李群论中的伴随表示与指数映射导数

Shuyong Chen

2023 年 10 月 13 日

1 简介

本文旨在深化对李群理论，尤其是伴随表示和指数映射导数的理解。本文将深入探讨伴随表示、雅可比矩阵和指数映射与对数映射的导数等几种数学概念的关系。在此基础上，详细分析了指数映射和对数映射的导数，揭示了它们与一阶线性近似之间的关系，并强调这些概念的几何直觉。最后以三维旋转群为例，通过具体的矩阵计算验证了理论结果，加深了对概念的理解。

2 李群知识快速回顾

李群是一种具有连续性质的群，它在一个光滑流形上定义，并且具有群结构。李群的一个典型例子是三维旋转群，表示三维空间中的旋转。

李代数是李群的局部线性化，它描述了李群在单位元附近的局部结构。李代数是一个向量空间，它具有一个叫做李括号的二元运算。李代数的一个典型例子是三维空间中的向量叉积。

2.1 李群的定义和性质

李群 $\mathcal{G}$ 是一个光滑流形，其元素满足群公理。一方面，可微或光滑流形 (*smooth manifold*) 是局部类似于线性空间的拓扑空间。流形的光滑性意味着在每个点上存在唯一的切空间。这个空间是一个线性或向量空间，我们可以在上面做微积分。另一方面，一个群 (*group*) $(\mathcal{G}, \circ)$ 是一个集合，$\mathcal{G}$，具有组合运算，$\circ$，这个算子，对于元素 $\mathcal{X}, \mathcal{Y}, \mathcal{Z} \in \mathcal{G}$，满足下列公理，

$$\text{封闭于 ‘}\circ\text{’} \quad : \mathcal{X} \circ \mathcal{Y} \in \mathcal{G} \tag{1}$$

$$\text{幺元 } \mathcal{E} \quad : \mathcal{E} \circ \mathcal{X} = \mathcal{X} \circ \mathcal{E} = \mathcal{X} \tag{2}$$

$$\text{逆元 } \mathcal{X}^{-1} \quad : \mathcal{X}^{-1} \circ \mathcal{X} = \mathcal{X} \circ \mathcal{X}^{-1} = \mathcal{E} \tag{3}$$

$$\text{结合性} \quad : (\mathcal{X} \circ \mathcal{Y}) \circ \mathcal{Z} = \mathcal{X} \circ (\mathcal{Y} \circ \mathcal{Z}) \tag{4}$$

在一个李群 (*Lie group*) 中，流形在每一点上看起来都是一样的 (例如在球面的表面上)，并且因此在任意点上的所有切空间都是相同的。群结构要求流形元素的组合保持在流形上 (封闭方程)，并且每个元素在流形中也有一个逆元 (逆元方程)。其中一个特别的元素是幺元 (幺元方程)，并因此有一个特殊的切空间是幺元处的正切，我们称之为李群的李代数。李群结合了光滑流形的局部性质，使我们能够利用群的全局性质进行微积分，从而实现远处对象的非线性组合。

2.2 李代数的定义和性质

李代数是一种与李群相关的数学结构，它是一个向量空间。李代数 $\mathfrak{g}$ 被称为代数，因为它被赋予了一个二元运算，即李括号 (*Lie bracket*) $[X, Y]$，其性质与 $\mathcal{G}$ 的群运算密切相关，满足以下性质：

1. 反对称性：对于任意的向量 x, y，有 $[x, y] = -[y, x]$。

2. Jacobi 恒等式：对于任意的向量 x, y, z，有 $[x,[y,z]] + [y,[z,x]] + [z,[x,y]] = 0$。

李代数的定义中，李括号是一个非常重要的概念，它描述了向量空间中两个向量之间的代数关系。李括号的反对称性意味着它是一个反对称的二元运算，而李恒等式和 Jacobi 恒等式则描述了李括号的代数性质，它们保证了李代数的结构是稳定的，并且可以通过李括号来推导出其他的代数性质。

2.3 切空间与李代数

给定 $\mathcal{X}(t)$ 为在李群流形 $\mathcal{M}$ 上移动的点，它的速度 $\dot{\mathcal{X}} = \partial\mathcal{X}/\partial t$ 属于 $\mathcal{X}$ 处与 $\mathcal{M}$ 正切的空间，我们标记为 $T_{\mathcal{X}}\mathcal{M}$。流形的光滑性，即没有边或尖峰，意味着在每个点上存在唯一的切空间。这种切空间的结构在任意地方都是相同的。

幺元 $T_{\mathcal{E}}\mathcal{M}$ 处的切空间称为 $\mathcal{M}$ 的李代数 (*Lie algebra*)，标记为 $\mathfrak{g}$，

$$\text{Lie algebra} \;:\; \mathfrak{g} \triangleq T_{\mathcal{E}}\mathcal{M} \;.$$

每一个李群都有一个关联的李代数。我们通过以下事实将李群与其李代数联系起来:

- 李代数 $\mathfrak{g}$ 是一个向量空间。因此，它的元素可以用 $\mathbb{R}^m$ 中的向量来标识 (*identified*)，其维数 m 是 $\mathcal{M}$ 的自由度。

- 指数映射 (*exponential map*)，$\exp : \mathfrak{g} \to \mathcal{M}$，将李代数的元素精确地转化为群的元素。对数映射 log 是其逆运算。

- 通过线性变换，在 $\mathcal{X}$ 处的切空间中的向量可以变换到幺元 $\mathcal{E}$ 处的切空间中。这种变换称为伴随 (*adjoint*) 变换。

李代数可以局部定义到一个切点 $\mathcal{X}$，建立局部坐标系于 $T_{\mathcal{X}}\mathcal{M}$。我们将用“帽子”修饰符 $\hat{}$ 来表示李代数的元素，例如 $\mathbf{v}^\wedge$ 表示速度或 $\boldsymbol{\tau}^\wedge = (\mathbf{v}t)^\wedge = \mathbf{v}^\wedge t$ 表示一般元素。还可以添加一个左上标来指定精确的切空间，例如 ${}^{\mathcal{X}}\mathbf{v}^\wedge \in T_{\mathcal{X}}\mathcal{M}$ 和 ${}^{\mathcal{E}}\mathbf{v}^\wedge \in T_{\mathcal{E}}\mathcal{M}$。

通过对群的约束方程 (3) (即逆元方程) 进行时间微分，李代数的结构可以被找到。对于乘法群，这将产生新的约束 $\mathcal{X}^{-1}\dot{\mathcal{X}} + \dot{\mathcal{X}^{-1}}\mathcal{X} = 0$，适用于正切于 $\mathcal{X}$ 的元素 (项 $\dot{\mathcal{X}^{-1}}$ 是其逆项的导数) 。因此，李代数的元素的形式是，

$$\mathbf{v}^\wedge = \mathcal{X}^{-1}\dot{\mathcal{X}} = -\dot{\mathcal{X}^{-1}}\mathcal{X}. \tag{5}$$

2.4 指数映射

指数映射 exp() 允许我们精确地将李代数的元素变换到群上，一般称为收回 (*retraction*) 操作。直观地说，exp() 将切元素缠绕在大圆弧或测地线 (*geodesic*) 跟随的流形上 (就像将弦缠绕在球上

一样)。逆映射是 log()，即展开操作。exp() 映射通过考虑流形上 $\mathcal{X} \in \mathcal{M}$ 的时间导数自然产生，如下所示。从方程 (5) 我们有，

$$\dot{\mathcal{X}} = \mathcal{X}\mathbf{v}^\wedge \tag{6}$$

将 $\mathbf{v}$ 做为常数，这是一个常微分方程 (ODE)，其解为

$$\mathcal{X}(t) = \mathcal{X}(0)\exp\left(\mathbf{v}^\wedge t\right) \tag{7}$$

即 $\exp\left(\mathbf{v}^\wedge t\right)$ 将李代数的元素 $\mathbf{v}^\wedge t$ 映射到群中。这被称为指数映射 (*exponential map*)。

为了给指数映射提供一个更通用的定义，让我们定义切增量 $\boldsymbol{\tau} \triangleq \mathbf{v}t \in \mathbb{R}^m$ 作为每一时刻的速度，这样我们就有 $\boldsymbol{\tau}^\wedge = \mathbf{v}^\wedge t \in \mathfrak{g}$ 做为李代数中的一个点。指数映射和它的逆 (对数映射)，现在可以写成，

$$\exp : \mathfrak{g} \to \mathcal{M} \quad ; \quad \boldsymbol{\tau}^\wedge \mapsto \mathcal{X} = \exp\left(\boldsymbol{\tau}^\wedge\right) \tag{8}$$

$$\log : \mathcal{M} \to \mathfrak{g} \quad ; \quad \mathcal{X} \mapsto \boldsymbol{\tau}^\wedge = \log(\mathcal{X}). \tag{9}$$

通过写出绝对收敛的泰勒级数，

$$\exp\left(\boldsymbol{\tau}^\wedge\right) = \mathcal{E} + \boldsymbol{\tau}^\wedge + \frac{1}{2}\boldsymbol{\tau}^{\wedge 2} + \frac{1}{3!}\boldsymbol{\tau}^{\wedge 3} + \cdots, \tag{10}$$

并且利用 $\boldsymbol{\tau}^\wedge$ 的幂的代数性质，得到了乘法群中的指数封闭形式。然后将这些求逆以找到对数映射。指数映射的关键性质是

$$\exp\left((t+s)\,\boldsymbol{\tau}^\wedge\right) = \exp\left(t\boldsymbol{\tau}^\wedge\right)\exp\left(s\boldsymbol{\tau}^\wedge\right) \tag{11}$$

$$\exp\left(t\boldsymbol{\tau}^\wedge\right) = \exp\left(\boldsymbol{\tau}^\wedge\right)^t \tag{12}$$

$$\exp\left(-\boldsymbol{\tau}^\wedge\right) = \exp\left(\boldsymbol{\tau}^\wedge\right)^{-1} \tag{13}$$

$$\exp\left(\mathcal{X}\boldsymbol{\tau}^\wedge\mathcal{X}^{-1}\right) = \mathcal{X}\exp\left(\boldsymbol{\tau}^\wedge\right)\mathcal{X}^{-1} \tag{14}$$

其中方程 (14) 很有用，在推导伴随矩阵和 Jacobian 矩阵时经常用到。

大写的 Exp 和 Log 映射是将向量元素 $\boldsymbol{\tau} \in \mathbb{R}^m\ (\cong T_{\mathcal{E}}\mathcal{M})$ 直接映射到元素 $\mathcal{X} \in \mathcal{M}$ 的快捷方式。我们有，

$$\mathrm{Exp} : \mathbb{R}^m \to \mathcal{M} \quad ; \quad \boldsymbol{\tau} \mapsto \mathcal{X} = \mathrm{Exp}(\boldsymbol{\tau}) \tag{15}$$

$$\mathrm{Log} : \mathcal{M} \to \mathbb{R}^m \quad ; \quad \mathcal{X} \mapsto \boldsymbol{\tau} = \mathrm{Log}(\mathcal{X}). \tag{16}$$

显然有

$$\mathcal{X} = \mathrm{Exp}(\boldsymbol{\tau}) \triangleq \exp\left(\boldsymbol{\tau}^\wedge\right) \tag{17}$$

$$\boldsymbol{\tau} = \mathrm{Log}(\mathcal{X}) \triangleq \log(\mathcal{X})^\vee. \tag{18}$$

2.5 加号和减号算子

加号和减号算子允许我们在 (弯曲的) 流形中的元素之间引入增量，并在 (平坦的) 切向量空间中表示它们。它们用 $\oplus$ 和 $\ominus$ 表示，将一个 Exp/Log 操作与一个结合操作组合在一起。由于组合的非交换性，它们根据操作数的顺序在右结合版本和左结合版本中定义。右结合算子是，

$$\mathrm{right}-\oplus : \mathcal{Y} = \mathcal{X} \oplus {}^{\mathcal{X}}\boldsymbol{\tau} \triangleq \mathcal{X} \circ \mathrm{Exp}\left({}^{\mathcal{X}}\boldsymbol{\tau}\right) \in \mathcal{M} \tag{19}$$

$$\mathrm{right}-\ominus : {}^{\mathcal{X}}\boldsymbol{\tau} = \mathcal{Y} \ominus \mathcal{X} \triangleq \mathrm{Log}\left(\mathcal{X}^{-1} \circ \mathcal{Y}\right) \in T_{\mathcal{X}}\mathcal{M}. \tag{20}$$

因为在 $\oplus$ 方程中 $\operatorname{Exp}\left({}^{\mathcal{X}}\boldsymbol{\tau}\right)$ 出现在组合的右手侧，${}^{\mathcal{X}}\boldsymbol{\tau}$ 属于 $\mathcal{X}$ 处的切空间：我们按照约定说这个 ${}^{\mathcal{X}}\boldsymbol{\tau}$ 是在 $\mathcal{X}$ 处的局部 (local) 坐标系中表示—我们标记到参考坐标系的左上标。

左结合算子是

$$\text{left}-\oplus: \mathcal{Y}={}^{\mathcal{E}}\boldsymbol{\tau}\oplus\mathcal{X}\triangleq\operatorname{Exp}\left({}^{\mathcal{E}}\boldsymbol{\tau}\right)\circ\mathcal{X}\in\mathcal{M} \tag{21}$$

$$\text{left}-\ominus: {}^{\mathcal{E}}\boldsymbol{\tau}=\mathcal{Y}\ominus\mathcal{X}\triangleq\operatorname{Log}\left(\mathcal{Y}\circ\mathcal{X}^{-1}\right)\in T_{\mathcal{E}}\mathcal{M}. \tag{22}$$

现在，在 $\oplus$ 方程中 $\operatorname{Exp}\left({}^{\mathcal{E}}\boldsymbol{\tau}\right)$ 出现在左侧，并且我们有 ${}^{\mathcal{E}}\boldsymbol{\tau}\in T_{\mathcal{E}}\mathcal{M}$：我们说这个 ${}^{\mathcal{E}}\boldsymbol{\tau}$ 是在全局 (global) 坐标系中表示。

请注意，虽然右结合和左结合算子 $\oplus$ 是按操作数顺序区分的，但符号 $\ominus$ 在上述方程中是不明确的。在本文中，我们默认表示局部扰动，因此我们默认使用 $\oplus$ 和 $\ominus$ 的右结合形式。

2.6 伴随和伴随矩阵

如果我们在方程 (19, 21) 中标识 $\mathcal{Y}$，我们就得到 ${}^{\mathcal{E}}\boldsymbol{\tau}\oplus\mathcal{X}=\mathcal{X}\oplus{}^{\mathcal{X}}\boldsymbol{\tau}$，它确定局部切元素和全局切元素之间的关系。我们用方程 (14, 19, 21) 来扩展它为

$$\begin{aligned}
\operatorname{Exp}\left({}^{\mathcal{E}}\boldsymbol{\tau}\right)\mathcal{X}&=\mathcal{X}\operatorname{Exp}\left({}^{\mathcal{X}}\boldsymbol{\tau}\right)\\
\exp\left({}^{\mathcal{E}}\boldsymbol{\tau}^{\wedge}\right)&=\mathcal{X}\exp\left({}^{\mathcal{X}}\boldsymbol{\tau}^{\wedge}\right)\mathcal{X}^{-1}=\exp\left(\mathcal{X}\,{}^{\mathcal{X}}\boldsymbol{\tau}^{\wedge}\mathcal{X}^{-1}\right)\\
{}^{\mathcal{E}}\boldsymbol{\tau}^{\wedge}&=\mathcal{X}\,{}^{\mathcal{X}}\boldsymbol{\tau}^{\wedge}\mathcal{X}^{-1}
\end{aligned}$$

伴随 因此，我们将 $\mathcal{M}$ 在 $\mathcal{X}$ 处的伴随 (*adjoint*)，记为 $\operatorname{Ad}_{\mathcal{X}}$，定义为

$$\operatorname{Ad}_{\mathcal{X}}:\mathfrak{g}\to\mathfrak{g};\ \boldsymbol{\tau}^{\wedge}\mapsto\operatorname{Ad}_{\mathcal{X}}\left(\boldsymbol{\tau}^{\wedge}\right)\triangleq\mathcal{X}\,{}^{\mathcal{X}}\boldsymbol{\tau}^{\wedge}\mathcal{X}^{-1}, \tag{23}$$

