R-RANSAC with Matrix Lie Groups

I. Introduction

This note assumes the reader has a basic understanding of the theory relating to Matrix Lie Groups and their corresponding Lie algebras. This includes the matrix exponential and logarithm map, the vector space isomorphic to specific Lie algebras, the adjoint of the group and the algebra, and the left and right Jacobians.

II. Notation

- Let G denote a matrix Lie group.
- Let $\mathfrak{g}$ denote the Lie algebra of G.
- Let $\boxplus : G \times \mathfrak{g} \to \mathfrak{g}$ be defined as $(g, u) \mapsto g \exp(u)$
- Let $\boxminus : G \times G \to \mathfrak{g}$ be defined as $(g_1, g_2) \mapsto \log\left(g_2^{-1} \bullet g_1\right)$ where the operator $\bullet$ is matrix multiplication.
- Let the adjoint of the group be denoted $Ad : G \times G \to G$ and defined as $(g_1, g_2) \mapsto g_1 g_2 g_1^{-1}$. A variant of the map is $Ad_{g_1} : G \to G$ defined as $(g_2) \mapsto g_1 g_2 g_1^{-1}$.
- Let the adjoint of the Lie algebra be denoted $ad : G \times \mathfrak{g} \to \mathfrak{g}$ and defined as $(g, u) \mapsto gug^{-1}$. A variant of the map is $ad_g : \mathfrak{g} \to \mathfrak{g}$ and is defined as $(u) \mapsto gug^{-1}$.
- Let $p(x)$ denote the probability of x, $p(x, y)$ the joint probability of x and y, and $p(x|y)$ the probability of x conditioned on y.

In this note we are assuming a discrete system whose measurements arrive at non-fixed time intervals. Without loss in generality, we assume that the initial state, denoted x_0, occurred at time t_0, and that some other future state x_k occurred at time t_k. Given two sequential states x_{k+1} and x_k we denote the time interval between them as $t_{k+1} - t_k = \delta_{k+1}$. This might seem confusing, but it is to allow flexibility and clean notation. We will also use the notation

III. Problem Description

We assume a discrete, constant velocity, time invariant system of the form

$$x_k = x_{k-1} \boxplus (\delta_k u + \delta_k w_k) \tag{1a}$$

$$y_k = x_k \boxplus v_k \tag{1b}$$

where $x_k \in G$ is the state of the system at time t_k, $u \in \mathfrak{g}$ is the velocity, $y_k \in G$ is the observed state at time t_k, $w_k \in \mathfrak{g}$ is zero mean, white noise, Gaussian process noise with covariance Q, and $v_k \in \mathfrak{g}$ is zero mean, white noise, Gaussian measurement noise with covariance R.

IV. Log Maximum Likelihood Estimate

Let $y_{1:k}$ denote the set containing the observations from t_1 to t_k, $x_{0:k}$ be the set containing the corresponding states from t_0 to t_k. The posterior probability is

$$p(x_{0:k}, u \mid y_{1:k}).$$

Using Bayes rule we get

$$p(x_{0:k}, u \mid y_{1:k}) = \frac{p(y_k \mid x_{0:k}, u, y_{1:k-1})\, p(x_{0:k}, u \mid y_{1:k-1})}{p(y_k \mid y_{1:k-1})}.$$

Since $p(y_k \mid y_{1:k-1})$ does not depend on the parameters we are trying to estimate, we can replace it with a constant $\frac{1}{\eta}$ so that the posterior becomes

$$\eta p(y_k \mid x_{0:k}, u, y_{1:k-1})\, p(x_{0:k}, u \mid y_{1:k-1}).$$

Under the assumption that the system meets the conditions to be a Markov chain, we can simplify the posterior to

$$\eta p(y_k \mid x_k)\, p(x_{0:k}, u \mid y_{1:k-1}) = \eta p(y_k \mid x_k)\, p(x_k \mid x_{k-1}, u)\, p(x_{0:k-1}, u \mid y_{1:k-1}).$$

