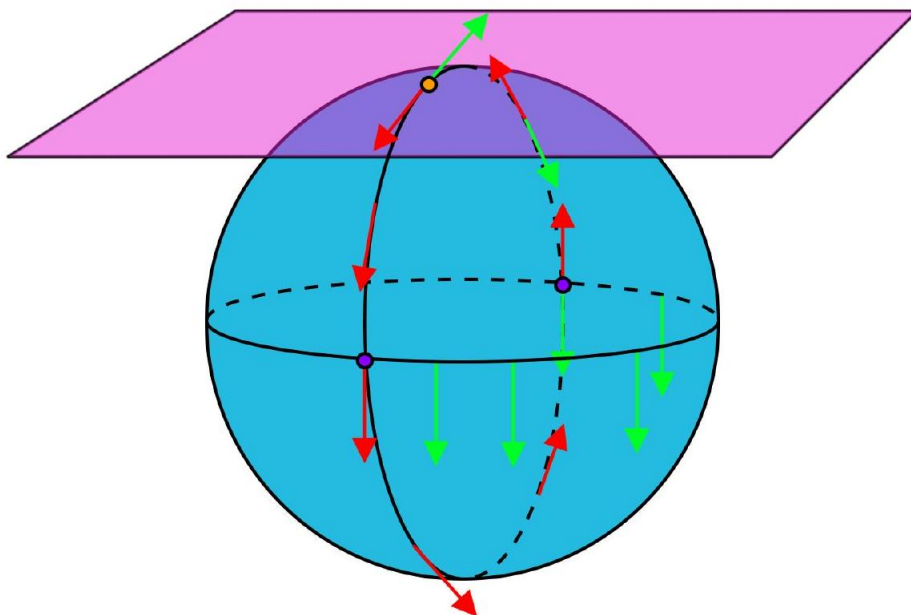


图 2: 图中, 位于北极点 $p \in M$ 的红色向量 v 沿着两条曲线平行传输。第一条是垂直于圆或本初子午线。注意, 当返回点 p 时, 向量 v 会回到其原始位置。与此相反, 当曲线到达第一个紫色点时, 向量 v 在途中返回其原始位置。与这种情况相反, 当曲线到达第一个紫色点时, 向量 v 在途中返回其原始位置。在绕过半个赤道后, 到达第二个紫色点, 然后沿着本初子午线返回。结果向量以绿色返回原始切空间, 但实际上等于 $-v$ 。这表明球体不是平的。

$t = 0$ 时的行为, 它只是一个在当前规定之下固定 p 的一个映射。此外, 如何获得微分同胚的一条路径 $(F_t)_p = \gamma(t)$, 其中 $\gamma(t)$ 表示一个向量 $v \in T_p M$?

一个非常自然的考虑因素是在一个流形上的流。给定一个单参数微分同胚子群 $F_t : \mathbb{R} \rightarrow \text{Diff}(M)$, 我们知道 $F_0 = \text{id}$, 因此, 在 $t = 0$ 时的行为, 以及理想情况下在 $t = 0$ 附近的局部行为, 必须得到合理的控制。

根据上述观察, 那么尝试将一个向量场相对于另一个向量场进行微分, 因为向量场定义了局部流。也就是说, 假设 X 和 Y 是在一个流形 M 上的向量场。为了简单起见, 我将假设它们对应的流确实是完整的, 尽管以下构造和论证可能适用于纯局部论证。

由于我们希望微分的向量场是 Y , 我们打算沿着向量场 X 进行微分, 设 $\theta : \mathbb{R} \times M \rightarrow M$ 是 X 的对应流映射。设 $p \in M$, 我们希望以与方程 (10) 类似的方式定义 Y 在 p 处的相对于 X 的导数。将这个导数标志为 $\mathcal{L}_X Y|_p$, 我们定义

$$\begin{aligned} \mathcal{L}_X Y|_p &:= \lim_{t \rightarrow 0} \frac{1}{t} \left(\left(d(\theta_t)_p \right)^{-1} (Y_{\theta_t(p)}) - Y_p \right) \\ &= \lim_{t \rightarrow 0} \frac{1}{t} \left(d(\theta_{-t})_{\theta_t(p)} (Y_{\theta_t(p)}) - Y_p \right), \end{aligned} \quad (11)$$