因此 ${}^{\mathcal{E}}\boldsymbol{\tau}^{\wedge}=\operatorname{Ad}_{\mathcal{X}}\left({}^{\mathcal{E}}\boldsymbol{\tau}^{\wedge}\right)$. 这定义了群在它自己的李代数上的伴随作用 (*adjoint action*)。伴随有两个有趣的 (并且很容易证明的) 性质，

$$\begin{aligned}
\text{Linear}:\ &\operatorname{Ad}_{\mathcal{X}}\left(a\boldsymbol{\tau}^{\wedge}+b\boldsymbol{\sigma}^{\wedge}\right)=a\operatorname{Ad}_{\mathcal{X}}\left(\boldsymbol{\tau}^{\wedge}\right)+b\operatorname{Ad}_{\mathcal{X}}\left(\boldsymbol{\sigma}^{\wedge}\right)\\
\text{Homomorphism}:\ &\operatorname{Ad}_{\mathcal{X}}\left(\operatorname{Ad}_{\mathcal{Y}}\left(\boldsymbol{\tau}^{\wedge}\right)\right)=\operatorname{Ad}_{\mathcal{X}\mathcal{Y}}\left(\boldsymbol{\tau}^{\wedge}\right).
\end{aligned}$$

伴随矩阵 因为 $\operatorname{Ad}_{\mathcal{X}}()$ 是线性的，我们可以找到一个等价的矩阵算子 $\mathbf{Ad}_{\mathcal{X}}$，它映射笛卡尔切向量 ${}^{\mathcal{E}}\boldsymbol{\tau}\cong{}^{\mathcal{E}}\boldsymbol{\tau}^{\wedge}$ 和 ${}^{\mathcal{X}}\boldsymbol{\tau}\cong{}^{\mathcal{X}}\boldsymbol{\tau}^{\wedge}$，

$$\mathbf{Ad}_{\mathcal{X}}:\mathbb{R}^m\to\mathbb{R}^m;{}^{\mathcal{X}}\boldsymbol{\tau}\mapsto{}^{\mathcal{E}}\boldsymbol{\tau}=\mathbf{Ad}_{\mathcal{X}}\,{}^{\mathcal{X}}\boldsymbol{\tau}, \tag{24}$$

我们称之为伴随矩阵 (*adjoint matrix*)。这可以通过将 $\vee$ 应用于方程 (23) 来计算，因此写为

$$\mathbf{Ad}_{\mathcal{X}}\boldsymbol{\tau}=\left(\mathcal{X}\boldsymbol{\tau}^{\wedge}\mathcal{X}^{-1}\right)^{\vee}, \tag{25}$$

然后扩展右手侧结合以标识伴随矩阵。伴随矩阵的其它性质是，

$$\mathcal{X}\oplus\boldsymbol{\tau}=\left(\mathbf{Ad}_{\mathcal{X}}\boldsymbol{\tau}\right)\oplus\mathcal{X} \tag{26}$$

$$\mathbf{Ad}_{\mathcal{X}^{-1}}=\mathbf{Ad}_{\mathcal{X}}{}^{-1} \tag{27}$$

$$\mathbf{Ad}_{\mathcal{X}\mathcal{Y}}=\mathbf{Ad}_{\mathcal{X}}\mathbf{Ad}_{\mathcal{Y}}. \tag{28}$$

请注意在方程 (27, 28) 中的左半部分通常比右半部分计算起来更方便快捷。我们将经常使用伴随矩阵将 $\mathcal{X}$ 处的切空间的向量线性变换为原点的切空间的向量，应用方程为 ${}^{\mathcal{E}}\boldsymbol{\tau}=\mathbf{Ad}_{\mathcal{X}}\,{}^{\mathcal{X}}\boldsymbol{\tau}$。

2.7 李群导数与 Jacobian 矩阵

Jacobian 矩阵 在向量微积分中，多个变量的向量值函数的**雅可比矩阵** (**Jacobian matrix**) 是其所有一阶偏导数的矩阵。

假设 $\mathbf{f}:\mathbf{R}^n \to \mathbf{R}^m$ 是一个函数，使得它的每个一阶偏导数都存在于 $\mathbf{R}^n$ 上。该函数将点 $\mathbf{x} \in \mathbf{R}^n$ 作为输入，并生成向量 $\mathbf{f}(\mathbf{x}) \in \mathbf{R}^m$ 作为输出。那么 $\mathbf{f}$ 的 Jacobian 矩阵被定义为一个 $m \times n$ 矩阵，记为 $\mathbf{J}$，其第 (i,j) 项为 $\mathbf{J}_{ij} = \frac{\partial f_i}{\partial x_j}$，或明确地

$$\mathbf{J} = \begin{bmatrix} \frac{\partial \mathbf{f}}{\partial x_1} & \cdots & \frac{\partial \mathbf{f}}{\partial x_n} \end{bmatrix} = \begin{bmatrix} \nabla^{\mathrm{T}} f_1 \\ \vdots \\ \nabla^{\mathrm{T}} f_m \end{bmatrix} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix}$$

其中 $\nabla^{\mathrm{T}} f_i$ 是梯度的转置 (行向量) 的第 i 个分量。

Jacobian 矩阵表示 $\mathbf{f}$ 在 $\mathbf{f}$ 的可微分的每个点上的微分。具体来说，如果 $\mathbf{f}(\mathbf{x})$ 在 $\mathbf{x}$ 处可微，如果 $\mathbf{h}$ 是一个列矩阵表示的位移向量，那么矩阵乘积 $\mathbf{J}(\mathbf{x}) \cdot \mathbf{h}$ 是另一个位移向量，即 $\mathbf{f}$ 在 $\mathbf{x}$ 的邻域内的变化的最佳线性近似。这意味着将 $\mathbf{y}$ 映射到 $\mathbf{f}(\mathbf{x}) + \mathbf{J}(\mathbf{x}) \cdot (\mathbf{y} - \mathbf{x})$ 的函数是 $\mathbf{f}(\mathbf{y})$ 对于所有靠近 $\mathbf{x}$ 的点的最佳线性逼近。该线性函数称为 $\mathbf{f}$ 在 $\mathbf{x}$ 处的**导数** (*derivative*) 或**微分** (*differential*)。

在李群中定义导数的各种方法中，我们主要关注 Jacobian 矩阵映射向量切空间的形式。这在工程应用很常见，因为在这些空间中，不确定性和增量可以被恰当而容易地定义。利用这些 Jacobian 矩阵，李群中的不确定性管理公式与向量空间中的不确定性管理公式基本相似。

向量空间上的 Jacobian 矩阵 对于多元函数 $f:\mathbb{R}^m \to \mathbb{R}^n$，Jacobian 矩阵定义为所有偏导数的 $n \times m$ 矩阵，

$$\mathbf{J} = \frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} \triangleq \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_m} \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_m} \end{bmatrix} \in \mathbb{R}^{n\times m}. \tag{29}$$

用下面的形式定义这个矩阵很方便。让我们分割 $\mathbf{J} = [\mathbf{j}_1 \cdots \mathbf{j}_m]$ 作为它的第 i 个列向量 $\mathbf{j}_i = \left[\frac{\partial f_1}{\partial x_i} \cdots \frac{\partial f_n}{\partial x_i}\right]^\top$，并让此列向量响应于

$$\mathbf{j}_i = \frac{\partial f(\mathbf{x})}{\partial x_i} \triangleq \lim_{h\to 0} \frac{f(\mathbf{x} + h\mathbf{e}_i) - f(\mathbf{x})}{h} \in \mathbb{R}^n, \tag{30}$$

其中 $\mathbf{e}_i$ 是 $\mathbb{R}^m$ 的自然基的第 i 个向量。至于分子，注意这个向量

$$\mathbf{v}_i(h) \triangleq f(\mathbf{x} + h\mathbf{e}_i) - f(\mathbf{x}) \quad \in \mathbb{R}^n \tag{31}$$

是当 $\mathbf{x}$ 在 $\mathbf{e}_i$ 方向上受到扰动时，$f(\mathbf{x})$ 的变化量，并且相应的 Jacobian 矩阵的列仅为 $\mathbf{j}_i = \partial \mathbf{v}_i(h)/\left.\partial h\right|_{h=0} = \lim_{h\to 0} \mathbf{v}_i(h)/h$。在本文中，为了方便起见，我们引入了紧凑形式，

$$\mathbf{J} = \frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} \triangleq \lim_{\mathbf{h}\to 0} \frac{f(\mathbf{x}+\mathbf{h}) - f(\mathbf{x})}{\mathbf{h}} \in \mathbb{R}^{n\times m}, \tag{32}$$

其中 $\mathbf{h} \in \mathbb{R}^m$，其用方程 (30) 计算所有列以形成方程 (29) 的定义。通过将分子扩展成 $\mathbf{h}$ 中的线性形式，并将左手侧标识为 Jacobian 矩阵，该形式可用于计算 Jacobian 矩阵，即机制方程，

$$\lim_{\mathbf{h}\to 0} \frac{f(\mathbf{x}+\mathbf{h}) - f(\mathbf{x})}{\mathbf{h}} = \cdots = \lim_{\mathbf{h}\to 0} \frac{\mathbf{J}\mathbf{h}}{\mathbf{h}} \triangleq \frac{\partial \mathbf{J}\mathbf{h}}{\partial \mathbf{h}} = \mathbf{J}. \tag{33}$$

最后请注意，对于 $\mathbf{h}$ 的小值，我们有线性近似值，

$$f(\mathbf{x}+\mathbf{h}) \underset{\mathbf{h}\to 0}{\longrightarrow} f(\mathbf{x}) + \frac{\partial f(\mathbf{x})}{\partial \mathbf{x}}\mathbf{h}. \tag{34}$$

李群上的右 Jacobian 矩阵 受上面标准导数定义方程 (32) 的启发，我们现在可以使用我们的 $\oplus$ 和 $\ominus$ 算子来定义作用于流形的函数 $f:\mathcal{M}\to\mathcal{N}$ 的 Jacobian 矩阵。使用右结合 (right-) 的 $\{\oplus,\ominus\}$ 代替 $\{+,-\}$，我们获得一个类似于标准导数的形式，

$$\frac{{}^{\mathcal{X}}Df(\mathcal{X})}{D\mathcal{X}} \triangleq \lim_{\boldsymbol{\tau}\to 0}\frac{f(\mathcal{X}\oplus\boldsymbol{\tau})\ominus f(\mathcal{X})}{\boldsymbol{\tau}}\in\mathbb{R}^{n\times m} \tag{35a}$$

$$= \lim_{\boldsymbol{\tau}\to 0}\frac{\operatorname{Log}\left(f(\mathcal{X})^{-1}\circ f(\mathcal{X}\circ\operatorname{Exp}(\boldsymbol{\tau}))\right)}{\boldsymbol{\tau}} \tag{35b}$$

$$= \left.\frac{\partial\operatorname{Log}\left(f(\mathcal{X})^{-1}\circ f(\mathcal{X}\circ\operatorname{Exp}(\boldsymbol{\tau}))\right)}{\partial\boldsymbol{\tau}}\right|_{\boldsymbol{\tau}=0}. \tag{35c}$$

我们把这种 Jacobian 矩阵称为 f 函数的右 Jacobian 矩阵。请注意方程 (35c) 只是在标准导数方程 (32) 中使用相当复杂的函数 $g(\boldsymbol{\tau})=\operatorname{Log}\left(f(\mathcal{X})^{-1}\circ f(\mathcal{X}\circ\operatorname{Exp}(\boldsymbol{\tau}))\right)$。将其写入方程 (35a) 中传达了更多的直觉：它是 $f(\mathcal{X})$ 相对于 $\mathcal{X}$ 的导数，只是我们表达为在切空间中的无穷小变化！实际上，由于使用右结合 (right-) 的 $\oplus$ 和 $\ominus$ 的操作，$\mathcal{X}$ 和 $f(\mathcal{X})$ 中的变量现在被表示为局部切空间中的向量，即分别在 $\mathcal{X}\in\mathcal{M}$ 和 $f(\mathcal{X})\in\mathcal{N}$ 处的正切。这个导数是一个适当的 Jacobian 矩阵 $\mathbb{R}^{n\times m}$ 线性地映射到局部 (local) 切空间 $T_{\mathcal{X}}\mathcal{M}\to T_{f(\mathcal{X})}\mathcal{N}$ (并且我们用一个局部‘$\mathcal{X}$’上标来标记导数)。就像在向量空间中一样，这个矩阵的列对应于方向导数。也就是对于向量

$$\boldsymbol{\sigma}_i(h)=f(\mathcal{X}\oplus h\mathbf{e}_i)\ominus f(\mathcal{X})\in\mathbb{R}^n \tag{36}$$

用方程 (31) 中的 $\mathbf{v}_i$ 来比较方程 (36) 的 $\boldsymbol{\sigma}_i$，其是当 $\mathcal{X}$ 沿着 $\mathbf{e}_i$ 方向变化时，$f(\mathcal{X})$ 的变化。其相应的 Jacobian 矩阵的列是

$$\mathbf{j}_i=\left.\frac{\partial\boldsymbol{\sigma}_i(h)}{\partial h}\right|_{h=0}.$$

注意，每当函数 f 从一个流形传递到另一个流形时，方程 (35a) 中的加号和减号算子必须被正确选择：$\oplus$ 对应定义域 (domain) $\mathcal{M}$，并且 $\ominus$ 对应陪域 (codomain) 或象 (image) $\mathcal{N}$。

对于 $\boldsymbol{\tau}$ 的小值，以下近似值适用，

$$f\left(\mathcal{X}\oplus{}^{\mathcal{X}}\boldsymbol{\tau}\right)\xrightarrow[{}^{\mathcal{X}}\boldsymbol{\tau}\to 0]{} f(\mathcal{X})\oplus\frac{{}^{\mathcal{X}}Df(\mathcal{X})}{D\mathcal{X}}\,{}^{\mathcal{X}}\boldsymbol{\tau}\in\mathcal{N}. \tag{37}$$

李群上的左 Jacobian 矩阵 导数也可以由左结合 (left-) 加号和减号算子定义，因此，

$$\frac{{}^{\mathcal{E}}Df(\mathcal{X})}{D\mathcal{X}} \triangleq \lim_{\boldsymbol{\tau}\to 0}\frac{f(\boldsymbol{\tau}\oplus\mathcal{X})\ominus f(\mathcal{X})}{\boldsymbol{\tau}}\in\mathbb{R}^{n\times m} \tag{38}$$

$$= \lim_{\boldsymbol{\tau}\to 0}\frac{\log\left(f(\operatorname{Exp}(\boldsymbol{\tau})\circ\mathcal{X})\circ f(\mathcal{X})^{-1}\right)}{\boldsymbol{\tau}}$$

$$= \left.\frac{\partial\log\left(f(\operatorname{Exp}(\boldsymbol{\tau})\circ\mathcal{X})\circ f(\mathcal{X})^{-1}\right)}{\partial\boldsymbol{\tau}}\right|_{\boldsymbol{\tau}=0},$$

我们称之为 f 函数的左 Jacobian 矩阵。请注意，现在 $\boldsymbol{\tau}\in T_{\mathcal{E}}\mathcal{M}$，并且分子属于 $T_{\mathcal{E}}\mathcal{N}$，因此左 Jacobian 矩阵是一个 $n\times m$ 的矩阵，映射了全局 (global) 切空间，$T_{\mathcal{E}}\mathcal{M}\to T_{\mathcal{E}}\mathcal{N}$，这是 $\mathcal{M}$ 和 $\mathcal{N}$ 的李代数 (并且我们用全局或原点‘$\mathcal{E}$’上标来标记导数)。对于 $\boldsymbol{\tau}$ 的小值，以下方程成立，

$$f\left({}^{\mathcal{E}}\boldsymbol{\tau}\oplus\mathcal{X}\right)\xrightarrow[{}^{\mathcal{E}}\boldsymbol{\tau}\to 0]{}\frac{{}^{\mathcal{E}}Df(\mathcal{X})}{D\mathcal{X}}\,{}^{\mathcal{E}}\boldsymbol{\tau}\oplus f(\mathcal{X})\quad\in\mathcal{N}. \tag{39}$$

我们可以从方程 (37, 39) 中展示左和右 Jacobian 矩阵由定义域 (domain) $\mathcal{M}$ 和象 (image)$\mathcal{N}$ 的伴随相关联的关系，

$$\frac{{}^{\mathcal{E}}Df(\mathcal{X})}{D\mathcal{X}}\mathbf{Ad}_{\mathcal{X}} = \mathbf{Ad}_{f(\mathcal{X})}\frac{{}^{\mathcal{X}}Df(\mathcal{X})}{D\mathcal{X}}. \tag{40}$$

