State Estimation - Assignment 7

Haniyeh Altafi

July 7, 2023

Identities

$$[\mathbf{a}]_\times \mathbf{b} = -[\mathbf{b}]_\times \mathbf{a}$$
$$[\mathbf{A}\mathbf{b}]_\times = \mathbf{A}[\mathbf{b}]_\times \mathbf{A}^\top$$

Q1: Left and Right derivatives on Lie Groups.
For the following function

$$f : SO(3) \to \mathbb{R}^3; \quad f(\mathbf{R}, \mathbf{p}) = \mathbf{R}\mathbf{p}$$

calculate the left and right derivatives of f with respect to $\mathbf{R}$, by applying the definitions of the left and right derivatives:

$$\frac{{}^R D f(\mathbf{R}, \mathbf{p})}{D\mathbf{R}}$$
$$\frac{{}^L D f(\mathbf{R}, \mathbf{p})}{D\mathbf{R}}$$

Q2: Jacobian.
Given a Lie group $\mathcal{M}$ with a composition operation $\circ$, and elements $\mathcal{X}, \mathcal{Y} \in \mathcal{M}$, calculate the derivative of $\mathcal{X} \circ \mathcal{Y}$ with respect to $\mathcal{Y}$, i.e.

$$\frac{{}^{\mathcal{Y}} D \mathcal{X} \circ \mathcal{Y}}{D\mathcal{Y}}$$

Q3: Adjoint Matrix Properties.
Given the Lie group of $M = SE(3)$ with the composition operation $\circ$, and elements $\mathcal{X}, \mathcal{Y} \in M$, show that

$$\mathbf{Ad}_{\mathcal{X}} \mathbf{Ad}_{\mathcal{Y}} = \mathbf{Ad}_{\mathcal{X} \circ \mathcal{Y}}$$

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

Answer:

Q1

For the function $f : \mathrm{SO}(3) \to \mathbb{R}^3$ defined as $f(R, p) = Rp$, we want to calculate the left and right derivatives of f with respect to R.

The left derivative of f with respect to R is denoted as $\frac{{}^L Df(\mathbf{R},\mathbf{p})}{D\mathbf{R}}$ and is defined as:

$$\frac{{}^L Df(\mathbf{R},\mathbf{p})}{D\mathbf{R}} = \frac{{}^L D\mathbf{R}\mathbf{p}}{D\mathbf{R}} = \lim_{\boldsymbol{\theta}\to 0} \frac{(\theta \oplus \boldsymbol{R})\mathbf{p} \ominus \mathbf{R}\mathbf{p}}{\boldsymbol{\theta}} = \lim_{\boldsymbol{\theta}\to 0} \frac{\mathrm{Exp}(\boldsymbol{\theta})\mathbf{R}\mathbf{p} - \mathbf{R}\mathbf{p}}{\boldsymbol{\theta}}$$

Expanding the first term using the distributive property of matrix multiplication, and by using the second Identity provided, we get:

$$\lim_{\boldsymbol{\theta}\to 0} \frac{(\mathbf{I} + [\boldsymbol{\theta}]_\times)\,\mathbf{R}\mathbf{p} - \mathbf{R}\mathbf{p}}{\boldsymbol{\theta}} = \lim_{\boldsymbol{\theta}\to 0} \frac{[\boldsymbol{\theta}]_\times \mathbf{R}\mathbf{p}}{\boldsymbol{\theta}} = \lim_{\boldsymbol{\theta}\to 0} \frac{-\boldsymbol{\theta}[\mathbf{R}\mathbf{p}]_\times}{\boldsymbol{\theta}} = -[\mathbf{R}\mathbf{p}]_\times = -\mathbf{R}[\mathbf{p}]_\times \mathbf{R}^\top \in \mathbb{R}^{3\times 3}$$

Therefore, the left derivative of f with respect to R is **-**$\mathbf{R}[\mathbf{p}]_\times \mathbf{R}^\top$.

Similarly, the right derivative of f with respect to R is denoted as $\frac{{}^R Df(\mathbf{R},\mathbf{p})}{D\mathbf{R}}$ and is defined as:

$$\frac{{}^R Df(\mathbf{R},\mathbf{p})}{D\mathbf{R}} = \frac{{}^R D\mathbf{R}\mathbf{p}}{D\mathbf{R}} = \lim_{\boldsymbol{\theta}\to 0} \frac{(\mathbf{R} \oplus \boldsymbol{\theta})\mathbf{p} \ominus \mathbf{R}\mathbf{p}}{\boldsymbol{\theta}} = \lim_{\boldsymbol{\theta}\to 0} \frac{\mathbf{R}\,\mathrm{Exp}(\boldsymbol{\theta})\mathbf{p} - \mathbf{R}\mathbf{p}}{\boldsymbol{\theta}}$$