Repeating the process recursively, we get

$$p\left(x_{0:k}, u \mid y_{1:k}\right) = \eta p\left(x_0\right) p\left(u\right) \prod_{j=1}^{k} p\left(x_j \mid x_{j-1}, u\right) p\left(y_j \mid x_j\right).$$

Maximizing the posterior is equivalent to minimizing the negative log posterior. The negative log posterior is

$$\log\left(-p\left(x_{0:k}, u \mid y_{1:k}\right)\right) = \log\left(-\eta\right) + \log\left(-p\left(x_0\right)\right) + \log\left(-p\left(u\right)\right) + \sum_{j=1}^{k} \log\left(-p\left(x_j \mid x_{j-1}, u\right)\right) + \sum_{j=1}^{k} \log\left(-p\left(y_j \mid x_j\right)\right).$$

A. Model Inversion

The system model described in (1) propagates the system forward in time. We could easily invert the model to propagate the system backwards in time to get the system

$$x_{k-1} = x_k \boxplus \left(-\delta_k u - \delta_k w_k\right) \tag{2a}$$

$$y_{k-1} = x_{k-1} \boxplus v_k. \tag{2b}$$

Since we are assuming that the velocity is constant, we can write the output as

$$y_{k-1} = x_k \boxplus \left(-\delta_k u - \delta_k w_k\right) \boxplus v_k.$$

In fact, we can write the output at any time as a function of the current state

$$y_{k-n} = x_k \boxplus \left(-\left(t_k - t_{k-n}\right)\left(u - w_k\right)\right) \boxplus v_k,$$

with n being an integer from the interval $0 \leq n \leq k$.

Using this model, the negative log posterior simplifies to

$$\log\left(-p\left(x_k, u\right)\right) = \sum_{j=1}^{k} \log\left(-p\left(y_j \mid x_k, u\right)\right),$$

thus we want to solve the LMLE problem

$$\arg\min_{x_k, u} \left(\sum_{j=1}^{k} \log\left(-p\left(y_j \mid x_k, u\right)\right)\right).$$

This requires knowing the likelihood $p\left(y_j \mid x_k, u\right)$ for all $j \in \{1, 2, \ldots, k\}$.

B. Likelihood

Let y_k denote the true observation, $\hat{y}_k$ denote the estimated observation, and $\Delta_{k-n} = t_{k-n} - t_k$, then

$$y_{k-n} = x_k \boxplus \left(\Delta_{k-n}\left(u - w_k\right)\right) \boxplus v_k,$$

and

$$\hat{y}_{k-n} = x_k \boxplus \left(\Delta_{k-n} u\right).$$

Under the assumption that the observation noise is small, we can approximate the the observation

$$y_{k-n} \approx x_k \boxplus \left(\Delta_{k-n}\left(u - w_k\right) + J_r\left(\Delta_{k-n}\left(u - w_k\right)\right)^{-1} v_k\right),$$

with J_r^{-1} being the inverse of the right Jacobian.

We define the error to be

$$z_{k-n} = y_{k-n} \boxminus \hat{y}_{k-n}.$$

Expanding it out gives us

$$\begin{aligned} z_{k-n} &= \log\left(\exp\left(-\Delta_{k-n} u\right) x_k^{-1} x_k \exp\left(\Delta_{k-n}\left(u - w_k\right) + J_r\left(\Delta_{k-n}\left(u - w_k\right)\right)^{-1} v_k\right)\right) \\ &= \log\left(\exp\left(-\Delta_{k-n} u\right) \exp\left(\Delta_{k-n}\left(u - w_k\right) + J_r\left(\Delta_{k-n}\left(u - w_k\right)\right)^{-1} v_k\right)\right). \end{aligned}$$

