CSCE 790: Neural Networks and Their Applications

AIISC and Dept. Computer Science and Engineering
Email: vignar@sc.edu

© Vignesh Narayanan

October 10, 2023

UNIVERSITY OF
South Carolina

What have we done so far?

- Neural networks - Parametric models for ML
- Multi-layer perceptrons
- Data sampled from some unknown set (space) is given
- Depending on the data, we formalize a learning problem
- We saw supervised learning - classification and regression
- Optimization of the NN parameters happens via training through back-propagation
- Numerical rounding errors and ill-conditioning, vanishing gradients, generalization errors..

Dynamical system (wiki)

- Dynamical system is a system in which a function describes the time dependence of a point in an ambient space
- At any given time, a dynamical system has a state representing a point in an appropriate state space
- This state is often given by a tuple of real numbers or by a vector in a geometrical manifold
- The evolution rule of the dynamical system is a function that describes what future states follow from the current state
- Often the function is deterministic, that is, for a given time interval only one future state follows from the current state
- However, some systems are stochastic, in that random events also affect the evolution of the state variables

Examples

- Discrete-time dynamical systems

$$x(t+1) = F(x(t)), \quad x(0) \in \mathbb{R}^n, \quad t \in \mathbb{Z}^+ = \{0, 1, 2, \ldots\}.$$

- Continuous-time dynamical systems

$$\frac{dx}{dt} = \dot{x}(t) = f(x(t)), \quad x(0) \in \mathbb{R}^n, \quad t \in [0, \infty).$$

Equilibrium point

Consider a (autonomous) dynamical system given by

$$\dot{x}(t) = f(x(t)), \quad x(0) \in \mathbb{R}^n, \quad t \in [0, \infty). \tag{1}$$

Definition

A point $x_e \in \mathbb{R}^n$ is called an *equilibrium point* of the system (3) if

$$x(t) = x_e \implies \dot{x}(t) = f(x_e) = 0, \quad \forall t^+. \tag{2}$$

Fixed point

Consider a (autonomous) dynamical system given by

$$x(t+1) = F(x(t)), \quad x(0) \in \mathbb{R}^n, \quad t \in \{0, 1, 2, \ldots\}. \tag{3}$$

Definition

A point $x_e \in \mathbb{R}^n$ is called an *fixed point* of the system (3) if

$$x(t) = x_e \implies x(t+1) - x(t) = F(x_e) = 0, \quad \forall t. \tag{4}$$

- States are quantities that informs on what the system is doing now
- Since the states are evolving with respect to time, we would like to know what the system will be doing after a long time – (steady state analysis)
- Fixed or Equilibrium points are important in the steady-state analysis of the system

Linear dynamic system

For a linear (time-invariant) system

$$\dot{x}(t) = Ax(t), \quad x(0) \in \mathbb{R}^n, \quad t \in [0, \infty) \tag{5}$$

- Equilibrium point
- $\dot{x}(t) = 0 \implies Ax(t) = 0$
- There can only be one equilibrium point in the nontrivial case (i.e., A is nonsingular) and it is at origin, i.e., $x_e = 0 \in \mathbb{R}^n$.

Steady-state analysis

Ways to understand what the system will be doing after a long time –

- Solve the differential equation $\dot{x}(t) = f(x(t))$ to find $x(t) = \int_t f(x(t))dt$ from some initial state
- Find qualitatively where the system states are evolving towards using graphical techniques

Challenges –

- Integration of the nonlinear function may not be solvable!
- For higher dimensional systems ($x(t) \in \mathbb{R}^n$ and $n > 2$), graphical approach cannot be used

Stability of an Equilibrium Point

Remark – Stability of a system always implies stability of an equilibrium point of the system!

- What is stability? - An important property of dynamic systems..
- Captures the '(in)sensitivity' of the system to (small) perturbations
- Region of attraction?

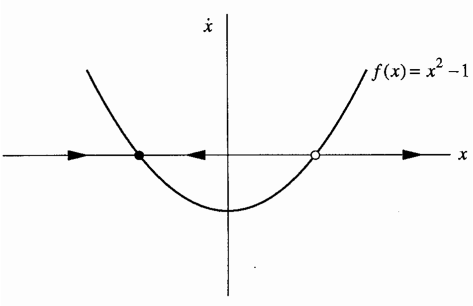


Figure: Stability analysis for the system $\dot{x}(t) = x^2(t) - 1$. Equilibrium points are $\{+1, -1\}$.