2.8 指数映射的导数

除了我们下面特别说明的所谓的左 Jacobian 矩阵 (*left Jacobian*)，这里扩展的所有 Jacobian 矩阵都是右 Jacobian 矩阵 (*right-Jacobian*)，即由方程 (35a) 所定义。通过遵循这里的提示，感兴趣的读者应该不会发现在扩展左 Jacobian 矩阵时有什么特别的困难。对于不愿意这样做的读者，方程 (40) 可用于此目的，因为

$$\frac{{}^{\mathcal{E}}Df(\mathcal{X})}{D\mathcal{X}} = \mathbf{Ad}_{f(\mathcal{X})}\frac{{}^{\mathcal{X}}Df(\mathcal{X})}{D\mathcal{X}}\mathbf{Ad}_{\mathcal{X}}{}^{-1}. \tag{41}$$

我们使用符号 $\mathbf{J}_{\mathcal{X}}^{f(\mathcal{X})} \triangleq \frac{Df(\mathcal{X})}{D\mathcal{X}}$ 和 $\mathbf{J}_{\mathcal{X}}^{\mathcal{Y}} \triangleq \frac{D\mathcal{Y}}{D\mathcal{X}}$ 。我们还注意到，$\mathbf{Ad}_{\mathcal{X}}{}^{-1}$ 应该由 $\mathbf{Ad}_{\mathcal{X}^{-1}}$ 实现。

流形 $\mathcal{M}$ 上的 Jacobian 矩阵 我们定义流形 $\mathcal{M}$ 的右 Jacobian 矩阵为 $\mathcal{X} = \mathrm{Exp}(\boldsymbol{\tau})$ 的右 Jacobian 矩阵，即，对于 $\boldsymbol{\tau} \in \mathbb{R}^m$，

$$\mathbf{J}_r(\boldsymbol{\tau}) \triangleq \frac{{}^{\boldsymbol{\tau}}D\mathrm{Exp}(\boldsymbol{\tau})}{D\boldsymbol{\tau}} \in \mathbb{R}^{m\times m}, \tag{42}$$

这由方程 (35a) 定义。右 Jacobian 矩阵将参数 $\boldsymbol{\tau}$ 的变化映射到 $\mathrm{Exp}(\boldsymbol{\tau})$ 处的局部 (*local*) 切空间中的变化，此即局部切空间的指数映射的导数。从方程 (35a) 这很容易证明，对于小的 $\delta\boldsymbol{\tau}$ 值，以下近似值成立，

$$\mathrm{Exp}(\boldsymbol{\tau}+\delta\boldsymbol{\tau}) \approx \mathrm{Exp}(\boldsymbol{\tau})\mathrm{Exp}\left(\mathbf{J}_r(\boldsymbol{\tau})\delta\boldsymbol{\tau}\right) \tag{43}$$

$$\mathrm{Exp}(\boldsymbol{\tau})\mathrm{Exp}(\delta\boldsymbol{\tau}) \approx \mathrm{Exp}\left(\boldsymbol{\tau}+\mathbf{J}_r^{-1}(\boldsymbol{\tau})\delta\boldsymbol{\tau}\right) \tag{44}$$

$$\mathrm{Log}(\mathrm{Exp}(\boldsymbol{\tau})\mathrm{Exp}(\delta\boldsymbol{\tau})) \approx \boldsymbol{\tau}+\mathbf{J}_r^{-1}(\boldsymbol{\tau})\delta\boldsymbol{\tau}. \tag{45}$$

作为补充，流形 $\mathcal{M}$ 的左 Jacobian 矩阵被定义为，

$$\mathbf{J}_l(\boldsymbol{\tau}) \triangleq \frac{{}^{\mathcal{E}}D\mathrm{Exp}(\boldsymbol{\tau})}{D\boldsymbol{\tau}} \in \mathbb{R}^{m\times m}, \tag{46}$$

使用左 Jacobian 矩阵方程 (38)，得出近似值

$$\mathrm{Exp}(\boldsymbol{\tau}+\delta\boldsymbol{\tau}) \approx \mathrm{Exp}\left(\mathbf{J}_l(\boldsymbol{\tau})\delta\boldsymbol{\tau}\right)\mathrm{Exp}(\boldsymbol{\tau}) \tag{47}$$

$$\mathrm{Exp}(\delta\boldsymbol{\tau})\mathrm{Exp}(\boldsymbol{\tau}) \approx \mathrm{Exp}\left(\boldsymbol{\tau}+\mathbf{J}_l^{-1}(\boldsymbol{\tau})\delta\boldsymbol{\tau}\right) \tag{48}$$

$$\mathrm{Log}(\mathrm{Exp}(\delta\boldsymbol{\tau})\mathrm{Exp}(\boldsymbol{\tau})) \approx \boldsymbol{\tau}+\mathbf{J}_l^{-1}(\boldsymbol{\tau})\delta\boldsymbol{\tau}. \tag{49}$$

左 Jacobian 矩阵将参数 $\boldsymbol{\tau}$ 的变化映射到全局 (*global*) 切空间或李代数中的变化，此即全局切空间的指数映射的导数。从方程 (43, 47) 我们可以把左 Jacobian 矩阵和右 Jacobian 矩阵用伴随联系起来，

$$\mathbf{Ad}_{\mathrm{Exp}(\boldsymbol{\tau})} = \mathbf{J}_l(\boldsymbol{\tau})\mathbf{J}_r^{-1}(\boldsymbol{\tau}). \tag{50}$$

此外，链式法则允许我们关联 $\mathbf{J}_r$ 和 $\mathbf{J}_l$ ，

$$\begin{aligned}\mathbf{J}_r(-\boldsymbol{\tau}) &\triangleq \mathbf{J}_{-\boldsymbol{\tau}}^{\mathrm{Exp}(-\boldsymbol{\tau})} = \mathbf{J}_{\boldsymbol{\tau}}^{\mathrm{Exp}(-\boldsymbol{\tau})}\mathbf{J}_{-\boldsymbol{\tau}}^{\boldsymbol{\tau}} = \mathbf{J}_{\boldsymbol{\tau}}^{\mathrm{Exp}(\boldsymbol{\tau})^{-1}}(-\mathbf{I}) \\ &= -\mathbf{J}_{\mathrm{Exp}(\boldsymbol{\tau})}^{\mathrm{Exp}(\boldsymbol{\tau})}\mathbf{J}_{\boldsymbol{\tau}}^{\mathrm{Exp}(\boldsymbol{\tau})} = \mathbf{Ad}_{\mathrm{Exp}(\boldsymbol{\tau})}\mathbf{J}_r(\boldsymbol{\tau}) \\ &= \mathbf{J}_l(\boldsymbol{\tau}).\end{aligned} \tag{51}$$

对于使用中的经典流形，如 $SO(3)$ 和 $SE(3)$，$\mathbf{J}_r, \mathbf{J}_r^{-1}, \mathbf{J}_l$ 和 $\mathbf{J}_l^{-1}$，存在封闭形式。

群作用 对于 $\mathcal{X} \in \mathcal{M}$ 和 $v \in \mathcal{V}$ 我们用方程 (35a) 定义

$$\mathbf{J}_{\mathcal{X}}^{\mathcal{X}\cdot v} \triangleq \frac{{}^{\mathcal{X}}D\mathcal{X}\cdot v}{D\mathcal{X}} \tag{52}$$

$$\mathbf{J}_{v}^{\mathcal{X}\cdot v} \triangleq \frac{{}^{v}D\mathcal{X}\cdot v}{Dv}. \tag{53}$$

由于群作用依赖于集合 $\mathcal{V}$，因此这些表达式不能通用化。

Log 映射的导数 对于 $\boldsymbol{\tau} = \mathrm{Log}(\mathcal{X})$，并根据方程 (45)，

$$\mathbf{J}_{\mathcal{X}}^{\mathrm{Log}(\mathcal{X})} = \mathbf{J}_r^{-1}(\boldsymbol{\tau}) \tag{54}$$

加号和减号的 Jacobian 矩阵 我们有

$$\mathbf{J}_{\mathcal{X}}^{\mathcal{X}\oplus\boldsymbol{\tau}} = \mathbf{J}_{\mathcal{X}}^{\mathcal{X}\circ(\mathrm{Exp}(\boldsymbol{\tau}))} = \mathbf{Ad}_{\mathrm{Exp}(\boldsymbol{\tau})}{}^{-1} \tag{55}$$

$$\mathbf{J}_{\boldsymbol{\tau}}^{\mathcal{X}\oplus\boldsymbol{\tau}} = \mathbf{J}_{\mathrm{Exp}(\boldsymbol{\tau})}^{\mathcal{X}\circ(\mathrm{Exp}(\boldsymbol{\tau}))}\mathbf{J}_{\boldsymbol{\tau}}^{\mathrm{Exp}(\boldsymbol{\tau})} = \mathbf{J}_r(\boldsymbol{\tau}) \tag{56}$$

并得到 $\mathcal{Z} = \mathcal{X}^{-1}\circ\mathcal{Y}$ 和 $\boldsymbol{\tau} = \mathcal{Y}\ominus\mathcal{X} = \mathrm{Log}(\mathcal{Z})$ 相关的 Jacobian 矩阵

$$\mathbf{J}_{\mathcal{X}}^{\mathcal{Y}\ominus\mathcal{X}} = \mathbf{J}_{\mathcal{Z}}^{\mathrm{Log}(\mathcal{Z})}\mathbf{J}_{\mathcal{X}^{-1}}^{\mathcal{Z}}\mathbf{J}_{\mathcal{X}}^{\mathcal{X}^{-1}} = -\mathbf{J}_l^{-1}(\boldsymbol{\tau}) \tag{57}$$

$$\mathbf{J}_{\mathcal{Y}}^{\mathcal{Y}\ominus\mathcal{X}} = \mathbf{J}_{\mathcal{Z}}^{\mathrm{Log}(\mathcal{Z})}\mathbf{J}_{\mathcal{Y}}^{\mathcal{Z}} = \mathbf{J}_r^{-1}(\boldsymbol{\tau}). \tag{58}$$

2.9 旋转群 S^3 与 SO(3)

单位四元数群 S^3 和特殊正交矩阵群 SO(3)，两者都可旋转 3 元向量。它们有同构的切空间，其元素可用 $\mathbb{R}^3$ 中的旋转向量标识，所以我们把它们放在一起研究。正是在这个空间 $\mathbb{R}^3$ 中，我们定义旋转率 $\boldsymbol{\omega} \triangleq \mathbf{u}\omega$、轴-角 $\boldsymbol{\theta} \triangleq \mathbf{u}\theta$，以及所有扰动和不确定性的向量。

四元数流形 S^3 是 SO(3) 的双倍覆盖，即 $\mathbf{q}$ 和 $-\mathbf{q}$ 表示相同的旋转 $\mathbf{R}$。第一个覆盖对应于正实部 $w > 0$ 的四元数。这两个群可以被认为是同构的第一覆盖。

2.9.1 Exp 与 Log 映射

Exp 与 Log 映射可以定义为 S^3 的四元数和 SO(3) 的旋转矩阵。对于四元数 $\mathbf{q} = (w, \mathbf{v}) \in \mathbb{H}$ 我们有

$$\mathbf{q} = \mathrm{Exp}(\theta\mathbf{u}) \triangleq \cos(\theta/2) + \mathbf{u}\sin(\theta/2) \quad \in \mathbb{H} \tag{59}$$

$$\theta\mathbf{u} = \mathrm{Log}(\mathbf{q}) \triangleq 2\mathbf{v}\frac{\arctan(\|\mathbf{v}\|, w)}{\|\mathbf{v}\|} \quad \in \mathbb{R}^3 \tag{60}$$

对于旋转矩阵我们有，

$$\mathbf{R} = \mathrm{Exp}(\theta\mathbf{u}) \triangleq \mathbf{I} + \sin\theta\,[\mathbf{u}]_\times + (1-\cos\theta)\,[\mathbf{u}]_\times^2 \quad \in \mathbb{R}^{3\times3} \tag{61}$$

$$\theta\mathbf{u} = \mathrm{Log}(\mathbf{R}) \triangleq \frac{\theta\left(\mathbf{R}-\mathbf{R}^\top\right)^\vee}{2\sin\theta} \quad \in \mathbb{R}^3, \tag{62}$$

其中 $\theta = \cos^{-1}\left(\frac{\mathrm{trace}(\mathbf{R})-1}{2}\right)$。

2.9.2 旋转作用

给定上述四元数和旋转矩阵的表达式，四元数在 3 元向量上的旋转作用是由双四元数积来完成的，

$$\mathbf{x}' = \mathbf{q}\mathbf{x}\mathbf{q}^* \tag{63}$$

当旋转矩阵使用单个矩阵积时，

$$\mathbf{x}' = \mathbf{R}\mathbf{x}. \tag{64}$$

两者相当于一个围绕轴 $\mathbf{u}$ 旋转角度 θ 弧度 (rad) 的右手旋转。从四元数 $\mathbf{q}$ 变换到旋转矩阵 $\mathbf{R}$ 的公式为

$$\mathbf{R}(\mathbf{q}) = \begin{bmatrix} w^2+x^2-y^2-z^2 & 2(xy-wz) & 2(xz+wy) \\ 2(xy+wz) & w^2-x^2+y^2-z^2 & 2(yz-wx) \\ 2(xz-wy) & 2(yz+wx) & w^2-x^2-y^2+z^2 \end{bmatrix}. \tag{65}$$

2.9.3 旋转群的 Jacobian 矩阵

由于我们定义的导数映射切向量空间，并且这些空间重叠于 S^3 和 SO(3) 的三维旋转流形，即，$\boldsymbol{\theta} = \mathrm{Log}(\mathbf{q}) = \mathrm{Log}(\mathbf{R})$，因此 Jacobian 矩阵独立于所使用的表示 ($\mathbf{q}$ 或 $\mathbf{R}$)。因此，我们考虑通用的 3D 旋转元素，并用无衬线字体 R 来标记它们。

伴随: 从方程 (25) 我们有

$$\mathbf{Ad}_{\mathsf{R}}\boldsymbol{\theta} = \left(\mathbf{R}\left[\boldsymbol{\theta}\right]_\times \mathbf{R}^\top\right)^\vee = \left(\left[(\mathbf{R}\boldsymbol{\theta})\right]_\times\right)^\vee = \mathbf{R}\boldsymbol{\theta}$$

因此

$$\mathbf{Ad}_{\mathsf{R}} = \mathbf{R}, \tag{66}$$

具体的，对于 $\mathbf{q}$ 有 $\mathbf{Ad}_{\mathbf{q}} = \mathbf{R}(\mathbf{q})$，参见方程 (65)，对于 $\mathbf{R}$ 有 $\mathbf{Ad}_{\mathbf{R}} = \mathbf{R}$。

右 Jacobian 矩阵与左 Jacobian 矩阵: 它们有封闭形式

$$\mathbf{J}_r(\boldsymbol{\theta}) = \mathbf{I} - \frac{1-\cos\theta}{\theta^2}\left[\boldsymbol{\theta}\right]_\times + \frac{\theta-\sin\theta}{\theta^3}\left[\boldsymbol{\theta}\right]_\times^2 \tag{67}$$

$$\mathbf{J}_r^{-1}(\boldsymbol{\theta}) = \mathbf{I} + \frac{1}{2}\left[\boldsymbol{\theta}\right]_\times + \left(\frac{1}{\theta^2} - \frac{1+\cos\theta}{2\theta\sin\theta}\right)\left[\boldsymbol{\theta}\right]_\times^2 \tag{68}$$

$$\mathbf{J}_l(\boldsymbol{\theta}) = \mathbf{I} + \frac{1-\cos\theta}{\theta^2}\left[\boldsymbol{\theta}\right]_\times + \frac{\theta-\sin\theta}{\theta^3}\left[\boldsymbol{\theta}\right]_\times^2 \tag{69}$$

$$\mathbf{J}_l^{-1}(\boldsymbol{\theta}) = \mathbf{I} - \frac{1}{2}\left[\boldsymbol{\theta}\right]_\times + \left(\frac{1}{\theta^2} - \frac{1+\cos\theta}{2\theta\sin\theta}\right)\left[\boldsymbol{\theta}\right]_\times^2 \tag{70}$$

其中我们可以观察到

$$\mathbf{J}_l = \mathbf{J}_r^\top, \quad \mathbf{J}_l^{-1} = \mathbf{J}_r^{-\top} \tag{71}$$

