

李群中的不确定表示与伴随矩阵

Shuyong Chen

2022 年 5 月 25 日

一题记：不确定是每个人考试时的魔咒。同桌的你，已经猜不出老师的问题。

1 简介

利用 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，其中伴随矩阵扮演着重要的角色。

2 前置知识

指数映射有如下重要性质

$$\exp((t+s)\tau^\wedge) = \exp(t\tau^\wedge) \exp(s\tau^\wedge) \quad (1)$$

$$\exp(t\tau^\wedge) = \exp(\tau^\wedge)^t \quad (2)$$

$$\exp(-\tau^\wedge) = \exp(\tau^\wedge)^{-1} \quad (3)$$

$$\exp(\mathcal{X}\tau^\wedge\mathcal{X}^{-1}) = \mathcal{X} \exp(\tau^\wedge) \mathcal{X}^{-1} \quad (4)$$

其中方程 (4) 很有用，在推导伴随矩阵和 Jacobian 矩阵时经常用到。

为方便指数映射的操作，我们需要定义大写指数映射。大写的 Exp 和 Log 映射是将向量元素 $\tau \in \mathbb{R}^m (\cong T_{\mathcal{E}}\mathcal{M})$ 直接映射到元素 $\mathcal{X} \in \mathcal{M}$ 的快捷方式。我们有，

$$\text{Exp} : \mathbb{R}^m \rightarrow \mathcal{M} ; \quad \tau \mapsto \mathcal{X} = \text{Exp}(\tau) \quad (5)$$

$$\text{Log} : \mathcal{M} \rightarrow \mathbb{R}^m ; \quad \mathcal{X} \mapsto \tau = \text{Log}(\mathcal{X}). \quad (6)$$

显然有，

$$\mathcal{X} = \text{Exp}(\tau) \triangleq \exp(\tau^\wedge) \quad (7)$$

$$\tau = \text{Log}(\mathcal{X}) \triangleq \log(\mathcal{X})^\vee. \quad (8)$$

对于不同流形的这些映射的实现的信息，参见参考文献。

为表达直觉，我们需要定义新的加号和减号算子。加号和减号算子允许我们在 (弯曲的) 流形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 (right-) 版本和左结合 (left-) 版本中定义。右结合 (right-) 算子是

$$\text{right-}\oplus: \mathcal{Y} = \mathcal{X} \oplus {}^{\mathcal{X}}\tau \triangleq \mathcal{X} \circ \text{Exp}({}^{\mathcal{X}}\tau) \in \mathcal{M} \quad (9)$$

$$\text{right-}\ominus: {}^{\mathcal{X}}\tau = \mathcal{Y} \ominus \mathcal{X} \triangleq \text{Log}(\mathcal{X}^{-1} \circ \mathcal{Y}) \in T_{\mathcal{X}}\mathcal{M}. \quad (10)$$

因为在方程 (9) 中 $\text{Exp}({}^{\mathcal{X}}\tau)$ 出现在组合的右侧， ${}^{\mathcal{X}}\tau$ 属于 \mathcal{X} 处的切空间，因此我们按照约定说这个 ${}^{\mathcal{X}}\tau$ 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 (left-) 算子是

$$\text{left-}\oplus: \mathcal{Y} = {}^{\varepsilon}\tau \oplus \mathcal{X} \triangleq \text{Exp}({}^{\varepsilon}\tau) \circ \mathcal{X} \in \mathcal{M} \quad (11)$$

$$\text{left-}\ominus: {}^{\varepsilon}\tau = \mathcal{Y} \ominus \mathcal{X} \triangleq \text{Log}(\mathcal{Y} \circ \mathcal{X}^{-1}) \in T_{\varepsilon}\mathcal{M}. \quad (12)$$

现在，在方程 (11) 中 $\text{Exp}({}^{\varepsilon}\tau)$ 在左侧，并且我们有 ${}^{\varepsilon}\tau \in T_{\varepsilon}\mathcal{M}$ ，因此我们说这个 ${}^{\varepsilon}\tau$ 是在全局 (global) 坐标系中表示。

注意，虽然右结合 (right-) 和左结合 (left-) 算子 \oplus 是按操作数顺序区分的，但符号 \ominus 在方程 (10) 和 (12) 中是不明确的。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 (right-) 形式。

3 伴随和伴随矩阵

如果我们在方程 (9, 11) 中标识 \mathcal{Y} ，我们就得到 ${}^{\varepsilon}\tau \oplus \mathcal{X} = \mathcal{X} \oplus {}^{\mathcal{X}}\tau$ ，它确定局部切元素和全局切元素之间的关系。我们用方程 (4, 9, 11) 来扩展它为

$$\begin{aligned} \text{Exp}({}^{\varepsilon}\tau) \mathcal{X} &= \mathcal{X} \text{Exp}({}^{\mathcal{X}}\tau) \\ \exp({}^{\varepsilon}\tau^{\wedge}) &= \mathcal{X} \exp({}^{\mathcal{X}}\tau^{\wedge}) \mathcal{X}^{-1} = \exp(\mathcal{X} {}^{\mathcal{X}}\tau^{\wedge} \mathcal{X}^{-1}) \\ {}^{\varepsilon}\tau^{\wedge} &= \mathcal{X} {}^{\mathcal{X}}\tau^{\wedge} \mathcal{X}^{-1} \end{aligned}$$

