

# 从零开始手写 VIO

## 第一讲 概述与课程介绍

贺一家、高翔、崔华坤

2019 年 10 月 10 日

① 课程内容与提要

② VIO 概述

③ 预备知识回顾

## Section 1

# 课程内容与提要

# 课程与内容提要

- 本课程专注于视觉 +IMU 融合定位的基础理论和实现，作为视觉 SLAM 的进阶课程
- 在本课程中，你将学习到的重点内容有：
  - IMU 的工作原理和噪声方程
  - 视觉与 IMU 紧耦合的基础理论
  - 从零开始实现 VIO 紧耦合优化器（仅基于 Eigen）

相比视觉 SLAM 基础课程，本次课程更加注重 IMU 融合的理论推导与实现。

# 课程与内容提要

## 讲师介绍：

贺一家，中科院自动化所博士，“白巧克力亦唯心”博主，研究方向包括视觉 SLAM、多传感器信息融合。CSDN 博客专家，发表 SCI 和 EI 论文数篇。



# 课程与内容提要

## 讲师介绍：

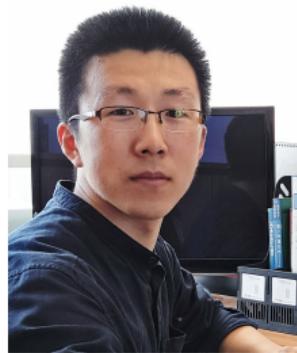
高翔，清华大学博士，慕尼黑工业大学博士后，《视觉 SLAM 十四讲》、《机器人学中的状态估计》作者、译者。曾在 RAS、IROS 等期刊会议发表论文。



# 课程与内容提要

## 讲师介绍：

崔华坤，北京工业大学物理学专业毕业，北京臻迪科技 SLAM 算法工程师，研究方向为基于多传感器融合的无人机定位导航。



# 课程与内容提要

- 课程时间安排：公式推导为主，代码为辅
  - 课程提要，约 1 小时
  - IMU 原理与模型，约 2.5 小时
  - 后端优化与 VIO BA，约 2.5 小时
  - 滑动窗口算法与 FEJ，约 2.5 小时
  - 后端求解系统与代码，约 2.5 小时
  - 前端介绍，约 1.5 小时
  - VIO 初始化，约 1.5 小时
  - VIO 系统回顾与展望，约 1 小时
- 课后作业时间安排：以编程为主，公式推导为辅，每周约 8 小时。

# 课程与内容提要

## 预备知识

- 数学：线性代数、微积分、概率论
- 编程：Linux、C++、OpenCV
- 英语：文献阅读
- 专业知识：2D/3D 计算机视觉、图像处理

## Section 2

# VIO 概述

# VIO 概述

VIO: (Visual-Inertial Odometry)

以视觉与 IMU 融合实现里程计

IMU (Inertial Measurement Unit), 惯性测量单元

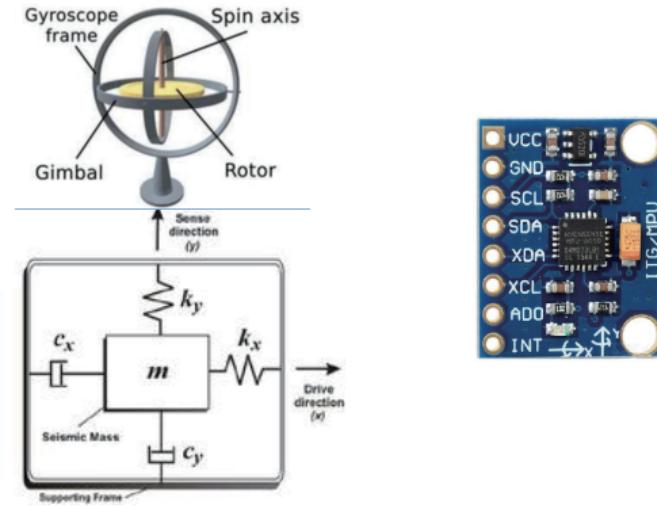
- 典型 6 轴 IMU 以较高频率 ( $\geq 100\text{Hz}$ ) 返回被测量物体的角速度与加速度
- 受自身温度、零偏、振动等因素干扰，积分得到的平移和旋转容易漂移