其中, 第二个方程来自事实: $\theta_t^{-1} = \theta_{-t}$ 。有几个事实需要验证, 但不会在这些笔记中提供, 最紧迫的是这个向量场确实是有定义的。对于那些对进一步细节感兴趣的人, 以及这些笔记的大部分符号和灵感来源, 我推荐你阅读 Jack Lee 的杰出微分几何著作 *Introduction to Smooth Manifolds* [2]。这本书是相对于这些材料的真正坚实的参考书, 他的对话式写作风格、详细论证和插图使这本书成为我向大家推荐的一本书。

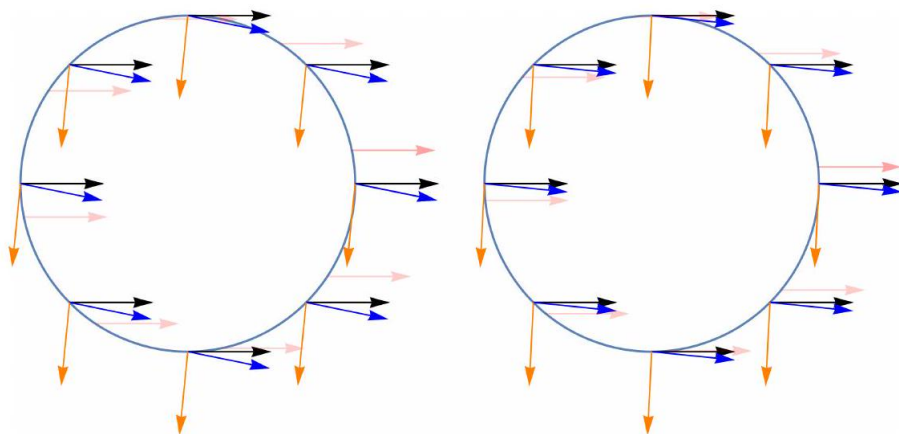


图 3: 这里向量场 $Y = \partial/\partial x$ 以黑色表示。在某一点的 X 的流由浅蓝色圆表示。向量场 $Y_{\theta_t(p)}$ 以浅红色表示。它在 X 流之下的拉回由蓝色表示, 注意, 这与 Y 不同。差商以橙色表示。左图为 $t = 0.2$, 右图为 $t = 0.1$ 。

在继续之前, 值得指出的是, 方程 (11) 中的定义与方程 (5) 中的定义有很大不同。回想一下, 通过将 p 转换为 $\gamma(t)$ 平移所诱导出的标识, 是关联 $T_{\gamma(t)}\mathbb{R}^n$ 和 $T_p\mathbb{R}^n$ 标识的方法。然而, $T_{\gamma(t)}\mathbb{R}^n$ 和 $T_p\mathbb{R}^n$ 之间的标识是由向量场的流诱导的。以下例子说明了这种差异。

例 1. 设 $M = \mathbb{R}^2$ 并定义

$$X = -y \frac{\partial}{\partial x} + x \frac{\partial}{\partial y} \text{ and } Y = \frac{\partial}{\partial x}. \quad (12)$$

设 $p = (x, y)$, 并让我们计算 $\mathcal{L}_X Y$ 。根据方程 (11), 我们首先计算 X 的流。事实上, 人们可以相对容易地证明, 相应的微分方程组给出为:

$$x' = -y \text{ and } y' = x \text{ where } (x(0), y(0)) = (x, y) \implies \theta_t(x, y) = \begin{pmatrix} \cos(t) & -\sin(t) \\ \sin(t) & \cos(t) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

也就是说, $Y_{\theta_t(p)} = \partial/\partial x|_{\theta_t(p)}$, 并因此

$$d(\theta_{-t})_{\theta_t(p)}(Y_{\theta_t(p)}) = d(\theta_{-t})_{\theta_t(p)}\left(\frac{\partial}{\partial x}\Big|_{\theta_t(p)}\right) = \cos(t)\frac{\partial}{\partial x}\Big|_p - \sin(t)\frac{\partial}{\partial y}\Big|_p.$$

