Supplementary Material : Rauch-Tung-Striebel Smoother on Lie groups

In this supplementary material, we provide the mathematical derivations of a generalization of the Rauch-Tung-Striebel smoother (RTS), also known as Extended Kalman Smoother, to the case where the state evolves on a matrix Lie group. Our solution relies on the link between the Gauss-Newton algorithm on Lie groups and the formalism of the concentrated Gaussian distribution on Lie groups. Our formalism yields closed-form equations for the smoothing of parameters produced by the Extended Kalman Filter on Lie groups (LG-EKF).

Please read section 4 in the submitted paper before reading this note.

I. Useful property

$$log_G^{\vee}\left(exp_G^{\wedge}\left(a\right)exp_G^{\wedge}\left(b\right)\right)=b+\varphi_G\left(b\right)a+O\left(\left\|a\right\|^2\right) \tag{1}$$

where

$$\varphi_G\left(b\right)=\sum_{n=0}^{\infty}\frac{B_n ad_G\left(b\right)^n}{n!}=Id_p+\frac{1}{2}ad_G\left(b\right)+\cdots \tag{2}$$

and $ad_G\left(b\right)a=\left[\left[b\right]_G^{\wedge}\left[a\right]_G^{\wedge}-\left[a\right]_G^{\wedge}\left[b\right]_G^{\wedge}\right]_G^{\vee}$. The B_n are the Bernoulli numbers. $\varphi_G\left(\cdot\right)$ is the inverse of the left Jacobian of G that we denote $\Phi_G\left(\cdot\right)$ (see supplementary material on LG-RBPS).

II. Derivation of the LG-RTS

A. Problem

$$\begin{aligned}p\left(x_{t+1},x_t|y_{1:T}\right)&=p\left(x_t|x_{t+1},y_{1:T}\right)p\left(x_{t+1}|y_{1:T}\right)\\&=p\left(x_t|x_{t+1},y_{1:t}\right)p\left(x_{t+1}|y_{1:T}\right)\\&=\frac{p\left(x_{t+1},x_t|y_{1:t}\right)}{p\left(x_{t+1}|y_{1:t}\right)}p\left(x_{t+1}|y_{1:T}\right)\end{aligned} \tag{3}$$

where

$$p\left(x_{t+1},x_t|y_{1:t}\right)=\mathcal{N}_G\left(\left[\begin{array}{c}x_{t+1}\\x_t\end{array}\right];\left[\begin{array}{c}\mu_{t+1|t}\\\mu_{t|t}\end{array}\right],\right. \tag{4}$$

$$\left.\Sigma=\left[\begin{array}{cc}Q_t+F_tP_{t|t}F_t^T & F_tP_{t|t}\\P_{t|t}F_t^T & P_{t|t}\end{array}\right]\right)$$

$$p\left(x_{t+1}|y_{1:t}\right)=\mathcal{N}_G\left(x_{t+1};\mu_{t+1|t},P_{t+1|t}=Q_t+F_tP_{t|t}F_t^T\right) \tag{5}$$

$$p\left(x_{t+1}|y_{1:T}\right)=\mathcal{N}_G\left(x_{t+1};\mu_{t+1|T},P_{t+1|T}\right) \tag{6}$$

In equations (4) and (5), the values of the parameters come from the propagation step of the LG-EKF while in eq.(6) they come from the LG-RTS smoother at time $t+1$.

Under the concentrated Gaussian assumption, the logarithm of (3) is :

$$l\left(x_{t+1},x_t\right)=\left\|\left[\begin{array}{c}log_G^{\vee}\left(x_{t+1}\mu_{t+1|t}^{-1}\right)\\log_G^{\vee}\left(x_t\mu_{t|t}^{-1}\right)\\log_G^{\vee}\left(x_{t+1}\mu_{t+1|t}^{-1}\right)\\log_G^{\vee}\left(x_{t+1}\mu_{t+1|T}^{-1}\right)\end{array}\right]\right\|_{\mathfrak{C}}^2 \tag{7}$$

where

$$\mathfrak{E} = \mathrm{diag}\left(\Sigma, -\left(Q_t + F_t P_{t|t} F_t^T\right), P_{t+1|T}\right) \tag{8}$$

and $\|\cdot\|_{\cdot}^2$ is the Mahalanobis distance.

We wish to find the minimum of (7), i.e we wish to estimate $\{\hat{x}_{t+1}, \hat{x}_t\} = \underset{x_{t+1}, x_t}{\mathrm{argmin}} \{l(x_{t+1}, x_t)\}$. To do so, we propose to apply a Gauss-Newton algorithm.