视觉 Visual Odometry

- 以图像形式记录数据，频率较低 (15 – 60Hz 居多)
- 通过图像特征点或像素推断相机运动

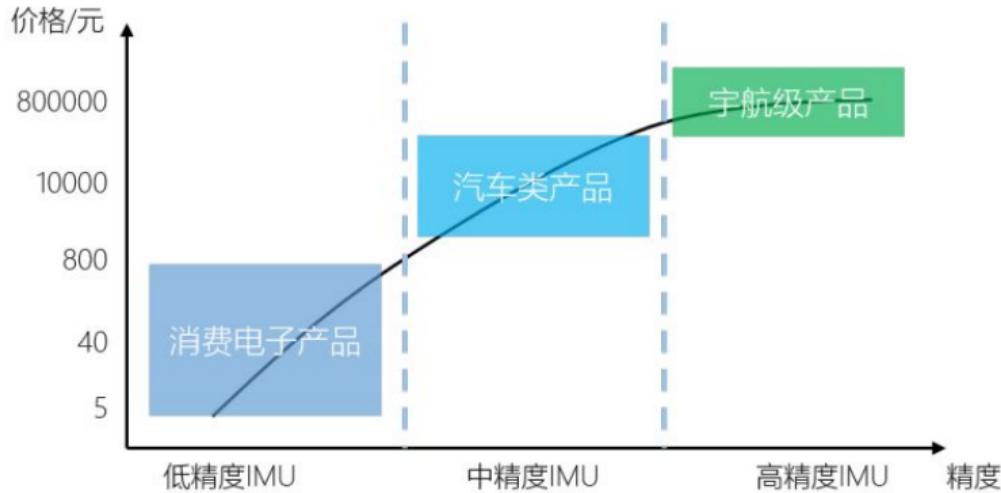
# VIO 概述

六自由度 IMU 本身由一个陀螺仪和一个加速度计组成，分别测量自身的角速度和加速度。



# VIO 概述

## IMU的精度/价格/使用场景



手机等电子产品多使用价格低廉的 MEMS IMU (如 MPU 6050)，自动驾驶类则多使用几万元的 IMU (如 Apollo 中使用的 Novatel SPAN-IGM-A1)<sup>1</sup>。

<sup>1</sup> 图片来自<https://zhuanlan.zhihu.com/p/32693377>

# VIO 概述

IMU 与视觉定位方案优势与劣势对比：

方案	IMU	视觉
优势	快速响应 不受成像质量影响 角速度普遍比较准确 可估计绝对尺度	不产生漂移 直接测量旋转与平移
劣势	存在零偏 低精度 IMU 积分位姿发散 高精度价格昂贵	受图像遮挡、运动物体干扰 单目视觉无法测量尺度 单目纯旋转运动无法估计 快速运动时易丢失

# VIO 概述

整体上，视觉和 IMU 定位方案存在一定互补性质：

- IMU 适合计算短时间、快速的运动；
- 视觉适合计算长时间、慢速的运动。

同时，可利用视觉定位信息来估计 IMU 的零偏，减少 IMU 由零偏导致的发散和累积误差；反之，IMU 可以为视觉提供快速运动时的定位。

# VIO 概述

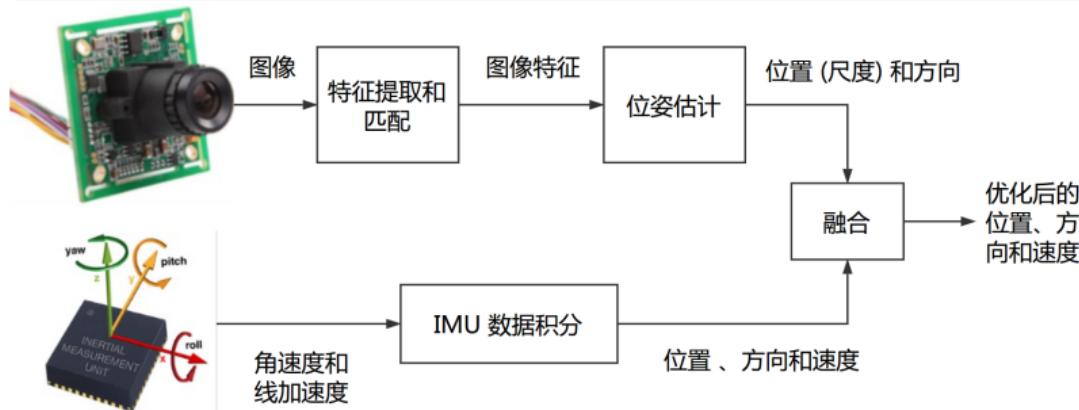
## IMU 数据可与多种定位方案融合

- 自动驾驶中通常用 IMU+GPS/差分 GPS/RTK 的融合定位方案，形成 GNSS-INS 组合导航系统，达到厘米级定位精度；
- 头戴式 AR/VR 头盔则多使用视觉 +IMU 的 VIO 定位系统，形成高帧率定位方案。