其中的负号是因为我们使用逆流, 即 θ_{-t} , 而不是 θ_t 。将这个结果代入方程 (11), 产生

$$\mathcal{L}_X Y = \lim_{t \rightarrow 0} \frac{1}{t} \left(\cos(t)\frac{\partial}{\partial x}\Big|_p - \sin(t)\frac{\partial}{\partial y}\Big|_p - \frac{\partial}{\partial x}\Big|_p \right) = \lim_{t \rightarrow 0} \frac{\cos(t) - 1}{t} \frac{\partial}{\partial x}\Big|_p - \frac{\sin(t)}{t} \frac{\partial}{\partial y}\Big|_p = -\frac{\partial}{\partial y}\Big|_p.$$

我们很容易在抽象的计算中再一次迷失。这个例子之所以引人入胜, 是因为通常我们认为向量场 $\partial/\partial x$ 是“常数”。在 $T_p\mathbb{R}^2$ 和 $T_q\mathbb{R}^2$ 的标准标识之下它是不变的, 这是由从 p 到 q 的平移所诱导的。然而, 我们并不是这样标识切空间的。在这里, 我们通过流出映射, 即 θ_t , 在 X 流的同一轨道上局部标识切空间。在我们的情况下, 这些是圆。现在, 当一个向量沿着一个圆移动时, 它似乎有理由改变其方向, 这正是正在发生的事情, 也是我们得到非平凡李导数的原因。图 3 展示了这一过程。

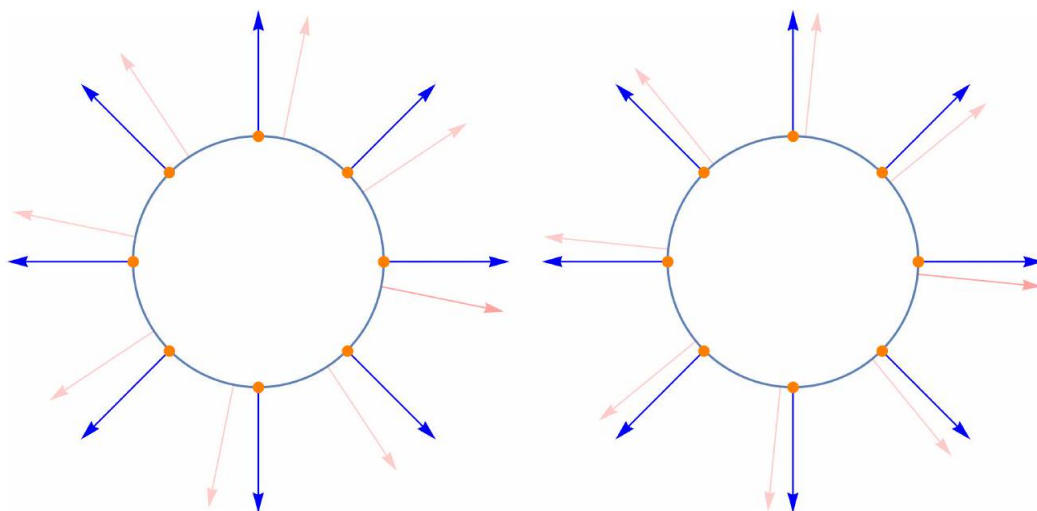


图 4: 这里向量场 $Y_{\theta_t(p)}$ 以红色表示, 其在 X 的流之下的拉回以蓝色表示。原始向量场 Y 以黑色表示, 但由于被蓝色的拉回所遮挡, 它们是相等的。差商由橙色的零向量表示。左图为 $t = 0.2$, 右图为 $t = 0.1$ 。

例 2. 设 $M = \mathbb{R}^2$, 并定义

$$X = -y \frac{\partial}{\partial x} + x \frac{\partial}{\partial y} \text{ and } Y = x \frac{\partial}{\partial x} + y \frac{\partial}{\partial y}. \quad (13)$$

