Concentration Bounds for Single Parameter Adaptive Control

Statistical Learning Theory Final Project

Heng-Sheng Chang

May 10, 2019

1 Intuition

Example 1.1 Consider a first order linear system $\dot{x} = ax + u$ where $a \in \mathbb{R}$ is fixed but unknown. Together with the adaptive control law

$$u = -kx, \quad \dot{k} = \gamma x^2, \quad \gamma > 0 \tag{1}$$

and let $x_1 = x$ and $x_2 = k$, we then have a closed-loop system

$$\dot{x}_1 = -(x_2 - a)x_1 \tag{2a}$$

$$\dot{x}_2 = \gamma {x_1}^2 \tag{2b}$$

where the equilibrium set $x_1 = 0$ is stable which can be shown by using LaSalle's theorem and choosing the Lyapunov function as

$$V(x_1, x_2) = \frac{{x_1}^2}{2} + \frac{(x_2 - b)^2}{2\gamma}, \quad b > a \tag{3}$$

△

From example 1.1 [2], we see that the system can be stablized by using the adaptive control law but the parameter a remain unknown. The purpose of this report is to analyse the transient dynamics of parameter estimation using an empirical risk minimization algorithm, the least square estimate.

2 Introduction

In this report, the transient response of parameter estimation of any control law are carried out using statistical concentration bounds. In addition, this paper also bounds the regret that is the difference between the loss of self-tunning adaptive controller and the loss of the best controller given full knowledge of the plant. Though in [4] it is limited to scalar system with an unknown state parameter, this method can be generalized to multivariable system with unknown states and control paramters as stated in [1] and [3].

The reason of finding a statistical bound on parameter estimation is that it helps to show the stability, robustness and performance of the control law that use the result of parameter estimation. The bound of the error of parameter estimation are used to be stated as the assumptions in the theories of analysising stability, robustness and performance of the system. Now, it provides an upper bound on the parameter estimation and thus it can be further applied in the theories of stability, robustness and performance of the system and might offer a better, more efficient or more robust control law.

3 Finite Time Estimation Error Bounds

To further simplify the question, we consider a scalar linear discrete time system

$$x_{k+1} = ax_k + u_k + w_k, \quad x_0 = 0 \tag{4}$$

where the parameter $a \in \mathbb{R}$ is fixed but as well unknown, similar in example 1.1 and $w_0, w_1, \ldots, w_t$ are independent zero mean sub-Guassian random variables with variance proxy σ^2. For any control input u_k, we have an emirical risk minimization algorithm to estimate the parameter a with the method of least square estimation:

$$\hat{a}_t = \arg\min_{\hat{a}\in\mathbb{R}} \sum_{k=0}^{t-1} (\hat{a}x_k - (x_{k+1} - u_k))^2 = \frac{\sum_{k=0}^{t-1} (x_{k+1} - u_k)x_k}{\sum_{k=0}^{t-1} {x_k}^2} \tag{5}$$

The first result in [4] gives probability bounds on the estimation error $\hat{a}_t - a$ which are stated as follows and graphed in Figure 1:

Theorem 3.1. *Define the dynamics of x_k and estimation $\hat{a}_t$ according to 4 and 5. Suppose for every k that u_k is a measurable function of $x_0, x_1, \ldots, x_k$, then for any $\varepsilon > 0$*

$$\mathsf{P}\{|\hat{a}_t - a| > \varepsilon\} \leq 2(1+\varepsilon^2)^{-t/2} \tag{6}$$

$$\mathsf{E}\left[(\hat{a}_t - a)^{2m}\right] \leq \frac{2^{m+1}m!}{(t-2)(t-4)\ldots(t-2m)}, \quad \forall\, t \geq 2m+1 \tag{7}$$

Proof. Shown in [4] Theorem 4. □

Remark 3.1. *Note that for any given error boundary $\varepsilon > 0$, the probability of error exceeding ε decays at least as fast as the exponential rate and it is independent of the control law determining u_k.*

