Non-Parametric Drift Model for Stochastic Differential Equations

Roland Hostettler, Filip Tronarp, Simo Särkkä

Department of Electrical Engineering and Automation
Aalto University
E-Mail: roland.hostettler@aalto.fi

April 13, 2018

We propose a reduced-rank Gaussian process model for the drift function in nonlinear stochastic differential equations

Stochastic Differential Equations

Nonlinear, time-varying SDE:

$$\mathrm{d}\mathbf{x}_t = \mathbf{f}(\mathbf{x}_t, \mathbf{u}_t, t)\mathrm{d}t + \mathrm{d}\boldsymbol{\beta}_t \tag{1}$$

where

- $\mathbf{x}_t \in \mathbb{R}^{N_x}$ is the state,
- $\mathbf{f}(\mathbf{x}_t, \mathbf{u}_t, t) = \begin{bmatrix} f_1(\mathbf{x}_t, \mathbf{u}_t, t) & \dots & f_L(\mathbf{x}_t, \mathbf{u}_t, t) \end{bmatrix}^\mathsf{T}$ are the *unknown* drift functions,
- $\mathbf{u}_t \in \mathbb{R}^{N_u}$ is a deterministic input, and
- $\boldsymbol{\beta}_t$ is Brownian motion with diffusion matrix $\mathbf{Q}$.

Gaussian Process Drift Model

Gaussian process drift model:

$$\mathrm{d}\mathbf{x}_t = \mathbf{f}(\mathbf{x}_t, \mathbf{u}_t, t)\mathrm{d}t + \mathrm{d}\boldsymbol{\beta}_t \tag{2a}$$

$$f_l(\mathbf{x}_t, \mathbf{u}_t, t) \sim \mathcal{GP}(m(\mathbf{x}_t, \mathbf{u}_t, t), k(\mathbf{x}_t, \mathbf{u}_t, t, \mathbf{x}'_t, \mathbf{u}'_t, t')) \tag{2b}$$

- This is non-Markovian (requires complete history of $\mathbf{x}_t$, $\mathbf{u}_t$, t)

Assumptions

1. Mean function is zero, $m(\mathbf{x}_t, \mathbf{u}_t, t) = 0$,
2. Covariance function is separable,

$$k(\mathbf{x}_t, \mathbf{u}_t, t, \mathbf{x}'_t, \mathbf{u}'_t, t') = k_S(\mathbf{x}_t, \mathbf{u}_t, \mathbf{x}'_t, \mathbf{u}'_t)k_T(t, t')$$

3. $k_T(t, t')$ is stationary, $k_T(t, t') = k_T(\tau)$ ($\tau = t' - t$)

Basis Function Expansion (1/2)

Decompose $f_l(\mathbf{x}_t, \mathbf{u}_t, t)$ such that

$$f_l(\mathbf{x}_t, \mathbf{u}_t, t) = \sum_{j=0}^{\infty} \alpha_{j,t} \psi_j(\mathbf{x}_t, \mathbf{u}_t) \tag{3}$$

- $\psi_j(\mathbf{x}_t, \mathbf{u}_t)$ are orthonormal eigenfunctions of $k_S(\mathbf{x}_t, \mathbf{u}_t, \mathbf{x}'_t, \mathbf{u}'_t)$

$$\langle \psi_i(\mathbf{x}_t, \mathbf{u}_t), \psi_j(\mathbf{x}_t, \mathbf{u}_{\boldsymbol{t}}) \rangle = \begin{cases} 1 & i = j, \\ 0 & i \neq j, \end{cases} \tag{4a}$$

$$\langle k_S(\mathbf{x}_t, \mathbf{u}_t, \mathbf{x}'_t, \mathbf{u}'_t), \psi_j(\mathbf{x}'_t, \mathbf{u}'_t) \rangle = \lambda_j \psi_j(\mathbf{x}_t, \mathbf{u}_t), \tag{4b}$$

- The coefficients are

$$\alpha_{j,t} = \langle f(\mathbf{x}_t, \mathbf{u}_t, t), \psi_j(\mathbf{x}_t, \mathbf{u}_t) \rangle. \tag{5}$$

Basis Function Expansion (2/2)

The time-varying coefficients are

$$\alpha_{j,t} = \langle f(\mathbf{x}_t, \mathbf{u}_t, t), \psi_j(\mathbf{x}_t, \mathbf{u}_t) \rangle. \tag{6}$$

