Stability Analysis of Discrete Time Recurrent Neural Networks

Jayant Singh and Dr. Nikita Barabanov

Department of Mathematics, North Dakota State University, Fargo, ND-58102

Introduction and Problem Setting

Consider an example of discrete time Recurrent Neural Network (RNN)

$$x_1^{k+1} = \tanh(W_1 x_1^k + V_n x_n^k + b_1)$$
$$x_2^{k+1} = \tanh(W_2 x_2^k + V_1 x_1^{k+1} + b_2) \qquad (1)$$
$$\dots$$
$$x_n^{k+1} = \tanh(W_n x_n^k + V_{n-1} x_{n-1}^{k+1} + b_n)$$

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. Analyze the problem of global asymptotic stability of the RNN described above.

Applications of RNN

The applications of RNN include, but are not limited to
(i) Voice recognition,
(ii) Pattern recognition, and
(iii) Modeling of nonlinear systems.

Theory of Absolute Stability

(i) Consider a discrete time single input single output system:

$$x^{k+1} = Ax^k + B\xi^k, \sigma^k = Cx^k, \xi^k = \varphi(\sigma^k) \qquad (2)$$

where, A is a matrix, B, C are vectors, $\varphi(\cdot)$ satisfies sector condition(i.e. $0 \leq \frac{\varphi(\sigma^k)}{\sigma^k} \leq \mu$ for some μ, and for all $\sigma^k \neq 0$.)
(ii) Next, analyze the stability of (2).

Liapunov Function

(i) Consider $V(x) = x^* Hx$, where $H = H^* > 0$.
Then, $V(x^{k+1}) - V(x^k) = (Ax^k + B\xi^k)^* H(Ax^k + B\xi^k) - (x^k)^* Hx^k$.
(ii) We want $V(x^{k+1}) - V(x^k) < 0$ for all $(x^k, \xi^k) \neq 0$. such that $\xi^k = \varphi(x^k)$, and $\varphi(\cdot)$ satisfies sector constraint.

Reformulated Problem

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

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

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

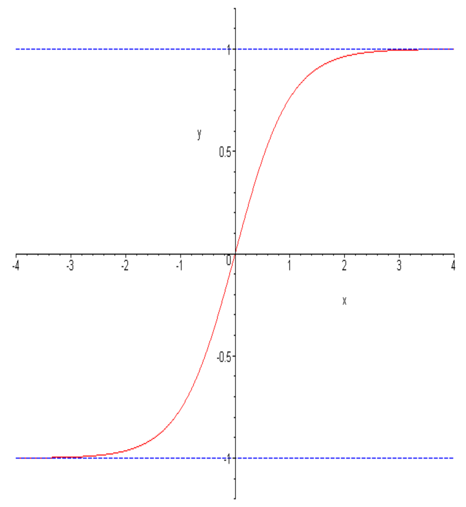
Solution: *Necessity*. $\mathbb{R}\text{e}(F((e^{i\omega}I - A)^{-1}Bw, w)) < 0$ for all $\omega \in [0, \pi]$ and $w \neq 0$, called the Frequency domain condition.

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

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

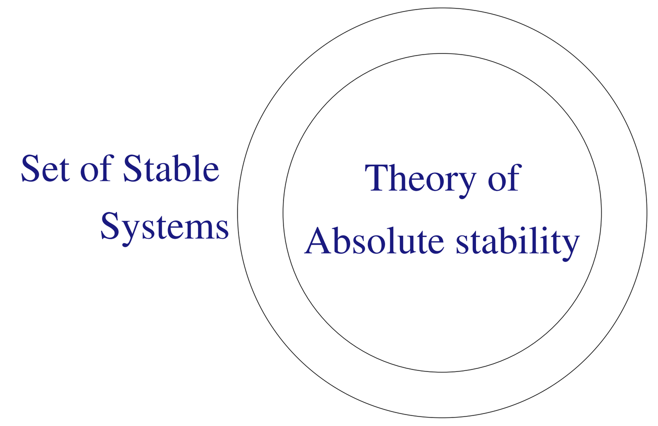
for all $(x, \xi) \neq 0$.

Applications to RNN



(i) In case of RNN, $\varphi(\cdot) = \tanh(\cdot)$ and
(ii) $F(\sigma) = \varphi(\sigma)(\sigma - \varphi(\sigma)) \geq 0$ is quadratic function.

Shortcomings in absolute stability approach



Remark. A more general stability criteria needs to be developed.

