Nonlinear Systems and Control

Mathematical Review & Contraction Mapping

Eugenio Schuster

schuster@lehigh.edu
Mechanical Engineering and Mechanics
Lehigh University

Euclidian Space

Set of n dimensional vectors

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix},$$

where x_i are real numbers. Denoted by $\mathcal{R}^n$.

$$x + y = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{bmatrix}, \quad \alpha x = \begin{bmatrix} \alpha x_1 \\ \alpha x_2 \\ \vdots \\ \alpha x_n \end{bmatrix}, \quad x^T y = \sum_{i=1}^{n} x_i y_i$$

Norms in $\mathcal{R}^n$

The norm $\|x\|$ is a real-valued function with the properties

1. $\|x\| \geq 0 \; \forall x \in \mathcal{R}^n$, with $\|x\| = 0 \iff x = 0$
2. $\|x + y\| \leq \|x\| + \|y\| \; \forall x, y \in \mathcal{R}^n$
3. $\|\alpha x\| = \alpha \|x\| \; \forall |\alpha| \in \mathcal{R}, x \in \mathcal{R}^n$

The p-norm is defined as

$$\begin{aligned} \|x\|_p &= \left(|x_1|^p + \ldots + |x_n|^p\right)^{\frac{1}{p}} \\ \|x\|_\infty &= \max_i |x_i| \end{aligned}$$

All p-norms are equivalent in the sense that

$$c_1 \|x\|_\alpha \leq \|x\|_\beta \leq c_2 \|x\|_\alpha \quad \forall x \in \mathcal{R}^n$$

For $p = 1, 2, \infty$, check c_1 and c_2 in Appendix A.

- In this course we will use notation $|x|_p$.
- Most of the time $|x|_2 \triangleq |x|$.
- We reserve $\|x\|_p$ for $\mathcal{L}_p$ norms on function spaces.

Hölder Inequality

$$|x^T y| \leq |x|_p |y|_q, \qquad \frac{1}{p} + \frac{1}{q} = 1$$

For $p = q = 2$, we get Cauchy-Schwartz inequality.

Mathematical Review

Consider now a matrix $A_{m \times n}$ and the linear mapping $y = Ax$. The induced p-norm of the matrix A is given by

$$|A|_p = \sup_{x \neq 0} \frac{|Ax|_p}{|x|_p} = \max_{|x|_p = 1} |Ax|_p$$

Then,

$$\begin{aligned} |A|_1 &= \max_j \sum_{i=1}^{m} |a_{ij}| && \text{columns} \\ |A|_\infty &= \max_i \sum_{j=1}^{n} |a_{ij}| && \text{rows} \\ |A|_2 &= \sqrt{\lambda_{max}(A^T A)} && \text{maximum eigenvalue} \end{aligned}$$

Converge of Sequences

A sequence $x_0, x_1, \ldots \in \mathcal{R}^n$, which is denoted $\{x_k\}$, converges to x if

$$|x_k - x| \to 0 \text{ as } k \to \infty$$

i.e., a sequence $x_0, x_1, \ldots \in \mathcal{R}^n$, which is denoted $\{x_k\}$, converges to x if

$$\forall \epsilon > 0, \quad \exists N \quad s.t. \quad |x_k - x| < \epsilon \quad \forall k \geq N$$

Mathematical Review

- x is an *accumulation point* of sequence $\{x_k\}$ if exists a subsequence of $\{x_k\}$ that converges to x, i.e., if exists infinite subset $\mathcal{K}$ of the nonnegative integers such that $\{x_k\}_{k\in\mathcal{K}} \to x$.

- A bounded sequence has at least one accumulation point in $\mathcal{R}^n$ (Bolzano-Weiestrass Theorem).

- A sequence of real numbers $\{r_k\}$ is said to be increasing (monotonically increasing or nondecreasing) if $r_k \leq r_{k+1}$. If $r_k < r_{k+1}$, it is said to be strictly increasing. Decreasing (monotonically decreasing or non increasing) and strictly decreasing sequences are defined similar with $r_k \geq r_{k+1}$.

- If a sequence is monotonic and bounded, then it has a limit. Decreasing $+$ bounded from below or increasing $+$ bounded from above $\to$ converges to real number.

