Gaussian Process based Model Predictive Control for Linear Time Varying Systems

IEEE 14^{th} International Workshop on Advanced Motion Control

Gang Cao

School of Engineering and Advanced Technology
Massey University, Auckland

April 23, 2016

MASSEY UNIVERSITY
TE KUNENGA KI PŪREHUROA
UNIVERSITY OF NEW ZEALAND

Outline

Abbreviations

- **GP: Gaussian process**
- **MPC: Model predictive control**
- **LTV: Linear time varying**
- **CG: Conjugate gradient**

Problem Formulation

Linear time-varying system:

$$\mathbf{x}_{k+1} = \mathbf{A}_k \mathbf{x}_k + \mathbf{B}_k \mathbf{u}_k + \mathbf{w}_k \tag{1}$$

- $\mathbf{x}_k \in \mathbb{R}^n$, $\mathbf{u}_k \in \mathbb{R}^m$, $\mathbf{w}_k \in \mathbb{R}^n$;
- State matrix $\mathbf{A}_k$ and Input matrix $\mathbf{B}_k$ are unknown;

Unconstrained MPC trajectory tracking problem:

$$\min \mathbf{V}_k^* = \min_{\mathbf{U}_k^*} \sum_{i=1}^{H} \left\{ \|\mathbf{x}_{k+i} - \mathbf{r}_{k+i}\|_{\mathbf{Q}}^2 + \|\mathbf{u}_{k+i-1}\|_{\mathbf{R}}^2 \right\} \tag{2}$$

- $\mathbf{U}_k^* = [\mathbf{u}_{k-1}^*, \cdots, \mathbf{u}_{k+H-1}^*]^T$;

Problem Formulation

Problems:

- Identification of unknown LTV system
- Solving the MPC problem effectively

Problem Formulation

Problems:

- Identification of unknown LTV system
- Solving the MPC problem effectively

Contributions:

- GP based MPC;
- Gradient based solution;

GP Learning of Unknown LTV

Gaussian Process Models:

GP Learning of Unknown LTV

Gaussian Process Models:

- Probabilistic data-driven model;

GP Learning of Unknown LTV

Gaussian Process Models:

- Probabilistic data-driven model;
- Specifying a squared exponential covariance function $\mathbf{K}(\mathbf{x}_{k,i}, \mathbf{x}_{k,j}, \boldsymbol{\theta})$;
  $\mathbf{K}(\tilde{\mathbf{x}}_i, \tilde{\mathbf{x}}_j) = \sigma_s^2 \exp(-\frac{1}{2}(\tilde{\mathbf{x}}_i - \tilde{\mathbf{x}}_j)^T \boldsymbol{\Lambda}(\tilde{\mathbf{x}}_i - \tilde{\mathbf{x}}_j)) + \sigma_n^2$

GP Learning of Unknown LTV

Gaussian Process Models:

- Probabilistic data-driven model;
- Specifying a squared exponential covariance function $\mathbf{K}(\mathbf{x}_{k,i}, \mathbf{x}_{k,j}, \boldsymbol{\theta})$;
- $\mathbf{x}_k^* \sim \mathcal{N}\{\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k\}$;

$$\boldsymbol{\mu}_k = \mathbf{K}(\tilde{\mathbf{x}}_k^*, \tilde{\mathbf{X}})(\mathbf{K}(\tilde{\mathbf{X}}, \tilde{\mathbf{X}}) + \sigma_n \mathbf{I})^{-1} \mathbf{X}_k \tag{3a}$$

$$\boldsymbol{\Sigma}_k = \mathbf{K}(\tilde{\mathbf{x}}_k^*, \tilde{\mathbf{x}}_k^*) - \mathbf{K}(\tilde{\mathbf{x}}_k^*, \tilde{\mathbf{X}})(\mathbf{K}(\tilde{\mathbf{X}}, \tilde{\mathbf{X}}) + \sigma_n \mathbf{I})^{-1} \mathbf{K}(\tilde{\mathbf{X}}, \tilde{\mathbf{x}}_k^*) \tag{3b}$$

GP Learning of Unknown LTV

Gaussian Process Models:

- Probabilistic data-driven model;
- Specifying a squared exponential covariance function $\mathbf{K}(\mathbf{x}_{k,i}, \mathbf{x}_{k,j}, \boldsymbol{\theta})$;
- $\mathbf{x}_k^* \sim \mathcal{N}\{\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k\}$;

$$\boldsymbol{\mu}_k = \mathbf{K}(\tilde{\mathbf{x}}_k^*, \tilde{\mathbf{X}})(\mathbf{K}(\tilde{\mathbf{X}}, \tilde{\mathbf{X}}) + \sigma_n \mathbf{I})^{-1} \mathbf{X}_k \tag{3a}$$

$$\boldsymbol{\Sigma}_k = \mathbf{K}(\tilde{\mathbf{x}}_k^*, \tilde{\mathbf{x}}_k^*) - \mathbf{K}(\tilde{\mathbf{x}}_k^*, \tilde{\mathbf{X}})(\mathbf{K}(\tilde{\mathbf{X}}, \tilde{\mathbf{X}}) + \sigma_n \mathbf{I})^{-1} \mathbf{K}(\tilde{\mathbf{X}}, \tilde{\mathbf{x}}_k^*) \tag{3b}$$

- Evaluation of model uncertainty $\rightarrow \boldsymbol{\Sigma}_k$;

GP Learning of Unknown LTV

Gaussian Process Models:

- Probabilistic data-driven model;
- Specifying a squared exponential covariance function $\mathbf{K}(\mathbf{x}_{k,i}, \mathbf{x}_{k,j}, \boldsymbol{\theta})$;
- $\mathbf{x}_k^* \sim \mathcal{N}\{\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k\}$;

$$\boldsymbol{\mu}_k = \mathbf{K}(\tilde{\mathbf{x}}_k^*, \tilde{\mathbf{X}})(\mathbf{K}(\tilde{\mathbf{X}}, \tilde{\mathbf{X}}) + \sigma_n \mathbf{I})^{-1}\mathbf{X}_k \tag{3a}$$

$$\boldsymbol{\Sigma}_k = \mathbf{K}(\tilde{\mathbf{x}}_k^*, \tilde{\mathbf{x}}_k^*) - \mathbf{K}(\tilde{\mathbf{x}}_k^*, \tilde{\mathbf{X}})(\mathbf{K}(\tilde{\mathbf{X}}, \tilde{\mathbf{X}}) + \sigma_n \mathbf{I})^{-1}\mathbf{K}(\tilde{\mathbf{X}}, \tilde{\mathbf{x}}_k^*) \tag{3b}$$

- Evaluation of model uncertainty $\rightarrow \boldsymbol{\Sigma}_k$;

Unknown LTV Learning:

- GP Inputs: $\tilde{\mathbf{x}}_k = (\mathbf{x}_k, \mathbf{u}_k)$;
- GP Outputs: $\delta\mathbf{x}_{k+1} = \mathbf{x}_{k+1} - \mathbf{x}_k$;
- Hyperparameters learning: Minimize the negative log of likelihood (CG);

Gaussian Process based Model Predictive Control for Linear Time Varying Systems

GP based MPC

Problem Reformulation:

$$\begin{aligned}\mathbb{E}\Big[\mathbf{V}_k^*\Big] &= \underbrace{\mathbb{E}\Big[\sum_{i=1}^{H}\big\{\,\|\mathbf{x}_{k+i}-\mathbf{r}_{k+i}\|_{\mathbf{Q}}^2+\|\mathbf{u}_{k+i-1}\|_{\mathbf{R}}^2\,\big\}\Big]}_{\text{Stochastic MPC}} \\ &= \underbrace{\sum_{i=1}^{H}\Big\{\,\|\boldsymbol{\mu}_{k+i}-\mathbf{r}_{k+i}\|_{\mathbf{Q}}^2+\|\mathbf{u}_{k+i-1}\|_{\mathbf{R}}^2+\mathrm{trace}\big(\mathbf{Q}\boldsymbol{\Sigma}_{k+i}\big)\Big\}}_{\text{Deterministic MPC}}\end{aligned} \tag{4}$$