B. Application of a Gauss-Newton algorithm

First, we linearise the term inside the norm of (7) in $\delta_{t+1}^{l+1/l} = 0$ and $\delta_t^{l+1/l} = 0$ where we defined $x_t = exp_G^{\wedge}\left(\delta_t^{l+1/l}\right) x_t^{(l)}$ and $x_{t+1} = exp_G^{\wedge}\left(\delta_{t+1}^{l+1/l}\right) x_{t+1}^{(l)}$. To do so, we define :

$$\delta_{t+1}^{l,t} = log_G^{\vee}\left(x_{t+1}^{(l)} \mu_{t+1|t}^{-1}\right) \tag{9}$$

$$\delta_t^{l,t} = log_G^{\vee}\left(x_t^{(l)} \mu_{t|t}^{-1}\right) \tag{10}$$

$$\delta_{t+1}^{l,T} = log_G^{\vee}\left(x_{t+1}^{(l)} \mu_{t+1|T}^{-1}\right) \tag{11}$$

Using (1) we obtain a new problem to minimize :

$$\underset{\delta_{t+1}^{l+1/l}, \delta_t^{l+1/l}}{\mathrm{argmin}} \left\{ \left\| e_l + J_l \begin{bmatrix} \delta_{t+1}^{l+1/l} \\ \delta_t^{l+1/l} \end{bmatrix} \right\|_{\mathfrak{E}}^2 \right\} \tag{12}$$

where

$$J_l = \begin{bmatrix} \varphi_G\left(\delta_{t+1}^{l,t}\right) & 0 \\ 0 & \varphi_G\left(\delta_t^{l,t}\right) \\ \varphi_G\left(\delta_{t+1}^{l,t}\right) & 0 \\ \varphi_G\left(\delta_{t+1}^{l,T}\right) & 0 \end{bmatrix} \text{ and } e_l = \begin{bmatrix} \delta_{t+1}^{l,t} \\ \delta_t^{l,t} \\ \delta_{t+1}^{l,t} \\ \delta_{t+1}^{l,T} \end{bmatrix} \tag{13}$$

We also define : $\varphi_G\left(\delta_{t+1}^{l,t}\right) = \mathcal{M}$, $\varphi_G\left(\delta_t^{l,t}\right) = \mathcal{S}$ and $\mathcal{L} = \varphi_G\left(\delta_{t+1}^{l,T}\right)$.

Then, the solution of (12) is :

$$\begin{bmatrix} \delta_{t+1}^{l+1/l} \\ \delta_t^{l+1/l} \end{bmatrix} = -\left(J_l^T \mathfrak{E}^{-1} J_l\right)^{-1} J_l^T \mathfrak{E}^{-1} \begin{bmatrix} \delta_{t+1}^{l,t} \\ \delta_t^{l,t} \\ \delta_{t+1}^{l,t} \\ \delta_{t+1}^{l,T} \end{bmatrix} \tag{14}$$

where

$$\mathfrak{E}^{-1} = \mathrm{diag}\left(\begin{bmatrix} Q_t^{-1} & -Q_t^{-1} F_t \\ -F_t^T Q_t^{-1} & \begin{matrix} F_t^T Q_t^{-1} F_t \\ +P_{t|t}^{-1} \end{matrix} \end{bmatrix}, -\begin{pmatrix} F_t P_{t|t} F_t^T \\ +Q_t \end{pmatrix}^{-1}, P_{t+1|T}^{-1}\right) \tag{15}$$

1) Derivation of $(J_l^T \mathfrak{E}^{-1} J_l)^{-1}$:

$$\left(J_l^T \mathfrak{E}^{-1} J_l\right)^{-1} = \begin{bmatrix} A' & B' \\ B'^T & D' \end{bmatrix} \tag{16}$$

It is possible to prove that :

$$A' = \mathcal{L}^{-1} P_{t+1|T} \mathcal{L}^{-T} \tag{17}$$

$$D' = \mathcal{S}^{-1}\left(P_{t|t} + L_t\left(\mathcal{M}\mathcal{L}^{-1} P_{t+1|T} \mathcal{L}^{-T} \mathcal{M}^T - P_{t+1|t}\right) L_t^T\right) \mathcal{S}^{-T} \tag{18}$$

and

$$B' = \mathcal{L}^{-1} P_{t+1|T} \mathcal{L}^{-T} \mathcal{M}^T L_t^T \mathcal{S}^{-T} \tag{19}$$

where $L_t = P_{t|t} F_t^T \left(Q_t + F_t P_{t|t} F_t^T\right)^{-1}$.