右结合的加号和减号： 对于 $\boldsymbol{\theta} = \mathsf{Q} \ominus \mathsf{R}$，我们有

$$\mathbf{J}_{\mathsf{R}}^{\mathsf{R}\oplus\boldsymbol{\theta}} = \mathbf{R}\left(\boldsymbol{\theta}\right)^{\top} \quad \mathbf{J}_{\boldsymbol{\theta}}^{\mathsf{R}\oplus\boldsymbol{\theta}} = \mathbf{J}_{r}\left(\boldsymbol{\theta}\right) \tag{72}$$

$$\mathbf{J}_{\mathsf{Q}}^{\mathsf{Q}\ominus\mathsf{R}} = \mathbf{J}_{r}^{-1}\left(\boldsymbol{\theta}\right) \quad \mathbf{J}_{\mathsf{R}}^{\mathsf{Q}\ominus\mathsf{R}} = -\mathbf{J}_{l}^{-1}\left(\boldsymbol{\theta}\right) \tag{73}$$

旋转作用： 我们有

$$\begin{aligned} \mathbf{J}_{\mathsf{R}}^{\mathsf{R}\cdot\mathsf{v}} \triangleq \lim_{\boldsymbol{\theta}\to 0} \frac{\left(\mathbf{R}\oplus\boldsymbol{\theta}\right)\mathbf{v} - \mathbf{R}\mathbf{v}}{\boldsymbol{\theta}} = \\ \frac{\mathbf{R}\mathrm{Exp}\left(\boldsymbol{\theta}\right)\mathbf{v} - \mathbf{R}\mathbf{v}}{\boldsymbol{\theta}} = \lim_{\boldsymbol{\theta}\to 0} \frac{\mathbf{R}\left(\mathbf{I} + \left[\boldsymbol{\theta}\right]_{\times}\right)\mathbf{v} - \mathbf{R}\mathbf{v}}{\boldsymbol{\theta}} \\ = \lim_{\boldsymbol{\theta}\to 0} \frac{\mathbf{R}\left[\boldsymbol{\theta}\right]_{\times}\mathbf{v}}{\boldsymbol{\theta}} = \lim_{\boldsymbol{\theta}\to 0} \frac{-\mathbf{R}\left[\mathbf{v}\right]_{\times}\boldsymbol{\theta}}{\boldsymbol{\theta}} = -\mathbf{R}\left[\mathbf{v}\right]_{\times} \end{aligned} \tag{74}$$

其中我们使用属性 $\mathrm{Exp}\left(\boldsymbol{\theta}\right) \approx \mathbf{I} + \left[\boldsymbol{\theta}\right]_{\times}$ 和 $\left[\mathbf{a}\right]_{\times}\mathbf{b} = -\left[\mathbf{b}\right]_{\times}\mathbf{a}$。另一个 Jacobian 矩阵为

$$\mathbf{J}_{\mathbf{v}}^{\mathsf{R}\cdot\mathsf{v}} \triangleq \lim_{\partial\mathbf{v}\to 0} \frac{\mathbf{R}\left(\mathbf{v} + \partial\mathbf{v}\right) - \mathbf{R}\mathbf{v}}{\partial\mathbf{v}} = \mathbf{R}. \tag{75}$$

2.10 补充说明

本章的内容主要来自文献 [1]。但是因为文献 [1] 太过于惜字如金，试图在 17 页的论文里讲述李群的入门知识，因此有很多重要的李群和李代数的知识没能加入到论文中。比如，李括号的内容就特意忽略，因此相关的李代数的伴随表示，以及指数映射的导数等等内容就只能在给出几个公式后快速带过。但是，伴随表示和指数映射的导数是李群理论中的两个核心内容。因此后面两个章节将补充这方面的内容，然后再用一章的示例，以旋转群 SO(3) 为例去理解这些知识。

3 伴随表示

在数学中，李群 G 的**伴随表示** (**adjoint representation**) 或**伴随作用** (**adjoint action**) 是一种将群元素表示为群的李代数的线性变换的方法，其中李代数被认为是向量空间。例如，如果 G 是 $GL(n,\mathbb{R})$，实 $n \times n$ 可逆矩阵的李群，则伴随表示是群同态，它将 $n \times n$ 可逆的矩阵 g 发送到 $\mathbb{R}^n$ 的所有线性变换的向量空间的自同态，该向量空间被定义为 $x \mapsto gxg^{-1}$。

对于任意李群，这种自然表示是通过共轭将 G 对其自身的作用线性化 (即取其微分) 而获得的。伴随表示可以定义为任意域上的线性代数群。

3.1 定义

设 G 为一个李群，并设

$$\Psi : G \to \mathrm{Aut}(G)$$

为映射 $g \mapsto \Psi_g$，其中 $\mathrm{Aut}(G)$ 为 G 的自同构群，并且 $\Psi_g : G \to G$ 由内部自同构 (共轭) 给出为

$$\Psi_g(h) = ghg^{-1}.$$

这个 Ψ 是一个李群同态。

对于在 G 中的每一个 g，定义 Ad_g 为 Ψ_g 在原点处的导数：

$$\mathrm{Ad}_g = \left(d\Psi_g\right)_e : T_eG \to T_eG$$

其中 d 是微分，并且 $\mathfrak{g} = T_eG$ 是在原点 e 处 (e 为群 G 的恒等元) 的切空间。由于 Ψ_g 是一个李群自同构，Ad_g 是李代数自同构；即 $\mathfrak{g}$ 到其自身的可逆线性变换，并保持李括号。而且，因为 $g \mapsto \Psi_g$ 是一个群同态，所以 $g \mapsto \mathrm{Ad}_g$ 也是一个群同态。因此，映射

$$\mathrm{Ad} : G \to \mathrm{Aut}(\mathfrak{g}), g \mapsto \mathrm{Ad}_g$$

是一个群表示，被称为 G 的**伴随表示** (**adjoint representation**)。

如果 G 是一般线性群 $\mathrm{GL}_n(\mathbb{C})$ (被称为浸入线性李群) 的浸入李子群，则李代数 $\mathfrak{g}$ 由矩阵组成，并且对于具有小的算子范数的矩阵 X，指数映射是矩阵指数 $\exp(X) = e^X$。因此，对于在 G 中的 g 以及在 $\mathfrak{g}$ 中的小 X ，取在 $t = 0$ 处 $\Psi_g(\exp(tX)) = ge^{tX}g^{-1}$ 的导数，我们得到：

$$\mathrm{Ad}_g(X) = gXg^{-1}$$

其中在右侧我们有矩阵的乘积。若 $G \subset \mathrm{GL}_n(\mathbb{C})$ 是一个闭子群 (即 G 一个是矩阵李群)，则该公式对于在 G 中的所有 g 和在 $\mathfrak{g}$ 中的所有 X 都有效。

简单地说，伴随表示是在 G 的单位元周围与 G 共轭作用相关联的各向同性表示 (isotropy representation)。

3.2 Ad 的导数

通过在恒等式处取导数，我们始终可以从李群 G 的表示形式传递到李代数的表示形式。

取伴随映射的导数

$$\mathrm{Ad} : G \to \mathrm{Aut}(\mathfrak{g})$$

在单位元处给出 G 的李代数 $\mathfrak{g} = \mathrm{Lie}(G)$ 的**伴随表示** (**adjoint representation**)：

$$\begin{aligned} \mathrm{ad} : \mathfrak{g} &\to \mathrm{Der}(\mathfrak{g}) \\ x &\mapsto \mathrm{ad}_x = d(\mathrm{Ad})_e(x) \end{aligned}$$

其中 $\mathrm{Der}(\mathfrak{g}) = \mathrm{Lie}(\mathrm{Aut}(\mathfrak{g}))$ 是 $\mathrm{Aut}(\mathfrak{g})$ 的李代数，其可以通过 $\mathfrak{g}$ 的导子代数标识。我们可以证明

$$\mathrm{ad}_x(y) = [x, y]$$

对于所有的 $x, y \in \mathfrak{g}$ 成立，其中右侧由向量场的李括号给出 (诱导)。实际上，回想一下，将 $\mathfrak{g}$ 视为在 G 上左不变向量场的李代数，在 $\mathfrak{g}$ 上的括号给出为：对于左不变向量场 X, Y，

$$[X, Y] = \lim_{t \to 0} \frac{1}{t} \left(d\varphi_{-t}(Y) - Y\right)$$

其中 $\varphi_t : G \to G$ 标识由 X 生成的流。结果表明，$\varphi_t(g) = g\varphi_t(e)$，这大致是因为两侧满足定义流的相同 ODE。也就是，$\varphi_t = R_{\varphi_t(e)}$，其中 R_h 标识与 $h \in G$ 的右乘法。另一方面，由于 $\Psi_g = R_{g^{-1}} \circ L_g$，根据链式规则，

$$\mathrm{Ad}_g(Y) = d\left(R_{g^{-1}} \circ L_g\right)(Y) = dR_{g^{-1}}\left(dL_g(Y)\right) = dR_{g^{-1}}(Y)$$

因为 Y 是左不变的。因此，

$$[X, Y] = \lim_{t \to 0} \frac{1}{t} \left(\mathrm{Ad}_{\varphi_t(e)}(Y) - Y\right)$$

这就是我们需要证明的。

因此，ad_x 与下面第 3.3 节“李代数的伴随表示”中的定义相同。Ad 和 ad 通过指数映射相关：具体地，$\mathrm{Ad}_{\exp(x)} = \exp(\mathrm{ad}_x)$ 对于在李代数中所有的 x 都成立。它是通过指数映射将李群与李代数同态联系起来的一般结果的一个后果。

如果 G 是一个浸入线性李群，则上述计算将简化，如前所述，$\mathrm{Ad}_g(Y) = gYg^{-1}$，并因此有 $g = e^{tX}$，

$$\mathrm{Ad}_{e^{tX}}(Y) = e^{tX}Ye^{-tX}.$$

取其在 $t = 0$ 处的导数，我们有：

$$\mathrm{ad}_X Y = XY - YX.$$

一般情况也可以由线性情形导出：其实，设 G' 为一个浸入线性李群，其李代数与 G 相同，则 Ad 在 G 的单位元处的导数与 G' 的导数重合；因此，在不丧失一般性的情况下，G 可以假定为 G'。

大写/小写符号在文献中广泛使用。因此，例如，在代数 $\mathfrak{g}$ 中的一个向量 x 生成在群 G 中的一个向量场 X。类似地，在 $\mathfrak{g}$ 中的向量的伴随映射 $\mathrm{ad}_x y = [x, y]$ 同态于在群 G 上 (其做为一个流形) 的向量场的李导数 $\mathrm{L}_X Y = [X, Y]$。

进一步的内容请参见“指数映射的导数”一文。

3.3 李代数的伴随表示

设 $\mathfrak{g}$ 为某个域上的李代数。给定李代数 $\mathfrak{g}$ 的一个元素 x，我们定义在 $\mathfrak{g}$ 上的 x 的伴随作用为映射

$$\mathrm{ad}_x : \mathfrak{g} \to \mathfrak{g} \quad \text{with} \quad \mathrm{ad}_x(y) = [x, y]$$

其对所有在 $\mathfrak{g}$ 中的 y 成立。它被称为**伴随自同态** (**adjoint endomorphism**) 或**伴随作用** (**adjoint action**)。(ad_x 通常也标志为 $\mathrm{ad}(x)$。) 由于括号是双线性的，这决定线性映射

$$\mathrm{ad} : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g}) = (\mathrm{End}(\mathfrak{g}), [\,,\,])$$

由 $x \mapsto \mathrm{ad}_x$ 给出。在 $\mathrm{End}(\mathfrak{g})$ 内，根据定义，括号由两个算子的交换子 (commutator) 给出为：

$$[T, S] = T \circ S - S \circ T$$

其中 $\circ$ 标识线性映射的组合。使用上述括号定义雅各比恒等式 (Jacobi identity)

$$[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0$$

采用形式

$$([\mathrm{ad}_x, \mathrm{ad}_y])(z) = \left(\mathrm{ad}_{[x,y]}\right)(z)$$

其中 x, y 和 z 是 $\mathfrak{g}$ 的任意元素。

最后一个恒等式表明 ad 是一个李代数同态；即从括号到括号的线性映射。因此，ad 是李代数的一个表示，并被称为代数 $\mathfrak{g}$ 的**伴随表示** (**adjoint representation**)。

如果 $\mathfrak{g}$ 是有限维的，并且选择了它的一个基 (basis)，则 $\mathfrak{gl}(\mathfrak{g})$ 是平方矩阵的李代数，并且组合对应于矩阵乘法。

ad 的核是 $\mathfrak{g}$ 的中心 (这只是对定义的重新表述)。另一方面，对于在 $\mathfrak{g}$ 中的每个元素 Z，线性映射 $\delta = \mathrm{ad}_z$ 服从莱布尼兹定律：

$$\delta([x,y]) = [\delta(x), y] + [x, \delta(y)]$$

这对于在代数中的所有 x 和 y 都成立 (雅可比恒等式的重述)。也就是说，ad_z 是一个导数，并且 $\mathfrak{g}$ 在 ad 下的象 (image) 是 $\mathrm{Der}(\mathfrak{g})$ 的一个子代数，$\mathfrak{g}$ 的所有导数的空间。

当 $\mathfrak{g} = \mathrm{Lie}(G)$ 是李群 G 的李代数时，ad 是 Ad 在 G 的单位元处的微分。

3.4 属性

下表总结了在定义中提到的各种映射的属性

$\Psi : G \to \mathrm{Aut}(G)$	$\Psi_g : G \to G$
李群同态： - $\Psi_{gh} = \Psi_g \Psi_h$	李群自同构： - $\Psi_g(ab) = \Psi_g(a)\Psi_g(b)$ - $(\Psi_g)^{-1} = \Psi_{g^{-1}}$
$\mathrm{Ad} : G \to \mathrm{Aut}(\mathfrak{g})$	$\mathrm{Ad}_g : \mathfrak{g} \to \mathfrak{g}$
李群同态： - $\mathrm{Ad}_{gh} = \mathrm{Ad}_g \mathrm{Ad}_h$	李代数自同构： - Ad_g 是线性的 - $(\mathrm{Ad}_g)^{-1} = \mathrm{Ad}_{g^{-1}}$ - $\mathrm{Ad}_g[x,y] = [\mathrm{Ad}_g x, \mathrm{Ad}_g y]$
$\mathrm{ad} : \mathfrak{g} \to \mathrm{Der}(\mathfrak{g})$	$\mathrm{ad}_x : \mathfrak{g} \to \mathfrak{g}$
李代数同态： - ad 是线性的 - $\mathrm{ad}_{[x,y]} = [\mathrm{ad}_x, \mathrm{ad}_y]$	李代数导数： - ad_x 是线性的 - $\mathrm{ad}_x[y,z] = [\mathrm{ad}_x y, z] + [y, \mathrm{ad}_x z]$

3.5 证明

本节我们证明 Ad 与 ad 通过指数映射相关的关系式，因为在后面证明指数映射的导数时会用到这个关系式。

命题 1. 对于任意 $X \in \mathrm{M}_n(\mathbb{R}) = \mathfrak{gl}(n, \mathbb{R})$，我们有

$$\mathrm{Ad}_{e^X} = e^{\mathrm{ad}_X} = \sum_{k=0}^{\infty} \frac{(\mathrm{ad}_X)^k}{k!};$$

也就是，

$$\begin{aligned} e^X Y e^{-X} &= e^{\mathrm{ad}_X} Y \\ &= Y + [X,Y] + \frac{1}{2!}[X,[X,Y]] + \frac{1}{3!}[X,[X,[X,Y]]] \\ &\quad + \cdots \end{aligned}$$

对于所有的$X, Y \in \mathrm{M}_n(\mathbb{R})$ 成立。

证明. 如前所述，$\mathrm{Ad}_g(Y) = gYg^{-1}$，并且有 $g = e^{tX}$，则

$$\mathrm{Ad}_{e^{tX}}(Y) = e^{tX} Y e^{-tX}.$$

取其在 $t=0$ 处的导数，我们有：

$$\begin{aligned}\frac{d}{dt}\left(e^{tX}Ye^{-tX}\right)\bigg|_{t=0} &= \frac{d}{dt}(\exp(tX)\cdot Y)\cdot\exp(-tX)+\exp(tX)\cdot Y\cdot\frac{d}{dt}\exp(-tX)\bigg|_{t=0}\\ &= X\exp(tX)\cdot Y\cdot\exp(-tX)-\exp(tX)\cdot Y\cdot X\exp(-tX)|_{t=0}\\ &= XY-YX\\ &= [X,Y]\\ &= \mathrm{ad}_X Y,\end{aligned}$$