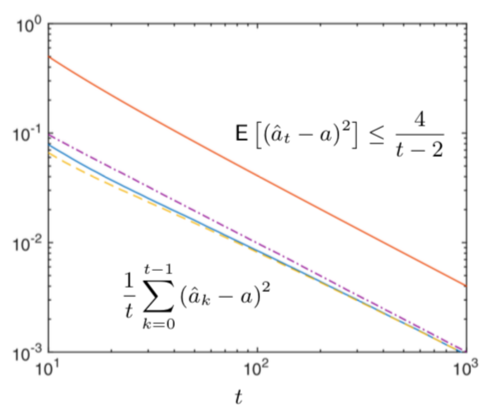

Figure 1: The upper red solid line is the upper bound $\dfrac{4}{t-2}$ of $\mathsf{E}\left[(\hat{a}_t - a)^2\right]$ by Theorem 3.1. The other 3 lines are the averages of empirical loss $\dfrac{1}{t}\sum_{k=0}^{t-1} (\hat{a}_k - a)^2$ over 10^5 simulations with the parameter $a = 0, 1, 2$ respectively.

4 A Regret Bound for Self-Tunning Control

Theorem 4.1. *Define the dynamics of x_k and estimation $\hat{a}_t$ according to 4 and 5 together with the feedback law $u_k = -\hat{a}_k x_k$ for $k \geq 2$ and with $x_0 = 0$. We have the following closed-loop system*

$$\begin{cases} x_1 = w_0, \quad x_2 = ax_1 + w_1 \\ x_{t+1} = (a - \hat{a}_t)x_t + w_t = \dfrac{\sum\limits_{k=1}^{t-1} w_k x_k}{\sum\limits_{k=1}^{t-1} {x_k}^2} x_t + w_t, \quad t \geq 2 \end{cases} \tag{8}$$

Then

$$\mathsf{E}\left[{x_t}^2\right] \leq \frac{704(t-2)^2}{(t-9)^3} + \sigma^2, \quad \forall t \geq 10 \tag{9}$$

Proof. Shown in [4] Theorem 6. □

Remark 4.1. *The feedback controller $u_k = -\hat{a}_k x_k$ is called the self-tunning control. If the plant is well known that means the parameter a is known, than the best feedback controller is $u_k = -ax_k$ with the cost $\mathsf{E}\left[x_t^2\right] = \mathsf{E}\left[{w_{t-1}}^2\right] = \sigma^2$. Therefore, the value $\dfrac{704(t-2)^2}{(t-9)^2}$ is an upperbound for the regret on self-tunning control law. The upper bound are graphed in Figure 2.*

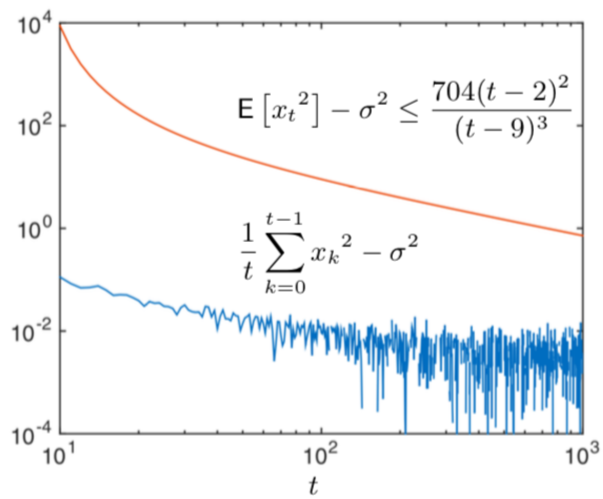

Figure 2: The upper red solid line is the upper bound $\dfrac{704(t-2)^2}{(t-9)^3}$ of the risk by Theorem 4.1. The other line is the average performance difference between the self-tunning control ($u_k = -\hat{a}_k x_k$) and the best controller given full knowldege of the plant ($u_k = -ax_k$) over 10^5 simulations with the parameters $a = 1$ and $\sigma^2 = 1$.