Under the assumption that $-\Delta_{k-n} w_k + J_r\left(\Delta_{k-n}\left(u - w_k\right)\right)^{-1} v_k$ is small, we can approximate the error to be

$$\begin{aligned} z_{k-n} &\approx \log\left(\exp\left(-\Delta_{k-n}u\right)\exp\left(\Delta_{k-n}u\right)\exp\left(J_r\left(\Delta_{k-n}u\right)\left(-\Delta_{k-n}w_k + J_r\left(\Delta_{k-n}\left(u-w_k\right)\right)^{-1} v_k\right)\right)\right) \\ &= J_r\left(\Delta_{k-n}u\right)\left(-\Delta_{k-n}w_k + J_r\left(\Delta_{k-n}\left(u-w_k\right)\right)^{-1} v_k\right). \end{aligned}$$

Since w_k can be thought of as a small deviation from u, we can use the Taylor series expansion up to the first order on the term $J_r\left(\Delta_{k-n}\left(u-w_k\right)\right)^{-1}$ to get

$$J_r\left(\Delta_{k-n}\left(u-w_k\right)\right)^{-1} \approx J_r\left(\Delta_{k-n}u\right)^{-1} - \left.\frac{\partial J_r^{-1}}{\partial u}\right| \Delta_{k-n}w_k,$$

thus the error is approximated as

$$\begin{aligned} z_{k-n} &\approx J_r\left(\Delta_{k-n}u\right)\left(-\Delta_{k-n}w_k + \left(J_r\left(\Delta_{k-n}u\right)^{-1} - \left.\frac{\partial J_r^{-1}}{\partial u}\right| \Delta_{k-n}w_k\right) v_k\right) \\ &= -\Delta_{k-n}J_r\left(\Delta_{k-n}u\right) w_k - J_r\left(\Delta_{k-n}u\right)\left(\left.\frac{\partial J_r^{-1}}{\partial u}\right| \Delta_{k-n}w_k\right) v_k + v_k \end{aligned}$$

Since z_{k-n} is a random variable, we need to calculate the mean and covariance in order to form the distribution. Under the assumption that $E\left[w_k v_k\right] = 0$ we get

$$\begin{aligned} E\left[z_{k-n}\right] &= E\left[\Delta_{k-n}J_r\left(\Delta_{k-n}u\right) w_k - J_r\left(\Delta_{k-n}u\right)\left(\left.\frac{\partial J_r^{-1}}{\partial u}\right| \Delta_{k-n}w_k\right) v_k + v_k\right] \\ &= E\left[v_k\right] - \Delta_{k-n}J_r\left(\Delta_{k-n}u\right) E\left[w_k\right] \\ &= 0. \end{aligned}$$

The covariance of the error is

$$\begin{aligned} E\left[z_{k-n} z_{k-n}^\top\right] &= R + J_r\left(\Delta_{k-n}u\right)\Delta_{k-n} Q J_r\left(\Delta_{k-n}u\right)\Delta_{k-n} \\ &= R + \bar{Q}. \end{aligned}$$

Thus

$$p\left(z_{k-n} \mid x_k, u\right) = \mathcal{N}\left(0, R + \bar{Q}\right).$$

$$\arg\min_{x_k, u} \sum_{n=1}^{k} z_{k-n}^\top \left(R + \bar{Q}\right)^{-1} z_{k-n}.$$

Since x_k is an element of the matrix Lie group, it does really support operations such as addition and scalar multiplications. To work around this, we optimize over an element in the Lie algebra $\tau_k = \log\left(x_k\right)$. Finally, the final form of the optimization is

$$\arg\min_{\tau_k, u} \sum_{n=1}^{k} z_{k-n}^\top \left(R + \bar{Q}\right)^{-1} z_{k-n}. \tag{3}$$

C. Seeding

Note that the optimization problem (3) is non-linear. Nonlinear optimization problems are slow relative to linear optimizations problems. Nonlinear optimization problems can converge faster if they are seeded smartly. In this subsection we present one possible scheme of seeding. We assume that we get two measurements at two distinct time steps y_k and y_{k-n} according to the model described in (2), then we can approximate x_k by setting it equal to y_k, and then we can approximate u using the equation

$$u \approx \frac{1}{\Delta_{k-n}} y_k \boxminus y_{k-n}.$$

If we wanted to use multiple measurement to estimate u, it would become a simple linear least squares method.