即

$$\mathrm{Ad}_{e^X}=e^{\mathrm{ad}_X}.$$

□

这个关系式表明，对于给定的李代数元素 X，它的指数映射 $\exp(X)$ 的伴随矩阵等于对应的李代数的伴随矩阵的指数映射。这个公式的意义在于，通过指数映射和伴随矩阵的关系，我们可以在李群和李代数之间进行转换和对应。也即，李群和李代数之间存在指数映射的对应关系，对于任意李代数 $\mathfrak{g}$ 中的元素 X，我们都可以找到对应的李群 $\mathcal{G}$ 中的元素 $\exp(X)$；同时，李代数和李群之间也存在着相似的伴随表达变换的概念。因此, 李代数元素 X 与李群元素 $\exp(X)$ 之间，其伴随变换应该存在某种一一对应关系。这两者之间的联系可以通过这个公式 $\mathrm{Ad}_{\exp(X)}=\exp\left(\mathrm{ad}_X\right)$ 来表达。

3.6 伴随表示的几何意义

伴随表示是我们理解和研究李群和李代数的重要工具。从几何的角度，伴随表示提供了关于李群和李代数内部运动方式的直观理解；从本质的角度，伴随表示揭示了李群和李代数的内在结构，它连接了群的自动性质和代数的李括号运算。

3.6.1 伴随表示的几何直觉

伴随表示揭示了李群和李代数的内部运动方式。具体来说，对于李群 G 中的元素 g 和 h，伴随表示 $\mathrm{Ad}(g)(h)=ghg^{-1}$ 描述了一个“共轭”操作：首先按照 g 进行变换，然后进行 h 的变换，最后再按照 g^{-1} 的方式变换回来，最终得到的总的变换，即是新的元素 $\mathrm{Ad}(g)(h)$。在几何上，这可以理解为 g 定义了一个“参考框架”的改变，h 是在新的参考框架下的变换。

对于李代数 $\mathfrak{g}$ 中的元素 X 和 Y，伴随表示 $\mathrm{ad}(X)(Y)=[X,Y]=XY-YX$ 描述了一个“对易”操作，这个操作测量了 X 和 Y 的“不对易程度”。在几何上，这可以理解为 X 和 Y 对应的无穷小变换是否可以交换顺序。

3.6.2 伴随表示的本质理解

伴随表示的本质在于它提供了一种理解和描述李群和李代数中的“内部”运动的方式。这种“内部”运动是李群和李代数的重要性质，是理解和研究李群和李代数的重要工具。

更深入地，伴随表示揭示了李群和李代数的本质结构。对于李群来说，伴随表示描述了群的“自动”结构，即群元素自身如何通过共轭操作作用在群上。这是群结构的一个重要方面，因为它描述了群的内部运动方式。

对于李代数来说，伴随表示描述了李代数的“李结构”，即李括号运算。这是李代数结构的一个重要方面，因为它描述了李代数中元素的交换性质，这反映了李群的局部运动方式。

从这个角度来看，伴随表示提供了一种理解李群和李代数内部结构的重要工具，它将群和代数的结构联系起来，并且揭示了它们的几何和代数性质。

4 指数映射的导数

在李群理论中，指数映射是从一个李群 G 的李代数 $\mathfrak{g}$ 到 G 的映射。如果 G 是矩阵李群，则指数映射可约化为矩阵指数。指数映射，标志为 $\exp:\mathfrak{g}\to G$，是解析的，并且具有这样一个导数 $\frac{d}{dt}\exp(X(t)):\mathrm{T}\mathfrak{g}\to \mathrm{T}G$，其中 $X(t)$ 是在李代数中的一个 C^1 路径，以及一个密切相关的微分 $d\exp:\mathrm{T}\mathfrak{g}\to\mathrm{T}G$。

在本文中，符号 $\exp(X)$ 和 e^X 将互换使用，以标志给定参数的指数，除非如前所述，符号具有专用的不同含义。为了在方程中具有更好的可读性，这里首选微积分形式的符号。另一方面，exp 样式有时对于内联公式更方便，并且在需要进行真正区分的少数情况下是必需的。

4.1 声明

指数映射的导数给出为：

$$\boxed{\frac{d}{dt}e^{X(t)}=e^{X(t)}\frac{1-e^{-\mathrm{ad}_X}}{\mathrm{ad}_X}\frac{dX(t)}{dt}.} \tag{76}$$

解释

- $X=X(t)$ 是在李代数中的一条 C^1 (连续可微) 路径，其导数为 $X'(t)=\frac{dX(t)}{dt}$。在不需要时可省略参数 t。

- ad_X 是李代数的线性变换，给出为 $\mathrm{ad}_X(Y)=[X,Y]$。它是李代数对其自身的伴随作用。

- 分数 $\frac{1-\exp(-\mathrm{ad}_X)}{\mathrm{ad}_X}$ 由幂级数给出为

$$\frac{1-e^{-\mathrm{ad}_X}}{\mathrm{ad}_X}=\sum_{k=0}^{\infty}\frac{(-1)^k}{(k+1)!}\left(\mathrm{ad}_X\right)^k. \tag{77}$$

  它由线性自同态的指数映射的幂级数推导出，如在矩阵幂运算中。

- 当 G 是矩阵李群时，所有指数的出现都由它们的幂级数展开式给出。

- 当 G 不是矩阵李群时，$\frac{1-\exp(-\mathrm{ad}_X)}{\mathrm{ad}_X}$ 仍然由其幂级数方程 (77) 给出，而公式中 exp 的另外两个项，现在是在李理论中的指数映射，指左不变向量场 X 的“时间-1”的流，即在一般情况下定义的李代数元素，关于李群 G 的分析流形。这仍然与矩阵情况下的公式完全相同。代数 $\mathfrak{g}$ 的元素与李群的元素 $\exp(X(t))$ 的左乘法被解释为应用左平移 $\mathrm{dL}_{\exp(X(t))}$ 的微分。

- 该公式适用于 exp 被认为是在 $\mathbb{R}$ 或 $\mathbf{C}$ 上的矩阵空间的映射的情况，请参见矩阵指数。当 $G=\mathrm{GL}(n,\mathbf{C})$ 或 $\mathrm{GL}(n,\mathbf{R})$ 时，这些概念正好重合。

为了计算在 X 处 exp 的微分 $d\exp$，$d\exp_X : \mathrm{T}\mathfrak{g}_X \to \mathrm{T}G_{\exp(X)}$，我们使用标准方程

$$d\exp_X Y = \frac{d}{dt}e^{Z(t)}\bigg|_{t=0}, Z(0) = X, Z'(0) = Y.$$

对于 $Z(t) = X + tY$，跟随在方程 (76) 之后，其结果立刻得出为

$$d\exp_X Y = e^X \frac{1 - e^{-\mathrm{ad}_X}}{\mathrm{ad}_X} Y. \tag{78}$$

特别地，$d\exp_0 : \mathrm{T}\mathfrak{g}_0 \to \mathrm{T}G_{\exp(0)} = \mathrm{T}G_e$ 是恒等式，因为 $\mathrm{T}\mathfrak{g}_X \simeq \mathfrak{g}$ (因为 $\mathfrak{g}$ 是向量空间) 并且 $\mathrm{T}G_e \simeq \mathfrak{g}$。

4.2 证明

下面给出的证明假设一个矩阵李群。这意味着从李代数到矩阵李群的指数映射由通常的幂级数给出，即矩阵指数化。只要 exp 的每一次出现都得到正确的解释，证明结论在一般情况下仍然成立。

证明大纲使用了参数化表达式相对于 s 微分的技巧

$$\Gamma(s,t) = e^{-sX(t)}\frac{\partial}{\partial t}e^{sX(t)}$$

以获得 Γ 的一阶微分方程，其可在 s 中的直接积分法求解。则该解为 $e^X\Gamma(1,\mathrm{t})$ 。

引理. 设 Ad 标志群在其李代数上的伴随作用。对于 $A \in G, X \in \mathfrak{g}$，其作用给出为 $\mathrm{Ad}_A X = AXA^{-1}$。一个 Ad 和 ad 之间的常用关系给出为[1]

$$\boxed{\mathrm{Ad}_{e^X} = e^{\mathrm{ad}_X}, X \in \mathfrak{g}.} \tag{79}$$

证明. 使用乘积规则两次寻找，

$$\frac{\partial\Gamma}{\partial s} = e^{-sX}(-X)\frac{\partial}{\partial t}e^{sX(t)} + e^{-sX}\frac{\partial}{\partial t}\left[X(t)e^{sX(t)}\right] = e^{-sX}\frac{dX}{dt}e^{sX}.$$

则通过上述方程 (79)，我们观察到

$$\frac{\partial\Gamma}{\partial s} = \mathrm{Ad}_{e^{-sX}}X' = e^{-\mathrm{ad}_{sX}}X'.$$

收集各项产出

$$\Gamma(1,t) = e^{-X(t)}\frac{\partial}{\partial t}e^{X(t)} = \int_0^1 \frac{\partial\Gamma}{\partial s}ds = \int_0^1 e^{-\mathrm{ad}_{sX}}X'ds.$$

使用形式幂级数展开指数，逐项积分，并最后识别方程 (77)，

$$\Gamma(1,t) = \int_0^1 \sum_{k=0}^{\infty}\frac{(-1)^k s^k}{k!}(\mathrm{ad}_X)^k\frac{dX}{dt}ds = \sum_{k=0}^{\infty}\frac{(-1)^k}{(k+1)!}(\mathrm{ad}_X)^k\frac{dX}{dt} = \frac{1-e^{-\mathrm{ad}_X}}{\mathrm{ad}_X}\frac{dX}{dt},$$

这就得出结果。 □

[1] 可在这里找到该恒等式的一个证明。这个关系很简单，根据李对应 (Lie correspondence) 关系，这就是一个李群的表示和它的李代数的表示之间的关系，因为 Ad 和 ad 都是 $\mathrm{ad} = d\mathrm{Ad}$ 的表示。

4.3 备注

注 1. 从参考资料中对比指数映射的导数公式，就会发现文献 [3,4,5] 的结论

$$D_{\exp}(X) = \sum_{i=0}^{\infty} \frac{(-1)^i \mathrm{ad}_X^i}{(i+1)!} \tag{80}$$

与文献 [2] 的方程 (35)

$$D_{\exp}(X) = \sum_{i=0}^{\infty} \frac{\mathrm{ad}_X^i}{(i+1)!} \tag{81}$$

不一致。这其实是选取全局坐标和局部坐标造成的差异。

本章的前两节的内容与文献 [3,4,5] 的内容一致，对应的是文献 [1] 中表示局部坐标系的右结合 (right-) 算子 $\oplus$ 和 $\ominus$，文献 [1] 的方程 (25,26)，即本文的方程 (19,20)，以及表示局部 X 处的导数，即右 Jacobian 矩阵，文献 [1] 的方程 (41a)，即本文的方程 (35a)。其结果，方程 (78)，对应于文献 [1] 的方程 (67)，即本文的方程 (42)。

文献 [2] 从其定义方程 (1) 开始，对应的是文献 [1] 中表示全局坐标系的左结合 (left-) 算子 $\oplus$ 和 $\ominus$，文献 [1] 的方程 (27,28)，即本文的方程 (21,22)，以及表示全局或原点处的导数，即左 Jacobian 矩阵，文献 [1] 的方程 (44)，即本文的方程 (38)。其结果，文献 [2] 的方程 (35)，对应于文献 [1] 的方程 (71)，即本文的方程 (46)。

全局表示与局部表示的导数之间的关系，参见文献 [1] 的方程 (75,76)，即本文的方程 (50,51)。

要证明全局的指数映射的导数也可使用上一节同样的方法，只需要将 e^{-sX} 和 e^{sX} 对调即可证明。文献 [2] 也提供了类似的证明方法。

另外，对于幂级数

$$\sum_{i=0}^{\infty} \frac{(-1)^i \mathrm{ad}_x^i}{(i+1)!}$$

通常习惯地被简写为

$$\frac{\mathrm{id} - e^{-\mathrm{ad}_x}}{\mathrm{ad}_x},$$

这样的写法更加简洁，也反映了这个级数与指数函数关系的本质。类似地，对于幂级数

$$\sum_{i=0}^{\infty} \frac{\mathrm{ad}_x^i}{(i+1)!}$$

也可以被简写为

$$\frac{e^{\mathrm{ad}_x} - \mathrm{id}}{\mathrm{ad}_x}.$$

注 2. 对数映射 log 是指数映射 exp 的逆函数。在函数的双射区域中，逆映射的导数是导数的逆映射。因此 log 的导数为

$$D_{\log}(x) = D_{\exp}^{-1}(x) \tag{82}$$

注 3. 文献 [1,2] 与文献 [3,4,5]，对于导数的定义方程还有一些微妙的差异。文献 [3,4,5] 是相对于一个实数，即相对于时间 $t = 0$ 的求导；而文献 [1,2] 是相对于一个李代数元素或向量元素的求导，即相对于一个状态量的求导。这两者都有各自的应用领域。例如，在使用 EKF/ESKF 对刚体的位姿

进行最优估计的算法中就得到应用。EKF/ESKF 的一次迭代分为两个阶段：时间更新和测量更新。在时间更新阶段，我们需要使用运动学方程或动力学方程，也就是使用时间相关的方程，因此使用的导数适用于文献 [3,4,5] 的定义。在测量更新阶段，我们认为用于校正的测量值是同步测量得到，只有状态量差异而没有时间差异，因此使用的导数适用于文献 [1,2] 的定义。

注 4. 我们想寻找一个 $n \times n$ 雅可比矩阵 $f'(\xi)$，以使得

$$f(\xi+\delta) \approx f(\xi) \exp\left(\widehat{f'(\xi)\delta}\right) \tag{83}$$

该方程即前面的近似方程 (43)。指数映射的导数告诉我们，对于任意李代数元素 $\hat{\xi} \in \mathfrak{g}$，存在一个线性映射 $d\exp_{\hat{\xi}} : T_{\hat{\xi}}\mathfrak{g} \to T_{\exp(\hat{\xi})}G$，给出为

$$d\exp_{\hat{\xi}} \hat{x} = \exp(\hat{\xi}) \frac{1-\exp(-\mathrm{ad}_{\hat{\xi}})}{\mathrm{ad}_{\hat{\xi}}} \hat{x} \tag{84}$$

其中 $\hat{x} \in T_{\hat{\xi}}\mathfrak{g}$，并且 $\mathrm{ad}_{\hat{\xi}}$ 它本身是一个线性映射，把 $\hat{x}$ 带到 $[\hat{\xi}, \hat{x}]$，即李括号。上面的公式很重要，因为它直接根据切向量来表达 $\mathfrak{g}$ 和 G，但是了解线性映射存在的事实更重要。方程 (84) 是一个切向量，我们看到它将 x 映射到局部坐标 y，如下所示：

$$\hat{y} = \frac{1-\exp(-\mathrm{ad}_{\hat{\xi}})}{\mathrm{ad}_{\hat{\xi}}} \hat{x}$$

这可被用来构造雅可比矩阵 $f'(\xi)$。

4.4 指数映射与导数的几何直觉

理解指数映射和导数的几何直觉能加深我们对李群和李代数结构的理解。下面我们从几何的角度来理解指数映射和导数的意义。

4.4.1 指数映射的几何直觉

指数映射的几何直觉可以从其在李群和李代数中的定义理解。在李代数中，元素 X 可以看作是无穷小变换，而指数映射 $\exp(X)$ 则生成了对应的有限变换。

指数映射 $\exp : \mathfrak{g} \to G$ 从李代数 $\mathfrak{g}$ 到李群 G 的映射，从几何上看，它是一个从切空间 (李代数) 到流形 (李群) 的映射。指数映射给出了从李群的单位元 e 出发，沿着给定方向 $X \in \mathfrak{g}$ “行进”的路径。具体地，对于 $t \in \mathbb{R}$，$\exp(tX)$ 给出了从 e 出发，沿着方向 X，“行走”距离 t 后到达的位置。

在几何上，这可以被视为在李群的流形上定义的一种“直线”或“测地线”。这个“直线”的特性是，它始终沿着 (由 X 给出的) 同一方向前进，即其是由 X 生成的单参数子群在 G 上的积分曲线。