Sets

- Set $S \subset \mathcal{R}^n$ is *open* if $\forall x \in S$, $\exists N(x,\epsilon) = \{z \in \mathcal{R}^n | |z - x| < \epsilon\}$ such that $N \subset S$, where $N(x,\epsilon)$ is an ϵ-neighborhood of x.
- Set $S \subset \mathcal{R}^n$ is *closed* if its complement in $\mathcal{R}^n$ is open. Equivalently, S is closed if and only if every convergent $\{x_k\}$ with elements in S converges to a point $x \in S$.
- Set S is *bounded* if exists $r > 0$ such that $|x| < r \ \forall x \in S$.
- Set S is *compact* if it is closed and bounded.
- Point p is *boundary point* of S if every neighborhood of p contains at least one point $\in S$ and one point $\notin S$.
- Set S is *convex* if $\forall x, y \in S$, $\forall \theta \in (0,1)$, $\theta x + (1-\theta)y \in S$.

Mathematical Review

Notation:

- δS denotes the boundary of S (set of boundary points)
- $S - \delta S$ denotes the interior of S
- And open set $\equiv$ interior ($\delta S = 0$)
- $\bar{S} = S \cup \delta S$ denotes the closure of S
- A closed set $\equiv$ closure

Additional Definitions:

- An open set S is *connected* if every pair of points in S can be joined by an arc lying in S.
- A set S is called a *region* if it is the union of an open connected set with some, none, or all boundary points.
- If none of the boundary points are included, the region is called an *open region* or *domain*.

Continuity

- Notation: $f : S_1 \rightarrow S_2$ - Function mapping S_1 into S_2.

- $f : \mathcal{R}^n \rightarrow \mathcal{R}^m$ is *continuous* ($\mathcal{C}^0$) at a point x if $f(x_k) \rightarrow f(x)$ whenever $x_k \rightarrow x$. Equivalently, if

$$\forall \epsilon > 0 \quad \exists \delta > 0 \quad s.t. \quad |x - y| < \delta \Rightarrow |f(x) - f(y)| < \epsilon$$

  We say that f is *uniformly continuous* on S if δ depends on ϵ and not on x.

- $f : \mathcal{R} \rightarrow \mathcal{R}^n$ is *piecewise continuous* on interval $J \subset \mathcal{R}$ if for every bounded subinterval $J_0 \subset J$, f is continuous $\forall x \in J_0$, except, possibly at a finite number of points x_0 where f may be discontinuous. Moreover, $\lim_{h \rightarrow 0} f(x_0 + h)$ and $\lim_{h \rightarrow 0} f(x_0 - h)$ exist. Function has finite jump at x_0.

Differentiability

- Scalar $f : \mathcal{R} \to \mathcal{R}$ is *differentiable* at x if

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

  exists.

- $f : \mathcal{R}^n \to \mathcal{R}^m$ is *continuously differentiable* at x ($\mathcal{C}^1$) if $\partial f_i / \partial x_j$ exists and are continuous at x.

Mathematical Review

- For a function $f : \mathcal{R}^n \to \mathcal{R}$,

$$\frac{\partial f}{\partial x} = \begin{bmatrix} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} & \cdots & \frac{\partial f}{\partial x_n} \end{bmatrix}$$

$$\nabla f(x) = \left(\frac{\partial f}{\partial x}\right)^T = \frac{\partial^T f}{\partial x} \qquad \text{Gradient}$$

- For a function $f : \mathcal{R}^n \to \mathcal{R}^m$,

$$\frac{\partial f}{\partial x} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix} \qquad \text{Jacobian}$$

Mean Value Theorem

Let $f : \mathcal{R}^n \to \mathcal{R}$ be continuously differentiable on $S \subset \mathcal{R}^n$. Let $x, y \in S$ be such that the line segment

$$L(x,y) = \{z | z = \theta x + (1-\theta) y, \quad 0 < \theta < 1\}$$

is in S. Then, $\exists z \in L(x,y)$ such that

$$f(y) - f(x) = \left. \frac{\partial f}{\partial x} \right|_{x=z} (y - x)$$

"The difference of the function can be represented as a linear function."

Example: $f(x) = x^2 \in \mathcal{C}^1$

Implicit Function Theorem

Assume $f(x,y) : \mathcal{R}^n \times \mathcal{R}^m \to \mathcal{R}^n$ is continuously differentiable ($\mathcal{C}^1$) on S. Let $f(x_0,y_0)=0$ and $\frac{\partial f}{\partial x}(x_0,y_0)$ be non-singular at a point $(x_0,y_0) \in S$. Then, exists neighborhoods U of x_0 and V of y_0 such that for each $y \in V$, the equation $f(x,y)=0$ has a unique solution $x \in U$, expressed as $x = g(y)$, $g \in \mathcal{C}^1$ (continuously differentiable).