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

$$\text{Ad}_{\mathcal{X}}: \mathfrak{m} \rightarrow \mathfrak{m}; \quad \tau^{\wedge} \mapsto \text{Ad}_{\mathcal{X}}(\tau^{\wedge}) \triangleq \mathcal{X} \tau^{\wedge} \mathcal{X}^{-1}, \quad (13)$$

因此 ${}^{\varepsilon}\tau^{\wedge} = \text{Ad}_{\mathcal{X}}({}^{\mathcal{X}}\tau^{\wedge})$ 。这定义了群在它自己的李代数上的伴随作用 (adjoint action)。

伴随矩阵: 因为 $\text{Ad}_{\mathcal{X}}()$ 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 ${}^{\varepsilon}\tau \cong {}^{\varepsilon}\tau^{\wedge}$ 和 ${}^{\mathcal{X}}\tau \cong {}^{\mathcal{X}}\tau^{\wedge}$ ，

$$\mathbf{Ad}_{\mathcal{X}}: \mathbb{R}^m \rightarrow \mathbb{R}^m; \quad {}^{\mathcal{X}}\tau \mapsto {}^{\varepsilon}\tau = \mathbf{Ad}_{\mathcal{X}} {}^{\mathcal{X}}\tau, \quad (14)$$

我们称之为伴随矩阵 (adjoint matrix)。这可以通过将 $^{\vee}$ 应用于方程 (13) 来计算，因此写为

$$\mathbf{Ad}_{\mathcal{X}}\tau = ({}^{\mathcal{X}}\tau^{\wedge} \mathcal{X}^{-1})^{\vee}, \quad (15)$$

然后扩展右侧结合以标识伴随矩阵。伴随矩阵有一些常用的性质

$$\mathcal{X} \oplus \tau = (\mathbf{Ad}_{\mathcal{X}}\tau) \oplus \mathcal{X} \quad (16)$$

$$\mathbf{Ad}_{\mathcal{X}^{-1}} = \mathbf{Ad}_{\mathcal{X}}^{-1} \quad (17)$$

$$\mathbf{Ad}_{\mathcal{X}\mathcal{Y}} = \mathbf{Ad}_{\mathcal{X}}\mathbf{Ad}_{\mathcal{Y}}. \quad (18)$$

请注意在方程 (17, 18) 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, 使用 ${}^{\varepsilon}\boldsymbol{\tau} = \mathbf{Ad}_{\mathcal{X}} \boldsymbol{\tau}$, 方程 (14)。

4 流形中的不确定性与协方差传播

我们定义局部扰动 $\boldsymbol{\tau}$ 为在切向量空间 $T_{\bar{\mathcal{X}}}\mathcal{M}$ 中围绕着点 $\bar{\mathcal{X}} \in \mathcal{M}$ 的扰动, 使用右结合 (right-) 的 \oplus 和 \ominus ,

$$\mathcal{X} = \bar{\mathcal{X}} \oplus \boldsymbol{\tau}, \quad \boldsymbol{\tau} = \mathcal{X} \ominus \bar{\mathcal{X}} \in T_{\bar{\mathcal{X}}}\mathcal{M}. \quad (19)$$

协方差矩阵可以通过标准期望算子 $\mathbb{E}[\cdot]$ 在 $\bar{\mathcal{X}}$ 处的切空间上正确定义,

$$\boldsymbol{\Sigma}_{\mathcal{X}} \triangleq \mathbb{E}[\boldsymbol{\tau}\boldsymbol{\tau}^{\top}] = \mathbb{E}\left[(\mathcal{X} \ominus \bar{\mathcal{X}})(\mathcal{X} \ominus \bar{\mathcal{X}})^{\top}\right] \in \mathbb{R}^{m \times m}, \quad (20)$$

这允许我们定义流形上的高斯变量, $\mathcal{X} \sim \mathcal{N}(\bar{\mathcal{X}}, \boldsymbol{\Sigma}_{\mathcal{X}})$ 。注意, 虽然我们写 $\boldsymbol{\Sigma}_{\mathcal{X}}$, 但协方差还是正切扰动 $\boldsymbol{\tau}$ 的协方差。由于 $T\mathcal{M}$ 的维度 m 与 \mathcal{M} 的自由度相匹配, 因此这些协方差被很好地定义。

扰动也可以在全局参考中表示, 即在原点 $T_{\varepsilon}\mathcal{M}$ 处的切空间中, 使用左结合 (left-) 的 \oplus 和 \ominus ,

$$\mathcal{X} = \boldsymbol{\tau} \oplus \bar{\mathcal{X}}, \quad \boldsymbol{\tau} = \mathcal{X} \ominus \bar{\mathcal{X}} \in T_{\varepsilon}\mathcal{M}. \quad (21)$$

这允许使用在方程 (20) 中的左结合 (left-) 减号的协方差矩阵的全局规范。因为“水平 (horizontal)”是一个全局规范。例如, 一个三维方向已知是在水平面中的旋转, 可以与协方差矩阵 ${}^{\varepsilon}\boldsymbol{\Sigma} = \text{diag}(\sigma_{\phi}^2, \sigma_{\theta}^2, \infty)$ 相关联。因此必须在全局参考中指定 ${}^{\varepsilon}\boldsymbol{\Sigma}$ 。

因为全局扰动和局部扰动由伴随方程联系起来, 它们的协方差的变换可以用方程

$${}^{\varepsilon}\boldsymbol{\Sigma}_{\mathcal{X}} = \mathbf{Ad}_{\mathcal{X}} \boldsymbol{\Sigma}_{\mathcal{X}} \mathbf{Ad}_{\mathcal{X}}^{\top}. \quad (22)$$