GP based MPC

Problem Reformulation:

$$\mathbb{E}\Big[\mathbf{V}_k^*\Big] = \underbrace{\mathbb{E}\Big[\sum_{i=1}^{H}\big\{\, \|\mathbf{x}_{k+i} - \mathbf{r}_{k+i}\|_{\mathbf{Q}}^2 + \|\mathbf{u}_{k+i-1}\|_{\mathbf{R}}^2 \,\big\}\Big]}_{\text{Stochastic MPC}}$$
$$= \underbrace{\sum_{i=1}^{H}\Big\{\, \|\boldsymbol{\mu}_{k+i} - \mathbf{r}_{k+i}\|_{\mathbf{Q}}^2 + \|\mathbf{u}_{k+i-1}\|_{\mathbf{R}}^2 + \mathrm{trace}\big(\mathbf{Q}\boldsymbol{\Sigma}_{k+i}\big)\Big\}}_{\text{Deterministic MPC}} \tag{4}$$

Implementation Issues:

- Uncertainty propagation → GP predictions at uncertain inputs;

GP based MPC

Problem Reformulation:

$$\begin{aligned}\mathbb{E}\left[\mathbf{V}_k^*\right] &= \underbrace{\mathbb{E}\left[\sum_{i=1}^{H}\left\{\|\mathbf{x}_{k+i}-\mathbf{r}_{k+i}\|_{\mathbf{Q}}^2+\|\mathbf{u}_{k+i-1}\|_{\mathbf{R}}^2\right\}\right]}_{\text{Stochastic MPC}} \\ &= \underbrace{\sum_{i=1}^{H}\left\{\|\boldsymbol{\mu}_{k+i}-\mathbf{r}_{k+i}\|_{\mathbf{Q}}^2+\|\mathbf{u}_{k+i-1}\|_{\mathbf{R}}^2+\operatorname{trace}\left(\mathbf{Q}\boldsymbol{\Sigma}_{k+i}\right)\right\}}_{\text{Deterministic MPC}}\end{aligned} \tag{4}$$

Implementation Issues:

- Uncertainty propagation → GP predictions at uncertain inputs;
- Effective Solutions → Gradient based optimization;

Gradient based optimization

Unconstrained optimization question:

$$\mathbf{U}_k^* = \arg\min_{\mathbf{U}_k^*} \mathbb{E}\left\{\mathbf{V}_k^*\right\} \tag{5a}$$

$$\text{s.t. } \boldsymbol{\mu}_{k+1} = \mathcal{G}(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k, \mathbf{u}_k); \tag{5b}$$

Gradient based optimization

Unconstrained optimization question:

$$\mathbf{U}_k^* = \arg\min_{\mathbf{U}_k^*} \mathbb{E}\left\{\mathbf{V}_k^*\right\} \tag{5a}$$

$$\text{s.t. } \boldsymbol{\mu}_{k+1} = \mathcal{G}(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k, \mathbf{u}_k); \tag{5b}$$

- Gradient based linear search;

$$\mathbf{U}_{k+1} = \mathbf{U}_k + \boldsymbol{\alpha} \cdot \underbrace{\frac{\partial \mathbb{E}\{\mathbf{V}_k\}}{\partial \mathbf{U}_k}}_{\text{Gradient}}; \tag{6}$$

$$\frac{\partial \mathbb{E}\{\mathbf{V}_k\}}{\partial \mathbf{U}_k} = \frac{\partial \mathbb{E}\{\mathbf{V}_k\}}{\partial \boldsymbol{\mu}_{k+1}} \frac{\partial \boldsymbol{\mu}_{k+1}}{\partial \mathbf{U}_k} + \frac{\partial \mathbb{E}\{\mathbf{V}_k\}}{\partial \boldsymbol{\Sigma}_{k+1}} \frac{\partial \boldsymbol{\Sigma}_{k+1}}{\partial \mathbf{U}_k} + \frac{\partial \mathbb{E}\{\mathbf{V}_k\}}{\partial \mathbf{U}_k} \tag{7}$$

Gradient based optimization

Unconstrained optimization question:

$$\mathbf{U}_k^* = \arg\min_{\mathbf{U}_k^*} \mathbb{E}\left\{\mathbf{V}_k^*\right\} \tag{5a}$$

$$\text{s.t. } \boldsymbol{\mu}_{k+1} = \mathcal{G}(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k, \mathbf{u}_k); \tag{5b}$$

