Stability Analysis of Discrete time Recurrent Neural Networks

Jayant Singh
Advisor: Dr. Nikita Barabanov

Department of Mathematics
North Dakota State University
Fargo,ND

Outline

- Problem Statement
- Applications of Recurrent Neural Networks (RNN)
- Absolute stability
- Novel Approach-Reduction of dissipativity domain
- Summary

Problem Setting

Consider an example of discrete time RNN

$$
\begin{aligned}
x_1^{k+1} &= \tanh(W_1 x_1^k + V_n x_n^k + b_1) \\
x_2^{k+1} &= \tanh(W_2 x_2^k + V_1 x_1^{k+1} + b_2) \\
&\ldots \\
x_n^{k+1} &= \tanh(W_n x_n^k + V_{n-1} x_{n-1}^{k+1} + b_n)
\end{aligned}
\tag{1}
$$

where x_n^k is the state vector of *nth* layer at step k,
W_n, V_n are weight matrices , and
b_n represents the bias vector.

Problem 1.

Analyse the problem of global asymptotic stability of the RNN described above. (i.e. analyse the problem of global asymptotic stability of equilibrium point of the RNN.)

RNN Applications

The applications of RNN include, but are not limited to

- Voice recognition,
- Pattern recognition, and
- Modelling of non-linear systems

- Consider a discrete time SISO system:

$$\begin{aligned} x^{k+1} &= Ax^k + B\xi^k, \sigma^k = Cx^k \\ \xi^k &= \varphi(\sigma^k) \end{aligned} \tag{2}$$

  where, A ,B, C are matrices , $\varphi(\cdot)$ satisfies some condition.
- Analyze the stability of (2)

Lyapunov function approach

- Consider $V(x) = x^* H x$, where $H = H^* > 0$.
  $\Rightarrow V(x^{k+1}) - V(x^k) = (Ax^k + B\xi^k)^* H (Ax^k + B\xi^k) - (x^k)^* H x^k$
- We want $V(x^{k+1}) - V(x^k) < 0$ for all $(x^k, \xi^k) \neq 0$.s.t. $\xi^k = \varphi(x^k)$,and $\varphi(\cdot)$satisfies some constraint.

Reformulated Problem

Problem 2.

Suppose F is a quadratic function. Moreover, assume there exists matrix L s.t. $A + BL$ is stable(i.e.(A,B) is stabilizable), and $F(x, Lx) \geq 0$. Find necessary and sufficient conditions for the existence of $H = H^* > 0$ s.t.

$$(Ax + B\xi)^* H(Ax + B\xi) - x^* Hx + F(x, \xi) < 0 \tag{3}$$

for all $(x, \xi) \neq 0$.

Solution: Necessary condition is given by $\mathbb{R}e(F((e^{i\omega} I - A)^{-1} Bw, w)) < 0$ for all $\omega \in [0, \pi]$ and $w \neq 0$, called the Frequency domain condition.
And sufficient condition is provided by Kalman Szegö Lemma which says that .....

Lemma 3.

Assume (A, B) is stabilizable. And $\mathbb{R}e(F((e^{i\omega}I - A)^{-1}Bw, w)) < 0$ for all $\omega \in [0, \pi]$ and $w \neq 0$. Then there exists $H = H^* > 0$ s.t.

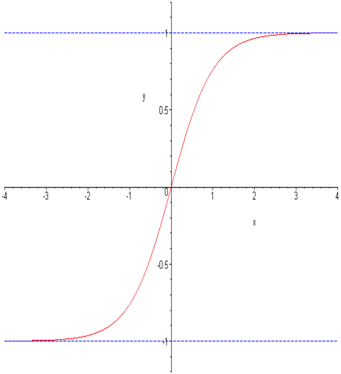
$$(Ax + B\xi)^* H(Ax + B\xi) - x^* Hx + F(x, \xi) < 0$$

for all $(x, \xi) \neq 0$.

As a consequence, there exists $H = H^* > 0$ such that $x^* Hx$ is a Lyapunov function.

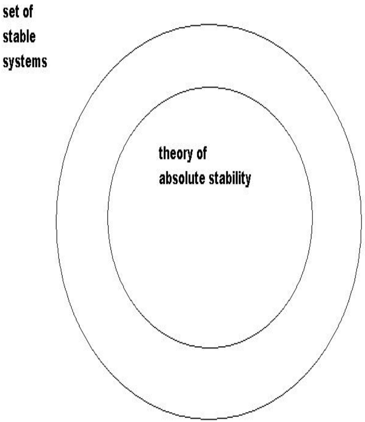
Application to RNN Stability

- In case of RNN, $\varphi(\cdot) = \tanh(\cdot)$,
- $0 \leq \frac{\tanh(\sigma)}{\sigma} \leq 1$(sector condition)
- $\varphi(\sigma)(\sigma - \varphi(\sigma)) \geq 0$ is the quadratic function



Shortcomings in absolute stability approach

- A more general stability criteria should be developed.