V. Constrained Problem for SE(n)

The LMLE presented in the previous section relied on the measurement being an element of the group. This is not always possible depending on the sensors available. In this section we assume that the system can be modeled on the manifold $SE(n)$, only the position of the system is observable, system has constant angular and translational velocity.

Let $x \in SE(n)$, then x can be written in matrix form as

$$x = \begin{bmatrix} R & P \\ 0_{3\times 1} & 1 \end{bmatrix},$$

where $R \in SO(n)$ represents the attitude and $P \in \mathbb{R}^n$ represents the position.

Let $u \in \mathfrak{se}(n)$, then u can be written in matrix form as

$$u = \begin{bmatrix} \omega & \rho \\ 0_{3\times 1} & 0 \end{bmatrix},$$

where $\omega \in \mathfrak{so}(n)$ represents the angular velocity in the body frame, and $\rho \in \mathbb{R}^n$ denotes the translational velocity in the body frame.

The system can be described as

$$x_k = x_{k-1} \boxplus (\delta_k u + \delta_k w_k) \tag{4a}$$
$$y_k = x_k C + v_k, \tag{4b}$$

where

$$C = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

so that $y_k \in \mathbb{R}^{n+1}$. We can write the output at time k as

$$\begin{aligned} y_k &= x_{k-1} \boxplus (\delta_k u + \delta_k w_k) C + v_k \\ &= x_{k-1} \exp(\delta_k u + \delta_k w_k) C + v_k. \end{aligned}$$

Since $w_k \in \mathfrak{se}(n)$ we can represent it as

$$w_k = \begin{bmatrix} \nu & \eta \\ 0 & 0 \end{bmatrix},$$

where $\omega \in \mathfrak{so}(n)$ and $\eta \in \mathbb{R}^n$.

Using the exponential map, we can write

$$x_{k-1} \exp(\delta_k u + \delta_k w_k)$$

as

$$\begin{aligned} &\begin{bmatrix} R_{k-1} & P_{k-1} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \exp(\delta_k(\omega_k + \nu_k)) & V(\delta_k(\omega_k + \nu_k))\delta_k(\rho_k + \eta_k) \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} R_{k-1}\exp(\delta_k(\omega_k + \nu_k)) & P_{k-1} + R_{k-1}V(\delta_k(\omega_k + \nu_k))\delta_k(\rho_k + \eta_k) \\ 0 & 1 \end{bmatrix} \end{aligned}$$

Let $v \in \mathfrak{so}(n)$, $\theta = \|v\|$, and I be the identity matrix, then

$$V(v) = I + \left(\frac{1 - \cos(\theta)}{\theta^2}\right) v + \left(\frac{\theta - \sin(\theta)}{\theta^3}\right) v^2.$$

Using the expanded from, we get that

$$x_{k-1}\exp(\delta_k u + \delta_k w_k) C = P_{k-1} + R_{k-1} V(\delta_k(\omega_k + \nu_k))\delta_k(\rho_k + \eta_k),$$

thus

$$y_k = P_{k-1} + R_{k-1} V(\delta_k(\omega_k + \nu_k))\delta_k(\rho_k + \eta_k) + v_k.$$

Notice that ν_k acts like a small deviation from ω_k. We can approximate $V\left(\delta_k\left(\omega_k+\nu_k\right)\right)$ by taking it's Taylor series expansion. The series form for $V\left(v\right)$ is