4.4.2 导数的几何直觉

导数的几何直觉可以通过考虑李群元素或李代数元素的微小变化来理解。在李群中，元素 g 的微小变化可以通过考虑 g 在单位元附近的运动来理解。在李代数中，元素 X 的微小变化可以通过考虑 X 在零元附近的运动来理解。

导数在几何上通常被理解为一个函数的局部线性近似。在李群和李代数的语境中，这一概念有两种主要的几何解释。

首先，指数映射的导数 $\frac{d}{dt}\exp(tX)$ 描述了如何沿着李群的“直线”前进。具体地，$\exp(tX)\cdot X$ 给出了在点 $\exp(tX)$ 处，沿着方向 X “行走”的速度。这直观地说明了如何在李群的流形上“前进”。

其次，对数映射的导数 $\frac{d}{dt}\log(g\exp(tX))$ 描述了如何在李代数中变化。具体地，$\mathrm{Ad}_{g^{-1}}(X)$ 给出了在点 g 处，对于李群元素的微小变化 $g\exp(tX)$，其对应的李代数元素如何变化。这提供了一种理解李代数的动态变化的方法。

这两种几何解释为我们提供了深入理解李群和李代数，以及它们之间关系的视角。这些视角在理解和应用李群和李代数的理论时都非常有用。

5 三维旋转群 SO(3) 及其李代数 so(3)

5.1 SO(3) 和 so(3) 的基本性质

5.1.1 SO(3) 作为李群

三维旋转群 SO(3) 被定义为 3×3 正交矩阵和旋转组成的群:

$$SO(3) = \{R \in \mathbb{R}^{3\times 3} | RR^T = I, \det(R) = 1\}$$

具有矩阵乘法的群运算。SO(3) 是一个紧致李群。

旋转群 SO(3) 具有一些重要的性质：

- 封闭性：两个旋转的组合仍然是一个旋转，即，如果 $R_1, R_2 \in SO(3)$，则 $R_1R_2 \in SO(3)$。
- 连续性：对于任意小的 $\epsilon > 0$，总存在一个足够小的 $\delta > 0$，使得对于所有满足 $|R - I| < \delta$ 的 $R \in SO(3)$，我们有 $|RR_0 - R_0| < \epsilon$，对于所有 $R_0 \in SO(3)$ 成立。这说明了旋转群 SO(3) 是一个连续的群。
- 结合性：旋转群 SO(3) 满足结合律，即对于任意的 $R_1, R_2, R_3 \in SO(3)$，我们有 $(R_1R_2)R_3 = R_1(R_2R_3)$。
- 单位元和逆元：旋转群 SO(3) 的单位元是 3×3 单位矩阵 I，即对于任意的 $R \in SO(3)$，我们有 $RI = IR = R$。每个元素的逆元是其转置，即对于任意的 $R \in SO(3)$，我们有 $RR^T = R^TR = I$。

5.1.2 so(3) 作为李代数

对应的李代数 so(3) 被定义为 3×3 的反对称矩阵:

$$so(3) = \{A \in \mathbb{R}^{3\times 3} | A^T = -A\}$$

其李括号运算为矩阵的交换子。满足 $A^T = -A$ (A^T 是 A 的转置) 条件的矩阵可以表示为：

$$A = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix},$$

其中，$(a_1, a_2, a_3)^{\mathrm{T}}$ 是一个三维向量，它以轴-角形式描述了旋转的轴和角度。这个向量的模长等于旋转的角度，向量的方向与旋转的轴成正比。

首先，so(3) 作为一个李代数，它的主要性质是满足李括号运算，即对于任意的 $X, Y \in so(3)$，有 $[X, Y] = XY - YX \in so(3)$。其次，so(3) 也是一个向量空间，它对于向量加法和标量乘法封闭。

此外，李代数 so(3) 还具有一些重要的性质：

1. 线性空间结构：李代数 so(3) 是一个线性空间，意味着两个李代数元素的线性组合仍然是一个李代数元素。

2. 李括号运算：李代数 so(3) 中定义了一个特殊的二元运算，称为李括号或者李括号运算。对于任意的 $A, B \in \mathfrak{so}(3)$，李括号定义为 $[A, B] = AB - BA$。这个运算满足反对称性 ($[A, B] = -[B, A]$)、双线性性 (每个变量分别线性) 和雅可比恒等式 ($[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0$)。

3. 与 SO(3) 的关系：李代数 so(3) 可以被视为旋转群 SO(3) 在单位元处的切空间。具体来说，给定一个 so(3) 中的元素 A，可以通过指数映射 $\exp(A)$ 找到对应的 SO(3) 中的元素。反过来，也可以通过对数映射 $\log(R)$ 找到给定的 SO(3) 中元素 R 对应的 so(3) 中的元素。

5.1.3 与三维空间的联系

SO(3) 作用于 $\mathbb{R}^3$ 给出三维空间的旋转变换。so(3) 与叉积运算对应。

5.2 三维向量的叉乘与斜对称矩阵

三维向量的叉乘和斜对称矩阵在旋转群 SO(3) 和李代数 so(3) 的理论中有着紧密的联系。

5.2.1 三维向量的叉乘

三维向量的叉乘是一个二元运算，对于两个三维向量 $a, b \in \mathbb{R}^3$，其叉乘定义为

$$a \times b = \begin{bmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{bmatrix},$$

其中，$a = (a_1, a_2, a_3)^T$, $b = (b_1, b_2, b_3)^T$。

叉乘操作具有以下性质：

- 反对称性：$a \times b = -b \times a$
- 分配律：$a \times (b + c) = a \times b + a \times c$
- 双线性：$(\alpha a + \beta b) \times c = \alpha(a \times c) + \beta(b \times c)$，其中 α, β 是标量。

叉乘操作与李代数 so(3) 有着紧密的联系。实际上，给定一个 so(3) 中的元素 A，其对应的向量 a 可以通过 $A = [a]_\times$ 得到，其中 $[a]_\times$ 是由向量 a 定义的斜对称矩阵。

5.2.2 斜对称矩阵

斜对称矩阵是一种特殊的方阵，它满足 $A^T = -A$。对于一个三维向量 $a \in \mathbb{R}^3$，我们可以构造一个与之对应的斜对称矩阵

$$[a]_\times = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix}.$$

这个斜对称矩阵的一个重要性质是，对于任意三维向量 $b \in \mathbb{R}^3$，都有

$$[a]_\times b = a \times b.$$

这表明，对于 $so(3)$ 中的元素 X，可以找到一个对应的三维向量 a，使得 X 的作用等价于 a 的叉乘。即，斜对称矩阵可以用来表示三维空间中的叉乘操作。

斜对称矩阵与李代数 so(3) 的关系非常紧密。实际上，李代数 so(3) 就是所有 3×3 斜对称矩阵的集合。这提供了一种从向量到旋转的方法，即通过将向量转化为斜对称矩阵，然后在斜对称矩阵上应用指数映射，我们可以得到一种表示旋转的方法。

5.3 SO(3) 和 so(3) 的伴随表示

5.3.1 SO(3) 的伴随矩阵表示

SO(3) 的伴随表示采用矩阵乘法的形式:

$$\mathrm{Ad}_{\mathbf{R}} A = \mathbf{R} A \mathbf{R}^{-1}$$

其中 $R \in \mathrm{SO}(3)$,$A \in \mathrm{so}(3)$。其结果已经在前面的方程 (66) 中给出

$$\mathrm{Ad}_{\mathbf{R}} = \mathbf{R}.$$

下面从其微分式进行证明。

在李群中，经常需要将一个切向量从一个元素周围的切空间中变换到另一个元素的切空间中。**伴随**执行此变换。一般来说，李群的一个很好的性质是这个变换是线性的。对于李群的一个元素 X，伴随写为 Ad_X：

$$\begin{aligned} \boldsymbol{\omega} \in \mathrm{so}(3), \mathbf{R} \in \mathrm{SO}(3) \\ \mathbf{R} \cdot \exp(\boldsymbol{\omega}) = \exp\left(\mathrm{Ad}_{\mathbf{R}} \cdot \boldsymbol{\omega}\right) \cdot \mathbf{R} \end{aligned} \tag{85}$$

首先，方程 (85) 中的特征重新排列如下：

$$\exp\left(\mathrm{Ad}_{\mathbf{R}} \cdot \boldsymbol{\omega}\right) = \mathbf{R} \cdot \exp(\boldsymbol{\omega}) \cdot \mathbf{R}^{-1}$$

然后，在不失一般性的情况下，对于 $t \in \mathbb{R}$，设 $\boldsymbol{\omega} = t \cdot \boldsymbol{v}$，在 $t = 0$ 时求相对于 t 的微分：

$$\begin{aligned} \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \exp\left(\mathrm{Ad}_{\mathbf{R}} \cdot t \cdot \boldsymbol{v}\right) &= \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \left[\mathbf{R} \cdot \exp(t \cdot \boldsymbol{v}) \cdot \mathbf{R}^{-1}\right] \\ \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \left[\mathbf{I} + \left(\mathrm{Ad}_{\mathbf{R}} \cdot t \cdot \boldsymbol{v}\right)_\times + O\left(t^2\right)\right] &= \mathbf{R} \cdot \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \left[\mathbf{I} + (t \cdot \boldsymbol{v})_\times + O\left(t^2\right)\right] \cdot \mathbf{R}^{-1} \\ \left(\mathrm{Ad}_{\mathbf{R}} \cdot \boldsymbol{v}\right)_\times &= \mathbf{R} \cdot \boldsymbol{v}_\times \cdot \mathbf{R}^{-1} \\ &= (\mathbf{R}\boldsymbol{v})_\times \\ \Longrightarrow \mathrm{Ad}_{\mathbf{R}} &= \mathbf{R} \end{aligned}$$

对于 SO(3) 的情况，群元素的伴随变换特别简单：它与用来表示元素的旋转矩阵相同。通过一个群元素旋转一个切向量，将其从元素右侧的切空间“移动”到左侧的切空间。

5.3.2 so(3) 的叉积表达式

so(3) 的伴随表示为叉积:

$$\mathrm{ad}_A B = A \times B$$

其中 $A, B \in \mathrm{so}(3)$, 对应到 $\mathbb{R}^3$ 向量的叉积。这个关系式可以从叉积的定义推导出来。

下面我们将验证 so(3) 的伴随表示 ad_A 的作用等价于向量的叉积，也就是说，我们要验证 $[A, B] = AB - BA$ 对应于 $a \times b$ 的反对称矩阵。

首先，对于李代数 so(3) 中的元素 $A, B \in \mathrm{so}(3)$，我们有

$$A = \begin{pmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & -b_3 & b_2 \\ b_3 & 0 & -b_1 \\ -b_2 & b_1 & 0 \end{pmatrix}.$$

接下来，我们计算 AB：

$$AB = \begin{pmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{pmatrix} \begin{pmatrix} 0 & -b_3 & b_2 \\ b_3 & 0 & -b_1 \\ -b_2 & b_1 & 0 \end{pmatrix} = \begin{pmatrix} -a_3b_3 - a_2b_2 & a_2b_1 & a_3b_1 \\ a_1b_2 & -a_1b_1 - a_3b_3 & a_3b_2 \\ a_1b_3 & a_2b_3 & -a_1b_1 - a_2b_2 \end{pmatrix}.$$

然后，我们计算 BA：

$$BA = \begin{pmatrix} 0 & -b_3 & b_2 \\ b_3 & 0 & -b_1 \\ -b_2 & b_1 & 0 \end{pmatrix} \begin{pmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{pmatrix} = \begin{pmatrix} -b_3a_3 - b_2a_2 & b_2a_1 & b_3a_1 \\ b_1a_2 & -b_1a_1 - b_3a_3 & b_3a_2 \\ b_1a_3 & b_2a_3 & -b_1a_1 - b_2a_2 \end{pmatrix}.$$

于是我们得到 $AB - BA$：

$$\begin{aligned} AB - BA &= \begin{pmatrix} 0 & a_2b_1 - b_2a_1 & a_3b_1 - b_3a_1 \\ a_1b_2 - b_1a_2 & 0 & a_3b_2 - b_3a_2 \\ a_1b_3 - b_1a_3 & a_2b_3 - b_2a_3 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & -(a_1b_2 - a_2b_1) & a_3b_1 - a_1b_3 \\ a_1b_2 - a_2b_1 & 0 & -(a_2b_3 - a_3b_2) \\ -(a_3b_1 - a_1b_3) & a_2b_3 - a_3b_2 & 0 \end{pmatrix}. \end{aligned}$$

最后，我们可以验证这个矩阵就是向量 $a \times b = (a_2b_3 - a_3b_2, a_3b_1 - a_1b_3, a_1b_2 - a_2b_1)^T$ 的反对称矩阵。所以，我们验证了 so(3) 的伴随表示 ad_A 的作用等价于向量的叉积，即

$$\mathrm{ad}_A = \begin{pmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{pmatrix}.$$

5.3.3 与叉积相关矩阵

进一步的，我们计算 $AB^T - BA^T$ 这种形式的矩阵。

对于两个三维向量 $A, B \in \mathbb{R}^3$，其中，$A = (a_1, a_2, a_3)^T$, $B = (b_1, b_2, b_3)^T$，其叉乘为

$$A \times B = \begin{bmatrix} a_2b_3 - a_3b_2 \\ a_3b_1 - a_1b_3 \\ a_1b_2 - a_2b_1 \end{bmatrix}$$

我们计算 $AB^T - BA^T$，

$$\begin{aligned} AB^T - BA^T &= \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \begin{bmatrix} b_1 & b_2 & b_3 \end{bmatrix} - \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix} \\ &= \begin{bmatrix} a_1b_1 & a_1b_2 & a_1b_3 \\ a_2b_1 & a_2b_2 & a_2b_3 \\ a_3b_1 & a_3b_2 & a_3b_3 \end{bmatrix} - \begin{bmatrix} b_1a_1 & b_1a_2 & b_1a_3 \\ b_2a_1 & b_2a_2 & b_2a_3 \\ b_3a_1 & b_3a_2 & b_3a_3 \end{bmatrix} \\ &= \begin{bmatrix} 0 & a_1b_2 - a_2b_1 & a_1b_3 - a_3b_1 \\ a_2b_1 - a_1b_2 & 0 & a_2b_3 - a_3b_2 \\ a_3b_1 - a_1b_3 & a_3b_2 - a_2b_3 & 0 \end{bmatrix} \end{aligned}$$

通过观察，我们可以看到：$AB^T - BA^T = -(A \times B)^\wedge$，即这两个矩阵的每一个元素都是相同的，只是符号相反。现在让我们考虑 $AB^T - BA^T$ 的元素，可以看到，矩阵 $AB^T - BA^T$ 的非对角线元素对应于向量 A 和 B 的叉积的各个分量。具体来说，

$$\begin{aligned} (AB^T - BA^T)_{23} &= a_2b_3 - a_3b_2 = (A \times B)_1 \\ (AB^T - BA^T)_{31} &= a_3b_1 - a_1b_3 = (A \times B)_2 \\ (AB^T - BA^T)_{12} &= a_1b_2 - a_2b_1 = (A \times B)_3 \end{aligned}$$

因此，矩阵 $AB^T - BA^T$ 可以被视为向量 A 和 B 的叉积的一种矩阵表示。

$AB^T - BA^T$ 这种形式的矩阵在许多应用领域中都有重要的作用。特别地，它在物理学、计算机图形学以及机器人学中有着广泛的应用。

1. 物理学：在物理学中，特别是在经典力学和量子力学中，$AB^T - BA^T$ 通常被用来描述系统的动力学。特别地，如果 A 和 B 是两个物理量，那么 $[A, B] = AB^T - BA^T$ 就被称为这两个量的对易子。例如，角动量算符的对易关系就是这种形式的，它在描述粒子的旋转动力学时起着关键作用。

2. 计算机图形学：在计算机图形学中，$AB^T - BA^T$ 的形式常常用于描述二维或三维空间的旋转。特别地，如果 A 和 B 是空间中的两个向量，那么 $AB^T - BA^T$ 就对应于 A 和 B 的叉积，这是一种描述旋转的重要方式。另外，这种矩阵也常常用于实现各种图形变换，如缩放、剪切和反射等。

3. 机器人学：在机器人学中，$AB^T - BA^T$ 的形式常常用于描述机器人关节的旋转。特别地，如果 A 和 B 是描述机器人关节位置的向量，那么 $AB^T - BA^T$ 就可以描述关节的旋转动力学，这在实现机器人的运动控制中非常重要。