Method of reduction of Dissipativity domain

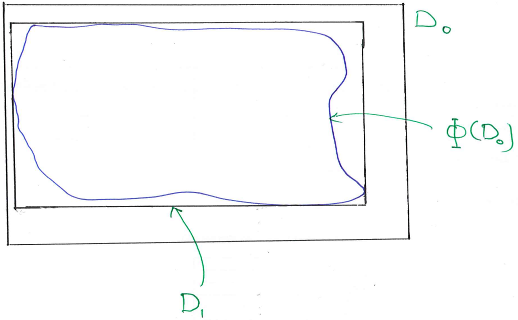
- $x^{k+1} = \phi(x^k)$,$\phi(\cdot)$ is bounded non-linear function.
- Construct $\{D_k\}$ such that $D_{k+1} \subsetneq D_k, \phi(D_k) \subset D_{k+1}$ then $x^k \in D_k$, provided that $x^0 \in D_0$. Thus if $D_k \to 0$, then $x^k \to 0$,as $k \to \infty$.
- $D_{k+1} = \{x \in D_k : f_{k+1,j}(x) \leq \alpha_{k+1,j}, j = 1 \ldots m_{k+1}\}$ where m defines the number of constraints, $f_{k,j}$ defines the linear function, $\alpha_{k+1,j} := \max_{x \in D_k} f_{k,j}(\phi(x))$.

Convex Lyapunov function and constrained optimization problem

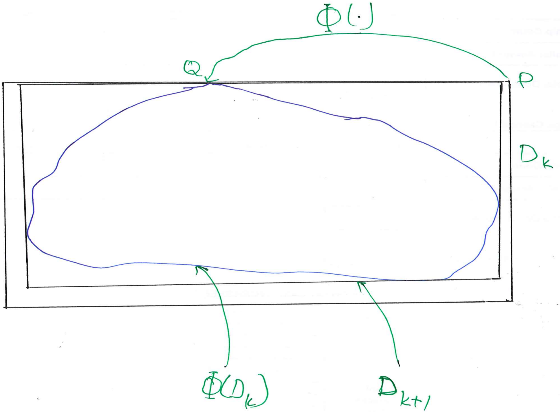
Theorem 4.

Define $\alpha_j^{k+1} = \max_{y \in D_k}(f_j(\phi(y)))$. *Assume system* $x^{k+1} = \phi(x^k)$ *has a convex Lyapunov function. Then there exists linear functions* $f_1, f_2, \ldots f_m$ *such that* $D_{k+1} = \{y : f_j(y) \leq \alpha_j^{k+1}, j = 1 \ldots m\}$, *and* $\{D_k\} \to 0$.

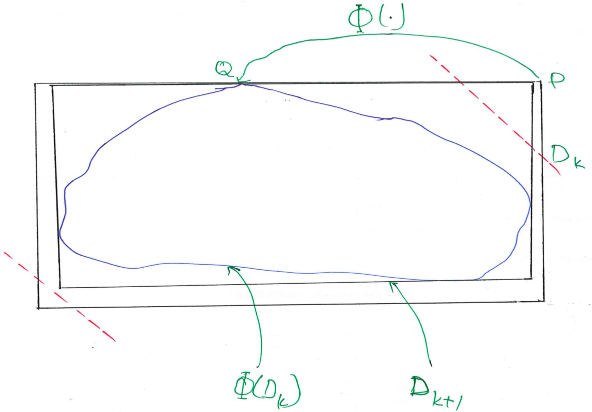
Construction of sets



Addition of new constraints

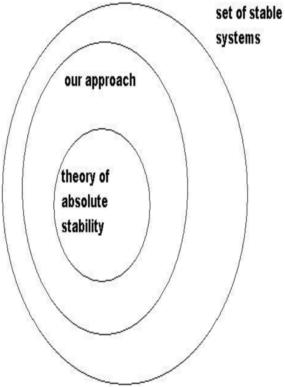


Addition of new constraints contd.



Discussion about new approach

- We assume existence of convex Lyapunov function .On the other hand, in theory of absolute stability approach, we considered the problem of existence of quadratic Lyapunov function(a type of convex function).



Problem Under Consideration

Problem 5.

Given the function $f(x) = \sum_{i=1}^{n} c_i \phi(x_i)$, *where* $c_i \neq 0$ *for all* i. *How to locate the points of local maxima for* $f(\cdot)$ *over a convex set(for our case it is a rectangle)?*

Summary

- The stability problem of discrete time RNN is studied
- Results from theory of absolute stability has been discussed
- The method of reduction of dissipativity domain has been presented
- Some applications of RNN have been mentioned