- Then:

$$\mathrm{Cov}\{\alpha_{i,t}, \alpha_{j,t'}\} = k_T(t, t')\lambda_j \delta_{ij}$$

$$k(\mathbf{x}_t, \mathbf{u}_t, t, \mathbf{x}'_t, \mathbf{u}'_t, t') = k_T(t, t') \sum_{j=0}^{\infty} \lambda_j \psi_j(\mathbf{x}_t, \mathbf{u}_t) \psi_j^*(\mathbf{x}'_t, \mathbf{u}'_t)$$

- Thus:

$$\alpha_{j,t} \sim \mathcal{GP}(0, k_T(t, t')\lambda_j)$$

- $\alpha_{j,t}$s are still non-Markovian

Spectral Decomposition

Coefficients:

$$\alpha_{j,t} \sim \mathcal{GP}(0, k_{\alpha,j}(t,t')) \tag{7}$$

with $k_{\alpha,j}(t,t') = k_T(t,t')\lambda_j = k_T(\tau)\lambda_j$
Power spectral density:

$$S_{\alpha,j}(\omega) = \mathcal{F}_\tau\{k_{\alpha,j}(\tau)\} \tag{8}$$

Decomposition:

$$S_{\alpha,j}(\omega) = q_{\alpha,j} H(\mathrm{i}\,\omega) H(\mathrm{i}\,\omega)^*. \tag{9}$$

Equivalent representation:

$$\mathrm{d}\mathbf{z}_{j,t} = \mathbf{A}_j \mathbf{z}_{j,t} \mathrm{d}t + \mathbf{B}_j \mathrm{d}\varepsilon_{j,t}, \tag{10a}$$

$$\alpha_{j,t} = \mathbf{C}_j \mathbf{z}_{j,t}, \tag{10b}$$

Model Summary

We can write the complete model as

$$\mathrm{d}\mathbf{z}_{j,t} = \mathbf{A}_j \mathbf{z}_{j,t} \mathrm{d}t + \mathbf{B}_j \mathrm{d}\varepsilon_{j,t}, \tag{11a}$$

$$\alpha_{j,t} = \mathbf{C}_j \mathbf{z}_{j,t}, \tag{11b}$$

$$\mathrm{d}x_l = \mathbf{\Psi}(\mathbf{x}_t, \mathbf{u}_t) \boldsymbol{\alpha}_{l,t} \mathrm{d}t + \mathrm{d}\beta_{l,t} \tag{11c}$$

where

- $\mathbf{\Psi}(\mathbf{x}_t, \mathbf{u}_t) = \begin{bmatrix} \psi_1(\mathbf{x}_t, \mathbf{u}_t) & \psi_2(\mathbf{x}_t, \mathbf{u}_t) & \dots & \psi_J(\mathbf{x}_t, \mathbf{u}_t) \end{bmatrix}$,
- $\boldsymbol{\alpha}_{l,t} = \begin{bmatrix} \alpha_{1,t} & \alpha_{2,t} & \dots & \alpha_{J,t} \end{bmatrix}^\mathsf{T}$.

Parameter Estimation

Objective: Estimate $\boldsymbol{\theta}$ given training data $\mathbf{y}_{1:N} = \{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_N\}$ from

$$\mathbf{y}_n = \mathbf{H}\mathbf{x}_{t_n} + \mathbf{v}_n, \tag{12}$$

Solution: Maximizing the posterior

$$\begin{aligned} p(\boldsymbol{\theta} \mid \mathbf{y}_{1:N}) &\propto p(\boldsymbol{\theta}) p(\mathbf{y}_{1:N} \mid \boldsymbol{\theta}) \\ &= p(\boldsymbol{\theta}) \prod_{n=1}^{N} p(\mathbf{y}_n \mid \mathbf{y}_{1:n-1}, \boldsymbol{\theta}) \end{aligned} \tag{13}$$

- Extended Kalman filter to approximate $p(\mathbf{y}_n \mid \mathbf{y}_{1:n-1}, \boldsymbol{\theta})$
- Euler discretization of the nonlinear continuous-time system