Expanding the first term using the distributive property of matrix multiplication, we get:

$$\lim_{\boldsymbol{\theta}\to 0} \frac{\mathbf{R}\,(\mathbf{I} + [\boldsymbol{\theta}]_\times)\,\mathbf{p} - \mathbf{R}\mathbf{p}}{\boldsymbol{\theta}} = \lim_{\boldsymbol{\theta}\to 0} \frac{\mathbf{R}[\boldsymbol{\theta}]_\times \mathbf{p}}{\boldsymbol{\theta}} = \lim_{\boldsymbol{\theta}\to 0} \frac{-\mathbf{R}[\mathbf{p}]_\times \boldsymbol{\theta}}{\boldsymbol{\theta}} = -\mathbf{R}[\mathbf{p}]_\times \in \mathbb{R}^{3\times 3}$$

Therefore, the right derivative of f with respect to R is **-**$\mathbf{R}[\mathbf{p}]_\times$.

So, the Left and Right derivatives are not the same.

Q2

We have this equation in the paper[1].

$$\mathbf{J}_{\mathcal{Y}}^{\mathcal{X}\circ\mathcal{Y}} \triangleq \frac{{}^{\mathcal{Y}}D\mathcal{X}\circ\mathcal{Y}}{D\mathcal{Y}} \in \mathbb{R}^{m\times m}$$

We can prove that :

$$\mathbf{J}_{\mathcal{Y}}^{\mathcal{X}\circ\mathcal{Y}} = \frac{{}^{\mathcal{Y}}D\mathcal{X}\circ\mathcal{Y}}{D\mathcal{Y}} = \lim_{\boldsymbol{\tau}\to 0} \frac{\mathcal{X}\circ(\mathcal{Y}\oplus\boldsymbol{\tau})\ominus\mathcal{X}\circ\mathcal{Y}}{\boldsymbol{\tau}}$$

$$\lim_{\boldsymbol{\tau}\to 0} \frac{\log\left[(\mathcal{X}\mathcal{Y})^{-1}(\mathcal{Y}\,\mathrm{Exp}(\boldsymbol{\tau})\mathcal{X})\right]}{\boldsymbol{\tau}} = \lim_{\boldsymbol{\tau}\to 0} \frac{\log\left[\mathcal{X}^{-1}\,\mathrm{Exp}(\boldsymbol{\tau})\mathcal{X}\right]}{\boldsymbol{\tau}}$$

$$= \lim_{\boldsymbol{\tau}\to 0} \frac{\log\left[\mathcal{X}^{-1}\mathcal{X}\mathcal{X}\right]}{\boldsymbol{\tau}} = \lim_{\boldsymbol{\tau}\to 0} \frac{\log[\mathcal{X}]}{\boldsymbol{\tau}} = \lim_{\boldsymbol{\tau}\to 0} \frac{\boldsymbol{\tau}}{\boldsymbol{\tau}} = \mathbf{I}$$

Q3

part a:

In this question, First, we are calculating the left side of the equation :

$$\mathbf{Ad}_{\mathcal{X}\circ\mathcal{Y}} = \left(\mathbf{M}_{\mathcal{X}\circ\mathcal{Y}}\boldsymbol{\tau}^{\wedge}\mathbf{M}_{\mathcal{X}\circ\mathcal{Y}}^{-1}\right)^{\vee}$$

$$\mathbf{M}_{\mathcal{X}\circ\mathcal{Y}} = \mathbf{M}_{\mathcal{X}}\mathbf{M}_{\mathcal{Y}} = \begin{bmatrix} \mathbf{R}_{\mathcal{X}}\mathbf{R}_{\mathcal{Y}} & \mathbf{t}_{\mathcal{X}} + \mathbf{t}_{\mathcal{Y}}\mathbf{R}_{\mathcal{X}} \\ \mathbf{0} & \mathbf{1} \end{bmatrix}$$

$$\mathbf{M}_{\mathcal{X}\circ\mathcal{Y}}^{-1} = \mathbf{M}_{\mathcal{X}}^{-1}\mathbf{M}_{\mathcal{Y}}{}^{-1} = \begin{bmatrix} \mathbf{R}_{\mathcal{X}}{}^{\top} & -\mathbf{R}_{\mathcal{X}}{}^{\top}\mathbf{t}_{\mathcal{X}} \\ \mathbf{0} & \mathbf{1} \end{bmatrix}\begin{bmatrix} \mathbf{R}_{\mathcal{Y}}{}^{\top} & -\mathbf{R}_{\mathcal{Y}}{}^{\top}\mathbf{t}_{\mathcal{Y}} \\ \mathbf{0} & \mathbf{1} \end{bmatrix}$$
$$= \begin{bmatrix} \mathbf{R}_{\mathcal{X}}{}^{\top}\mathbf{R}_{\mathcal{Y}}{}^{\top} & -\mathbf{R}_{\mathcal{X}}{}^{\top}\mathbf{R}_{\mathcal{Y}}{}^{\top}\mathbf{t}_{\mathcal{Y}} - \mathbf{R}_{\mathcal{X}}{}^{\top}\mathbf{t}_{\mathcal{X}} \\ \mathbf{0} & \mathbf{1} \end{bmatrix}$$