协方差的传播通过函数 $f: \mathcal{M} \rightarrow \mathcal{N}; \mathcal{X} \mapsto \mathcal{Y} = f(\mathcal{X})$ 只需要用 Jacobian 矩阵方程来线性化点 $\mathcal{Y} = f(\mathcal{X})$ 附近 $\boldsymbol{\tau}$ 小值的近似方程 $f(\mathcal{X} \oplus \boldsymbol{\tau}) \xrightarrow{\boldsymbol{\tau} \rightarrow 0} f(\mathcal{X}) \oplus \frac{Df(\mathcal{X})}{D\mathcal{X}} \boldsymbol{\tau}$ 以获得熟悉的公式,

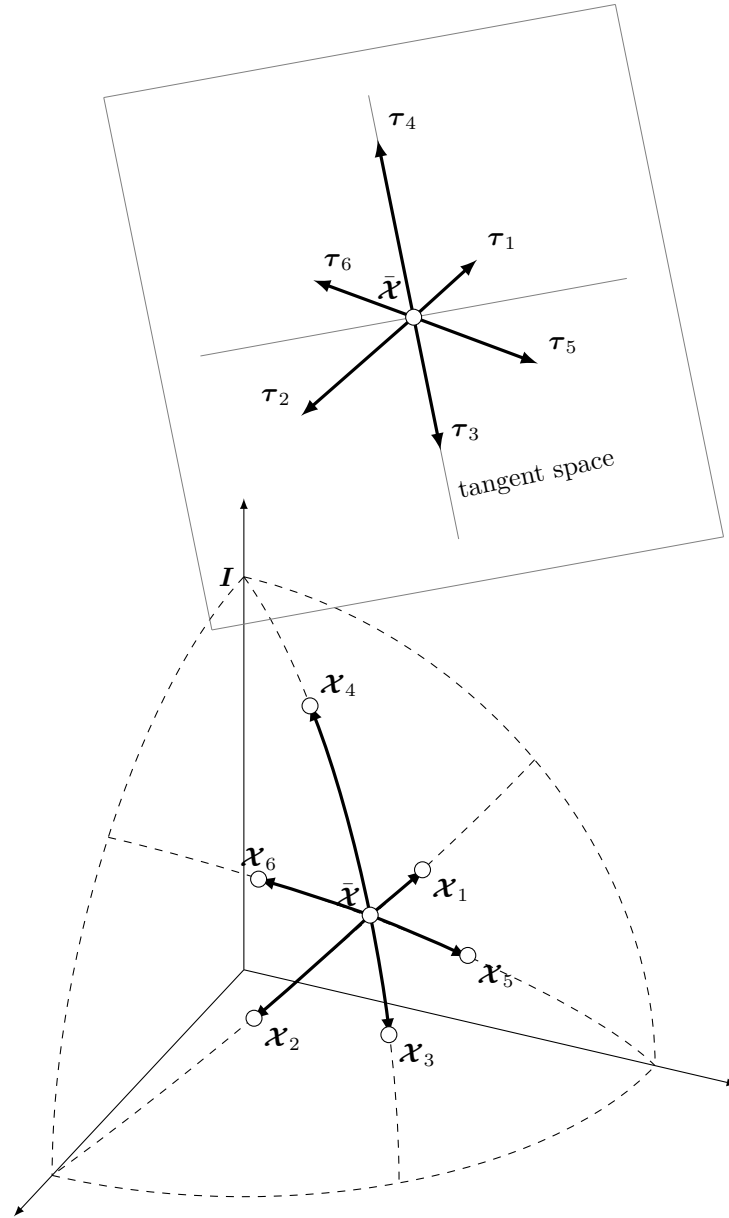
$$\boldsymbol{\Sigma}_{\mathcal{Y}} \approx \frac{Df}{D\mathcal{X}} \boldsymbol{\Sigma}_{\mathcal{X}} \frac{Df^{\top}}{D\mathcal{X}} \in \mathbb{R}^{n \times n}. \quad (23)$$

5 全局上的流形均值

在上一节中出现的 $\bar{\mathcal{X}}$, 是流形上的均值元素。那么我们该如何理解与获得这个流形上的均值元素 $\bar{\mathcal{X}}$? 和向量空间类似, 只是这里我们使用 \ominus 算子。假设在某个时刻我们测量得到 n 个流形元素 \mathcal{X}_i , 则均值元素 $\bar{\mathcal{X}}$ 有特性

$$\arg \min \sum_{i=1}^n \|(\mathcal{X}_i \ominus \bar{\mathcal{X}})\|^2 = \arg \min \sum_{i=1}^n (\boldsymbol{\tau}_i^{\top} \boldsymbol{\tau}_i)$$

我们可以用 $\text{SO}(3)$ 的均值从几何上直观理解这个问题。找来一个篮球, 用笔随便找一个地方均匀点几个点, 这就是我们抽样得到的随机旋转。然后我们估摸着在这几个点的中心区域点一个点, 这就是我们猜测的平均旋转。我们手上只有直尺, 量不了球面上两点的距离。但我们手上还有一些柔软的琴弦 (string), 于是我们用琴弦从猜测点出发, 到其中一个随机点的时候用手按住标记, 然后拉直了琴弦在直尺上看数据, 这就是从猜测点到随机点的距离。把所有的距离相加就是总距离。