Recap: Stability

Definition (Asymptotic Stability)

An equilibrium point x_e is locally asymptotically stable (AS) if there exists a closed and bounded set $S \subset \mathbb{R}^n$ such that, for every $x_0 \in S$, one has $\|x(t) - x_e\| \to 0$ as $t \to \infty$. In other words, the state $x(t)$ converges to x_e. If $S = \mathbb{R}^n$, then the stability is global, i.e., with $x_0 \in \mathbb{R}^n$, the states converge to the equilibrium point.

Definition (Lyapunov Stability (SISL))

An equilibrium point x_e is stable in the sense of Lyapunov (SISL) if for every $\varepsilon > 0$ there exists a $\delta(\varepsilon) > 0$ such that $\|x_0 - x_e\| < \delta$ implies that $\|x(t) - x_e\| < \varepsilon$ for $t \geq t_0$.

Definition (Boundedness)

An equilibrium point x_e is said to be uniformly ultimately bounded if there exists a closed and bounded set $S \subset \mathbb{R}^n$ so that for all $x_0 \in S$ there exists a bound B and a time $T(B, x_0)$ such that $\|x(t) - x_e\| \leq B$ for all $t \geq t_0 + T$.

Illustrations

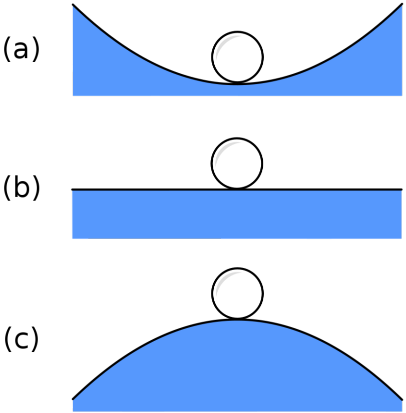


Figure: Types of stability- Illustration via ball with dissipative friction in a gravitational field. (a) Asymptotic stability (b) Stability in the sense of Lyapunov (c) Unstable behavior. **

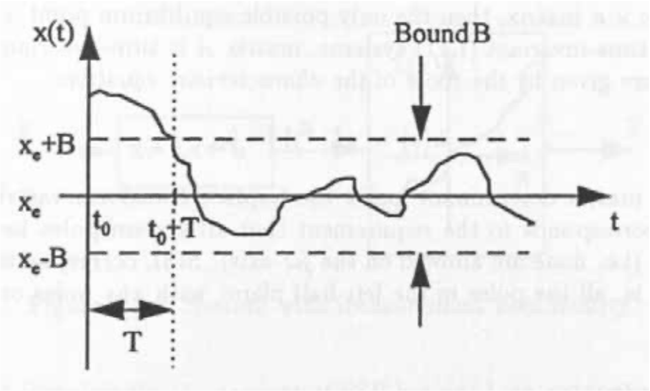


Figure: Boundedness- Illustration via state trajectory of a generic dynamical system (Lewis, 2008).

Lyapunov Stability Theory

- Lyapunov stability theory is a powerful tool for nonlinear stability analysis and control design
- Named after Russian Mathematician Aleksandr Mikhailovich Lyapunov and proposed in his thesis work "The General Problem of Stability of Motion" at Kharkov University in 1892
- The analyze stability of an equilibrium point through a "Lyapunov function"
- Fundamentally, a Lyapunov function can be considered as a proxy to "**generalized Energy function**"

Lyapunov Stability Theory - Preliminaries

- Key idea is to quantify 'the energy function' and see if it is 'increasing or decreasing' as the system evolves
- Components/Requirements :
  1. A function that can be used as a Lyapunov function candidate and
  2. Ability to check increase of decrease in function value
- We can compare 'objects' if they have a scalar numerical representation!
- Consequently:
  1. Lyapunov function should be a scalar function, i.e., $L(x(t)) : \mathbb{R}^n \to \mathbb{R}$, where $x(t)$ is the state of a dynamical system.
  2. We need the notion of positive definite (PD), negative definite (ND), positive semi-definite (PSD) and negative semi-definite (NSD) functions..