$$\mathbf{Ad}_{\mathcal{X}\circ\mathcal{Y}} = \left(\begin{bmatrix} \mathbf{R}_{\mathcal{X}}\mathbf{R}_{\mathcal{Y}} & \mathbf{t}_{\mathcal{X}} + \mathbf{t}_{\mathcal{Y}}\mathbf{R}_{\mathcal{X}} \\ \mathbf{0} & \mathbf{1} \end{bmatrix}\begin{bmatrix} [\boldsymbol{\theta}]_{\times} & \boldsymbol{\rho} \\ \mathbf{0} & 0 \end{bmatrix}\begin{bmatrix} \mathbf{R}_{\mathcal{X}}{}^{\top}\mathbf{R}_{\mathcal{Y}}{}^{\top} & -\mathbf{R}_{\mathcal{X}}{}^{\top}\mathbf{R}_{\mathcal{Y}}{}^{\top}\mathbf{t}_{\mathcal{Y}} - \mathbf{R}_{\mathcal{X}}{}^{\top}\mathbf{t}_{\mathcal{X}} \\ \mathbf{0} & \mathbf{1} \end{bmatrix}\right)^{\vee}$$

$$= \left(\begin{bmatrix} \mathbf{R}_{\mathcal{X}}\mathbf{R}_{\mathcal{Y}}[\boldsymbol{\theta}]_{\times}\mathbf{R}_{\mathcal{X}}{}^{\top}\mathbf{R}_{\mathcal{Y}}{}^{\top} & -\mathbf{R}_{\mathcal{X}}\mathbf{R}_{\mathcal{Y}}[\boldsymbol{\theta}]_{\times}\mathbf{R}_{\mathcal{X}}{}^{\top}\mathbf{R}_{\mathcal{Y}}{}^{\top}\mathbf{t}_{\mathcal{Y}} - \mathbf{R}_{\mathcal{X}}\mathbf{R}_{\mathcal{Y}}[\boldsymbol{\theta}]_{\times}\mathbf{R}_{\mathcal{X}}{}^{\top}\mathbf{t}_{\mathcal{X}} + \boldsymbol{\rho}\mathbf{R}_{\mathcal{X}}\mathbf{R}_{\mathcal{Y}} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\right)^{\vee}$$

$$= \left(\begin{bmatrix} [\mathbf{R}_{\mathcal{X}}\mathbf{R}_{\mathcal{Y}}\boldsymbol{\theta}]_{\times} & -[\mathbf{R}_{\mathcal{X}}\mathbf{R}_{\mathcal{Y}}\boldsymbol{\theta}]_{\times}\mathbf{t}_{\mathcal{Y}} - \mathbf{R}_{\mathcal{Y}}[\mathbf{R}_{\mathcal{X}}\boldsymbol{\theta}]_{\times}\mathbf{t}_{\mathcal{X}} + \boldsymbol{\rho}\mathbf{R}_{\mathcal{X}}\mathbf{R}_{\mathcal{Y}} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\right)^{\vee}$$
$$= \left(\begin{bmatrix} [\mathbf{R}_{\mathcal{X}}\mathbf{R}_{\mathcal{Y}}\boldsymbol{\theta}]_{\times} & +\mathbf{R}_{\mathcal{X}}\mathbf{R}_{\mathcal{Y}}\boldsymbol{\theta}[\mathbf{t}_{\mathcal{Y}}]_{\times} + \mathbf{R}_{\mathcal{Y}}\mathbf{R}_{\mathcal{X}}\boldsymbol{\theta}[\mathbf{t}_{\mathcal{X}}]_{\times} + \boldsymbol{\rho}\mathbf{R}_{\mathcal{X}}\mathbf{R}_{\mathcal{Y}} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\right)^{\vee}$$