以与上述示例类似的方式, 可以进行计算以得出 Y 在 X 的流之下是不变的, 因此这种情况下的李导数实际上为 0。值得验证这些计算, 但下面提供了一个图片来演示这一点。

例 3. 设 $M = \mathbb{R}^2$, 并定义

$$X = (x + y) \frac{\partial}{\partial x} + y \frac{\partial}{\partial y} \text{ and } Y = x \frac{\partial}{\partial y}. \quad (14)$$

这里是一个例子, 其中向量场 Y 沿着直线 $x = 0$ 消失 (为零), 但其沿着 Y 轴的导数通常是非零的。同样, 没有太大的困难, 人们可以验证, X 的流由下式给出为:

$$\theta_t(x, y) = \begin{pmatrix} e^t & te^t \\ 0 & e^t \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \Rightarrow Y_{\theta_t(p)} = (e^t x + te^t y) \frac{\partial}{\partial y} \Big|_{\theta_t(p)}.$$

计算差商的极限, 产生

$$\begin{aligned} \mathcal{L}_X Y &= \lim_{t \rightarrow 0} \frac{1}{t} \left((e^t x + te^t y) \left(-te^{-t} \frac{\partial}{\partial x} \Big|_p + e^{-t} \frac{\partial}{\partial y} \Big|_p \right) - x \frac{\partial}{\partial y} \Big|_p \right) \\ &= \lim_{t \rightarrow 0} - (x + ty) \frac{\partial}{\partial x} \Big|_p + ye^{-t} \frac{\partial}{\partial y} \Big|_p = -x \frac{\partial}{\partial x} \Big|_p + y \frac{\partial}{\partial y} \Big|_p, \end{aligned}$$

注意, 除了在原点处之外, 在直线 $x = 0$ 上确实不消失 (为零)。下面的图 5 展示了沿着 X 的两条不同积分曲线的计算过程。

应该提到的是, 还有另一种表述, 即向量场在一个流形 M 上作用在光滑实值函数代数上的李导数。具体来说, 给定在一个流形 M 上的两个向量 X 和 Y , 可以将 X 和 Y 的李括号定义为向

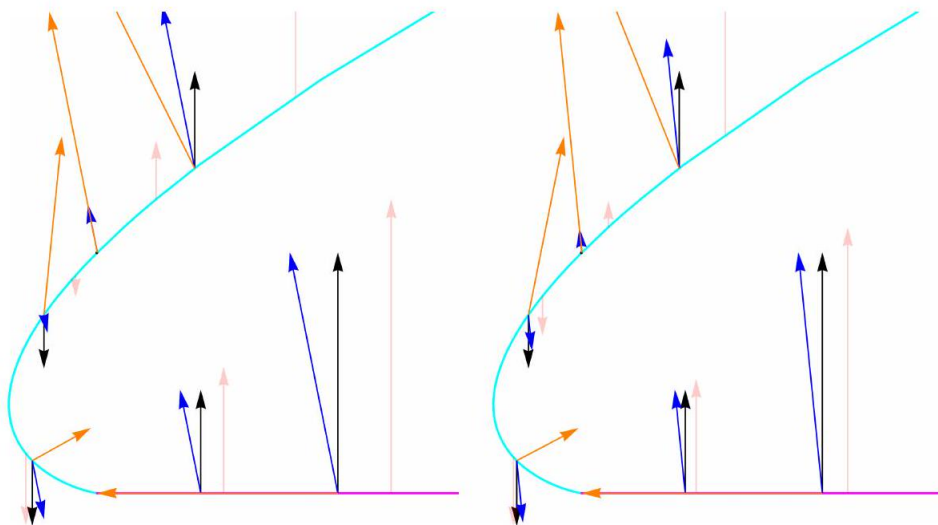


图 5: 向量场 $Y = x\partial/\partial y$ 沿着向量场 $X = (x+y)\partial/\partial x + y\partial/\partial y$ 的两条积分曲线分别用黑色和浅红色表示。浅红色的向量场是在 $\theta_t(p)$ 处求值的 Y , 其在 X 流之下的拉回用蓝色表示。差商用橙色表示。注意, 有一处橙色和蓝色向量互为标量倍数。这一点对应于 Y 消失 (为零) 的点, 但其李导数不为零。左图为 $t = 0.2$, 右图为 $t = 0.1$ 。

