Learning to Control using a Convex Combination of Controllers

Athindran Ramesh Kumar and Peter J. Ramadge

Department of Electrical and Computer Engineering
Princeton University
ACC 2021
arkumar@princeton.edu

April 14, 2021

Motivation - Conservatism based on confidence

Conservatism in the face of uncertainty

How do we learn to drive? Play it safe until we understand how the car behaves.

- Start with a conservative controller
- Transition to a aggressive controller - based on current model uncertainty
- Will a convex combination of controllers work? Agarwal et al. 2019; Singh et al. 1994
- Tune the weights - devise a scheme

CVM control - Learning to switch

Problem Setup

We assume linear system with state measurable

$$x_{t+1} = A_t x_t + B_t u_t$$

- Only estimates $\hat{A}_t$, $\hat{B}_t$ with $A_t = \hat{A}_t + \triangle_{At}$ and $B_t = \hat{B}_t + \triangle_{Bt}$
- Have state-feedback controllers K_1, K_2
- One of them will aggressively focus on tracking the reference excitation, other has higher stability margin and robustness
- Key problem: What does it mean to convex combine K_1, K_2?

Convex combination of state feedback controllers

Bad news

- Spectral radius non-convex non-smooth
- Stability not guaranteed

Good news - still a lot of structure in the problem for LTI-SISO systems

$$\begin{aligned} K_3 &= \alpha K_1 + (1-\alpha)K_2 \\ L_3(i\omega) &= K_3(i\omega I - A)^{-1}B \\ &= \alpha L_1 + (1-\alpha)L_2 \qquad (1) \\ p_3(z) &= \det(zI - A) + K_3 \text{adj}(zI - A)B \\ &= \alpha p_1(z) + (1-\alpha)p_2(z) \\ &= \alpha p_1(z)\left(1 + \frac{1-\alpha}{\alpha}\frac{p_2(z)}{p_1(z)}\right) \qquad (2) \end{aligned}$$

Root Locus and Nyquist Plot

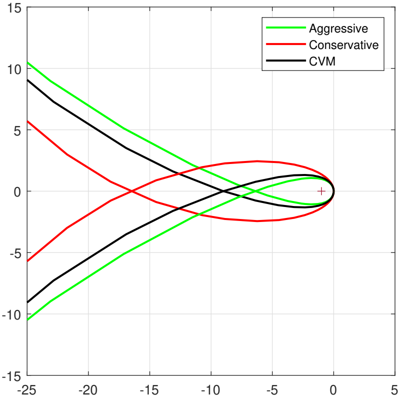


Figure: Nyquist plot of CVM control

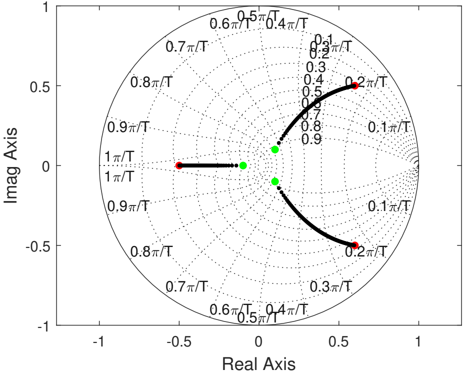


Figure: Root locus plot varying α

Notation

$f : \mathbb{R}^{n \times n} \to \mathbb{R}$ is twice differentiable at $M \in \mathbb{R}^{n \times n}$
Gradient $\nabla f(M)$ of f such that $Df(M)[H] = \langle \nabla f(M), H \rangle$

Derivative of eigenvalue

λ_1 -nonrepeated non-zero maximal eigenvalue of M with $Mu_1 = \lambda_1 u_1$
v_1 - eigenvector of M^* for eigenvalue $\bar{\lambda}_1$.
Geometric multiplicity of λ_1 one
-> $\rho(M) = |\lambda_1(M)|$ infinitely differentiable[1]:

$$D\lambda_1(M)[H] = \frac{v_1^* H u_1}{v_1^* u_1} \tag{3}$$

Our approach to ensure stability - **use gradients of the spectral radius**

[1] Magnus 1985.

