Derivations for GFX node

Wouter Kouw

November 23, 2020

1 Problem

Suppose we have the following linear second-order differential equation:

$$x''(t) + \theta_2 x'(t) + \theta_1 x(t) = \eta u(t) + w(t) \tag{1}$$

Reduce this to a multivariate first-order differential equation with the following substitutions:

$$\begin{aligned} z_1(t) &= x(t) && (2a) \\ z_2(t) &= x'(t)\,, && (2b) \end{aligned}$$

which produces:

$$\begin{aligned} z_1'(t) &= z_2(t) && (3a) \\ z_2'(t) &= -\theta_2 z_2(t) - \theta_1 z_1(t) + \eta u(t) + w(t)\,. && (3b) \end{aligned}$$

We can re-write this into a matrix form:

$$\begin{bmatrix} z_1'(t) \\ z_2'(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\theta_1 & -\theta_2 \end{bmatrix} \begin{bmatrix} z_1(t) \\ z_2(t) \end{bmatrix} + \begin{bmatrix} 0 \\ \eta \end{bmatrix} u(t) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} w(t)\,, \tag{4}$$

where $z_t = \begin{bmatrix} z_1(t) & z_2(t) \end{bmatrix}^\top$.

1.1 Discretization

We can perform an approximate discretization using Euler-Maruyama, at non-overlapping time points t_k and t_{k+1}. The Wiener process $w(t)$ becomes:

$$w(t) \approx \frac{\beta_{t_{k+1}} - \beta_{t_k}}{\Delta t_k}\,, \tag{5}$$

where $\Delta t_k = t_{k+1} - t_k$. I rename the increment $\beta_{t_{k+1}} - \beta_{t_k}$ to be w_t, which, if the time-points t_{k+1} and t_k do not overlap, follows a Gaussian distribution (Def. 4.1 [1]):

$$w_t \sim \mathcal{N}(0, \tau^{-1} \Delta t_k)\,. \tag{6}$$

Henceforth, we will assume all Δt_k are equal and use subscripts $t+1$ and t. The states $z'(t)$ are approximated as $(z_{t+1} - z_t)/\Delta t$. The control signal $u(t)$ is observed and directly maps to u_t. Using this discretization, we get:

$$\left(\begin{bmatrix} z_{1,t+1} \\ z_{2,t+1} \end{bmatrix} - \begin{bmatrix} z_{1,t} \\ z_{2,t} \end{bmatrix}\right) / \Delta t = \begin{bmatrix} 0 & 1 \\ -\theta_1 & -\theta_2 \end{bmatrix} \begin{bmatrix} z_{1,t} \\ z_{2,t} \end{bmatrix} + \begin{bmatrix} 0 \\ \eta \end{bmatrix} u_t + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \frac{w_t}{\Delta t} \tag{7a}$$

$$\begin{bmatrix} z_{1,t+1} \\ z_{2,t+1} \end{bmatrix} - \begin{bmatrix} z_{1,t} \\ z_{2,t} \end{bmatrix} = \begin{bmatrix} 0 & \Delta t \\ -\theta_1 \Delta t & -\theta_2 \Delta t \end{bmatrix} \begin{bmatrix} z_{1,t} \\ z_{2,t} \end{bmatrix} + \begin{bmatrix} 0 \\ \eta \Delta t \end{bmatrix} u_t + \begin{bmatrix} 0 \\ 1 \end{bmatrix} w_t \tag{7b}$$

$$\underbrace{\begin{bmatrix} z_{1,t+1} \\ z_{2,t+1} \end{bmatrix}}_{z_{t+1}} = \underbrace{\begin{bmatrix} 1 & \Delta t \\ -\theta_1 \Delta t & -\theta_2 \Delta t + 1 \end{bmatrix}}_{A(\theta)} \underbrace{\begin{bmatrix} z_{1,t} \\ z_{2,t} \end{bmatrix}}_{z_t} + \underbrace{\begin{bmatrix} 0 \\ \eta \Delta t \end{bmatrix}}_{B(\eta)} u_t + \begin{bmatrix} 0 \\ 1 \end{bmatrix} w_t \tag{7c}$$

The above discrete-time state transition can be cast to a Gaussian distribution as:

$$z_{t+1} \sim \mathcal{N}(A(\theta) z_t + B(\eta) u_t, Q)\,. \tag{8}$$

for

$$Q = \begin{bmatrix} 0 & 0 \\ 0 & \tau^{-1} \end{bmatrix}, \; A(\theta) = \begin{bmatrix} 1 & \Delta t \\ -\theta_1 \Delta t & -\theta_2 \Delta t + 1 \end{bmatrix}, \text{ and } B(\eta) = \begin{bmatrix} 0 \\ \eta \Delta t \end{bmatrix}. \tag{9}$$

The expectation of the inverse of Q is obtained by a noise injection:

$$\mathbb{E}_{q(\tau)} Q^{-1} = \mathbb{E}_{q(\tau)} \begin{bmatrix} 0 & 0 \\ 0 & \tau^{-1} \end{bmatrix}^{-1} \approx \mathbb{E}_{q(\tau)} \begin{bmatrix} \epsilon & 0 \\ 0 & \tau^{-1} \end{bmatrix}^{-1} = \begin{bmatrix} \epsilon^{-1} & 0 \\ 0 & \tau \end{bmatrix} = m_Q\,. \tag{10}$$

(wk) It might be important to have a covariance matrix Q without the 0 on the diagonal. That depends on whether the noise should also apply to the substitution in Equation 3a.