量场, 它在 M 上作用于光滑实值函数的代数, 即对于所有的 $f \in C^\infty(M)$, 有 $(XY - YX)f = X(Yf) - Y(Xf)$ 。

虽然以下对李导数的表述在计算上更为便捷, 且在理论上更易于处理, 但它有一个严重的缺点, 那就是没有任何‘明显’的几何解释, 至少据我所知是这样。实际上, 在我看来, 上述表述甚至定义一个向量场都远非显而易见。然而, 它使得人们能够相对容易地计算李导数, 并且有一个直接的坐标表示。如果 X 和 Y 在坐标中局部地由 $X = X^i\partial/\partial x^i$ 和 $Y = Y^j\partial/\partial x^j$ 给出, 则这两个向量场的李括号为

$$\mathcal{L}_X Y = [X, Y] = \left(X^i \frac{\partial Y^j}{\partial x^i} - Y^j \frac{\partial X^i}{\partial x^j} \right) \frac{\partial}{\partial x^i}. \quad (15)$$

这一定义和关于流的导数公式的一致性尚不清楚, 但你仍然可以在 Jack Lee 的文章 [2] 中找到证明。

3 对易子的导数

正如摘要中承诺的那样, 我们最后提供一个证明, 在 M 上的向量场 X 和 Y 的李导数由它们的流的对易子的二阶导数给出。设 X 和 Y 的流由 θ 和 ϕ 给出。选取一个点 $p \in M$, 和一个以 p 为中心的坐标贴片, $(V, \psi): V \rightarrow \mathbb{R}^n$ 。假设这个坐标贴片足够小, 以至于有一个盒子, 盒子的大小为 $(-\epsilon, \epsilon)^2$, 其中 θ 和 ϕ 在围绕 p 的 V 中的每个地方都有定义。这个规定是为了保证流在有限时间内不会爆炸, 并且在它们离开开集 V 的范围内才会变得没有定义。由于这个讨论是纯局部的, 我们可以假定事实上将 V 标识为以原点为中心的 \mathbb{R}^n 的一个开子集。

考虑映射 $(\phi_{-t}\theta_{-t}\phi_t\theta_t)(0): (-\epsilon, \epsilon) \rightarrow \mathbb{R}^n$ 。如果我们围绕原点对 $\theta: (-\epsilon, \epsilon) \times V \rightarrow V$ 和

$\phi_t : (-\epsilon, \epsilon) \times V \rightarrow V$ 都进行泰勒展开, 我们有如下结果

$$\begin{aligned}\theta(t, x) &= \theta(0, 0) + \frac{\partial \theta}{\partial t}(0, 0)t + \frac{\partial \theta}{\partial x}(0, 0)x + \frac{1}{2} \left(\frac{\partial^2 \theta}{\partial t^2}(0, 0)t^2 \right. \\ &\quad \left. + \frac{\partial^2 \theta}{\partial t \partial x}(0, 0)tx + \frac{\partial^2 \theta}{\partial x \partial t}(0, 0)tx + x^T \frac{\partial^2 \theta}{\partial x^2}(0, 0)x \right) + \dots \\ &= 0 + \frac{\partial \theta}{\partial t}(0, 0)t + \frac{\partial \theta}{\partial x}(0, 0)x + \frac{1}{2} \frac{\partial^2 \theta}{\partial t^2}(0, 0)t^2 + \frac{\partial^2 \theta}{\partial x \partial t}(0, 0)tx + \frac{1}{2} x^T \frac{\partial^2 \theta}{\partial x^2}(0, 0)x + \dots\end{aligned}\quad (16)$$