- **Our method** - Use gradients to tune spectral radius to desired ρ_d
- A_t, B_t be the system at time t
- $K_t = (1-\alpha_t)K_1 + \alpha_t K_2$
- Estimates $\hat{A}_t, \hat{B}_t$ with $\triangle_{At}, \triangle_{Bt}$ the additive errors in the estimates
- v_K, u_K eigenvectors of $(\hat{A}_t - \hat{B}_t K_{t-1})^T$ and $\hat{A}_t - \hat{B}_t K_{t-1}$ for the eigenvalue with the maximum radius

$$\rho_t \approx \left|\lambda_1(\hat{A}_t - \hat{B}_t K_{t-1}) + D_A(\lambda_1)[\triangle_{At}] + D_B(\lambda_1)[\triangle_{Bt}] + \frac{d\lambda_1}{d\alpha}(\triangle\alpha_t)\right|$$

$$\approx \left|\lambda_1(\hat{A}_t - \hat{B}_t K_{t-1}) + \frac{v_K^H \triangle_{At} u_K}{v_K^H u_K} + \frac{v_K^H \triangle_{Bt} K_{t-1} u_K}{v_K^H u_K}\right.$$

$$\left. + \frac{v_K^H B_{et}(K_2 - K_1) u_K}{v_K^H u_K}(\alpha_t - \alpha_{t-1})\right|$$

$$\lesssim \rho_{t-1} + s_K \|\triangle_{A,t}\|_2 + s_K \|K_{t-1}\|_2 \|\triangle_{B,t}\|_2 + s_\alpha(\alpha_t - \alpha_{t-1})$$

Therefore,

$$\rho_t \leq \rho_{t-1} + s_K \|\triangle_{At}\|_2 + s_K \|K_{t-1}\|_2 \|\triangle_{Bt}\|_2 + s_\alpha(\alpha_t - \alpha_{t-1}) \tag{4}$$

where $s_K = \frac{\|v_K^H\| \|u_K\|}{|v_K^H u_K|}$, $s_\alpha = \text{Re}\left(\frac{\bar{\lambda}_K}{|\lambda_K|} \frac{v_K^H \hat{B}_t(K_2 - K_1)u_K}{v_K^H u_K}\right)$

Let $\|\triangle_{At}\|_2 \leq \delta_{At}$ and $\|\triangle_{Bt}\|_2 \leq \delta_{Bt}$

- Compute an aggressive controller K_1 and a robust controller K_2 at a lower frequency
- At each time perform the following update (η_t - learning rate):

$$\rho_c = \rho_{t-1} + s_K \delta_{At} + s_K \|K_{t-1}\| \delta_{Bt} \tag{5}$$

$$\alpha_t = \alpha_{t-1} + \eta_t s_\alpha (\rho_d - \rho_c) \tag{6}$$

$$K_t = (1 - \alpha_t) \times K_1 + \alpha_t \times K_2 \tag{7}$$

Theorem

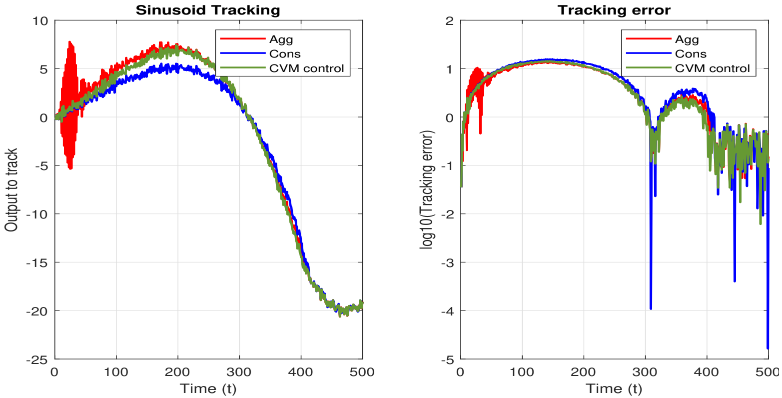
Assume that the model estimates and the generated controllers are bounded, $\max\{\|K_2\|, \|K_1\|\} \leq c_k$, $\max\{\|\hat{B}_t\|, \|B_t\|\} \leq c_b$. *Then,*