刚才那个猜测点是随便猜的。我们按照某种梯度下降方法继续猜下一个点，直到找到总距离最小的那个点。在这期间，我们用琴弦量弧段，就是量测地线。琴弦拉直了用直尺量，就是 $\text{Log}()$ 操作，或者说展开 (unwrap) 操作，将球面 (流形) 上的元素映射到了向量空间。在向量空间中，距离相加就得到了总距离。因为局部切空间不同，各个点映射到切空间中的向量也不同，所以总距离就有变化。而平均旋转，就是总距离最小的那个点，但我们没有解析解，因为各处局部切空间无法直接相互比较。所以要求全局最小总距离，就只能靠某种方法试探和逼近。

随着切点的不同，在以切点为原点建立的切空间中，各个映射点围成的多面体的体积/面积也不相同。我们需要在流形上试探找到一个点，在该点上建立的切空间中，切点/原点正好是多面体的几何中心，也就是平均向量最小，

$$\left\| \sum_{i=1}^n (x_i \ominus \bar{x}) \right\|^2 = 0$$

这样从原点出发到各个映射点的总距离最小，这时各个映射点围成的多面体的体积/面积也最小。这

实际上也是施瓦兹不等式的一个应用。但是该点没有解析解，只能靠搜索找到。

因此在现有估计算法中，很少使用这种算法，一是速度慢，二是我们不可能有完全的数据集合，时变的传感器数据是在每个采样时刻陆续到达。因此我们实际上是采用在局部切空间中的递归算法求解均值元素。

6 递归求解均值元素

我们再次用 $SO(3)$ 的均值从几何上直观理解递归求解均值的算法。

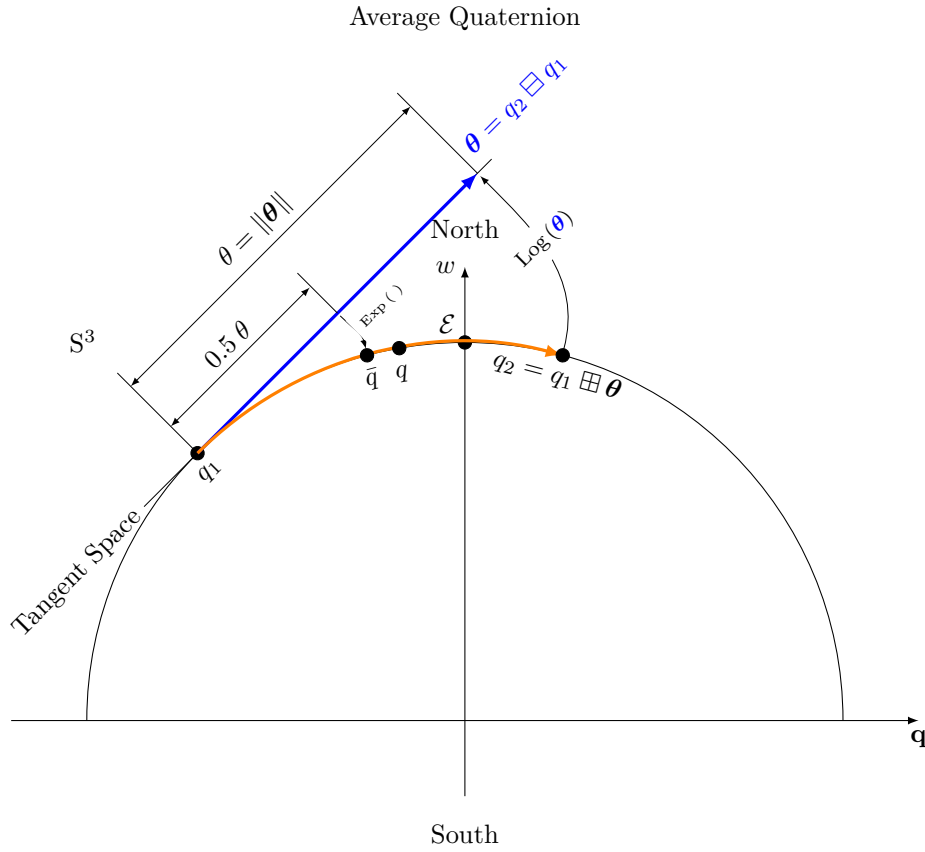
在某个时刻我们通过传感器测量得到两个姿态，用单位四元数表示为 q_1 和 q_2 ，两者关系为 $q_2 = q_1 \otimes r$ ，其中 r 为两者之间的差值四元数，所以有 $r = q_1^{-1} \otimes q_2$ 。我们需要把差值四元数 r 向量化，于是我们选择在 q_1 的切空间中计算两者的差值向量 $\theta = q_2 \boxminus q_1$ 。其计算方程为

$$\begin{aligned} r &= \begin{bmatrix} w_r \\ \mathbf{q}_r \end{bmatrix} \\ &= \begin{bmatrix} \cos(\|\theta\|/2) \\ \frac{\theta}{\|\theta\|} \sin(\|\theta\|/2) \end{bmatrix}. \\ \|\theta\| &= |2 \arccos(w_r)|. \\ \theta &= \mathbf{q}_r \frac{\|\theta\|}{\sin(\|\theta\|/2)}. \end{aligned}$$

差值向量 θ 实际上是旋转向量。姿态从 q_1 变换到 q_2 ，所走路径是向量 θ 收回到流形上所形成的测地线。如果我们取一半的路径，就是两者的平均值。因此新的差值四元数 \bar{r} 为

$$\bar{r} = \begin{bmatrix} \cos(\|\theta\|/4) \\ \frac{\theta}{\|\theta\|} \sin(\|\theta\|/4) \end{bmatrix}.$$