# 融合方案

## 松耦合

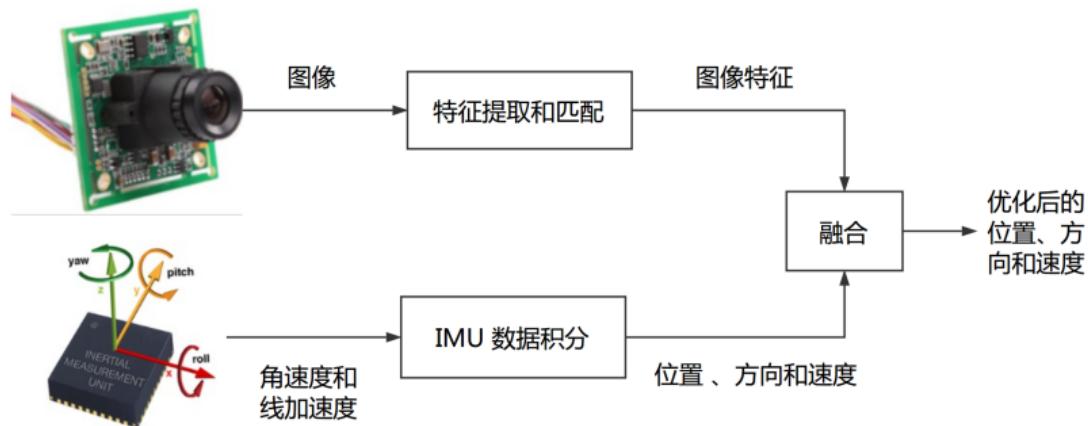
将 IMU 定位与视觉/GNSS 的位姿直接进行融合，融合过程对二者本身不产生影响，作为后处理方式输出。典型方案为卡尔曼滤波器。



# 融合方案

## 紧耦合

融合过程本身会影响视觉和 IMU 中的参数（如 IMU 的零偏和视觉的尺度）。典型方案为 MSCKF 和非线性优化。



# VIO 概述

## 为什么要使用紧耦合

- 单纯凭（单目）视觉或 IMU 都不具备估计 Pose 的能力：视觉存在尺度不确定性、IMU 存在零偏导致漂移；
- 松耦合中，视觉内部 BA 没有 IMU 的信息，在整体层面来看不是最优的。
- 紧耦合可以一次性建模所有的运动和测量信息，更容易达到最优。

# VIO 概述

## 本课程要探讨的问题

- IMU 的测量数据表达了系统的什么状态，受哪些噪声影响？
- 如何建立一个带有 IMU 测量信息和视觉特征点信息的非线性优化问题并进行求解？
- 该问题随着时间将发生怎样的演变？

## Section 3

# 预备知识回顾

# 预备知识回顾

数学符号约定：

- 普通变量： $a, b, c$
- 矩阵和向量： $\mathbf{A}, \mathbf{B}, \mathbf{v}$
- 集合： $\mathbb{R}, \mathbb{Z}$
- 特殊集合： $\mathcal{F}, \mathcal{G}$
- 希腊字母和向量： $\alpha, \boldsymbol{\alpha}$
- 李代数： $\mathfrak{so}(3), \mathfrak{se}(3)$

# 预备知识回顾

## 三维刚体运动：

- 我们通常在机器人/车辆上定义各种坐标系，如：
  - 世界坐标系 W；
  - IMU 坐标系 I；
  - 相机坐标系 C；
- 坐标系之间变换关系由一个  $SE(3)$  给出。如 I 到 W 系的变换矩阵为： $T_{WI}$ ：

$$T_{WI} = \begin{bmatrix} R_{WI} & t_{WI} \\ \mathbf{0}^T & 1 \end{bmatrix} \in \mathbb{R}^{4 \times 4}. \quad (1)$$

- $R_{WI}$  为  $3 \times 3$  的旋转矩阵， $t_{WI}$  为平移向量。
- $T_{WI}$  右乘一个 I 系下（齐次）坐标，将得到该点 W 系下坐标。

