Primer on stability and convergence

System:

$$\dot{x} = f(x,u)$$

Is unknown

Task: learn f via an RNN.

Model: $\hat{f}$ common notation of an estimate

Assume 1-layer net

$$\dot{\hat{x}} = \theta^T \varphi(\hat{x}, u)$$

State estimate

weights

activation function

Via an approximation theorem,

$$f(x,u) = \theta^{*T} \varphi(x,u) + e(x,u)$$

Ideal weights

State estimation error:

$$\tilde{x} := \hat{x} - x$$

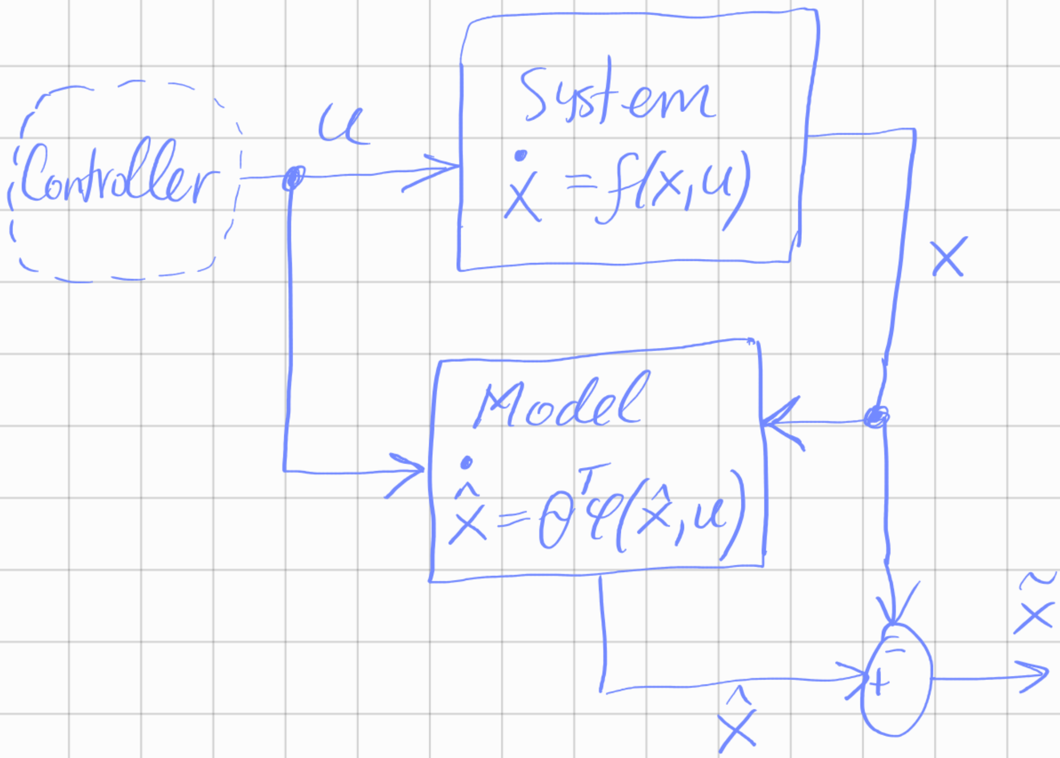
$$\tilde{\theta} := \theta - \theta^*$$

Let's learn the weights via a gradient descent rule:

$$\dot{\theta} = -\alpha\, \varphi(\hat{x}, u)\, \tilde{x}^T$$

$M \times 1 \quad 1 \times n$

learning rate



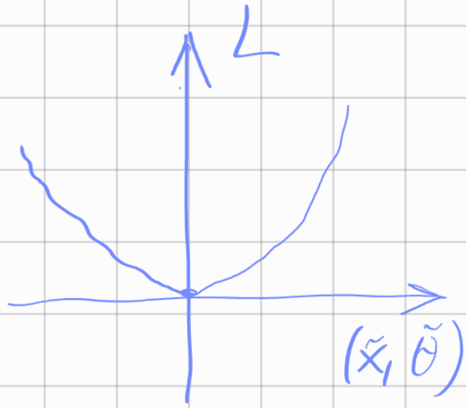
Question: how do the state and weight estimation behave under the chosen update rule?

Candidate Lyapunov function (a.k.a. „abstract energy function"):

$$L(\tilde{x}, \tilde{\theta}) := \frac{1}{2}\tilde{x}^T\tilde{x} + \frac{1}{2\alpha}\operatorname{tr}\left(\tilde{\theta}^T\tilde{\theta}\right)$$

Let's examine the properties of L:

- L is scalar
- $L(0, 0) = 0$
- $L \geq 0 \quad \forall \; \tilde{x}, \tilde{\theta}$



Idea: if L decreases, then

$$\|\tilde{x}\| \longrightarrow 0, \quad \|\tilde{\theta}\| \longrightarrow 0$$

So, our objective is to try to achieve this

$$\text{"} \dot{L} < 0 \text{"} \leftarrow \text{need this}$$

Let's work it out.