Bouc–Wen Oscillator: Model

Model structure:

$$\mathrm{d}x_{1,t} = x_{2,t}\mathrm{d}t \tag{14a}$$

$$\mathrm{d}x_{2,t} = f(\mathbf{x}_t, u_t, t)\mathrm{d}t + \mathrm{d}\beta_t. \tag{14b}$$

Covariance function:

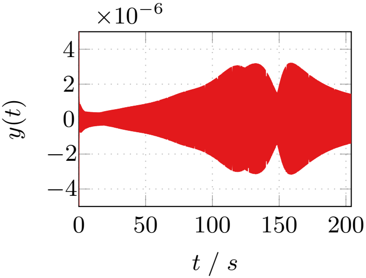
$$k(\mathbf{x}_t, u_t, t, \mathbf{x}_t', u_t', t') = (k_{\mathsf{SE}}(\mathbf{x}_t, \mathbf{x}_t') + k_{\mathsf{SE}}(u_t, u_t'))k_{\mathsf{OU}}(t, t'), \tag{15}$$

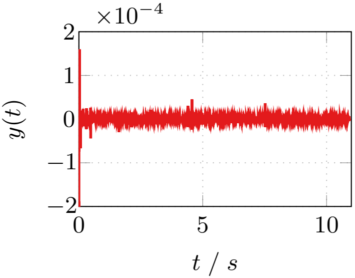
- Fourier basis functions $\psi_j(x) = \frac{1}{\sqrt{\gamma}} \exp\left(\frac{\mathrm{i}\, j 2\pi x}{\gamma}\right)$
- Truncate at 25 ($\mathbf{x}_t$) and 5 (u_t) eigenfunctions

Bouc–Wen Oscillator: Prediction Results

Prediction RMSE:

- Training: 2.65×10^{-5}
- Multisine validation: 0.580×10^{-5}
- Swept sine validation: 0.096×10^{-5}





Cascaded Tanks: Model

Model:

$$\mathrm{d}x_{1,t} = f_1(x_{1,t}, u_t, t)\mathrm{d}t + \mathrm{d}\beta_{1,t} \quad (16a)$$

$$\mathrm{d}x_{2,t} = f_2(x_{1,t}, x_{2,t}, t)\mathrm{d}t + \mathrm{d}\beta_{2,t} \quad (16b)$$

Covariance function:

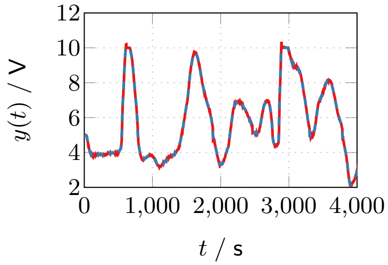
$$k(\mathbf{x}, \mathbf{x}', t, t') = k_{\mathsf{SE}}(\mathbf{x}, \mathbf{x}')k_{\mathsf{OU}}(t, t'). \quad (17)$$

- Remaining parameters the same as for the Bouc–Wen example

Cascaded Tanks: Results

Prediction RMSE:

- Training set RMSE: 51.5×10^{-3}
- Validation set RMSE: 57.6×10^{-3}.



?? Measured; **??** Predicted

Summary

- Proposed model:

$$\mathrm{d}\mathbf{x}_t = \mathbf{f}(\mathbf{x}_t, \mathbf{u}_t, t)\mathrm{d}t + \mathrm{d}\boldsymbol{\beta}_t$$
$$f_l(\mathbf{x}_t, \mathbf{u}_t, t) \sim \mathcal{GP}(0, k_S(\mathbf{x}_t, \mathbf{u}_t, \mathbf{x}'_t, \mathbf{u}'_t)k_T(t, t'))$$

- Basis function expansion and spectral decomposition yield:

$$\begin{aligned}
\mathrm{d}\mathbf{z}_{j,t} &= \mathbf{A}_j\mathbf{z}_{j,t}\mathrm{d}t + \mathbf{B}_j\mathrm{d}\varepsilon_{j,t}, \\
\alpha_{j,t} &= \mathbf{C}_j\mathbf{z}_{j,t}, \\
\mathrm{d}x_l &= \boldsymbol{\Psi}(\mathbf{x}_t, \mathbf{u}_t)\boldsymbol{\alpha}_{l,t}\mathrm{d}t + \mathrm{d}\beta_{l,t}
\end{aligned}$$

- Promising results if the assumptions are satisfied