- Gradient based linear search;

$$\mathbf{U}_{k+1} = \mathbf{U}_k + \boldsymbol{\alpha} \cdot \underbrace{\frac{\partial \mathbb{E}\{\mathbf{V}_k\}}{\partial \mathbf{U}_k}}_{\text{Gradient}}; \tag{6}$$

$$\frac{\partial \mathbb{E}\{\mathbf{V}_k\}}{\partial \mathbf{U}_k} = \frac{\partial \mathbb{E}\{\mathbf{V}_k\}}{\partial \boldsymbol{\mu}_{k+1}} \frac{\partial \boldsymbol{\mu}_{k+1}}{\partial \mathbf{U}_k} + \frac{\partial \mathbb{E}\{\mathbf{V}_k\}}{\partial \boldsymbol{\Sigma}_{k+1}} \frac{\partial \boldsymbol{\Sigma}_{k+1}}{\partial \mathbf{U}_k} + \frac{\partial \mathbb{E}\{\mathbf{V}_k\}}{\partial \mathbf{U}_k} \tag{7}$$

- End when $\mathbb{E}\{\mathbf{V}_k\}_{\mathbf{U}_k} \leq \epsilon$;

Gradient based optimization

Unconstrained optimization question:

$$\mathbf{U}_k^* = \arg\min_{\mathbf{U}_k^*} \mathbb{E}\left\{\mathbf{V}_k^*\right\} \tag{5a}$$

$$\text{s.t. } \boldsymbol{\mu}_{k+1} = \mathcal{G}(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k, \mathbf{u}_k); \tag{5b}$$

- Gradient based linear search;

$$\mathbf{U}_{k+1} = \mathbf{U}_k + \boldsymbol{\alpha} \cdot \underbrace{\frac{\partial \mathbb{E}\{\mathbf{V}_k\}}{\partial \mathbf{U}_k}}_{\text{Gradient}}; \tag{6}$$

$$\frac{\partial \mathbb{E}\{\mathbf{V}_k\}}{\partial \mathbf{U}_k} = \frac{\partial \mathbb{E}\{\mathbf{V}_k\}}{\partial \boldsymbol{\mu}_{k+1}} \frac{\partial \boldsymbol{\mu}_{k+1}}{\partial \mathbf{U}_k} + \frac{\partial \mathbb{E}\{\mathbf{V}_k\}}{\partial \boldsymbol{\Sigma}_{k+1}} \frac{\partial \boldsymbol{\Sigma}_{k+1}}{\partial \mathbf{U}_k} + \frac{\partial \mathbb{E}\{\mathbf{V}_k\}}{\partial \mathbf{U}_k} \tag{7}$$

- End when $\mathbb{E}\{\mathbf{V}_k\}_{\mathbf{U}_k} \leq \epsilon$;
- Multi-Start with different initial values;

Numerical Simulation

2-Inputs-2-Outputs Numerical Example:

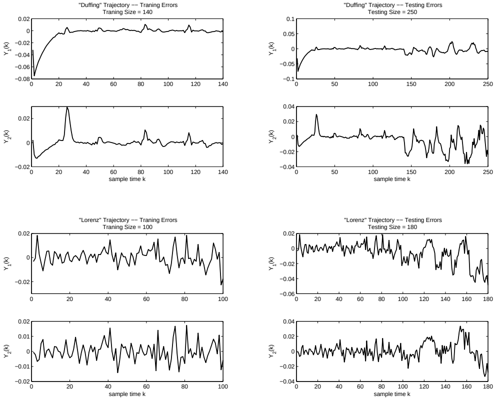
$$\dot{\mathbf{x}} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \mathbf{x} + \begin{bmatrix} a(t) & 1 \\ b(t) & 0 \end{bmatrix} \mathbf{u} + \mathbf{w} \qquad (8)$$

- $a(t) = 1 + \sin(2\pi t/1500)$ and $b(t) = \cos(2\pi t/1500)$;
- $\mathbf{w} \sim \mathcal{N}(0, 0.01)$;
- 2 trajectories: "Duffing" and "Lorenz";
- 50 repeats;
- Observations generated by using a linear MPC;
- Compare to the Nelder-Mead method (derivative-free);