Gronwall(-Bellman) Inequality

Let

- $\lambda : [a,b] \to \mathcal{R}$ be continuous
- $\mu : [a,b] \to \mathcal{R}$ be continuous and nonnegative

If a continuous function $y : [a,b] \to \mathcal{R}$ satisfies

$$y(t) \leq \lambda(t) + \int_a^t \mu(s)y(s)ds$$

for $a \leq t \leq b$, then on the same interval

1. $y(t) \leq \lambda(t) + \int_a^t \lambda(s)\mu(s)e^{\int_s^t \mu(\tau)d\tau}ds$
2. If $\lambda(t) \equiv \lambda$ (constant), $y(t) \leq \lambda e^{\int_a^t \mu(\tau)d\tau}$
3. If, in addition, $\mu(t) \equiv \mu \geq 0$ (constant), $y(t) \leq \lambda e^{\mu(t-a)}$

Proof: Let

- $z(t) = \int_a^t \mu(s)y(s)ds$
- $v(t) = z(t) + \lambda(t) - y(t) \geq 0 \iff y(t) \leq \lambda(t) + z(t)$

Then,

$$\dot{z} = \mu(t)y(t) = \mu(t)z(t) + \mu(t)\lambda(t) - \mu(t)v(t)$$

This is a scalar linear state equation with the state transition function

$$\phi(t,s) = e^{\int_s^t \mu(\tau)d\tau}$$

Since $z(a) \equiv 0$ (see definition of $z(t)$ above), we have

$$z(t) = \int_a^t \phi(t,s)[\mu(s)\lambda(s) - \mu(s)v(s)]ds$$

The term $\int_a^t \phi(t,s)\mu(s)v(s)$ is nonnegative ($\Rightarrow -\int_a^t \phi(t,s)\mu(s)v(s) \leq 0$). Therefore,

$$z(t) \leq \int_a^t \phi(t,s)\mu(s)\lambda(s)ds$$

Mathematical Review

We can write

$$z(t) \leq \int_a^t e^{\int_s^t \mu(\tau)d\tau} \mu(s)\lambda(s)ds$$

Since $y(t) \leq \lambda(t) + z(t)$, this completes the proof for case 1, i.e.

$$y(t) \leq \lambda(t) + \int_a^t e^{\int_s^t \mu(\tau)d\tau} \mu(s)\lambda(s)ds$$

In the special case when $\lambda(t) \equiv \lambda$, we have

$$\begin{aligned}
\int_a^t e^{\int_s^t \mu(\tau)d\tau} \mu(s)\lambda(s)ds = \lambda \int_a^t e^{\int_s^t \mu(\tau)d\tau} \mu(s)ds &= -\lambda \int_a^t \frac{d}{ds}\left\{ e^{\int_s^t \mu(\tau)d\tau} \right\} ds \\
&= -\lambda e^{\int_s^t \mu(\tau)d\tau} \Big|_{s=a}^{s=t} \\
&= -\lambda + \lambda e^{\int_a^t \mu(\tau)d\tau}
\end{aligned}$$

This proves Case 2, i.e.

$$y(t) \leq \lambda - \lambda + \lambda e^{\int_a^t \mu(\tau)d\tau} = \lambda e^{\int_a^t \mu(\tau)d\tau}$$

The proof of Case 3 (both λ and μ constants) follows by integration.

Contraction Mapping

Fixed Point

Consider a mapping $T(\cdot)$, i.e. an equation of the form $x = T(x)$. A point x^* is said to be a *fixed point* of T if

$$x^* = T(x^*)$$

since T leaves x^* invariant. A classical idea for finding a fixed point is the successive approximation method (also known as Picard iteration):

$$x_{k+1} = T(x_k)$$

The contraction mapping theorem gives sufficient conditions under which there is a fixed point x^* of $x = T(x)$ and the sequence $\{x_k\}$ defined above (successive approximation method) converges to x^*. This theorem holds not only when T is a mapping between *Euclidian* spaces but also when T is a mapping between *Banach* spaces $\subset \mathcal{R}^n$. It is useful to consider the contraction mapping theorem in this more general setting. Therefore, *Banach* spaces need to be introduced first.

Linear Vector Space

Assume $x, y \in \mathcal{X}$ (linear vector space) and $\alpha, \beta \in \mathcal{R}$.