# 预备知识回顾

约定：

- 当某个量表达坐标系的转换关系时，写在右下脚标，例如  $T_{WB}$ 。
- 当表达矢量在某坐标系中取的坐标时，写在右上角标，如  $v^w$  表达速度矢量在 World 系坐标。
- I 系也称为 Body 系。
- 定义明确时，有时会省略该脚标，我们会直接谈论  $R, t$  这样的量。
- 不刻意区分齐次和非齐次坐标，因为在程序中可以自动完成转换，且无歧义。
- 默认以  $T_{WI}$  表达并存储 IMU 的定位信息，而不是  $T_{IW}$ 。二者实际互为逆，存储哪一类区别不大，视习惯而定。
- 同理， $T_{WI}$  的平移部分可直接视作 IMU 在世界中的坐标，从而进行绘图或可视化操作。

# 预备知识回顾

四元数：

- 旋转矩阵  $R$  亦可用四元数  $q$  描述。
- 四元数  $q$  有一个实部和三个虚部。我们把实部写在前：

$$\mathbf{q} = [q_0, q_1, q_2, q_3]^\top \text{ 或 } \mathbf{q} = [w, x, y, z]^\top \quad (2)$$

其中  $q_0$  为实部， $[q_1, q_2, q_3]^\top$  为虚部。因为实部为标量，虚部为矢量，所以也可记为：

$$\mathbf{q} = [s, \mathbf{v}]^\top. \quad (3)$$

其中  $s$  为标量， $\mathbf{v}$  为虚部的矢量。

# 预备知识回顾

四元数之间可以进行乘法运算：

$$\begin{aligned}
 \mathbf{q}_a \otimes \mathbf{q}_b = & w_a w_b - x_a x_b - y_a y_b - z_a z_b \\
 & + (w_a x_b + x_a w_b + y_a z_b - z_a y_b) i \\
 & + (w_a y_b - x_a z_b + y_a w_b + z_a x_b) j \\
 & + (w_a z_b + x_a y_b - y_a x_b + z_a w_b) k.
 \end{aligned} \tag{4}$$

或：

$$\mathbf{q}_a \otimes \mathbf{q}_b = \left[ s_a s_b - \mathbf{v}_a^\top \mathbf{v}_b, s_a \mathbf{v}_b + s_b \mathbf{v}_a + \mathbf{v}_a \times \mathbf{v}_b \right]^\top. \tag{5}$$

此外，四元数可类似复数，定义加减、模长、逆、共轭等运算，不一一展开。

# 预备知识回顾

单位四元数可表达任意三维旋转，且无奇异性。

四元数和角轴的转换关系：

假设某个旋转运动的旋转轴为单位向量  $\mathbf{u}$ ，绕该轴旋转的角度为  $\theta$ ，那么它对应的单位四元数为：

$$\mathbf{q} = \begin{bmatrix} \cos \frac{\theta}{2} \\ \mathbf{u} \sin \frac{\theta}{2} \end{bmatrix} \quad (6)$$

当旋转一段微小时间，即角度趋于 0 时，容易有：

$$\Delta \mathbf{q} = \begin{bmatrix} \cos \frac{\delta\theta}{2} \\ \mathbf{u} \sin \frac{\delta\theta}{2} \end{bmatrix} \approx \begin{bmatrix} 1 \\ \mathbf{u} \frac{\delta\theta}{2} \end{bmatrix} = \begin{bmatrix} 1 \\ \frac{1}{2}\delta\theta \end{bmatrix} \quad (7)$$

其中  $\delta\theta$  的方向表示旋转轴，模长表示旋转角度。

# 预备知识回顾：对时间求导

角速度：

$$\omega = \lim_{\Delta t \rightarrow 0} \frac{\delta \theta}{\Delta t} \quad (8)$$

四元数时间导数：

$$\begin{aligned}
 \dot{\mathbf{q}} &\triangleq \lim_{\Delta t \rightarrow 0} \frac{\mathbf{q}(t + \Delta t) - \mathbf{q}(t)}{\Delta t} \\
 &= \lim_{\Delta t \rightarrow 0} \frac{\mathbf{q} \otimes \Delta \mathbf{q} - \mathbf{q}}{\Delta t} \\
 &= \lim_{\Delta t \rightarrow 0} \frac{\mathbf{q} \otimes \left( \begin{bmatrix} 1 \\ \frac{1}{2}\delta\theta \end{bmatrix} - \begin{bmatrix} 1 \\ \mathbf{0} \end{bmatrix} \right)}{\Delta t} \\
 &= \mathbf{q} \otimes \begin{bmatrix} 0 \\ \frac{1}{2}\boldsymbol{\omega} \end{bmatrix}
 \end{aligned} \quad (9)$$