2) Derivation of $\delta_{t+1}^{l+1/l}$: If $x_{t+1}^{(l)} = \mu_{t+1|T}$ then from (11), $\delta_{t+1}^{l,T} = 0$ and consequently $\mathcal{L} = Id$. From (14) and using (13), (15) and (16) we obtain :

$$\begin{aligned}\delta_{t+1}^{l+1/l} &= -P_{t+1|T}\mathcal{M}^T \\ &\left(\left(\left(Q_t^{-1} - \left(Q_t + F_t P_{t|t} F_t^T\right)^{-1}\right) - L_t^T F_t^T Q_t^{-1}\right)\delta_{t+1}^{l,t}\right. \\ &\left.+\left(-Q_t^{-1}F_t + L_t^T\left(F_t^T Q_t^{-1} F_t + P_{t|t}^{-1}\right)\right)\delta_t^{l,t}\right)\end{aligned} \tag{20}$$

It is possible to prove that the term in $\delta_{t+1}^{l,t}$ is null as well as the term in $\delta_t^{l,t}$. Thus, by initializing $x_{t+1}^{(0)} = \mu_{t+1|T}$, we proved :

$$\delta_{t+1}^{l+1/l} = 0 \tag{21}$$

for any l. Consequently : $\delta_{t+1}^{l,T} = 0$ and $\mathcal{L} = Id$.

3) Derivation of $\delta_t^{l+1/l}$: From (14) and using (13), (15) and (16) as well as $\delta_{t+1}^{l,T} = 0$ and $\mathcal{L} = Id$, we obtain :

$$\begin{aligned}\delta_t^{l+1/l} = -\mathcal{S}^{-1}\Big\{&\Big(L_t\mathcal{M}P_{t+1|T}\mathcal{M}^T\Big(Q_t^{-1} - \left(Q_t + F_t P_{t|t}F_t^T\right)^{-1}\Big) \\ &-\left(P_{t|t} + L_t\left(\mathcal{M}P_{t+1|T}\mathcal{M}^T - P_{t+1|t}\right)L_t^T\right)F_t^T Q_t^{-1}\Big)\delta_{t+1}^{l,t} \\ &+\Big(\Big(P_{t|t} + L_t\left(\mathcal{M}P_{t+1|T}\mathcal{M}^T - P_{t+1|t}\right)L_t^T\Big)\Big(F_t^T Q_t^{-1}F_t + P_{t|t}^{-1}\Big) \\ &-L_t\mathcal{M}P_{t+1|T}\mathcal{M}^T Q_t^{-1}F_t\Big)\delta_t^{l,t}\Big\}\end{aligned} \tag{22}$$

It is possible to prove that the term in $\delta_t^{l,t}$ equals $-\mathcal{S}^{-1}\delta_t^{l,t}$ while the term in $\delta_{t+1}^{l,t}$ equals $\mathcal{S}^{-1}L_t\delta_{t+1}^{l,t}$. Thus we have :

$$\delta_t^{l+1/l} = \mathcal{S}^{-1}\left(L_t\delta_{t+1}^{l,t} - \delta_t^{l,t}\right) \tag{23}$$

C. Algorithm

By initializing $x_{t+1}^{(0)} = \mu_{t+1|T}$, we obtained (eq.(21) and eq.(23)) :

$$\begin{bmatrix} \delta_{t+1}^{l+1/l} \\ \delta_t^{l+1/l} \end{bmatrix} = \begin{bmatrix} 0 \\ \mathcal{S}^{-1}\left(L_t\delta_{t+1}^{l,t} - \delta_t^{l,t}\right) \end{bmatrix} \tag{24}$$

Thus :

$$\delta_t^{l+1,t} = log_G^{\vee}\left(x_t^{(l+1)}\mu_{t|t}^{-1}\right) \simeq \delta_t^{l,t} + \mathcal{S}\delta_t^{l+1/l} = L_t\delta_{t+1}^{l,t} \tag{25}$$

Consequently, we have the following update equation :

$$\begin{aligned}x_t^{(l+1)} &= exp_G^{\wedge}\left(\delta_t^{l+1,t}\right)\mu_{t|t} = exp_G^{\wedge}\left(L_t\delta_{t+1}^{l,t}\right)\mu_{t|t} \\ &= exp_G^{\wedge}\left(L_t log_G^{\vee}\left(\mu_{t+1|T}\mu_{t+1|t}^{-1}\right)\right)\mu_{t|t}\end{aligned} \tag{26}$$

The solution of our problem is reached after a single iteration. Thus : $\hat{x}_t = exp_G^{\wedge}\left(L_t log_G^{\vee}\left(\mu_{t+1|T}\mu_{t+1|t}^{-1}\right)\right)\mu_{t|t}$.

Then we set $\mu_{t|T} = \hat{x}_t$ and approximate the covariance matrix with D' :

$$P_{t|T} = \mathcal{S}^{-1}\left(P_{t|t} + L_t\left(\mathcal{M}P_{t+1|T}\mathcal{M}^T - P_{t+1|t}\right)L_t^T\right)\mathcal{S}^{-T} \tag{27}$$

Remark : $\hat{x}_{t+1} = \mu_{t+1|T}$ consequently A' equals $P_{t+1|T}$ as expected.