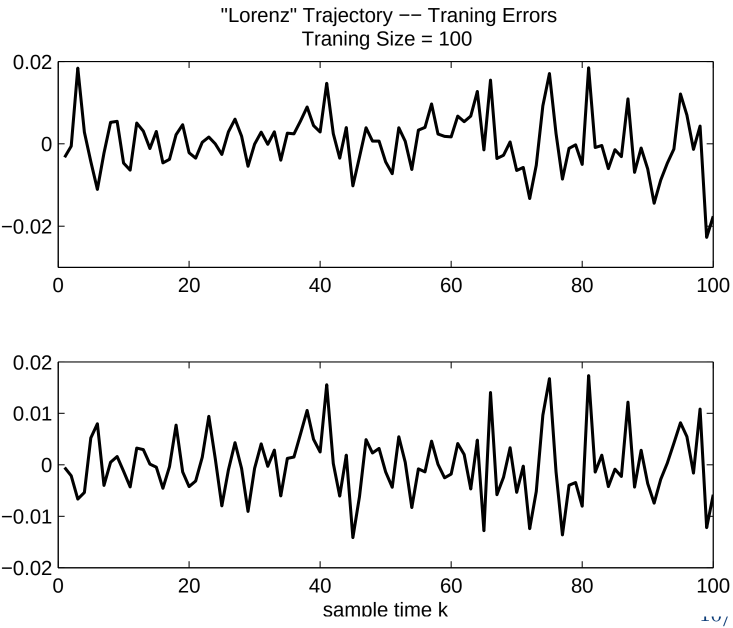
Numerical Simulation

Modelling Results:

	Training MSE	Testing MSE
“Duffing”	2.2338×10^{-4}	3.4091×10^{-4}
“Lorenz”	8.7979×10^{-5}	4.0674×10^{-4}



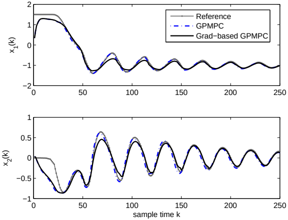
Numerical Simulation



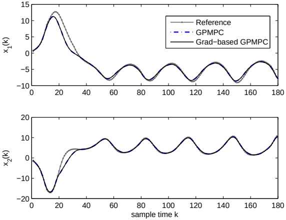
Numerical Simulation

Control Results:

	Approximately Required Time(s)	
	Grad-GPMPC	GPMPC
"Duffing"	30	70
"Lorenz"	16	37

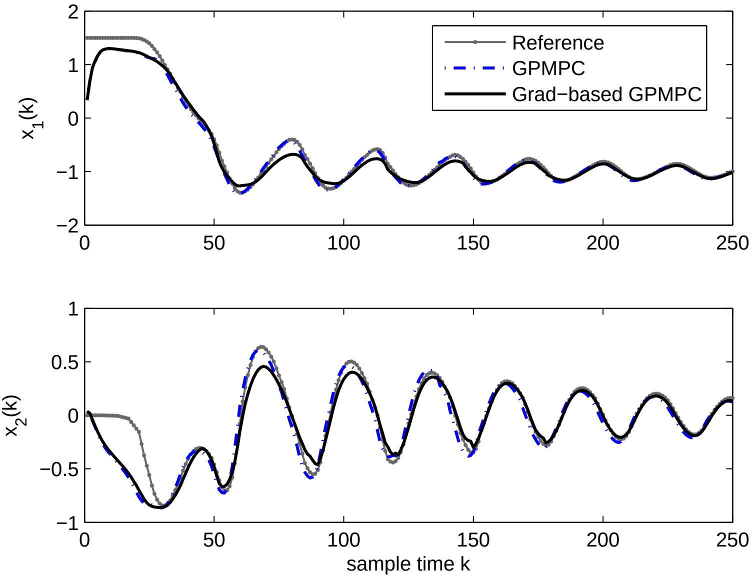


(i) "Duffing" Trajectory



(j) "Lorfenz" Trajectory

Numerical Simulation



Conclusion

Conclusions:

- Proposed GP based MPC performs well in tracking problems;
- Gradient based solution works faster;

Conclusion

Conclusions:

- Proposed GP based MPC performs well in tracking problems;
- Gradient based solution works faster;

Future Works:

- Constrained MPC for Nonlinear unknown systems;
- Guaranteed stability and robustness against uncertainties;

Thanks!

Questions?