- $x + y \in \mathcal{X}$
- $x + y = y + x$
- $(x + y) + z = x + (y + z)$
- $\exists$ zero vector $0 \in \mathcal{X}$ such that $x + 0 = x \ \forall x \in \mathcal{X}$
- $\alpha x \in \mathcal{X}$
- $0 \cdot x = 0, 1 \cdot x = x$
- $(\alpha\beta)x = \alpha(\beta x)$
- $\alpha(x + y) = \alpha x + \alpha y$
- $(\alpha + \beta)x = \alpha x + \beta x$

Normed Linear Vector Space

A linear space $\mathcal{X}$ is normed if for each $x \in \mathcal{X}$, there is a real-valued norm $\|x\|$ that satisfies

- $\|x\| > 0 \ \forall x \in \mathcal{X} - \{0\}$, $\|0\| = 0$.
- $\|x + y\| \leq \|x\| + \|y\| \ \forall x, y \in \mathcal{X}$.
- $\|\alpha x\| = |\alpha| \|x\| \ \forall x \in \mathcal{X}, \alpha \in \mathcal{R}$.

Convergence

A sequence $\{x_k\} \in \mathcal{X}$, a normed linear space, converges to $x \in \mathcal{X}$ if $\|x_k - x\| \to 0$ as $k \to \infty$.

Closed Set

A set $S \subset \mathcal{X}$ is closed if and only if every convergent sequence with elements in S has its limit in S.

Contraction Mapping

Cauchy Sequence

A sequence $\{x_k\} \in \mathcal{X}$, a normed linear space, is said to be a *Cauchy sequence* if

$$\|x_k - x_m\| \to 0 \text{ as } k, m \to \infty.$$

Every convergent sequence is Cauchy, but not vice versa.

Complete Set

A normed linear space $\mathcal{X}$ is complete if every Cauchy sequence in $\mathcal{X}$ converges to a vector in $\mathcal{X}$.

Banach Space

A complete normed linear space is a *Banach* space.

Contraction Mapping

Example (B.1): Continuous functions $f : [a, b] \to \mathcal{R}^n$, denoted by $\mathcal{C}[a, b]$, with norm

$$\|x\|_{\mathcal{C}} = \max_{t \in [a,b]} \|x(t)\|$$

where the right-hand norm is any p-norm in $\mathcal{R}^n$. To prove that $\mathcal{C}[a, b]$ is a Banach space, we need to prove

1. $\mathcal{C}[a, b]$ is a linear space.
2. $\mathcal{C}[a, b]$ together with $\|x\|_{\mathcal{C}}$ is a normed linear space.
3. $\mathcal{C}[a, b]$ together with $\|x\|_{\mathcal{C}}$ is a complete normed linear space
    - Every Cauchy sequence in $\mathcal{C}[a, b]$ converges to a vector in $\mathcal{C}[a, b]$

Proof:

1. With the sum defined as $(x + y)(t) = x(t) + y(t)$, the scalar multiplication as $(\alpha x)(t) = \alpha x(t)$, and the zero vector as the function that is identically zero on $[a, b]$, $\mathcal{C}[a, b]$ is a linear space over $\mathcal{R}$.
2. It is easy to show that the defined norm satisfies the three properties that are needed to conclude that $\mathcal{C}[a, b]$ is a normed linear space.

Contraction Mapping

3. This is the tricky part ...
   - *Pointwise convergence*: Suppose that $\{x_k\}$ is a Cauchy sequence in $\mathcal{C}[a, b]$. For each fixed $t \in [a, b]$,

   $$\|x_k(t) - x_m(t)\| \leq \|x_k - x_m\|_\mathcal{C} \to 0 \quad \text{as } k, m \to \infty$$

     So $\{x_k(t)\}$ is a Cauchy sequence in $\mathcal{R}^n$. But $\mathcal{R}^n$ with any p-norm is complete because convergence implies componentwise convergence and $\mathcal{R}$ is complete. Therefore, there is a real vector $x(t)$ to which the sequence converges: $x_k(t) \to x(t)$. This proves pointwise convergence.
   - *Uniform convergence in* $t \in [a, b]$: Given $\epsilon > 0$, choose N such that $\|x_k - x_m\|_\mathcal{C} < \epsilon/2$ for $k, m > N$. Then, for $k > N$

   $$\|x_k(t) - x(t)\| \leq \|x_k(t) - x_m(t)\| + \|x_m(t) - x(t)\| \leq \|x_k - x_m\|_\mathcal{C} + \|x_m(t) - x(t)\|$$

     By choosing m sufficiently large (which may depend on t), each term on the right-hand side can be made smaller than $\epsilon/2$; so $\|x_k(t) - x(t)\| < \epsilon$ for $k > N$. Hence, $\{x_k(t)\}$ converges to $x(t)$, uniformly in $t \in [a, b]$.