$$V = I + \frac{v}{2!} + \frac{v^2}{3!} + \frac{v^3}{4!} + \cdots ,$$

thus

$$\begin{aligned}
\frac{\partial V}{\partial v} &= \frac{I}{2!} + \frac{2v}{3!} + \frac{3v^2}{4!} + \frac{4v^3}{5!} + \cdots \\
&= \frac{I}{2} + \sum_{i=0}^{\infty} \frac{\left(2i+2\right) v^{2i+1}}{\left(2i+3\right)!} + \sum_{i=0}^{\infty} \frac{\left(2i+3\right) v^{2i+2}}{\left(2i+4\right)!} \\
&= \frac{I}{2} + \sum_{i=0}^{\infty} \frac{\left(2i+2\right)\left(-1\right)^i \theta^{2i}}{\left(2i+3\right)!} v + \sum_{i=0}^{\infty} \frac{\left(2i+3\right)\left(-1\right)^i \theta^{2i}}{\left(2i+4\right)!} v^2 .
\end{aligned}$$

Notice that

$$\sum_{i=0}^{\infty} \frac{\left(2i+2\right)\left(-1\right)^i \theta^{2i}}{\left(2i+3\right)!}$$

is similar to the Taylor Series of the $\sin$ function. We are going to manipulate it to get a closed form by first multiplying it by $\left(-1\right)\left(-1\right)\frac{\theta^2}{\theta^2}$ and adding and subtracting $\frac{\sin(\theta)}{\theta^3}$ to it.

$$\begin{aligned}
\left(-1\right)\left(-1\right)\frac{\theta^3}{\theta^3}\sum_{i=0}^{\infty} \frac{\left(2i+2\right)\left(-1\right)^i \theta^{2i}}{\left(2i+3\right)!} &= \frac{-1}{\theta^3}\sum_{i=0}^{\infty} \frac{\left(2i+2\right)\left(-1\right)^{i+1} \theta^{2i+3}}{\left(2i+3\right)!} \\
&= \frac{\sin\left(\theta\right)}{\theta^3} - \frac{\sin\left(\theta\right)}{\theta^3} - \frac{1}{\theta^3}\sum_{i=0}^{\infty} \frac{\left(2i+2\right)\left(-1\right)^{i+1} \theta^{2i+3}}{\left(2i+3\right)!} \\
&\frac{\sin\left(\theta\right)}{\theta^3} - \frac{1}{\theta^3}\left(\sum_{i=0}^{\infty} \frac{\left(-1\right)^i x^{2i+1}}{\left(2n+1\right)!} + \sum_{i=0}^{\infty} \frac{\left(2i+2\right)\left(-1\right)^{i+1} \theta^{2i+3}}{\left(2i+3\right)!}\right) \\
&= \frac{\sin\left(\theta\right)}{\theta^3} - \frac{1}{\theta^3}\theta\sum_{i=0}^{\infty} \frac{\left(-1\right)^n \theta^{2i}}{2i!} \\
&= \frac{\sin\left(\theta\right)}{\theta^3} - \frac{1}{\theta^2}\cos\left(\theta\right) .
\end{aligned}$$

Next we are going to do a similar process to

$$\sum_{i=0}^{\infty} \frac{\left(2i+3\right)\left(-1\right)^i \theta^{2i}}{\left(2i+4\right)!}$$

in order to get a closed form solution. We are going to first multiply it by $\frac{\theta^4}{\theta^4}$ and then adding and subtracting $\frac{\cos(\theta)}{\theta^4}$ to it

$$\begin{aligned}
\frac{\theta^4}{\theta^4}\sum_{i=0}^{\infty} \frac{\left(2i+3\right)\left(-1\right)^i \theta^{2i}}{\left(2i+4\right)!} &= \frac{1}{\theta^4}\sum_{i=0}^{\infty} \frac{\left(2i+3\right)\left(-1\right)^i \theta^{2i+4}}{\left(2i+4\right)!} \\
&= \frac{-\cos\left(\theta\right)}{\theta^4} + \frac{\cos\left(\theta\right)}{\theta^4} + \frac{1}{\theta^4}\sum_{i=0}^{\infty} \frac{\left(2i+3\right)\left(-1\right)^i \theta^{2i+4}}{\left(2i+4\right)!} \\
&= \frac{-\cos\left(\theta\right)}{\theta^4} + \frac{1}{\theta^4}\left(\sum_{i=0}^{\infty} \frac{\left(-1\right)^i \theta^{2i}}{2i!} + \sum_{i=0}^{\infty} \frac{\left(2i+3\right)\left(-1\right)^i \theta^{2i+4}}{\left(2i+4\right)!}\right) \\
&= \frac{-\cos\left(\theta\right)}{\theta^4} + \frac{1}{\theta^4}\left(1 + \frac{\theta^2}{2} - \theta\sin\left(\theta\right)\right) \\
&= \frac{1-\cos\left(\theta\right)}{\theta^4} - \frac{\sin\left(\theta\right)}{\theta^3} + \frac{1}{2\theta^2} .
\end{aligned}$$