Lyapunov Stability Theory - Preliminaries

Definition (PD, ND, PSD, NSD)

Let $L(x) : \mathbb{R}^n \to \mathbb{R}$ be a scalar function such that $L(0) = 0$, and S be a compact subset (closed and bounded subset) of $\mathbb{R}^n$. Then, $L(x)$ is said to be

1. Locally positive definite if $L(x) > 0$, when $x \neq 0$ and for any $x \in S$
2. Locally positive semi-definite if $L(x) \geq 0$, $\forall x \in S$
3. Locally negative definite if $L(x) < 0$, when $x \neq 0$ and $\forall x \in S$
4. Locally negative semi-definite if $L(x) \leq 0$, $\forall x \in S$

The definition can be extended to characterize the function in the entire real domain, i.e., when $S = \mathbb{R}^n$, we replace 'local' with 'global' in the definition.

Lyapunov Function

Consider the (autonomous) nonlinear dynamical system described by

$$\dot{x}(t) = f(x(t)), \quad x(0) \in \mathbb{R}^n, \quad t \in [0, \infty). \tag{6}$$

Definition (Lyapunov function)

A function $L(x(t)) : \mathbb{R}^n \to \mathbb{R}$ with continuous partial derivatives is said to be a **Lyapunov function** for the system (6) if for some compact set $S \in \mathbb{R}^n$, one has locally:

- $L(x(t))$ is positive definite, i.e., $L(x) > 0$ for $x \neq 0$ and for all $x \in S$
- $\dot{L}(x(t))$ is negative semi-definite, i.e., $\dot{L}(x(t)) = (\frac{\partial L}{\partial x})' \frac{dx}{dt} \leq 0$

Lyapunov functions: Remarks

- If we indeed view the Lyapunov function as a generalized 'Energy function', we require the function to satisfy two requirements:
- $L(x(t)) > 0$ for $x \neq 0$, for all $x \in S$
- This implies - finite energy everywhere except at origin
- $\dot{L}(x(t)) \leq 0$
- The second requirement implies that the change in Lyapunov function with respect to time is negative, i.e., the finite energy is non-increasing
- This definition helps understand stability of the system - (Does it specify the associated equilibrium point?)

Main Results

Theorem (SISL)

If there exists a Lyapunov function for the system

$$\dot{x}(t) = f(x(t)), \quad x(0) \in \mathbb{R}^n, \quad t \in [0, \infty), \tag{7}$$

then the equilibrium point is stable in the sense of Lyapunov (SISL).

Theorem (AS)

If there exists a Lyapunov function $L(x(t)) : \mathbb{R}^n \to \mathbb{R}$ *for the system (7) which satisfies* $\dot{L}(x(t)) < 0$ *for* $x \in S \subset \mathbb{R}^n$*, then the equilibrium point is asymptotically stable (AS).*

- If the Lyapunov function is supported on $S = \mathbb{R}^n$ and if $L(x(t)) \to \infty$ as $\|x(t)\| \to \infty$, then the equilibrium point is globally SISL.
- Global AS requires conditions in Theorem (AS) to hold for $S = \mathbb{R}^n$

Example 1

Consider the system:

$$\begin{aligned} \dot{x}_1(t) &= x_1(t)x_2^2(t) + x_1(t)(x_1^2(t) + x_2^2(t) - 3) \\ \dot{x}_2(t) &= -x_1^2(t)x_2(t) + x_2(t)(x_1^2(t) + x_2^2(t) - 3). \end{aligned} \tag{8}$$

Is the equilibrium point at origin stable?

Example 2

Consider the system:

$$\begin{aligned}\dot{x}_1(t) &= x_1(t)x_2^2(t) - x_1(t)(x_1^2(t) + x_2^2(t))\\ \dot{x}_2(t) &= -x_1^2(t)x_2(t) - x_2(t)(x_1^2(t) + x_2^2(t)).\end{aligned} \tag{9}$$

Is the equilibrium point at origin stable?