Contraction Mapping

We finally need to prove that: i- $x(t)$ is continuous, ii- $\{x_k\}$ converges to x in the norm of $\mathcal{C}[a,b]$.

- *Continuity*: Consider

$$\|x(t+\delta)-x(t)\| \leq \|x(t+\delta)-x_k(t+\delta)\|+\|x_k(t+\delta)-x_k(t)+\|x_k(t)-x(t)\|$$

Since $\{x_k(t)\}$ converges uniformly to $x(t)$, given any $\epsilon > 0$, we can choose k large enough to make both the first and third terms on the right-hand side less than $\epsilon/3$. Because $x_k(t)$ is continuous, we can choose δ small enough to make the second term less than $\epsilon/3$. Therefore, $x(t)$ is continuous.

- *Convergence of $\{x_k\}$ to x in the norm of $\mathcal{C}[a,b]$*: Direct consequence of the uniform convergence.

Contraction Mapping Theorem

Let S be a closed subset of a Banach space $\mathcal{X}$ and let T be a mapping that maps S into S. Suppose that

$$\|T(x) - T(y)\| \leq \rho\|x - y\| \quad \forall x, y \in S, 0 \leq \rho < 1.$$

We say in this case that T is a contraction over S. Then,

- there exist a unique vector $x^* \in S$ satisfying $x^* = T(x^*)$.
- the point x^* can be obtained by the method of successive approximation, starting from any arbitrary initial vector in S.

Proof: Select an arbitrary $x_1 \in S$ and define the sequence $\{x_k\}$ by the formula $x_{k+1} = T(x_k)$ (successive approximation). Since T maps S into S, $x_k \in S$ for all $k \geq 1$.

1. The first step is to prove that $\{x_k\}$ is Cauchy. We have

$$\|x_{k+1} - x_k\| = \|T(x_k) - T(x_{k-1})\| \\ \leq \rho\|x_k - x_{k-1}\| \leq \rho^2\|x_{k-1} - x_{k-2}\| \leq \cdots \leq \rho^{k-1}\|x_2 - x_1\|$$

It follows that

$$\begin{aligned} \|x_{k+r} - x_k\| &\leq \|x_{k+r} - x_{k+r-1}\| + \|x_{k+r-1} - x_{k_r-2}\| + \cdots + \|x_{k+1} - x_k\| \\ &\leq [\rho^{k+r-2} + \rho^{k+r-3} + \cdots \rho^{k-1}]\|x_2 - x_1\| \\ &\leq \rho^{k-1} \sum_{i=0}^{r-1} \rho^i \|x_2 - x_1\| \\ &\leq \rho^{k-1} \sum_{i=0}^{\infty} \rho^i \|x_2 - x_1\| = \frac{\rho^{k-1}}{1-\rho}\|x_2 - x_1\| \end{aligned}$$

The right-hand side tends to zero as $k \to \infty$. Thus, the sequence is Cauchy.

Contraction Mapping

2. The second step is to prove that $\{x_k\} \to x^* \in \mathcal{X}$ as $k \to \infty$. This arises from the fact that $\mathcal{X}$ is a Banach space. Moreover, since S is closed, $x^* \in S$.
3. The third step is to show that x^* satisfies $x^* = T(x^*)$. For any $x_k = T(x_{k-1})$, we have

$$\begin{aligned} \|x^* - T(x^*)\| \leq \|x^* - x_k\| + \|x_k - T(x^*)\| &= \|x^* - x_k\| + \|T(x_{k-1}) - T(x^*)\| \\ &\leq \|x^* - x_k\| + \rho\|x_{k-1} - x^*\| \end{aligned}$$

   By choosing k large enough, the right-hand side of the inequality can be made arbitrarily small. Thus, $\|x^* - T(x^*)\| = 0$; that is, $x^* = T(x^*)$.
4. The last step is to show that x^* is the unique fixed point of T in S. Suppose that both x^* and y^* are fixed points. Then,

$$\|x^* - y^*\| = \|T(x^*) - T(y^*)\| \leq \rho\|x^* - y^*\|$$

   Since $\rho < 1$, we have $x^* = y^*$.