Combining everything together we get

$$\frac{\partial V}{\partial v} = \frac{I}{2} + \left(\frac{\sin(\theta)}{\theta^3} - \frac{1}{\theta^2}\cos(\theta)\right) v + \left(\frac{1-\cos(\theta)}{\theta^4} - \frac{\sin(\theta)}{\theta^3} + \frac{1}{2\theta^2}\right) v^2.$$

We can now approximate the observed position

$$y_k \approx P_{k-1} + R_{k-1} V\left(\delta_k\left(\omega_k\right)\right)\delta_k\left(\rho_k + \eta_k\right) + \frac{\partial V}{\partial v}\left(\delta\omega_k\right)\delta_k \nu_k \delta_k\left(\rho_k + \eta_k\right) + v_k.$$

We will make one last simplification to get the final output by assuming that $\delta_k^2 v_k \eta_k$ is negligible relative to everything else.

$$\begin{aligned} y_k &\approx P_{k-1} + R_{k-1} V\left(\delta_k \omega_k\right)\delta_k\left(\rho_k + \eta_k\right) + \frac{\partial V}{\partial v}\left(\delta_k\omega_k\right)\delta_k^2 \nu_k \rho_k + v_k \\ &= P_{k-1} + R_{k-1} V\left(\delta_k \omega_k\right)\delta_k\left(\rho_k + \eta_k\right) - \frac{\partial V}{\partial v}\left(\delta_k\omega_k\right)\delta_k^2 \rho_k^\wedge \nu_k^\vee + v_k \end{aligned}$$

Calculating the mean and covariance under the assumption that $E\left[\nu_k \eta_k^\top\right] = 0$ yields

$$\begin{aligned} E\left[y_k \mid x_{k-1}\right] &= P_{k-1} + R_{k-1} V\left(\delta_k\left(\omega_k\right)\right)\delta_k \rho_k \\ \operatorname{cov}\left(y_k, y_k \mid x_{k-1}\right) &= \delta_k^2 R_{k-1}^\top V\left(\delta_k\omega_k\right) E\left[\eta_k\eta_k^\top\right] V\left(\delta_k\omega_k\right)^\top R_{k-1}^\top + \delta_k^4 \frac{\partial V}{\partial v}\left(\delta_k\omega_k\right)\rho_k^\wedge E\left[\nu_k^\vee\left(\nu_k^\vee\right)^\top\right]\left(\rho_k^\wedge\right)^\top \frac{\partial V}{\partial v}\left(\delta_k\omega_k\right)^\top + E\left[v_k v_k^\top\right. \end{aligned}$$

In order to calculate the LMLE, we need $p\left(y_{k-n} \mid x_k\right)$. This easily follows from the results above.

$$\begin{aligned} E\left[y_{k-n} \mid x_k\right] &= P_k - \Delta_{k-n} R_k V\left(-\Delta_{k-n}\omega\right)\rho \\ \operatorname{cov}\left(y_{k-n}, y_{k-n} \mid x_k\right) &= \Delta_{k-n}^2 R_k V\left(-\Delta_{k-n}\omega_k\right) E\left[\eta_k\eta_k^\top\right] V\left(-\Delta_{k-n}\omega_k\right)^\top R_k^\top + \Delta_{k-n}^4 \frac{\partial V}{\partial v}\left(-\Delta_{k-n}\omega_k\right)\rho_k^\wedge E\left[\nu_k^\vee\left(\nu_k^\vee\right)^\top\right]\left(\rho_k^\wedge\right)^\top \frac{\partial}{\partial} \\ &= \bar{Q} + R. \end{aligned}$$