因此，$AB^T - BA^T$ 这种形式的矩阵在许多应用领域中都非常重要，它提供了一种强大的工具来描述和分析各种旋转和动力学系统。

5.3.4 关系式的验证

可以直接验证这两种伴随表示满足关系式:

$$\mathrm{Ad}_{\exp(A)} = \exp(\mathrm{ad}_A)$$

矩阵指数和叉积具有这一对应关系。

SO(3) 和 so(3) 的伴随表示都可用矩阵运算明确表达, 并满足一般的对应关系式。

5.4 指数映射的导数

5.4.1 罗德里格斯公式

将斜对称矩阵转化为旋转矩阵的指数映射，只是在生成元线性组合上的矩阵指数：

$$\begin{aligned}\exp\left(\boldsymbol{\omega}_{\times}\right) &\equiv \exp\begin{pmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{pmatrix} \\ &= \mathbf{I} + \boldsymbol{\omega}_{\times} + \frac{1}{2!}\boldsymbol{\omega}_{\times}^2 + \frac{1}{3!}\boldsymbol{\omega}_{\times}^3 + \cdots\end{aligned}$$

成对地写这些项，我们有：

$$\exp\left(\boldsymbol{\omega}_{\times}\right) = \mathbf{I} + \sum_{i=0}^{\infty}\left[\frac{\boldsymbol{\omega}_{\times}^{2i+1}}{(2i+1)!} + \frac{\boldsymbol{\omega}_{\times}^{2i+2}}{(2i+2)!}\right]$$

现在我们可以利用斜对称矩阵的一个特性：

$$\boldsymbol{\omega}_{\times}^3 = -\left(\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{\omega}\right)\cdot\boldsymbol{\omega}_{\times}$$

首先将这个特征扩展到一般情况：

$$\begin{aligned}\theta^2 &\equiv \boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{\omega} \\ \boldsymbol{\omega}_{\times}^{2i+1} &= (-1)^i\theta^{2i}\boldsymbol{\omega}_{\times} \\ \boldsymbol{\omega}_{\times}^{2i+2} &= (-1)^i\theta^{2i}\boldsymbol{\omega}_{\times}^2\end{aligned}$$

现在我们可以对指数映射级数进行因子化，并识别系数中的泰勒展开式：

$$\begin{aligned}\exp\left(\boldsymbol{\omega}_{\times}\right) &= \mathbf{I} + \left(\sum_{i=0}^{\infty}\frac{(-1)^i\theta^{2i}}{(2i+1)!}\right)\boldsymbol{\omega}_{\times} + \left(\sum_{i=0}^{\infty}\frac{(-1)^i\theta^{2i}}{(2i+2)!}\right)\boldsymbol{\omega}_{\times}^2 \\ &= \mathbf{I} + \left(1 - \frac{\theta^2}{3!} + \frac{\theta^4}{5!} + \cdots\right)\boldsymbol{\omega}_{\times} + \left(\frac{1}{2!} - \frac{\theta^2}{4!} + \frac{\theta^4}{6!} + \cdots\right)\boldsymbol{\omega}_{\times}^2 \\ &= \mathbf{I} + \left(\frac{\sin\theta}{\theta}\right)\boldsymbol{\omega}_{\times} + \left(\frac{1-\cos\theta}{\theta^2}\right)\boldsymbol{\omega}_{\times}^2 \qquad (86)\end{aligned}$$

方程 (86) 就是大家熟悉的罗德里格斯公式，其中 $\theta = \|\boldsymbol{\omega}\|$。该公式对应于前面给出的方程 (61)。指数映射产生一个由 $\boldsymbol{\omega}$ 给出旋转轴并旋转 θ 弧度的旋转。

指数映射可以用对数求逆，即从 SO(3) 到 so(3) 的映射：

$$
\begin{aligned}
\mathbf{R} &\in \mathrm{SO}(3) \\
\theta &= \arccos\left(\frac{\operatorname{tr}(\mathbf{R})-1}{2}\right) \\
\ln(\mathbf{R}) &= \frac{\theta}{2\sin\theta}\cdot\left(\mathbf{R}-\mathbf{R}^{\mathrm{T}}\right)
\end{aligned}
$$

然后向量 $\boldsymbol{\omega}$ 从 $\ln(\mathbf{R})$ 的非对角元素中提取。该公式对应于前面给出的方程 (62)。

对于方程 (86)，我们定义为局部坐标系到全局坐标系的旋转，根据 $\mathbf{R}^{-1}=\mathbf{R}^{\mathrm{T}}$，则全局坐标系到局部坐标系的旋转为

$$
\mathbf{R}^{\mathrm{T}}=\mathbf{I}-\left(\frac{\sin\theta}{\theta}\right)\boldsymbol{\omega}_{\times}+\left(\frac{1-\cos\theta}{\theta^{2}}\right)\boldsymbol{\omega}_{\times}^{2} \tag{87}
$$

5.4.2 局部坐标系的导数

代数 $\mathfrak{so}(3)$ 的元素为 3×3 斜对称矩阵，且伴随表示相同：

$$
\begin{aligned}
\boldsymbol{\omega} &\in \Re^{3} \\
\boldsymbol{\omega}_{\times} &= \begin{pmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{pmatrix} \in \mathfrak{so}(3) \\
\mathrm{ad}_{\boldsymbol{\omega}} &= \boldsymbol{\omega}_{\times} \\
\mathrm{ad}_{\boldsymbol{\omega}}^{3} &= -\|\boldsymbol{\omega}\|^{2}\cdot\mathrm{ad}_{\boldsymbol{\omega}}
\end{aligned}
$$

根据前面给出的指数映射 exp 的求导方程 (80)，我们可以求出 SO(3) 在局部坐标系的导数。由于 ad 的高阶次幂折回到低阶次幂，因此我们可以收集系列中的项：

$$
\begin{aligned}
D_{\exp}(\boldsymbol{\omega}) &= \mathbf{I}-\left(\sum_{i=0}^{\infty}\frac{(-1)^{i}\cdot\|\boldsymbol{\omega}\|^{2i}}{(2i+2)!}\right)\cdot\mathrm{ad}_{\boldsymbol{\omega}}+\left(\sum_{i=0}^{\infty}\frac{(-1)^{i}\cdot\|\boldsymbol{\omega}\|^{2i}}{(2i+3)!}\right)\cdot\mathrm{ad}_{\boldsymbol{\omega}}^{2} \\
&= \mathbf{I}-\left(\frac{1-\cos\|\boldsymbol{\omega}\|}{\|\boldsymbol{\omega}\|^{2}}\right)\cdot\boldsymbol{\omega}_{\times}+\left(\frac{1-\frac{\sin\|\boldsymbol{\omega}\|}{\|\boldsymbol{\omega}\|}}{\|\boldsymbol{\omega}\|^{2}}\right)\cdot\boldsymbol{\omega}_{\times}^{2}
\end{aligned} \tag{88}
$$

该方程就是前面的方程 (67) 给出的右 Jacobian 矩阵 $\mathbf{J}_r(\boldsymbol{\omega})$。注意，该矩阵同样是在文献 [1] 的方程 (174) 中给出 SE(3) 矩阵的 $\mathbf{V}$ 矩阵的转置矩阵 $\mathbf{V}^{\top}$。

根据方程 (82)，在 exp 和 log 的双射区域中，log 的导数是

$$
D_{\log}(\boldsymbol{\omega})=D_{\exp}^{-1}(\boldsymbol{\omega}) \tag{89}
$$

根据前面的方程 (68) 给出的右 Jacobian 矩阵的逆矩阵 $\mathbf{J}_r{}^{-1}(\boldsymbol{\omega})$，我们有

$$
\begin{aligned}
D_{\log}(\boldsymbol{\omega}) &= \mathbf{J}_r{}^{-1}(\boldsymbol{\omega}) \\
&= \mathbf{I}+\frac{1}{2}\boldsymbol{\omega}_{\times}+\left(\frac{1}{\theta^{2}}-\frac{1+\cos\theta}{2\theta\sin\theta}\right)\boldsymbol{\omega}_{\times}^{2}
\end{aligned}
$$

5.4.3 全局坐标系的导数

根据前面给出的指数映射 exp 的求导方程 (81)，我们可以求出 SO(3) 在全局坐标系的导数：

$$
\begin{aligned}
D_{\exp}(\boldsymbol{\omega}) &= \mathbf{I} + \left(\sum_{i=0}^{\infty} \frac{(-1)^i \cdot \|\boldsymbol{\omega}\|^{2i}}{(2i+2)!}\right) \cdot \mathrm{ad}_{\boldsymbol{\omega}} + \left(\sum_{i=0}^{\infty} \frac{(-1)^i \cdot \|\boldsymbol{\omega}\|^{2i}}{(2i+3)!}\right) \cdot \mathrm{ad}_{\boldsymbol{\omega}}^2 \\
&= \mathbf{I} + \left(\frac{1-\cos\|\boldsymbol{\omega}\|}{\|\boldsymbol{\omega}\|^2}\right) \cdot \boldsymbol{\omega}_\times + \left(\frac{1-\frac{\sin\|\boldsymbol{\omega}\|}{\|\boldsymbol{\omega}\|}}{\|\boldsymbol{\omega}\|^2}\right) \cdot \boldsymbol{\omega}_\times^2
\end{aligned} \tag{90}
$$

该方程就是前面的方程 (69) 给出的左 Jacobian 矩阵 $\mathbf{J}_l(\boldsymbol{\omega})$。注意，该矩阵同样是在文献 [1] 的方程 (174) 中给出 SE(3) 矩阵的 $\mathbf{V}$ 矩阵。并且根据前面的方程 (70) 给出的左 Jacobian 矩阵的逆矩阵 $\mathbf{J}_l{}^{-1}(\boldsymbol{\omega})$，我们有

$$
\begin{aligned}
D_{\log}(\boldsymbol{\omega}) &= \mathbf{J}_l{}^{-1}(\boldsymbol{\omega}) \\
&= \mathbf{I} - \frac{1}{2}\boldsymbol{\omega}_\times + \left(\frac{1}{\theta^2} - \frac{1+\cos\theta}{2\theta\sin\theta}\right)\boldsymbol{\omega}_\times^2
\end{aligned}
$$

因此，观察全局坐标系的导数和局部坐标系的导数之间的关系，由方程 (71) 给出为

$$\mathbf{J}_l = \mathbf{J}_r^\top, \quad \mathbf{J}_l^{-1} = \mathbf{J}_r^{-\top}.$$

5.4.4 姿态的导数与瞬时角速度

刚体的姿态 $\mathbf{R}$ 为李群流形 SO(3) 上移动的点，其变化速度 $\dot{\mathbf{R}} = \partial\mathbf{R}/\partial t$ 属于在当前点 $\mathbf{R}$ 处与 SO(3) 正切的空间。假设球面上的向量以角速度 $\boldsymbol{\omega}$ 匀速旋转，根据约束条件

$$\mathbf{R}\mathbf{R}^{\mathrm{T}} = \mathbf{I}$$

两边取导数

$$\frac{\mathrm{d}}{\mathrm{d}t}\left\{\mathbf{R}\mathbf{R}^{\mathrm{T}}\right\} = \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{I} \to \dot{\mathbf{R}}\mathbf{R}^{\mathrm{T}} + \mathbf{R}\dot{\mathbf{R}}^{\mathrm{T}} = 0$$

这就是一个斜对称矩阵

$$\dot{\mathbf{R}}\mathbf{R}^{\mathrm{T}} = -\left(\dot{\mathbf{R}}\mathbf{R}^{\mathrm{T}}\right)^{\mathrm{T}}$$

我们定义

$$[\boldsymbol{\omega}]_\times = \dot{\mathbf{R}}\mathbf{R}^{\mathrm{T}}$$

则旋转矩阵的导数 (旋转运动的基本方程) 为：

$$\dot{\mathbf{R}} = [\boldsymbol{\omega}]_\times \mathbf{R}$$

当 $\mathbf{R}(0) = \mathbf{I}$ 时的解为

$$\mathbf{R} = e^{[\boldsymbol{\omega}]_\times t}$$

其中 $[\boldsymbol{\omega}]_\times$ 表示在切空间中的角速度。

我们在笛卡尔空间测量得到的角速度 $\boldsymbol{\omega}$ 为向量

$$\boldsymbol{\omega} = \begin{bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{bmatrix}$$

关连到 $[\boldsymbol{\omega}]_\times$ 这个斜对称矩阵则为

$$[\boldsymbol{\omega}]_\times = \begin{bmatrix} 0 & -\omega_z & \omega_y \\ \omega_z & 0 & -\omega_x \\ -\omega_y & \omega_x & 0 \end{bmatrix}$$

旋转在切空间中的导数，也就是在切空间中的角速度，对于三维空间为

$$\dot{\mathbf{R}} = \begin{cases} [\boldsymbol{\omega}]_\times & \text{if } \mathbf{R} = \mathbf{I} \\ [\boldsymbol{\omega}]_\times \mathbf{R} & \text{exp map} \end{cases}$$

这是在全局 (global) 坐标系中的表示。对于局部 (local) 坐标系的表示为

$$\dot{\mathbf{R}} = \mathbf{R}\,[\boldsymbol{\omega}]_\times \in \mathrm{SO}(3)$$

因为角速度 $\boldsymbol{\omega}$ 涉及到在全局 (global) 坐标系或在局部 (local) 坐标系中测量的问题，所以我们需要对符号使用上下标进行更明晰地标识。对于从机体帧 (body frame) b 到导航帧 (navigation frame) n 的 3D 旋转 $\mathbf{R}_b^n$，我们有空间角速度 (spatial angular velocity) $\boldsymbol{\omega}_{nb}^n$，其在导航帧中测量，

$$[\boldsymbol{\omega}_{nb}^n]_\times \triangleq \dot{\mathbf{R}}_b^n \left(\mathbf{R}_b^n\right)^T = \dot{\mathbf{R}}_b^n \mathbf{R}_n^b$$

并且有机体的角速度 $\boldsymbol{\omega}_{nb}^b$，其在机体帧内测量：

$$\left[\boldsymbol{\omega}_{nb}^b\right]_\times \triangleq \left(\mathbf{R}_b^n\right)^T \dot{\mathbf{R}}_b^n = \mathbf{R}_n^b \dot{\mathbf{R}}_b^n$$

我们可以通过共轭将这些斜对称矩阵从导航帧变换到机体帧，

$$\left[\boldsymbol{\omega}_{nb}^b\right]_\times = \mathbf{R}_n^b \left[\boldsymbol{\omega}_{nb}^n\right]_\times \mathbf{R}_b^n$$

但因为伴随表示满足

$$\mathrm{Ad}_{\mathbf{R}}\,[\boldsymbol{\omega}]_\times \triangleq \mathbf{R}\,[\boldsymbol{\omega}]_\times \mathbf{R}^T = [\mathbf{R}\boldsymbol{\omega}]_\times$$

我们甚至可以更容易地将空间角速度和机体角速度变换为 3 维向量：

$$\boldsymbol{\omega}_{nb}^b = \mathbf{R}_n^b \boldsymbol{\omega}_{nb}^n$$

上述的表示有些繁琐，我们可以进行一些简化，并且有时候导航帧与空间帧 (space frame) 通用，所以我们有

$$[\boldsymbol{\omega}_s]_\times \triangleq \dot{\mathbf{R}}_b^n \left(\mathbf{R}_b^n\right)^{-1}$$
$$[\boldsymbol{\omega}_b]_\times \triangleq \left(\mathbf{R}_b^n\right)^{-1} \dot{\mathbf{R}}_b^n$$

并且通常情况下，旋转矩阵 $\mathbf{R}$ 在没有上下标指示的情况下指的是 $\mathbf{R}_b^n$，而角速度 $\boldsymbol{\omega}$ 在没有上下标指示的情况下指的是 $\boldsymbol{\omega}_s$。于是对于角速度在坐标系中的变换

$$\boldsymbol{\omega}_s = \mathbf{R}\boldsymbol{\omega}_b$$
$$\boldsymbol{\omega}_b = \mathbf{R}^{-1}\boldsymbol{\omega}_s$$

可以这么理解

$$\boldsymbol{\omega}_s = \mathrm{Ad}_{\mathbf{R}}\boldsymbol{\omega}_b$$
$$\boldsymbol{\omega}_b = \mathrm{Ad}_{\mathbf{R}^{-1}}\boldsymbol{\omega}_s$$