并且四元数的平均值为 $\bar{q} = q_1 \otimes \bar{r}$ 。



根据常识，均值 \bar{q} 是很接近真值 q 的地方。为了提高精度，我们再考虑 q_1 和 q_2 各自的不确定度，用协方差表示，计算各自信任系数，就形成了互补滤波。再考虑各种传感器数据陆续到达的情况，最终的递归算法的思路是：我们已知机体的 $k-1$ 时刻的姿态及其协方差，在 k 时刻机体受到外力扰动，姿态发生变化，我们计算新的姿态及其协方差，这是时间更新阶段；在测量更新阶段，我们有校正传感器数据及其协方差，我们由此计算新的均值姿态以及新的协方差。后一步又称校正步骤，随着校正传感器增多，则均值姿态逐步逼近真值。

在卡尔曼滤波器这样的递归算法中，协方差 Σ_x 的传播通过时间更新和测量更新步骤中在不同的局部坐标系中计算，而最优估计值 \bar{x} 可以认为是由多种传感器的测量值以其噪声为权重进行校正后再进行计算的均值元素。

对于在局部切空间中的协方差 Σ_x ，在卡尔曼滤波器中称为 \mathbf{P} 矩阵，表示最小二乘法的解的不确定性。我们知道协方差矩阵 \mathbf{P} 所代表的概率分布在空间中张成一个椭球，其中矩阵的特征向量指示数据扩散方向，而特征值，即方差，指示扩散程度，矩阵中的协方差项指示椭球在原点坐标系中的旋转角度。在当前时刻，均值元素 \bar{x} 既是局部切空间中的原点，也是 \mathbf{P} 矩阵所表示的概率中心，就是真实值最有可能存在的地方。在每一个穿过概率中心的直线上取概率密度，得到的就是真实值在各个方向上的高斯分布。随着时间的推移，随着 \mathbf{P} 矩阵的传播，其概率中心偏离原点，而误差向量就是从切点/原点指向概率中心，新时刻下真实值最有可能存在的地方，因此误差向量就是在当前局部切空间中的最优估计。将误差向量从切空间收回到流形并作用到当前估计值上，则当前估计值就沿测地线移动到最优估计点上。

7 局部切空间的协方差变换

我们对 $\text{SO}(3)$ 上的局部切空间的协方差变换感兴趣，因为在每个局部切空间中的协方差应以当前线性化点为概率中心，如果当前姿态点向前变化，则新的协方差也要以新的姿态点为概率中心。如果我们原有姿态为 $\mathcal{X} = \mathbf{R}$ ，受到干扰向量 $\boldsymbol{\theta}$ 的扰动，新的姿态点为 $\mathcal{Y} = \mathbf{R} \oplus \boldsymbol{\theta}$ 。我们在 \mathcal{X} 点处有协方差 ${}^{\mathcal{X}}\boldsymbol{\Sigma}_{\mathcal{X}}$ ，我们需要将其变换到 \mathcal{Y} 点处的协方差 ${}^{\mathcal{Y}}\boldsymbol{\Sigma}_{\mathcal{X}}$ ，根据方程 (22) 我们有

$$\begin{aligned} {}^{\mathcal{Y}}\boldsymbol{\Sigma}_{\mathcal{X}} &= \mathbf{Ad}_{(\mathcal{Y} \ominus \mathcal{X})} {}^{\mathcal{X}}\boldsymbol{\Sigma}_{\mathcal{X}} \mathbf{Ad}_{(\mathcal{Y} \ominus \mathcal{X})}^{\top} \\ &= \mathbf{Ad}_{\mathcal{X}^{-1}} \mathbf{Ad}_{\mathcal{Y}} {}^{\mathcal{X}}\boldsymbol{\Sigma}_{\mathcal{X}} (\mathbf{Ad}_{\mathcal{X}^{-1}} \mathbf{Ad}_{\mathcal{Y}})^{\top}. \end{aligned}$$

对于流形上的导数，我们有

$$\frac{{}^{\mathcal{X}}Df(\mathcal{X})}{D\mathcal{X}} \triangleq \lim_{\boldsymbol{\theta} \rightarrow 0} \frac{f(\mathcal{X} \oplus \boldsymbol{\theta}) \ominus f(\mathcal{X})}{\boldsymbol{\theta}},$$

并且对于 $\boldsymbol{\theta}$ 的小值，以下近似值适用，

$$f(\mathcal{X} \oplus {}^{\mathcal{X}}\boldsymbol{\theta}) \xrightarrow{{}^{\mathcal{X}}\boldsymbol{\theta} \rightarrow 0} f(\mathcal{X}) \oplus \frac{{}^{\mathcal{X}}Df(\mathcal{X})}{D\mathcal{X}} {}^{\mathcal{X}}\boldsymbol{\theta},$$

因此对于 $\text{SO}(3)$ 元素经过各种化简，我们得到方程 (23) 的具体形式

$${}^{\mathcal{Y}}\boldsymbol{\Sigma}_{\mathcal{X}} \approx \mathbf{J}_{\boldsymbol{\theta}}^{\mathbf{R} \oplus \boldsymbol{\theta}} {}^{\mathcal{X}}\boldsymbol{\Sigma}_{\mathcal{X}} (\mathbf{J}_{\boldsymbol{\theta}}^{\mathbf{R} \oplus \boldsymbol{\theta}})^{\top}, \quad (24)$$