Once again, we can represent R_k in terms on an element in the Lie algebra so that $R_k = \exp\left(\tau_k\right)$. Thus the LMLE is

$$\arg\min_{P_k, \tau_k, \omega_k, \rho_k} \sum_{n=1}^{k} \left(y_{k-n} - \hat{y}_{k-n}\right)^\top \left(\bar{Q} + R\right)^{-1} \left(y_{k-n} - \hat{y}_{k-n}\right).$$

The question remains under what conditions is the system observable.

A. *Observability*

In order to calculate the observability of the system, we are going to model the system in an alternate, continuous-time, equivalent form. Let the state of the system at time t be

$$x = \begin{bmatrix} \tau \\ P \\ \omega \\ \rho_x \end{bmatrix},$$

where $R = \exp\left(\tau^\wedge\right)$ is the attitude, P is the position, ω is the angular velocity in the body frame, and ρ_x is the translational velocity along the x-axis. We assume that $\rho_y = \rho_z = 0$. The kinematics for $SE\left(3\right)$ is

$$\begin{bmatrix} \dot{R} & \dot{P} \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} \exp\left(\tau^\wedge\right) & P \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \omega^\wedge & \rho \\ 0 & 0 \end{bmatrix},$$

whose solution is

$$\begin{bmatrix} \exp\left(\tau^\wedge\right)\exp\left(\delta\omega^\wedge\right) & P + \delta R V\left(\delta\omega^\wedge\right)\rho \\ 0 & 1 \end{bmatrix}.$$

An equivariant flow on the Lie algebra for the rotation is

$$\log\left(\exp\left(\tau^\wedge\right)\exp\left(\delta\omega^\wedge\right)\right),$$

taking the derivative with respect to time yields

$$\left(J_r^{-1}\left(\tau^\wedge\right)\omega\right)^\wedge.$$

Thus, the state derivative is

$$\dot{x} = \begin{bmatrix} J_r^{-1}\left(\tau^\wedge\right)\omega \\ \exp\left(\tau^\wedge\right)\rho \\ 0 \\ 0 \end{bmatrix}.$$

We are assuming that we can only observe the position of the system, thus

$$y = Cx,$$

where

$$C = \begin{bmatrix} 0_{3\times3} & I_{3\times3} & 0_{3\times3} & 0_{3\times1} \end{bmatrix}.$$

The observability of a nonlinear system can only be determined locally using the Lie derivatives. Let

$$\begin{aligned} \dot{x} &= f\left(x\right), \\ y &= h\left(x\right), \end{aligned}$$

then the observability matrix is calculated as follows:

$$A = \begin{bmatrix} L_f^0 h \\ L_f^1 h \\ L_f^2 h \\ L_f^3 h \\ \vdots \\ L_f^n h \end{bmatrix},$$

and

$$\mathcal{O} = \frac{\partial A}{\partial x}.$$

If the matrix $\mathcal{O}$ is full rank after, then the system is locally observable. We are going to calculate the first four terms of the matrix A

$$A = \begin{bmatrix} P \\ \exp\left(\tau^\wedge\right)\rho \\ G\left(\tau,\rho\right) J_r^{-1}\left(\alpha\right)\omega \\ L_f^3 h \end{bmatrix}.$$

The calculations become very tedious by hand, so we used MATLAB to calculate $L_f^3 h$ and the matrix $\mathcal{O}$ for us. We found that the system is observable under the conditions that $\rho_x \neq 0$, and ω_y or ω_z are non zero.