# 预备知识回顾

除了利用四元数求导，亦可利用李代数进行旋转求导。

使用旋转矩阵  $R$  时，角速度为  $\omega$ ，那么  $R$  相对于时间的导数可写作：

$$\dot{R} = R \omega^{\wedge} \quad (10)$$

该式被称为泊松公式 (Poisson's equation)，其中  $^{\wedge}$  为反对称矩阵算子：

$$\omega^{\wedge} = \begin{bmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{bmatrix} \quad (11)$$

在本课程中亦记作： $\omega_{\times}$  或者  $[\omega]_{\times}$ ，表达含义相同。

# 预备知识回顾

$\mathfrak{so}(3)$  导数：

在优化带有旋转的函数时，通常计算一个增量  $\phi \in \mathfrak{so}(3)$ ，然后用它更新当前估计值：

$$\mathbf{R} \leftarrow \mathbf{R} \exp(\phi^\wedge). \quad (12)$$

其中  $\exp$  为  $\mathfrak{so}(3)$  至  $\text{SO}(3)$  上的指数映射。

本课程习惯为右乘，实际当中左右乘等价，仅为习惯上的差别。

注：(i) 不同的  $\mathbf{R}$  函数，具体的导数形式也不同。(ii) 在程序中，不必区分  $\mathbf{R}$  是以矩阵存储或是以四元数存储，只需按照该式更新即可。

# 预备知识回顾

常见的一些雅可比（以自变量为  $R$  举例）：  
旋转点的左扰动雅可比：

$$\begin{aligned}
 \frac{\partial (\mathbf{R}\mathbf{p})}{\partial \varphi} &= \lim_{\varphi \rightarrow 0} \frac{\exp(\varphi^\wedge) \exp(\phi^\wedge) \mathbf{p} - \exp(\phi^\wedge) \mathbf{p}}{\varphi} \\
 &= \lim_{\varphi \rightarrow 0} \frac{(\mathbf{I} + \varphi^\wedge) \exp(\phi^\wedge) \mathbf{p} - \exp(\phi^\wedge) \mathbf{p}}{\varphi} \\
 &= \lim_{\varphi \rightarrow 0} \frac{\varphi^\wedge \mathbf{R}\mathbf{p}}{\varphi} = \lim_{\varphi \rightarrow 0} \frac{-(\mathbf{R}\mathbf{p})^\wedge \varphi}{\varphi} = -(\mathbf{R}\mathbf{p})^\wedge.
 \end{aligned} \tag{13}$$

# 预备知识回顾

旋转点的右扰动雅可比：

$$\begin{aligned}
 \frac{\partial (\mathbf{R}\mathbf{p})}{\partial \varphi} &= \lim_{\varphi \rightarrow 0} \frac{\mathbf{R} \exp(\varphi^\wedge) \mathbf{p} - \mathbf{R}\mathbf{p}}{\varphi} \\
 &= \lim_{\varphi \rightarrow 0} \frac{\mathbf{R}(\mathbf{I} + \varphi^\wedge) \mathbf{p} - \mathbf{R}\mathbf{p}}{\varphi} \\
 &= \lim_{\varphi \rightarrow 0} \frac{\mathbf{R}\varphi^\wedge \mathbf{p}}{\varphi} = \lim_{\varphi \rightarrow 0} \frac{-\mathbf{R}\mathbf{p}^\wedge \varphi}{\varphi} = -\mathbf{R}\mathbf{p}^\wedge.
 \end{aligned} \tag{14}$$

# 预备知识回顾

旋转连乘的雅可比：

$$\begin{aligned}
 \frac{d \ln (\mathbf{R}_1 \mathbf{R}_2)^\vee}{d \mathbf{R}_2} &= \lim_{\phi \rightarrow 0} \frac{\ln (\mathbf{R}_1 \mathbf{R}_2 \exp(\phi^\wedge))^\vee - \ln (\mathbf{R}_1 \mathbf{R}_2)^\vee}{\phi} \\
 &= \lim_{\phi \rightarrow 0} \frac{\ln (\mathbf{R}_1 \mathbf{R}_2)^\vee + \mathbf{J}_r^{-1} \phi - \ln (\mathbf{R}_1 \mathbf{R}_2)^\vee}{\phi} \quad (15) \\
 &= \mathbf{J}_r^{-1} (\ln (\mathbf{R}_1 \mathbf{R}_2)^\vee)
 \end{aligned}$$