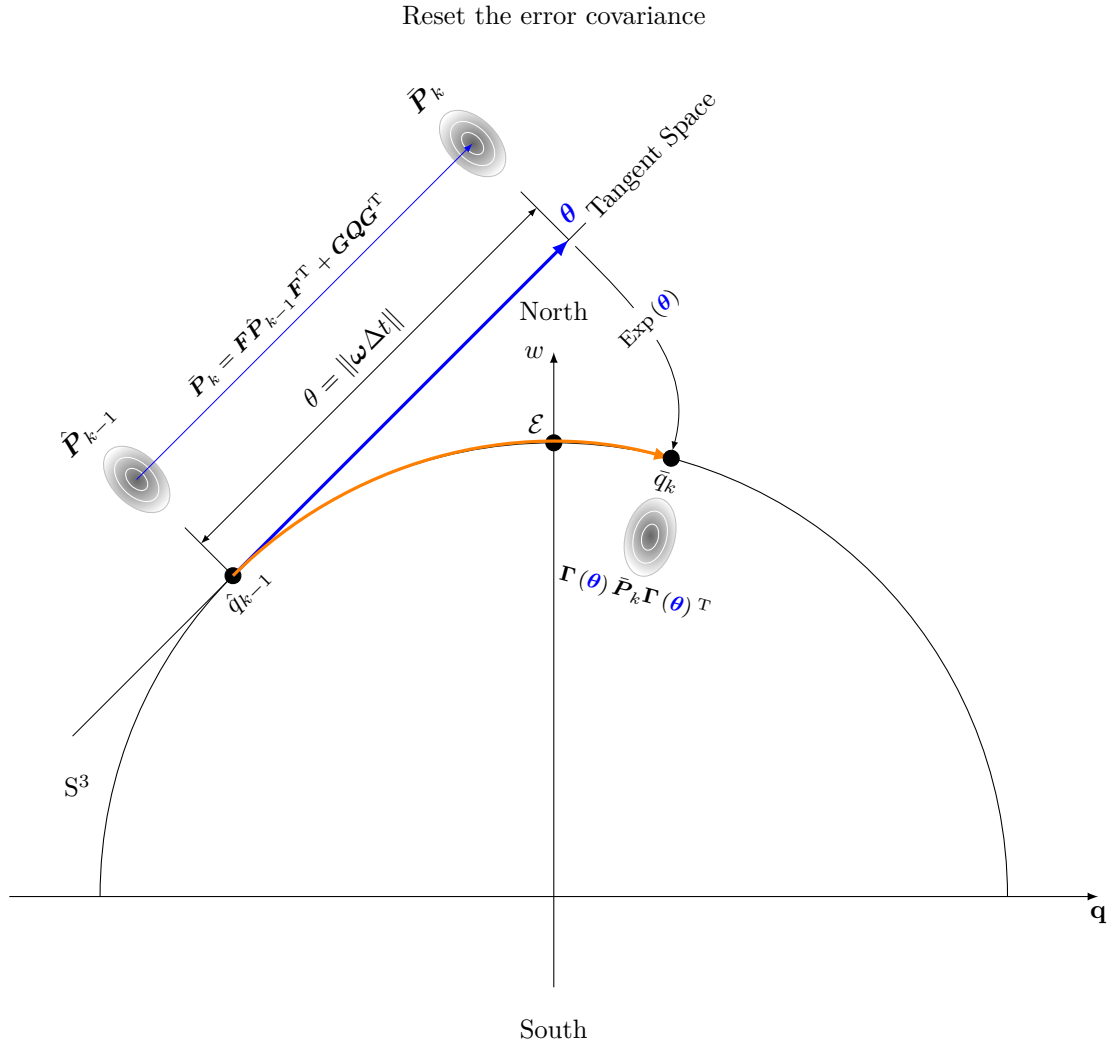
并且对于旋转 \mathbf{R} 有

$$\begin{aligned} \mathbf{J}_{\boldsymbol{\theta}}^{\mathbf{R} \oplus \boldsymbol{\theta}} &= \mathbf{J}_r(\boldsymbol{\theta}) \\ &= \mathbf{I} - \frac{1 - \cos \theta}{\theta^2} [\boldsymbol{\theta}]_{\times} + \frac{\theta - \sin \theta}{\theta^3} [\boldsymbol{\theta}]_{\times}^2 \\ \mathbf{J}_r(\boldsymbol{\theta})^{\top} &= \mathbf{J}_l(\boldsymbol{\theta}) \\ &= \mathbf{I} + \frac{1 - \cos \theta}{\theta^2} [\boldsymbol{\theta}]_{\times} + \frac{\theta - \sin \theta}{\theta^3} [\boldsymbol{\theta}]_{\times}^2. \end{aligned}$$

习惯上协方差的变换矩阵用符号 $\boldsymbol{\Gamma}$ 表示

$$\begin{aligned} \boldsymbol{\Gamma}(\boldsymbol{\theta}) &= \mathbf{J}_r(\boldsymbol{\theta}) \\ {}^{\mathcal{Y}}\boldsymbol{\Sigma}_{\mathcal{X}} &= \boldsymbol{\Gamma}(\boldsymbol{\theta}) {}^{\mathcal{X}}\boldsymbol{\Sigma}_{\mathcal{X}} \boldsymbol{\Gamma}(\boldsymbol{\theta})^{\top} \end{aligned}$$

在卡尔曼滤波器中的协方差矩阵的变换示意如下：



但是，一般情况下 IMU 传感器采样率很高， θ 都很小， $\theta \rightarrow 0$ ，所以 $\Gamma(\theta) \rightarrow \mathbf{I}$ ，在大多数实现中都省略局部切空间的协方差变换这一步，以提高运算速度。但是理解这一步的几何意义对于理解李群的局部切空间变换十分有帮助。