B. Seeding

When seeding the algorithm , we are going to ignore the probability aspect of the system. We also assume that the measurement y_k is available, that the body velocity is $\rho = \left[x, 0, 0\right]^\top$ with x being positive, and that we have at least three distinct measurements. Under these assumptions, the system will either move in a straight line if the angular velocity is 0 or orbit if it is non zero.

$$\begin{aligned} z_1 &= \left(y_k - y_{k-1}\right)/\Delta_{k-1} = R_k V\left(\Delta_{k-1}\omega\right)\rho \\ z_2 &= \left(y_k - y_{k-2}\right)/\Delta_{k-2} = R_k V\left(\Delta_{k-2}\omega\right)\rho \\ z_3 &= \left(y_k - y_{k-3}\right)/\Delta_{k-3} = R_k V\left(\Delta_{k-3}\omega\right)\rho \end{aligned}$$

VI. Derivative

Let

$$v = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}, \quad p = \begin{bmatrix} p_1 \\ p_2 \\ p_3 \end{bmatrix}, \quad \theta = \sqrt{v^\top v}.$$

Consider the Rodriquez formula for the exponential matrix

$$\exp\left(v^\wedge\right) = I + \frac{\sin(\theta)}{\theta} v^\wedge + \frac{1-\cos(\theta)}{\theta^2}\left(v^\wedge\right)^2,$$

then

$$\frac{d\exp\left(v^\wedge\right)p}{dv} = \alpha(\theta)(-p)^\wedge + \beta(\theta) v^\wedge p v^\top + \gamma(\theta) G + \delta(\theta)\left(v^\wedge\right)^2 p v^\top,$$

with

$$\begin{aligned}
\alpha(\theta) &= \frac{\sin(\theta)}{\theta}, \\
\beta(\theta) &= \frac{\cos(\theta)}{\theta^2} - \frac{\sin(\theta)}{\theta^3}, \\
\gamma(\theta) &= \frac{1-\cos(\theta)}{\theta^2}, \\
\delta(\theta) &= \frac{\sin(\theta)}{\theta^3} - 2\frac{1-\cos(\theta)}{\theta^4} \\
G &= \begin{bmatrix} v_2p_2 + v_3p_3 & -2v_2p_1 + v_1p_2 & -2v_3p_1 + v_1p_3 \\ v_2p_1 - 2v_1p_2 & v_1p_1 + v_3p_3 & -2v_3p_2 + v_2p_3 \\ v_3p_1 - 2v_1p_3 & v_3p_2 - 2v_2p_3 & v_1p_1 + v_2p_2 \end{bmatrix}
\end{aligned}$$

VII. SE(2) Observability

Assume the state

$$x = \begin{bmatrix} \alpha \\ P \\ \omega \\ \rho \end{bmatrix},$$

where $R = \exp\left(\alpha^\wedge\right)$, P is the position, ω is the angular velocity and ρ is the body x velocity. The derivative is

$$\dot{x} = \begin{bmatrix} \omega \\ \cos(\alpha)\rho \\ \sin(\alpha)\rho \\ 0 \\ 0 \end{bmatrix}.$$

Then the observability matrix is

$$O = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ -\sin(\alpha)\rho & 0 & 0 & 0 & \cos(\alpha) \\ \cos(\alpha)\rho & 0 & 0 & 0 & \sin(\alpha) \\ -\cos(\alpha)\rho\omega & 0 & 0 & -\sin(\alpha)\rho & -\sin(\alpha)\omega \\ \sin(\alpha)\rho\omega & 0 & 0 & -\cos(\alpha)\rho & -\cos(\alpha)\omega \end{bmatrix},$$

which is full rank provided that $\rho, \omega \neq 0$. Thus the system is locally observable provided $\rho, \omega \neq 0$.

VIII. Formulas

$$x_{k+1} = f\left(x_k, u\right) + w_k$$
$$y_{k+1} = h\left(x_{k+1}\right) + v_k$$

$$\arg\max_{x_0, u} p\left(x_0\right) \prod_{j=0}^{m} p\left(y_j \mid x_j\right)$$