$$= \begin{bmatrix} [\mathbf{t}_{\mathcal{Y}}]_{\times}\mathbf{R}_{\mathcal{X}}\mathbf{R}_{\mathcal{Y}}\boldsymbol{\theta} + [\mathbf{t}_{\mathcal{X}}]_{\times}\mathbf{R}_{\mathcal{X}}\mathbf{R}_{\mathcal{Y}}\boldsymbol{\theta} + \mathbf{R}_{\mathcal{X}}\mathbf{R}_{\mathcal{Y}}\boldsymbol{\rho} \\ \mathbf{R}_{\mathcal{X}}\mathbf{R}_{\mathcal{Y}}\boldsymbol{\theta} \end{bmatrix} = \begin{bmatrix} \mathbf{R}_{\mathcal{X}}\mathbf{R}_{\mathcal{Y}} & [\mathbf{t}_{\mathcal{Y}}]_{\times}\mathbf{R}_{\mathcal{X}}\mathbf{R}_{\mathcal{Y}} + [\mathbf{t}_{\mathcal{X}}]_{\times}\mathbf{R}_{\mathcal{X}}\mathbf{R}_{\mathcal{Y}} \\ \mathbf{0} & \mathbf{R}_{\mathcal{X}}\mathbf{R}_{\mathcal{Y}} \end{bmatrix}\begin{bmatrix} \boldsymbol{\rho} \\ \boldsymbol{\theta} \end{bmatrix}$$

where we used $[\mathbf{R}\boldsymbol{\theta}]_{\times} = \mathbf{R}[\boldsymbol{\theta}]_{\times}\mathbf{R}^{\top}$ and $[\mathbf{a}]_{\times}\mathbf{b} = -[\mathbf{b}]_{\times}\mathrm{a}$. So the adjoint matrix is

$$\begin{bmatrix} \mathbf{R}_{\mathcal{X}}\mathbf{R}_{\mathcal{Y}} & [\mathbf{t}_{\mathcal{Y}}]_{\times}\mathbf{R}_{\mathcal{X}}\mathbf{R}_{\mathcal{Y}} + [\mathbf{t}_{\mathcal{X}}]_{\times}\mathbf{R}_{\mathcal{X}}\mathbf{R}_{\mathcal{Y}} \\ \mathbf{0} & \mathbf{R}_{\mathcal{X}}\mathbf{R}_{\mathcal{Y}} \end{bmatrix}$$

Here, we are calculating the other side of the equation.

$$\mathbf{Ad}_{\mathcal{X}}\mathbf{Ad}_{\mathcal{Y}} = \begin{bmatrix} \mathbf{R}_{\mathcal{X}} & [\mathbf{t}_{\mathcal{X}}]_{\times}\mathbf{R}_{\mathcal{X}} \\ 0 & \mathbf{R}_{\mathcal{X}} \end{bmatrix}\begin{bmatrix} \mathbf{R}_{\mathcal{Y}} & [\mathbf{t}_{\mathcal{Y}}]_{\times}\mathbf{R}_{\mathcal{Y}} \\ 0 & \mathbf{R}_{\mathcal{Y}} \end{bmatrix} = \begin{bmatrix} \mathbf{R}_{\mathcal{X}}\mathbf{R}_{\mathcal{Y}} & [\mathbf{t}_{\mathcal{Y}}]_{\times}\mathbf{R}_{\mathcal{X}}\mathbf{R}_{\mathcal{Y}} + [\mathbf{t}_{\mathcal{X}}]_{\times}\mathbf{R}_{\mathcal{X}}\mathbf{R}_{\mathcal{Y}} \\ \mathbf{0} & \mathbf{R}_{\mathcal{X}}\mathbf{R}_{\mathcal{Y}} \end{bmatrix}$$

So, we have proved the equation below:

$$\mathbf{Ad}_{\mathcal{X}}\mathbf{Ad}_{\mathcal{Y}} = \mathbf{Ad}_{\mathcal{X}\circ\mathcal{Y}}$$

part b:

From the definition, we have :

$$\mathbf{Ad}_{\mathcal{X}} = \left[\mathcal{X}\tau^{\wedge}\mathcal{X}^{-1}\right]^{\vee}$$

$$\mathbf{Ad}_{\mathcal{X}}^{-1} = \left[\left[\mathcal{X}\tau^{\wedge}\mathcal{X}^{-1}\right]^{\vee}\right]^{-1} = \left[\left[\mathcal{X}\tau^{\wedge}\mathcal{X}^{-1}\right]^{-1}\right]^{\vee} = \left[\mathcal{X}^{-1}\tau^{\wedge}\mathcal{X}\right]^{\vee}$$

From the definition, we have :

$$\mathbf{Ad}_{\mathcal{X}^{-1}} = \left[\mathcal{X}^{-1}\tau^{\wedge}\mathcal{X}\right]^{\vee}$$

So, we have proved that this equation is true:

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

References

[1] Joan Sola, Jeremie Deray, and Dinesh Atchuthan. A micro lie theory for state estimation in robotics. *arXiv preprint arXiv:1812.01537*, 2018.