注释：方程 (24) 在文献 [4] 中通过 BCH 方程的一阶形式得到严格证明。但是在文献 [4] 中关于 \mathbf{J}_l 和 \mathbf{J}_r 的关系有一处理解错误。两者正确的关系是转置关系， $\mathbf{J}_l = \mathbf{J}_r^\top$ ，而他们错误理解为求逆关系， $\mathbf{J}_l = \mathbf{J}_r^{-1}$ ，并且旋转向量的动力学方程 (7, 8) 也理解错误。不过修正了这些错误后还是得到相同的方程，不影响最终结论。

8 流形中的不确定性

方程 (20) 直接定义的是当前线性化点处局部切空间中的不确定性。有时候我们需要使用在流形中的不确定性。

收回 (retraction) 是定义流形不确定性的基本概念。因为直接在流形上定义不确定性比较麻烦，并且难以计算。一般的想法是，我们可以定义切空间上的分布，并使用收回将它们映射回流形上。例如，我们可以在以 \mathcal{X} 为中心的切空间中定义一个零均值的高斯变量 $\eta \sim \text{Gaussian}(\mathbf{0}_{n \times 1}, \Sigma)$ ，并使用收回：

$$\begin{aligned} \text{Gaussian}(\mathcal{X}, \boldsymbol{\eta}) &:= \text{retract}_{\mathcal{X}}(\boldsymbol{\eta}) \\ &= \mathcal{X} \text{Exp}(\boldsymbol{\eta}) \end{aligned}$$

其中 $\boldsymbol{\eta}$ 为协方差矩阵 \mathbf{P} 对角线上的方差向量，因为非对角线上的协方差项只影响旋转，而不影响扩散程度。通过收回，我们可以在以 \mathcal{X} 为中心的切空间上定义高斯分布，并将它们映射回流形上，从而在具有均值 \mathcal{X} 和协方差 $\text{Cov}(\boldsymbol{\eta}) = \boldsymbol{\Sigma}$ 的流形上构造高斯分布。通过这样的定义，我们可以重新获取当前线性化点的高斯分布。

与上面一致，我们选择在右侧的收回作用。也是为了统一使用右乘的形式，根据各个公式的特点，我们将使用李群伴随的概念，将应用在右侧的增量或校正，与左侧的增量或校正联系起来。这样的性质将允许我们用代数方法处理李群定义的不确定性，并获得不同协方差变换的表达式。我们将集中讨论 3D 位姿，即 $\text{SE}(3)$ ，因为它具有广泛的适用性，但是类似的定义也应该适用于其它李群，因为它们主要依赖于伴随的定义。

传播不确定性有位姿组合 (Pose Composition)、位姿求逆 (Pose Inverse) 和相对位姿 (Relative Pose) 这三种操作。假设位姿是独立的或是相关的，有两种类似的推导。对于位姿相关的公式的推导，需要用到 BCH 公式的四阶交叉项的展开式，复杂很多，精度会有提高，参见参考文献 [2,5]。最后可以根据精度和计算量的要求进行选择。