1. *If* $\rho_d < \rho_c$ *and* $\delta_{At} \leq \epsilon_a$, $\delta_{Bt} \leq \epsilon_b$ *and* $\eta_t < c$ *then* $\rho_t \leq \rho_{t-1}$. *Also, if* $s_\alpha \neq 0$, $\rho_t < \rho_{t-1}$.
2. *If* $\rho_d > \rho_c$ *and* $\delta_{At} \leq \epsilon_a$, $\delta_{Bt} \leq \epsilon_b$ *and* $\eta_t < c$ *then* $\rho_t \geq \rho_{t-1}$. *Also, if* $s_\alpha \neq 0$, $\rho_t > \rho_{t-1}$.

Proof idea: Make assumptions to ensure differentiability since spectral radius is not convex. Use the second order taylor expansion and bound the second derivative.
Meaning of the theorem: If the algorithm is not stuck in a local minima, it will move in the desired direction.

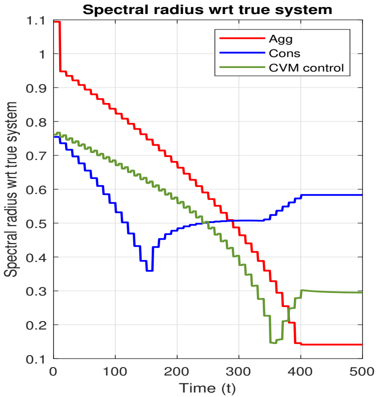
CVM control - Numerical Study

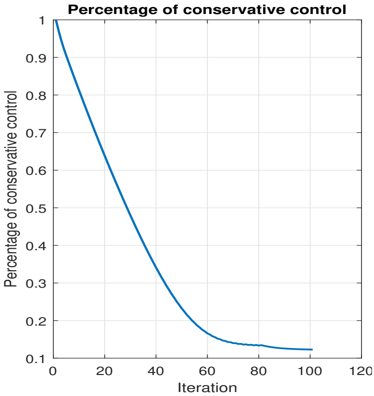
Experiment 1 - 3 dimensional SISO LTV system. Artificially vary the estimates from some initial offset to the true values



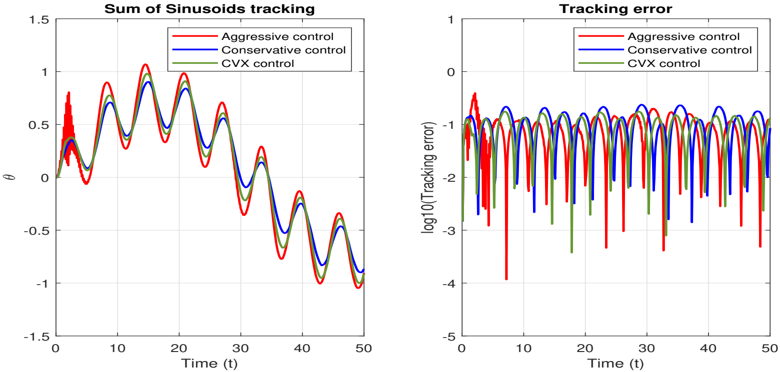
Experiment 1

Experiment 1 - 3 dimensional SISO LTV system. Artificially vary the estimates from some initial offset to the true values





Spring-beam - Nonlinear simulation with Euler's approximation.



Refer paper for more experiments on MIMO systems

Conclusions

Major takeaways:

- Adaptive control algorithm to transition from conservative to aggressive control
- Approximate linear models - require robust conservative controller
- Usefulness of method - use machine learning to refine system model
- Use model certainty to tune controller
- Can mimic human conservatism in controllers