5 Extensions

5.1 Unknown Control Parameter

The results can be extended from an unknown state parameter to an unknown control parameter. Consider the system

$$x_{k+1} = x_k + bu_k + w_k, \quad x_0 = 0 \tag{10}$$

We can estimate the parameter b with least square estimation and it gives us

$$\hat{b}_t = \frac{\sum\limits_{k=0}^{t-1}(x_{k+1} - x_k)u_k}{\sum\limits_{k=0}^{t-1} {u_k}^2} \tag{11}$$

There should have similar results related to theorem 3.1 and theorem 4.1.

5.2 Multivariables System

To extend the system from scalar to multivariables system, we need a restriction on input control law [1], to use least square estimation and have a error bound on the error of estimation. Consider the system

$$x_{t+1} = Ax_t + Bu_t + w_t, \quad x_0 = 0 \tag{12}$$

where u_t and w_t are i.i.d $\mathcal{N}(0, {\sigma_u}^2 I_p)$ and $\mathcal{N}(0, {\sigma_w}^2 I_n)$ random vectors. Then the least square estimator

$$(\hat{A}_t, \hat{B}_t) := \underset{(\hat{A}, \hat{B})}{\arg\min} \frac{1}{2} \sum_{k=0}^{t-1} \|\hat{A}x_t + \hat{B}u_t - x_{t+1}\|^2 \tag{13}$$

will have the error bound with probability at least $1 - \delta$

$$\|\hat{A}_t - A\| \leq \frac{16\sigma_w \sqrt{(n+2p)\log(18/\delta)}}{\sqrt{\min \lambda(\sigma_u^2 G_t G_t^* + \sigma_w^2 F_t F_t^*)}} \tag{14}$$

and

$$\|\hat{B}_t - B\| \leq \frac{16\sigma_w}{\sigma_u} \sqrt{(n+2p)\log(18/\delta)} \tag{15}$$

where G_t and F_t are defined as

$$G_t := \begin{bmatrix} A^{t-1}B & A^{t-2}B & \dots & AB & B \end{bmatrix} \text{ and } F_t := \begin{bmatrix} A^{t-1} & A^{t-2} & \dots & I_n \end{bmatrix} \tag{16}$$

That is to say

$$\mathsf{P}\left\{ \left| \hat{\theta}_t(i,j) - \theta(i,j) \right| > \varepsilon \right\} \leq \delta(\varepsilon, t) \tag{17}$$

where $\theta(i,j) := [A\ B]_{i,j}$. These results are proved and stated as Proposition 1.1 in [1] and in Section 4 in [3].

6 Conclusions and Future Works

These papers show analysis on parameter estimation. [4] shows error bounds on the estimation of sigle state parameter, finite error moment bound in finite time steps and regret upper bound of the self-tunning controller. This unknown state parameter estimation can also be extended to unknown control parameter and even to multivariables system with unknown state and input parameters, like in [1] and [3]. Though the later two papers restricts the control law to some guassian i.i.d. input, they still provide a concentration bound on parameter estimation.

At the end of my presentation, Prof. Hajek proposed a suggestion on doing analysis on the worst case scenario which should be good direction for the future research. In addition, the concentration bounds for the general input has not yet been carried out, so this might be another direction worth trying.

References

[1] Sarah Dean, Horia Mania, Nikolai Matni, Benjamin Recht, and Stephen Tu. On the sample complexity of the linear quadratic regulator. *arXiv preprint arXiv:1710.01688*, 2017.

[2] Hassan K Khalil and Jessy W Grizzle. *Nonlinear systems*, volume 3. Prentice hall Upper Saddle River, NJ, 2002.

[3] Dhruv Malik, Ashwin Pananjady, Kush Bhatia, Koulik Khamaru, Peter L Bartlett, and Martin J Wainwright. Derivative-free methods for policy optimization: Guarantees for linear quadratic systems. *arXiv preprint arXiv:1812.08305*, 2018.

[4] Anders Rantzer. Concentration bounds for single parameter adaptive control. In *2018 Annual American Control Conference (ACC)*, pages 1862–1866. IEEE, 2018.