最终 SO(3) 和 SE(3) 的伴随矩阵的作用关系见下表：(screw-table2.pdf)

<table>
<tr><th colspan="2">Space frame</th><th colspan="3">Transformation</th><th colspan="2">Body frame</th></tr>
<tr><th>Vector
(Cartesian space)</th><th>Lie algebra
(Tangent space)</th><th>from $\{b\}$ to $\{s\}$</th><th>Adjoint matrix</th><th>from $\{s\}$ to $\{b\}$</th><th>Lie algebra
(Tangent space)</th><th>Vector
(Cartesian space)</th></tr>
<tr><td>Spatial angular velocity
$\boldsymbol{\omega}_s$</td><td>$\begin{aligned}\dot{\mathbf{R}}\mathbf{R}^{-1} &= [\boldsymbol{\omega}_s\times] \\ \dot{\mathbf{R}} &= [\boldsymbol{\omega}_s\times]\,\mathbf{R} \\ \mathbf{R} &= \exp\left([\boldsymbol{\omega}_s\times]\cdot\Delta t\right)\end{aligned}$</td><td>$\begin{aligned}[\boldsymbol{\omega}_s\times] &= \mathbf{R}\,[\boldsymbol{\omega}_b\times]\,\mathbf{R}^{-1} \\ &= [(\mathbf{R}\boldsymbol{\omega}_b)\times] \\ \boldsymbol{\omega}_s &= \mathbf{R}\boldsymbol{\omega}_b\end{aligned}$</td><td>$\begin{aligned}[\mathrm{Ad}_{\mathbf{R}}] &= \mathbf{R} \\ \boldsymbol{\omega}_s &= \mathbf{R}\boldsymbol{\omega}_b \\ [\mathrm{Ad}_{\mathbf{R}^{-1}}] &= \mathbf{R}^{-1} \\ \boldsymbol{\omega}_b &= \mathbf{R}^{-1}\boldsymbol{\omega}_s\end{aligned}$</td><td>$\begin{aligned}[\boldsymbol{\omega}_b\times] &= \mathbf{R}^{-1}\,[\boldsymbol{\omega}_s\times]\,\mathbf{R} \\ &= \left[\left(\mathbf{R}^{-1}\boldsymbol{\omega}_s\right)\times\right] \\ \boldsymbol{\omega}_b &= \mathbf{R}^{-1}\boldsymbol{\omega}_s\end{aligned}$</td><td>$\begin{aligned}\mathbf{R}^{-1}\dot{\mathbf{R}} &= [\boldsymbol{\omega}_b\times] \\ \dot{\mathbf{R}} &= \mathbf{R}\,[\boldsymbol{\omega}_b\times] \\ \mathbf{R}^{-1} &= \exp\left([\boldsymbol{\omega}_b\times]\cdot\Delta t\right)\end{aligned}$</td><td>Body angular velocity
$\boldsymbol{\omega}_b$</td></tr>
<tr><td>Spatial velocity
$\boldsymbol{\mathcal{V}}_s = \begin{bmatrix}\boldsymbol{\omega}_s \\ \mathbf{v}_s\end{bmatrix}$</td><td>$\begin{aligned}\dot{\mathbf{T}}\mathbf{T}^{-1} &= \left[\boldsymbol{\mathcal{V}}_s^{\wedge}\right] \\ &= \left[\begin{array}{c|c}[\boldsymbol{\omega}_s\times] & \mathbf{v}_s \\ \hline \mathbf{0}^{\mathrm{T}} & 0\end{array}\right] \\ \dot{\mathbf{T}} &= \left[\boldsymbol{\mathcal{V}}_s^{\wedge}\right]\mathbf{T} \\ \mathbf{T} &= \exp\left(\left[\boldsymbol{\mathcal{V}}_s^{\wedge}\right]\cdot\Delta t\right)\end{aligned}$</td><td>$\begin{aligned}\left[\boldsymbol{\mathcal{V}}_s^{\wedge}\right] &= \mathbf{T}\left[\boldsymbol{\mathcal{V}}_b^{\wedge}\right]\mathbf{T}^{-1} \\ \boldsymbol{\mathcal{V}}_s &= [\mathrm{Ad}_{\mathbf{T}}]\,\boldsymbol{\mathcal{V}}_b\end{aligned}$</td><td>$\begin{aligned}[\mathrm{Ad}_{\mathbf{T}}] &= \begin{bmatrix}\mathbf{R} & \mathbf{0} \\ [\mathbf{r}\times]\,\mathbf{R} & \mathbf{R}\end{bmatrix} \\ \boldsymbol{\mathcal{V}}_s &= [\mathrm{Ad}_{\mathbf{T}}]\,\boldsymbol{\mathcal{V}}_b \\ [\mathrm{Ad}_{\mathbf{T}^{-1}}] &= \begin{bmatrix}\mathbf{R}^{-1} & \mathbf{0} \\ -\mathbf{R}^{-1}[\mathbf{r}\times] & \mathbf{R}^{-1}\end{bmatrix} \\ \boldsymbol{\mathcal{V}}_b &= [\mathrm{Ad}_{\mathbf{T}^{-1}}]\,\boldsymbol{\mathcal{V}}_s\end{aligned}$</td><td>$\begin{aligned}\left[\boldsymbol{\mathcal{V}}_b^{\wedge}\right] &= \mathbf{T}^{-1}\left[\boldsymbol{\mathcal{V}}_s^{\wedge}\right]\mathbf{T} \\ \boldsymbol{\mathcal{V}}_b &= [\mathrm{Ad}_{\mathbf{T}^{-1}}]\,\boldsymbol{\mathcal{V}}_s\end{aligned}$</td><td>$\begin{aligned}\mathbf{T}^{-1}\dot{\mathbf{T}} &= \left[\boldsymbol{\mathcal{V}}_b^{\wedge}\right] \\ &= \left[\begin{array}{c|c}[\boldsymbol{\omega}_b\times] & \mathbf{v}_b \\ \hline \mathbf{0}^{\mathrm{T}} & 0\end{array}\right] \\ \dot{\mathbf{T}} &= \mathbf{T}\left[\boldsymbol{\mathcal{V}}_b^{\wedge}\right] \\ \mathbf{T}^{-1} &= \exp\left(\left[\boldsymbol{\mathcal{V}}_b^{\wedge}\right]\cdot\Delta t\right)\end{aligned}$</td><td>Body velocity
$\boldsymbol{\mathcal{V}}_b = \begin{bmatrix}\boldsymbol{\omega}_b \\ \mathbf{v}_b\end{bmatrix}$</td></tr>
</table>

由上表中公式得出，空间角速度 $\boldsymbol{\omega}_s$ 不依赖于机体坐标系的选择。同样，机体角速度 $\boldsymbol{\omega}_b$ 也不依赖参考坐标系的选择。虽然 $\mathbf{R}$ 和 $\dot{\mathbf{R}}$ 各自都依赖于坐标系 $\{s\}$ 和 $\{b\}$，但是 $\dot{\mathbf{R}}\mathbf{R}^{-1}$ 与 $\{b\}$ 无关，$\mathbf{R}^{-1}\dot{\mathbf{R}}$ 与 $\{s\}$ 无关。

在上表中的微分算子 $[\boldsymbol{\omega}]_\times = \dot{\mathbf{R}}\mathbf{R}^{-1}$ 可计算得出生成元，即

$$G_i = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, G_j = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, G_k = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

则做为 so(3) 的一个元素的角速度 $\boldsymbol{\omega}$ 表示为生成元的线性组合：

$$\omega_x G_i + \omega_y G_j + \omega_z G_k \in \mathrm{so}(3)$$

由此看出，在空间坐标系中的微分算子和所选择的机体坐标系无关。并且生成元 G_i, G_j, G_k 与四元数的 3 个虚数 i, j, k 在切空间中的矩阵表示关系密切，

$$i \to \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad j \to \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}, \quad k \to \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

因为上述性质，这个斜对称矩阵 $[\boldsymbol{\omega}]_\times$ 又被称为速度张量 (velocity tensor)，而旋转矩阵 $\mathbf{R}$ 又被称为正交固有张量 (orthogonal proper tensor)。因为一个刚体只有一个角速度 $\boldsymbol{\omega}$，因此如何在刚体运动中找到与其上任意点 p 无关的速度不变量 (velocity invariant)，或称不变性向量 (invariant vector)，则是旋量理论 (The theory of screws) 或空间向量代数 (Spatial Vector Algebra) 研究的初始问题。因为斜对称矩阵和旋转矩阵有许多简单易解的性质，因此有些问题可以转换成对偶向量用罗德里格斯公式进行求解，这就需要对偶李代数的知识，比如对刚体位姿 SE(3) 的计算，或者是研究高阶运动学的高阶加速度向量场的问题。

5.5 小结

5.5.1 几何理解伴随表示

从几何的角度看，$SO(3)$ 的伴随表示描述的是一个给定的旋转如何影响其他旋转。在 $SO(3)$ 中，如果我们有两个旋转 R 和 g，其中 g 可以是以旋转表示的机体姿态，则伴随表示 $\mathrm{Ad}(R)(g) = RgR^{-1}$ 描述的就是先按照 R 旋转，然后进行 g 的旋转，最后再按照 R^{-1} 旋转回来，得到的新的旋转，或机体新的姿态。

如果对 $so(3)$ 的一个元素 Y 应用李群的伴随表示，即 $\mathrm{Ad}_R Y$，则其对应的伴随表示 $\mathrm{Ad}_R : so(3) \to so(3)$ 描述了在 R 的作用下，so(3) 中的元素如何变换。具体来说，伴随表示描述了在给

定旋转 R 的影响下，刚体的角速度向量 (在 $so(3)$ 中表示) 的变化。也就是说，如果一个刚体首先经历一个角速度为 ω 的旋转，然后再经历一个旋转 R，则刚体的最终角速度向量将会被伴随表示 Ad_R 变换，变为 $\mathrm{Ad}_R(\omega)$。

在 $so(3)$ 中，如果我们有两个无穷小旋转 X 和 Y，则李代数的伴随表示 $\mathrm{ad}(X)(Y) = [X,Y]$ 描述的就是 X 对 Y 的扰动效应。

5.5.2 几何理解指数映射和对数映射的导数

从几何的角度看，指数映射的导数描述的是无穷小旋转的“速度”，而对数映射的导数描述的是有限旋转的瞬时角速度。

对于 $so(3)$ 中的元素 X，指数映射的导数 $\frac{d}{dt}\big|_{t=0} \exp(tX)$ 是 X 自身，表明 X 就是对应的无穷小旋转的“速度”。这反映了无穷小旋转的线性性：无穷小旋转的“速度”与其大小成正比。

对于 $SO(3)$ 中的元素 R，对数映射的导数 $\frac{d}{dt}\big|_{t=0} \log(Re^{tX})$ 是 X 自身，表明 X 就是对应的有限旋转的瞬时角速度。这反映了有限旋转的非线性性：有限旋转的瞬时角速度不仅取决于旋转的大小，还取决于旋转的方向和旋转轴的位置。

指数映射的导数在几何上可以理解为描述无穷小旋转的速度。具体来说，对于给定的 $R \in SO(3)$，其对应的指数映射的导数 $D\exp_R : so(3) \to T_R SO(3)$ 描述了在 R 的作用下，$so(3)$ 中的元素如何变换。

具体来说，从几何的角度来看，指数映射的导数描述了在给定旋转 R 的影响下，刚体的角速度向量 (在 $so(3)$ 中表示) 的变化。也就是说，如果一个刚体首先经历一个角速度为 ω 的旋转，然后再经历一个旋转 R，则刚体的最终角速度向量将会被指数映射的导数 $D\exp_R$ 变换，其值为 $D\exp_R(\omega)$。

相似的，对数映射的导数在几何上可以理解为描述有限旋转的瞬时角速度。具体来说，对于给定的 $R \in SO(3)$，其对应的对数映射的导数 $D\log_R : T_R SO(3) \to so(3)$ 描述了在 R 的作用下，$SO(3)$ 中的元素如何变换。

同样的，从几何的角度来看，对数映射的导数描述了在给定旋转 R 的影响下，刚体的旋转矩阵 (在 $SO(3)$ 中表示) 的变化。也就是说，如果一个刚体首先经历一个旋转 R，然后再经历一个角速度为 ω 的旋转，则刚体的最终旋转矩阵将会被对数映射的导数 $D\log_R$ 变换，其值为 $D\log_R(\dot{R})$。

6 总结

本章将从李群的对称表示性质进行总结。李群的对称表示是描述李群元素如何作用在某一几何空间上，使得空间的几何性质保持不变。例如，旋转群 $SO(3)$ 的元素可以看作是作用在三维欧几里得空间 $\mathbb{R}^3$ 上的旋转，使得空间的欧几里得距离保持不变。

在本文涉及的内容中，利用的李群对称表示性质，主要有以下几个方面:

1. 李群的伴随表示 (Ad)：伴随表示是李群的一个重要表示，它的定义依赖于李群的对称性。对于李群 G 中的元素 g，它的伴随表示 Ad_g 是一个从李代数 $\mathfrak{g}$ 到自身的线性映射，定义为：

$$\mathrm{Ad}_g = \mathrm{d}(\mathrm{Int}_g)$$

   其中，Int_g 是 g 的内部自同构，它是一个从 G 到 G 的映射，定义为 $\mathrm{Int}_g(h) = ghg^{-1}$。内部自同构依赖于李群的对称性，它可以被看作是在 G 上的一个“转动”。

2. 李代数的伴随表示 (ad)：该伴随表示是李代数的一个表示，它可以从李群的伴随表示推导得出。对于李代数 $\mathfrak{g}$ 中的元素 X，它的李代数伴随表示 ad_X 是一个从 $\mathfrak{g}$ 到自身的线性映射，定义为：
$$\mathrm{ad}_X(Y) = [X, Y]$$
其中，$[X, Y]$ 是李括号，它是李代数的一个基本运算。李代数的伴随表示依赖于李代数的对称性，它可以被看作是在 $\mathfrak{g}$ 上的一个"转动"。并且李代数的伴随表示 (ad) 可以由李群的伴随表示 (Ad) 导出，其定义如下：
$$\mathrm{ad}_X = \frac{d}{dt}\,\mathrm{Ad}_{\exp(tX)}\Big|_{t=0}$$
其中，$\exp(tX)$ 是李群的指数映射，$\frac{d}{dt}$ 是对参数 t 的导数。这个公式体现了李群和李代数之间的联系，即指数映射和对数映射。

3. 伴随表示之间的关系式：李群和李代数的伴随表示之间通过指数映射有关系式 $\mathrm{Ad}_{\exp X} = \exp(\mathrm{ad}_X)$。这个关系依赖于指数映射所反映的李群对称表示的性质。它保证了局部线性结构和全局流形结构之间的表示一致性。

4. 李群上的指数映射：指数映射 $\exp : \mathfrak{g} \to G$ 将李代数 $\mathfrak{g}$ 映射到李群 G。它使用了李群在其李代数上的对称表示的性质, 因为指数映射是李群的对称表示 (在单位元附近的线性逼近) 对应的指数函数。从几何视角, 指数映射基于切空间到流形的指数收敛, 实现了局部线性结构到全局流形结构的映射。对数映射 $\log : G \to \mathfrak{g}$ 定义为指数映射的反函数，将李群映射到李代数。这同样也是一种对称表示。

从几何的角度来看，伴随表示可以被理解为李群的切向空间上的一个线性变换，它描述了如何通过李群的元素"扭曲"或"旋转"切向空间。这种"扭曲"或"旋转"的效果可以通过左乘和右乘映射的相互作用来实现，这正是李群的对称性质的体现。对于李代数的伴随表示，可以把它理解为在"无穷小"的层面上描述这种"扭曲"或"旋转"的效果。

更进一步的，它们之间还有一个公式 $\mathrm{Ad}_{\exp X} = (T_e \exp) \circ (\mathrm{ad}_X) \circ (T_e \log)$，它描述了李群的伴随表示和李代数的伴随表示之间的关系，以及它们与李群的指数映射和对数映射之间的关系。这个公式的几何直觉可以理解为：通过李群的"转动"(李群的伴随表示) 和李代数的"转动"(李代数的伴随表示)，我们可以从李代数"移动"到李群，然后再从李群"移动"回李代数。

7 参考资料

1. A micro Lie theory for state estimation in robotics

2. Derivative of the Exponential Map

3. The Derivation of the Exponential Map of Matrices

4. Derivative of the exponential map

5. Derivatives and Differentials