Particle-based Gaussian process optimization for input design in nonlinear dynamical models

Patricio E. Valenzuela[1], Johan Dahlin[2], Cristian R. Rojas[1] and Thomas B. Schön[3]

Abstract

- A Gaussian process optimization method for input design is presented.
- The method is suitable for nonlinear models.

Input design problem

Nonlinear state-space model:

$$\begin{aligned} x_t | x_{t-1} &\sim f_\theta(x_t | x_{t-1}, u_{t-1}), \\ y_t | x_t &\sim g_\theta(y_t | x_t, u_t), \\ x_0 &\sim \mu_\theta(x_0), \end{aligned}$$

Objective: Design $u_{1:T} := (u_1, \ldots, u_T) \in \mathcal{C}^T$ optimizing

$$\mathcal{I}_F^{\theta_0, \text{av}}(u_{1:T}) := \frac{1}{T} \mathbf{E}_u \left\{ \mathcal{I}_F^{\theta_0}(u_{1:T}) \right\} ,$$

where

$$\begin{aligned} \mathcal{I}_F^{\theta_0}(u_{1:T}) &:= \mathbf{E}\left\{ \mathcal{S}(\theta_0) \mathcal{S}^\top(\theta_0) | \, u_{1:T} \right\} , \\ \mathcal{S}(\theta_0) &:= \nabla \, \ell_\theta(y_{1:T})|_{\theta=\theta_0} , \\ \ell_\theta(y_{1:T}) &:= \log p_\theta(y_{1:T} | u_{1:T}) \, . \end{aligned}$$

Problem: Find an input signal $u_{1:T}^{\text{opt}} \in \mathcal{C}^T$ as

$$u_{1:T}^{\text{opt}} := \arg \max_{u_{1:T} \in \mathcal{C}^T} h(\mathcal{I}_F^{\theta_0, \text{av}}(u_{1:T})) \, ,$$

where $h \colon \mathbb{R}^{m \times m} \to \mathbb{R}$ is a matrix nondecreasing function.

Challenges:

- $\mathcal{I}_F^{\theta_0, \text{av}}(u_{1:T})$ unavailable in closed form.

  $\Rightarrow$ Estimate $\mathcal{I}_F^{\theta_0, \text{av}}(u_{1:T})$ using particle methods.

- Tractable parametrization for $u_{1:T}$.

  $\Rightarrow$ Restrict $\{u_t\}$ to, e.g., stationary AR process or stationary Markov process.

- $h(\mathcal{I}_F^{\theta_0, \text{av}}(u_{1:T}))$ difficult to optimize.

  $\Rightarrow$ Use Gaussian process optimization technique.

GPO in input design

(i) Given $u_{1:T}^{(k)}$, compute an estimate of $h(\mathcal{I}_F^{\theta_0, \text{av}}(u_{1:T}^{(k)}))$, denoted by $\widehat{h}_k$.

(ii) Estimate $h(\mathcal{I}_F^{\theta_0, \text{av}}(u_{1:T}))$ given $\mathcal{D}_k := \{u_{1:T}^{(j)}, \widehat{h}_j\}_{j=0}^k$ and $h(\mathcal{I}_F^{\theta_0, \text{av}}(\cdot)) \sim \mathcal{GP}\left(m(\cdot), \, \kappa(\cdot, \, \cdot)\right)$.

(iii) Given $h(\mathcal{I}_F^{\theta_0, \text{av}}(u_{1:T}))$, generate $u_{1:T}^{(k+1)}$ as

$$\begin{aligned} u_{1:T}^{(k+1)} &= \arg \max_{u_{1:T} \in \mathcal{C}^T} \mathbf{E}\left\{ I(u_{1:T}) \, | \mathcal{D}_k \right\} , \\ I(u_{1:T}) &:= \max \left\{ 0, \, h(\mathcal{I}_F^{\theta_0, \text{av}}(u_{1:T})) - \mu_{\max} - \xi \right\} , \\ \mu_{\max} &:= \max_{u_{1:T} \in \mathbf{u}_{1:T}^{(k)}} \mu(u_{1:T} | \mathcal{D}_k) \, . \end{aligned}$$

$\mu(\cdot | \mathcal{D}_k)$: Posterior mean of $h(\mathcal{I}_F^{\theta_0, \text{av}}(\cdot))$ given $\mathcal{D}_k$.

Example

$$\begin{aligned} x_{t+1} | x_t &\sim \mathcal{N}\left(\frac{1}{\gamma + x_t^2} + u_t, 0.1^2 \right), \\ y_t | x_t &\sim \mathcal{N}\left(\beta x_t^2, 1^2 \right), \end{aligned}$$

$\theta = \{\gamma, \beta\}$, $T = 10^3$, $h(\cdot) = \log \det(\cdot)$ and $\theta_0 = \{2, 0.8\}$.
$\{u_t\}$: Markov process with $n_m = 1$ and three cases for $\mathcal{C}$:

- Case 1: $\mathcal{C} = \{-1, 1\}$.
- Case 2: $\mathcal{C} = \{-1, 0, 1\}$.
- Case 3: $\mathcal{C} = \{-1, -1/3, 1/3, 1\}$.

Results:

Input	Binary	opt. Case 1	opt. Case 2	opt. Case 3
h^{opt}	4.11	4.11	4.15	4.44

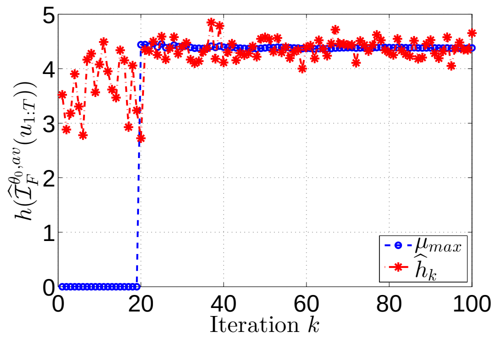

Figure 1. $\widehat{h}_k$ and $\mu_{\max}$ at iteration k for Case 3.

Future work

- Include model uncertainty.
- Online procedure to design $u_{1:T}$.

References

B. Shahriari, K. Swersky, Z. Wang, R.P. Adams, and N. de Freitas, "Taking the human out of the loop: A review of Bayesian optimization," Proceedings of the IEEE, vol. 104, no. 1, pp. 148–175, 2016.

(1) Department of Automatic Control, School of Electrical Engineering, KTH, Stockholm, Sweden.
(2) Department of Electrical Engineering, Linköping University, Linköping, Sweden
(3) Department of Information Technology, Uppsala University, Uppsala, Sweden