这里我使用了非常不精确的符号表示法。特别地, 表达式 $(\partial^2 \theta / \partial x^2)(0, 0)$ 应该标志一个向量-值函数在点 $(0, 0)$ 处的二阶导数。如果 θ 在 \mathbb{R} 中取值, 你会得到从 $\mathbb{R}^m \rightarrow \mathbb{R}$ 的一个映射的标准泰勒展开, 并且这个量表示 Hessian 矩阵。但在我们的例子中, 这是一个张量, 尽管如此, 通过在 $\theta = (\theta^1, \theta^2, \dots, \theta^n)$ 的坐标展开中对每个 θ^i 进行泰勒展开, 可以恢复它。不幸的是, 我不知道对此是否有一个标准的记号, 所以我只是自己编造的, 但如果你需要任何令人信服的东西, 只要考虑从 \mathbb{R}^2 到 \mathbb{R}^2 的两两情况, 并写出细节, 看看它到底是什么。

我们可以在方程 (16) 中计算导数, 因为它们简化了很多。回想一下, θ 的第一个输入是时间坐标, 而第二个输入是空间坐标。

我们有以下方程:

$$\frac{\partial \theta}{\partial t}(0, 0) = \frac{\partial}{\partial t} \Big|_{t=0} \theta(t, 0) = X(0), \quad (17)$$

$$\frac{\partial \theta}{\partial x}(0, 0) = \frac{\partial}{\partial x} \Big|_{x=0} \theta(0, x) = \text{id}, \quad (18)$$

$$\frac{\partial^2 \theta}{\partial t^2}(0, 0) = \frac{\partial^2}{\partial t^2} \Big|_{t=0} \theta(t, 0) = \frac{\partial}{\partial t} \Big|_{t=0} X_{\theta_t(0)} = \frac{\partial X}{\partial x}(0)X(0), \quad (19)$$

$$\frac{\partial^2 \theta}{\partial x \partial t}(0, 0) = \frac{\partial}{\partial x} \Big|_{x=0} \frac{\partial}{\partial t} \Big|_{t=0} \theta(t, x) = \frac{\partial}{\partial x} \Big|_{x=0} X(x) = \frac{\partial X}{\partial x}(0), \quad (20)$$

$$\frac{\partial^2 \theta}{\partial x^2}(0, 0) = \frac{\partial^2}{\partial x^2} \Big|_{x=0} \theta(0, x) = \frac{\partial^2}{\partial x^2} \Big|_{x=0} x = 0. \quad (21)$$

方程 (17) 立即从向量场与诱导流的关系得出, 即向量场随时间变化。方程 (18) 是因为 $\theta(0, x)$ 是在空间上的恒等映射, 所以它的导数是在切空间上的恒等映射。方程 (19) 也是流的性质的推导结果, 其中最后一个方程通过将链式法则应用于 $(X \circ \theta)(t, 0)$ 得到。方程 (20) 直接得到, 或多或少是通过定义得到的。方程 (21) 再次通过以下事实的推导而得到, 即 $\theta(0, x)$ 是在空间上的恒等映射, 因此它的二阶导数为零。为了确信方程 (21), 考虑函数 $(x_1, \dots, x_n) \mapsto (x_1, \dots, x_n)$ 的二阶导数数组会有所帮助。一阶导数可以被视为恒等矩阵, 对具有常数值矩阵的分量进行微分应得出零。将方程 (17)-(21) 的结果结合起来, 提供了 θ 在 $(0, 0)$ 处的二阶泰勒展开式。

$$\theta(t, x) = x + t \left(X(0) + \frac{\partial X}{\partial x}(0)x \right) + \frac{t^2}{2} \frac{\partial X}{\partial x}(0)X(0) + \dots \quad (22)$$

同样的结果适用于 ϕ ,

$$\phi(t, y) = y + t \left(Y(0) + \frac{\partial Y}{\partial x}(0)y \right) + \frac{t^2}{2} \frac{\partial Y}{\partial x}(0)X(0) + \dots \quad (23)$$