$$\dot{L} = \tilde{x}^T \dot{\tilde{x}} + \frac{1}{\alpha} tr\left(\tilde{\theta}^T \dot{\tilde{\theta}}\right)$$

$$= \tilde{x}^T\left(\hat{\theta}^T \varphi(\hat{x}, u) - \theta^{*T} \varphi(x, u) - e(x, u)\right)$$

$$+ \frac{1}{\alpha} tr\left(\tilde{\theta}^T \dot{\hat{\theta}}\right)$$

How? $\rightarrow$ $= \tilde{x}^T\left(\tilde{\theta}^T \varphi(\hat{x}, u) + \theta^{*T}\left(\varphi(\hat{x}, u) - \varphi(x, u)\right) - e\right)$

$$+\frac{1}{\alpha}\operatorname{tr}(\tilde{\theta}^T\dot{\theta})$$

$$= \boxed{\tilde{x}^T\tilde{\theta}^T\varphi(\hat{x},u)} + \tilde{x}^T\theta^{*T}(\varphi(\hat{x},u)-\varphi(x,u)) - \tilde{x}^T e + \boxed{\frac{1}{\alpha}\operatorname{tr}(\tilde{\theta}^T\dot{\theta})}$$

Recall: $\dot{\theta} = -\alpha\,\varphi(\hat{x},u)\tilde{x}^T$

So, $\frac{1}{\alpha}\operatorname{tr}(\tilde{\theta}^T\dot{\theta}) = -\operatorname{tr}(\underset{n\times M}{\tilde{\theta}^T}\,\underset{M\times 1}{\varphi(\hat{x},u)}\,\underset{1\times n}{\tilde{x}^T})$

$$= -\operatorname{tr}(\underset{1\times n}{\tilde{x}^T}\,\underset{n\times M}{\tilde{\theta}^T}\,\underset{M\times 1}{\varphi(\hat{x},u)})$$

$$= \boxed{-\tilde{x}^T\tilde{\theta}^T\varphi(\hat{x},u)}$$

They get cancelled out!

So, we end up having:

$$\dot{L} = \tilde{x}^T\theta^{*T}(\varphi(\hat{x},u)-\varphi(x,u)) - \tilde{x}^T e$$

Let's assume φ is Lipschitz-cont.

That mean

$$\varphi(\hat{x},u)-\varphi(x,u) \le \underset{\uparrow}{\operatorname{Lip}(\varphi)}\,\|\tilde{x}\|$$

Lipschitz constant of φ

$\boxed{F \text{ is Lip.-cont. if } |F(x)-F(y)| \le \operatorname{Lip}(F)\|x-y\|}$

So, with this at hand, we can write:

$$\tilde{x}^T \theta^{*T}(\varphi(\hat{x},u) - \varphi(x,u)) \leq$$
$$\frac{1}{2}\|\theta^*\|^2 \|\tilde{x}\|^2 + \frac{1}{2} Lip^2(\varphi) \|\tilde{x}\|^2$$

How did we get this?

Hint: $a^T b \leq \frac{1}{2}\|a\|^2 + \frac{1}{2}\|b\|^2$

Check:

$$\underbrace{\tilde{x}^T \theta^{*T}}_{\substack{1\times n \;\; n\times M \\ a}} \underbrace{(\varphi(\hat{x},u) - \varphi(x,u))}_{\substack{M\times 1 \\ b}}$$

Next, (say, $\|\theta^*\|^2 =: \bar{\theta}^{*2} = const$)

$$\dot{L} \leq \frac{1}{2}\bar{\theta}^{*2}\|\tilde{x}\|^2 + \frac{1}{2} Lip^2(\varphi)\|\tilde{x}\|^2 +$$
$$\frac{1}{2}\|\tilde{x}\|^2 + \frac{1}{2}\|e\|^2$$

Assume that x, u were in a compact

Then, $\|e\| \leq \varepsilon = const$

Let's group up terms and write:

$$\dot{L} \leq \left(\frac{1}{2}\bar{\theta}^{*2} + \frac{1}{2}\mathrm{Lip}^2(\varphi) + \frac{1}{2}\right)\|\tilde{x}\|^2 + \frac{1}{2}\varepsilon^2$$

At this point, there is nothing we can do to make if $\dot{L} < 0$,

but we can improve our model like so:

$$\dot{\hat{x}} = \theta^T\varphi(\hat{x}, u) \underbrace{- \lambda\tilde{x}}_{\text{a correction term to help out}}, \quad \lambda > 0$$

Then, repeating all the steps we've done so far would give:

$$\dot{L} \leq \left(\frac{1}{2}\bar{\theta}^{*2} + \frac{1}{2}\mathrm{Lip}^2(\varphi) + \frac{1}{2}\right)\|\tilde{x}\|^2 + \frac{1}{2}\varepsilon^2 - \lambda\|\tilde{x}\|^2$$

Hints: Take

$$\lambda > \frac{1}{2}\bar{\theta}^{*2} + \frac{1}{2}\mathrm{Lip}^2(\varphi) + \frac{1}{2} + 1$$

and so

$$\dot{L} \leq -\|\tilde{x}\|^2 + \frac{1}{2}\varepsilon^2$$

Questions: What can we say about the time evolution of $\|\tilde{x}\|$?
And $\|\tilde{\Theta}\|$?

Assignees:

Oussama Alyounes