	Pose Composition	Pose Inverse	Relative Pose
Independent	$\mathbf{T}_{ab} = \exp(\boldsymbol{\xi}_{ab}^\wedge) \bar{\mathbf{T}}_{ab} \quad \bar{\mathbf{T}}_{ab} \in \text{SE}(3), \boldsymbol{\xi}_{ab}^\wedge \in \mathfrak{se}(3)$		
	$\begin{aligned} \boldsymbol{\xi}_{ij} &\sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_{ij}) \\ \boldsymbol{\xi}_{jk} &\sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_{jk}) \end{aligned}$	$\boldsymbol{\xi}_{ij} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_{ij})$	$\begin{aligned} \boldsymbol{\xi}_{ij} &\sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_{ij}) \\ \boldsymbol{\xi}_{jk} &\sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_{jk}) \end{aligned}$
	$\begin{aligned} \bar{\mathbf{T}}_{ik} &\triangleq \bar{\mathbf{T}}_{ij} \bar{\mathbf{T}}_{jk} \\ \boldsymbol{\Sigma}_{ik} &\approx \boldsymbol{\Sigma}_{ij} + \text{Ad}_{\bar{\mathbf{T}}_{ij}} \boldsymbol{\Sigma}_{jk} \text{Ad}_{\bar{\mathbf{T}}_{ij}}^\top \end{aligned}$	$\begin{aligned} \bar{\mathbf{T}}_{ji} &\triangleq \bar{\mathbf{T}}_{ij}^{-1} \\ \boldsymbol{\Sigma}_{ji} &\approx \text{Ad}_{\bar{\mathbf{T}}_{ij}^{-1}} \boldsymbol{\Sigma}_{ij} \text{Ad}_{\bar{\mathbf{T}}_{ij}^{-1}}^\top \end{aligned}$	$\begin{aligned} \bar{\mathbf{T}}_{jk} &\triangleq \bar{\mathbf{T}}_{ij}^{-1} \bar{\mathbf{T}}_{ik} \\ \boldsymbol{\Sigma}_{jk} &\approx \text{Ad}_{\bar{\mathbf{T}}_{ij}^{-1}} \boldsymbol{\Sigma}_{ij} \text{Ad}_{\bar{\mathbf{T}}_{ij}^{-1}}^\top + \text{Ad}_{\bar{\mathbf{T}}_{ij}^{-1}} \boldsymbol{\Sigma}_{ik} \text{Ad}_{\bar{\mathbf{T}}_{ij}^{-1}}^\top \end{aligned}$
Relative	$\mathbf{T}_{ab} = \exp(\boldsymbol{\xi}_{ab}^\wedge) \bar{\mathbf{T}}_{ab} \quad \bar{\mathbf{T}}_{ab} \in \text{SE}(3), \boldsymbol{\xi}_{ab}^\wedge \in \mathfrak{se}(3)$		
	$\begin{aligned} \boldsymbol{\xi} &= \begin{bmatrix} \boldsymbol{\xi}_{ij}^\top, \boldsymbol{\xi}_{jk}^\top \end{bmatrix}^\top \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}) \\ \boldsymbol{\Sigma} &= \begin{bmatrix} \boldsymbol{\Sigma}_{ij} & \boldsymbol{\Sigma}_{ij,jk} \\ \boldsymbol{\Sigma}_{ij,jk}^\top & \boldsymbol{\Sigma}_{jk} \end{bmatrix} \end{aligned}$	$\boldsymbol{\xi}_{ij} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_{ij})$	$\begin{aligned} \boldsymbol{\xi} &= \begin{bmatrix} \boldsymbol{\xi}_{ij}^\top, \boldsymbol{\xi}_{jk}^\top \end{bmatrix}^\top \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}) \\ \boldsymbol{\Sigma} &= \begin{bmatrix} \boldsymbol{\Sigma}_{ij} & \boldsymbol{\Sigma}_{ij,jk} \\ \boldsymbol{\Sigma}_{ij,jk}^\top & \boldsymbol{\Sigma}_{jk} \end{bmatrix} \end{aligned}$
	$\begin{aligned} \bar{\mathbf{T}}_{ik} &\triangleq \bar{\mathbf{T}}_{ij} \bar{\mathbf{T}}_{jk} \\ \boldsymbol{\Sigma}_{ik} &\approx \boldsymbol{\Sigma}_{ij} + \text{Ad}_{\bar{\mathbf{T}}_{ij}} \boldsymbol{\Sigma}_{jk} \text{Ad}_{\bar{\mathbf{T}}_{ij}}^\top + \\ &\quad + \boldsymbol{\Sigma}_{ij,jk} \text{Ad}_{\bar{\mathbf{T}}_{ij}}^\top + \text{Ad}_{\bar{\mathbf{T}}_{ij}} \boldsymbol{\Sigma}_{ij,jk}^\top \end{aligned}$	$\begin{aligned} \bar{\mathbf{T}}_{ji} &\triangleq \bar{\mathbf{T}}_{ij}^{-1} \\ \boldsymbol{\Sigma}_{ji} &\approx \text{Ad}_{\bar{\mathbf{T}}_{ij}^{-1}} \boldsymbol{\Sigma}_{ij} \text{Ad}_{\bar{\mathbf{T}}_{ij}^{-1}}^\top \end{aligned}$	$\begin{aligned} \bar{\mathbf{T}}_{jk} &\triangleq \bar{\mathbf{T}}_{ij}^{-1} \bar{\mathbf{T}}_{ik} \\ \boldsymbol{\Sigma}_{jk} &\approx \text{Ad}_{\bar{\mathbf{T}}_{ij}^{-1}} \boldsymbol{\Sigma}_{ij} \text{Ad}_{\bar{\mathbf{T}}_{ij}^{-1}}^\top + \text{Ad}_{\bar{\mathbf{T}}_{ij}^{-1}} \boldsymbol{\Sigma}_{ik} \text{Ad}_{\bar{\mathbf{T}}_{ij}^{-1}}^\top + \\ &\quad - \text{Ad}_{\bar{\mathbf{T}}_{ij}^{-1}} \boldsymbol{\Sigma}_{ij,jk} \text{Ad}_{\bar{\mathbf{T}}_{ij}^{-1}}^\top - \text{Ad}_{\bar{\mathbf{T}}_{ij}^{-1}} \boldsymbol{\Sigma}_{ij,jk}^\top \text{Ad}_{\bar{\mathbf{T}}_{ij}^{-1}}^\top \end{aligned}$

用位姿点上的概率分布，根据位姿组合 (Pose Composition)、位姿求逆 (Pose Inverse) 和相对位姿 (Relative Pose) 这三种操作，我们就可以求出相关位姿点周围的概率分布。

9 总结

在李群中，我们是在当前线性化点处的局部切空间中直接定义的不确定性，这需要利用 Jacobian 矩阵。再将其收回 (retraction) 到流形上就在当前线性化点周围的流形上形成了流形上的不确

定性。时变的估计系统要不断地向前推进当前线性化点及其局部切空间，新的协方差矩阵也要以新的当前线性化点为概率中心，这需要利用伴随矩阵。同时，为了在位姿组合 (Pose Composition)、位姿求逆 (Pose Inverse) 和相对位姿 (Relative Pose) 这三种操作中保持右乘形式，即组合式的位姿在左侧，组合式的概率分布在右侧，以便进行分析，则需要大量利用伴随矩阵。

10 参考文献

1. Lie Groups for 2D and 3D Transformations - 2017
2. Characterizing the Uncertainty of Jointly Distributed Poses in the Lie Algebra - 2019
3. A micro Lie theory for state estimation in robotics - 2020
4. Full-Order Solution to the Attitude Reset Problem for Kalman Filtering of Attitudes - 2020
5. Reducing the uncertainty about the uncertainties - 2021