Method of reduction of Dissipativity domain

(i) $x^{k+1} = \phi(x^k)$, $\phi(\cdot)$ is bounded non-linear function.
(ii) 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 $k \to \infty$.
(iii) $D_{k+1} := \{x \in D_k : f_{k+1,j}(x) \leq \alpha_{k+1,j}, j = 1 \dots m_{k+1}\}$ where m defines the number of constraints, $f_{k,j}$ defines the linear function, $\alpha_{k+1,j} := \max_{x \in D_k} f_{k,j}(\phi(x))$.

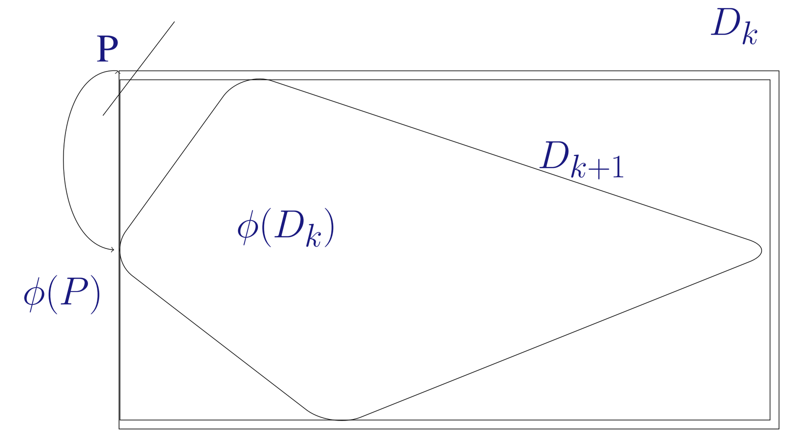
Convex Liapunov function and constrained optimization

Theorem: 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 Liapunov function.Then there exist linear functions $f_1, f_2, \dots f_m$ such that $D_{k+1} = \{y : f_j(y) \leq \alpha_j^{k+1}, j = 1 \dots m\}$, and $\{D_k\} \to 0$.

Algorithm

(i) $D_k := \{x :| x |\leq \alpha_j^k,\ j = 1 \dots m\}$ when $k = 0$.
(ii) Find $\max_{x \in D_k} \langle l_j, (\varphi(x)) \rangle := \alpha_j^{k+1}$ for all j and define $D_{k+1} = \{y : \langle l_j, y \rangle \leq \alpha_j^{k+1}\}$.
(iii) If $\max_j(\alpha_j^k - \alpha_j^{k+1}) > \varepsilon > 0$, increase k by 1 and go to step (ii) and repeat. Here ε is sufficiently small threshold.

Addition of New Constraints



Problem Under Consideration

Problem: 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)?

To this end, we have shown the following result.

Theorem (-): Consider the hyperplane $P = \{x : lx = b\}$ where l is a unit normal vector and $b \in \mathbb{R}$. Suppose the function $f(x) = \sum_{i=1}^n c_i \phi(x_i)$ where $c_i \neq 0$ for all i, defined on P.
Suppose the function $\phi(\cdot)$ satisfies the following conditions:

(i) $\phi(\cdot) \in C^2, \phi(-x) = -\phi(x), \phi'(x) > 0, x\phi''(x) < 0$, for all $x \neq 0$, and $\lim_{x \to \infty} \phi(x) < \infty$. Denote $\psi(\cdot) = (\phi'(\cdot))^{-1}$.

(ii) $x(\ln |\psi'(x)|)'$ is a monotonically increasing function of x.

(iii) Set $h(\beta q_j) = \frac{\psi'(\beta q_j)}{\psi'(\beta q_n)}$. Then $\frac{d}{d\beta}\left[\frac{h'(\beta q_j)}{h'(\beta q_l)}\right] \neq 0$, where $q_j < q_n < q_l$.
(iv) For all $p > q$, we have $\frac{d}{d\beta}\Big(\frac{\psi(\beta p)}{\psi(\beta q)}\Big) < 0$.

(v) For all $x > 0$, we have $\frac{d}{dx}\Big(x\frac{d}{dx}\Big(\frac{\psi(x)}{x\psi'(x)}\Big)\Big) \geq 0$.
Then, the function $f(x)$ has at most one point of local maximum on the hyperplane, P.

Discussion

Remark. We check the necessary condition for existence of convex Liapunov function. This gives a more general stability criteria.