The matrix A can be constructed from a vector $\theta = \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix}$ as:

$$A(\theta) = S + s\theta^\top, \quad S = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix}, \quad s = \begin{bmatrix} 0 \\ -\Delta t \end{bmatrix}. \tag{11}$$

1.2 Recognition model

I choose the following recognition factors:

$$z_{t-1} \sim \mathcal{N}(m_{z_{t-1}}, V_{z_{t-1}}) \tag{12a}$$
$$z_t \sim \mathcal{N}(m_{z_t}, V_{z_t}) \tag{12b}$$
$$\theta \sim \mathcal{N}(m_\theta, V_\theta) \tag{12c}$$
$$\eta \sim \mathcal{N}(m_\eta, v_\eta) \tag{12d}$$
$$\tau \sim \Gamma(a_\tau, b_\tau)\,. \tag{12e}$$

2 Messages

In computing the messages, the following results are useful:

$$\mathbb{E}_{q(\theta)} A(\theta) = \mathbb{E}_{q(\theta)} \left[S + s\theta^\top \right] = S + s m_\theta^\top = A(m_\theta) \tag{13a}$$

$$\mathbb{E}_{q(\eta)} B(\eta) = \mathbb{E}_{q(\eta)} \begin{bmatrix} 0 \\ \Delta t \end{bmatrix} \eta = \begin{bmatrix} 0 \\ \Delta t \end{bmatrix} m_\eta = B(m_\eta) \, . \tag{13b}$$

Message to z_t

$$\begin{aligned}
\overrightarrow{\nu}(z_t) &= \mathbb{E}_{q(z_{t-1})q(\theta)q(\eta)q(\tau)} \log \mathcal{N}(z_t \mid A(\theta) z_{t-1} + B(\eta) u_t, Q) && \text{(14a)} \\
&\propto -\frac{1}{2} \mathbb{E}\big(z_t - A(\theta) z_{t-1} - B(\eta) u_t\big)^\top Q^{-1} \big(z_t - A(\theta) z_{t-1} - B(\eta) u_t\big) && \text{(14b)} \\
&\propto -\frac{1}{2} \mathbb{E}\big[\underbrace{z_t^\top Q^{-1} z_t}_{①} - \underbrace{(A(\theta) z_{t-1})^\top Q^{-1} z_t}_{②} - \underbrace{(B(\eta) u_t)^\top Q^{-1} z_t}_{③} && \text{(14c)} \\
&\qquad - \underbrace{z_t^\top Q^{-1} A(\theta) z_{t-1}}_{④} - \underbrace{z_t Q^{-1} B(\eta) u_t}_{⑤} \big] \, . && \text{(14d)}
\end{aligned}$$

With the terms:

$$\begin{aligned}
① &= z_t^\top m_Q z_t && \text{(15a)} \\
② &= (A(m_\theta) m_{z_{t-1}})^\top m_Q z_t && \text{(15b)} \\
③ &= (B(m_\eta) u_t)^\top m_Q z_t && \text{(15c)} \\
④ &= z_t^\top m_Q A(m_\theta) m_{z_{t-1}} && \text{(15d)} \\
⑤ &= z_t^\top m_Q B(m_\eta) u_t \, . && \text{(15e)}
\end{aligned}$$

Plugging those back in:

$$\begin{aligned}
\overrightarrow{\nu}(z_t) &\propto -\frac{1}{2} \big[z_t^\top \underbrace{m_Q}_{\Phi} z_t - \big(A(m_\theta) m_{z_{t-1}} + B(m_\eta) u_t\big) m_Q z_t \\
&\qquad - z_t^\top m_Q \underbrace{\big(A(m_\theta) m_{z_{t-1}} + B(m_\eta) u_t\big)}_{\phi} \big] && \text{(16)} \\
&\sim \mathcal{N}(\phi, \Phi^{-1}) \, . && \text{(17)}
\end{aligned}$$

Message to z_{t-1}

Message to θ

Message to η

Message to τ

References

[1] Simo Särkkä and Arno Solin. *Applied stochastic differential equations*, volume 10. Cambridge University Press, 2019.