将方程 (22) 和 (23) 组合起来, 我们可以计算出在 $0 \in V$ 处的流的对易子。由于这些计算相当复杂, 我使用了 Mathematica, 关于确切的代码, 请参见下面的图 6。特别地, 这产生

$$\phi_{-t}\theta_{-t}\phi_t\theta_t(0) = t^2 \left(\frac{\partial Y}{\partial x}(0)X(0) - \frac{\partial X}{\partial x}(0)Y(0) \right) + \dots \quad (24)$$

```

In[21]:= ClearAll["Global`*"]
(*X is X(0) and A is the spatial derivative of X, (dX/dx)(0).
Higher order terms excluded as they will vanish upon evaluation.*)
theta[t_, x_] := x + t*(X + A*x) + ((t^2)/2)*A*x;
(*Y is Y(0) and B is the spatial derivative of Y, (dY/dx)(0).
Higher order terms excluded as they will vanish upon evaluation.*)
phi[s_, y_] := y + s*(Y + B*y) + ((s^2)/2)*B*y;
(*Defining the commutator of the flows*)
comm[s_, t_] := phi[-s, theta[-t, phi[s, theta[t, 0]]]];

(*Takes the derivative of the commutator with respect to t
then evaluates at t=0*)
Expand[comm[t, t]]

D[D[Expand[comm[t, t]], t], t] /. t -> 0

Out[25]= B t^2 X - 1/2 A^2 t^3 X - 1/2 A B t^3 X - B^2 t^3 X + 1/2 A B^2 t^4 X + 1/2 A^2 B^2 t^5 X - A t^2 Y + 1/2 A B t^3 Y - 1/2 B^2 t^3 Y + 1/2 A B^2 t^4 Y

Out[26]= 2 B X - 2 A Y

```

图 6: 用于计算对易子二阶导数的 Mathematica 代码。

其中剩余的项都是关于 t 的高阶项, 因此在二阶导数在零处评估时消失 (为零)。对公式 (25) 进行微分并在 $t = 0$ 处评估, 产生

$$\left. \frac{d^2}{dt^2} \right|_{t=0} (\phi_{-t} \theta_{-t} \phi_t \theta_t)(0) = 2 \left(\frac{\partial Y}{\partial x}(0) X(0) - \frac{\partial X}{\partial x}(0) Y(0) \right). \quad (25)$$

在系数为 2 的情况下, 方程 (25) 中的右手侧在坐标中为

$$\frac{\partial Y}{\partial x}(0) X(0) - \frac{\partial X}{\partial x}(0) Y(0) = \left(\frac{\partial Y^j}{\partial x^i}(0) X^i(0) - \frac{\partial X^j}{\partial x^i}(0) Y^i(0) \right) e_j, \quad (26)$$

这是方程 (15) 的形式。证明到此结束。

可能值得注意的是, 只要我们考虑二阶导数, 你可能会对此提出反对意见。当然, 一条曲线 $\gamma : I \rightarrow M$ 的一阶导数位于切向量空间 $T_{\gamma(0)}M$ 中, 但二阶导数却不在其中。我们之所以将二阶导数视为切向量空间的一个元素, 原因非常微妙。假设你有一条曲线 $\gamma : I \rightarrow M$, 使得 $\gamma'(0) = 0$ 。一般来说, 曲线的切向量通过 $\gamma'(0)f = (f \circ \gamma)'(0)$ 作用于光滑函数 $C^\infty(M)$ 的代数上, 其中 $f \circ \gamma : I \rightarrow \mathbb{R}$ 是一个光滑函数。因为一阶导数为零, 我们实际上可以通过 $\gamma''(0)f = (f \circ \gamma)''(0)$ 定义 $\gamma''(0)$ 在 $C^\infty(M)$ 上的一个作用。一般来说, 这并不满足一个导子的标准, 但是因为 $\gamma'(0) = 0$, 所以它满足了。值得写出技术细节以确保这一点, 但是一旦这得到验证, 我们实际上就可以确定标识 $\gamma''(0) \in T_p M$ 。在我们的案例中, 通过方程 (24) 可以看出, 这个条件得到了满足, 因为对易子没有一阶项。

References

- [1] Timothy Goldberg. What is a connection and what is it good for? 2008. URL <http://pi.math.cornell.edu/~goldberg/Notes/AboutConnections.pdf>.
- [2] John Lee. Smooth Manifolds Second Edition. pages 227-233, 2000.