Example 3

Consider the system:

$$
\begin{aligned}
\dot{x}_1(t) &= x_1(t)x_2^2(t) - x_1(t) \\
\dot{x}_2(t) &= -x_1^2(t)x_2(t).
\end{aligned} \tag{10}
$$

Is the equilibrium point at origin stable?

Example 4

Consider the system:

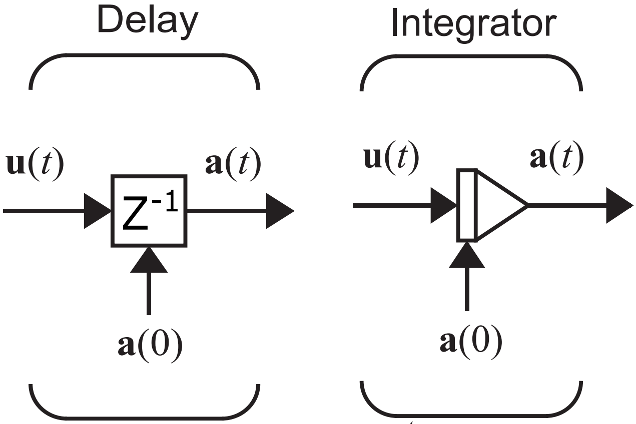
$$\dot{x}(t) = x^2(t) - 1. \tag{11}$$

Are the equilibrium points of this system stable?

Dynamic Neural Networks

- Dynamic networks have memory
- Dynamic networks are networks that contain delays (or integrators, for continuous-time networks) and that operate on a sequence of inputs
- In other words, the ordering of the inputs is important to the operation of the network
- These dynamic networks can have purely feed-forward connections, like the adaptive filters, or they can also have some feedback (recurrent) connections, like the Hopfield network
- Their response at any given time will depend not only on the current input, but on the history of the input sequence

Fundamental Memory Storage Units



$$\mathbf{a}(t) = \mathbf{u}(t-1) \qquad \mathbf{a}(t) = \int_0^t \mathbf{u}(\tau)\, d\tau + \mathbf{a}(0)$$

Figure: Fundamental units storing memory in discrete time and continuous time

Feedback

- Output of an element in the system influences in part the input applied to that particular element, giving rise to one of more closed paths for transmission of signals around the system
- Recurrent networks:

$$y_k(n) = A\dot{x}_j(n)$$

$$\dot{x}_j(n) = x_j(n) + By_k(n)$$

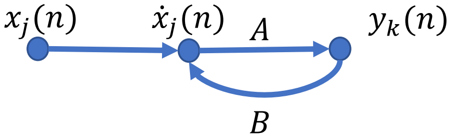


Figure: Recurrent network illustrated via signal flow graph

$$y_k(n) = A\dot{x}_j(n)$$

$$\dot{x}_j(n) = x_j(n) + By_k(n)$$

$$y_k(n) = A[x_j(n) + By_k(n)]$$

$$y_k(n) - ABy_k(n) = Ax_j(n)$$

$$y_k(n) = \frac{A}{1 - AB}x_j(n) \rightarrow \text{Closed-loop operator}$$

Example

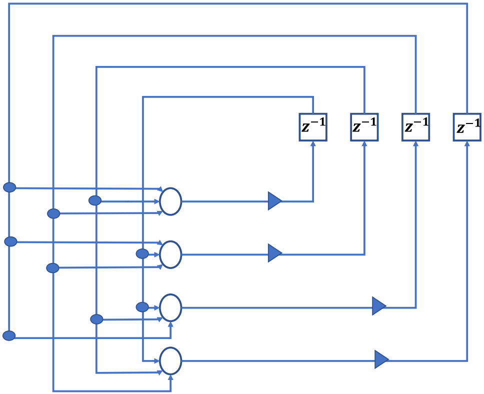


Figure: Example: Recurrent network with no self-loop and no hidden neuron

$$y_1(k) = y_2(k-1) + y_3(k-1) + y_4(k-1)$$
$$y_2(k) = y_1(k-1) + y_3(k-1) + y_4(k-1)$$
$$y_3(k) = y_1(k-1) + y_2(k-1) + y_4(k-1)$$
$$y_4(k) = y_1(k-1) + y_2(k-1) + y_3(k-1)$$