其中用到了对  $\forall \mathbf{R}$  :

$$\ln (\mathbf{R} \exp(\phi^\wedge))^\vee = \ln (\mathbf{R})^\vee + \mathbf{J}_r^{-1} \phi \quad (16)$$

和  $\mathbf{J}_r^{-1}$  为  $\text{SO}(3)$  上的右雅可比：

$$\mathbf{J}_r^{-1}(\theta \boldsymbol{\omega}) = \frac{\theta}{2} \cot \frac{\theta}{2} \mathbf{I} + \left(1 - \frac{\theta}{2} \cot \frac{\theta}{2}\right) \boldsymbol{\omega} \boldsymbol{\omega}^T + \frac{\theta}{2} \boldsymbol{\omega}^\wedge. \quad (17)$$

# 预备知识回顾

旋转连乘的雅可比：

$$\begin{aligned}
 \frac{d \ln (\mathbf{R}_1 \mathbf{R}_2)^\vee}{d \mathbf{R}_1} &= \lim_{\phi \rightarrow 0} \frac{\ln (\mathbf{R}_1 \exp(\phi^\wedge) \mathbf{R}_2)^\vee - \ln (\mathbf{R}_1 \mathbf{R}_2)^\vee}{\phi} \\
 &= \lim_{\phi \rightarrow 0} \frac{\ln \left( \mathbf{R}_1 \mathbf{R}_2 \exp \left( (\mathbf{R}_2^\top \phi)^\wedge \right) \right)^\vee - \ln (\mathbf{R}_1 \mathbf{R}_2)^\vee}{\phi} \\
 &= \mathbf{J}_r^{-1}(\ln(\mathbf{R}_1 \mathbf{R}_2)^\vee) \mathbf{R}_2^\top
 \end{aligned} \tag{18}$$

这里用到了  $\text{SO}(3)$  的伴随性质：

$$\mathbf{R}^\top \exp(\phi^\wedge) \mathbf{R} = \exp \left( \left( \mathbf{R}^\top \phi \right)^\wedge \right) \tag{19}$$

# 预备知识回顾

有关  $SE(3)$ : 由于  $SE(3)$  李代数性质复杂, 在 VIO 中, 我们通常使用  $SO(3) + t$  的形式表达旋转和平移。对平移部分使用矢量更新而非  $SE(3)$  上的更新。

# 习题

## 1. VIO 文献阅读

阅读 VIO 相关综述文献如<sup>a</sup>，回答以下问题：

- 视觉与 IMU 进行融合之后有何优势？
- 有哪些常见的视觉 +IMU 融合方案？有没有工业界应用的例子？
- 在学术界，VIO 研究有哪些新进展？有没有将学习方法用到 VIO 中的例子？

你也可以对自己感兴趣的方向进行文献调研，阐述你的观点。

<sup>a</sup> Jianjun Gui et al. "A review of visual inertial odometry from filtering and optimisation perspectives". In: *Advanced Robotics* 29.20 (2015), 1289–1301. ISSN: 0169-1864. DOI: {10.1080/01691864.2015.1057616}.

# 习题

## 2. 四元数和李代数更新

课件提到了可以使用四元数或旋转矩阵存储旋转变量。当我们用计算出来的  $\omega$  对某旋转更新时，有两种不同方式：

$$\begin{aligned} \mathbf{R} &\leftarrow \mathbf{R} \exp(\boldsymbol{\omega}^\wedge) \\ \text{或 } \mathbf{q} &\leftarrow \mathbf{q} \otimes \left[1, \frac{1}{2}\boldsymbol{\omega}\right]^\top \end{aligned} \tag{20}$$

请编程验证对于小量  $\boldsymbol{\omega} = [0.01, 0.02, 0.03]^\top$ ，两种方法得到的结果非常接近，实践当中可视为等同。因此，在后文提到旋转时，我们并不刻意区分旋转本身是  $\mathbf{q}$  还是  $\mathbf{R}$ ，也不区分其更新方式为上式的哪一种。

# 习题

## 3. 其他导数

使用右乘  $\mathfrak{so}(3)$ , 推导以下导数:

$$\frac{d(\mathbf{R}^{-1}\mathbf{p})}{d\mathbf{R}} \quad (21)$$

$$\frac{d \ln (\mathbf{R}_1 \mathbf{R}_2^{-1})^\vee}{d\mathbf{R}_2} \quad (22)$$