EE222 Course Notes

Anmol Parande

Spring 2022 - Professors Shankar Shastry and Koushil Srinath

Disclaimer: These notes reflect EE222 when I took the course (Spring 2022). They may not accurately reflect current course content, so use at your own risk. If you find any typos, errors, etc, please raise an issue on the GitHub repository.

Contents

1 Real Analysis

Definition 1 *The extended real line is the set*

$$\{-\infty\} \cup \mathbb{R} \cup \{\infty\}.$$

Definition 2 *The supremum of a set $S \subset \mathbb{R}$ is a value $a \in \mathbb{R}_e$ such that $\forall s \in S,\ s \leq a$ and if $b \in \mathbb{R}_e$ such that $\forall s \in S,\ s \leq b$, then $a \leq b$.*

Supremum is essentially the "least upper bound" in a set. It always exists, and is called $\sup S$. The opposite of supremum is the infinimum.

Definition 3 *The infinimum of a set $S \subset \mathbb{R}$ is a value $a \in \mathbb{R}_e$ such that $\forall s \in S,\ s \geq a$ and if $b \in \mathbb{R}_e$ such that $\forall s \in S,\ s \geq b$, then $a \geq b$.*

The infinimum is the "greatest upper bound". Like the supremum, it always exists, and it is denoted $\inf S$. Supremum and Infinimum can be applied to scalar function $f : S \to \mathbb{R}$ by letting

$$\sup_{x \in S} f(x) = \sup\{f(x) | x \in S\}.$$

1.1 Norms

Definition 4 *Let V be a vector space of $\mathbb{R}$, then $\|\cdot\| : V \to \mathbb{R}$ is a norm if $\forall \boldsymbol{x}, \boldsymbol{y} \in V, \alpha \in \mathbb{R}$,*

$$\|\boldsymbol{x}\| \geq 0, \qquad \boldsymbol{x} = 0 \Leftrightarrow \|\boldsymbol{x}\| = 0, \qquad \|\alpha \boldsymbol{x}\| = |\alpha| \|\boldsymbol{x}\|, \qquad \|\boldsymbol{x} + \boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\|.$$

Definition 5 *A normed space $(V, \|\cdot\|)$ is a vector space which is equipped with a norm $\|\cdot\| : V \to \mathbb{R}$.*

If we have an operator A which takes vectors from normed space $(X, \|\cdot\|_X)$ and outputs vectors in normed space $(Y, \|\cdot\|_Y)$, then we can define another norm on the vector space of operators from $X \to Y$.

Definition 6 *Let $A : X \to Y$ be an operator between normed spaces $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$, then the induced norm of A is*

$$\|A\|_i = \sup_{\|\boldsymbol{x}\|_X \neq 0} \frac{\|A\boldsymbol{x}\|_Y}{\|\boldsymbol{x}\|_X}$$

The induced norm can be thought of as the maximum gain of the operator.

Definition 7 *Two norms $\|\cdot\|$ and $|||\cdot|||$ on a vector space V are said to be equivalent if $\exists k_1, k_2 > 0$ such that*

$$\forall \boldsymbol{x} \in V,\ k_1\|\boldsymbol{x}\| \leq |||\boldsymbol{x}||| \leq k_2\|\boldsymbol{x}\|$$

If V is a finite dimensional vector space if and only if all norms of V are equivalent.

1.2 Sets

Definition 8 *Let $(V, \|\cdot\|)$ be a normed space, $a \in \mathbb{R}$, $a > 0$, $\boldsymbol{x}_0 \in V$, then the open ball of radius a centered around x_0 is given by*

$$B_a(\boldsymbol{x}_0) = \{\boldsymbol{x} \in V \mid \|\boldsymbol{x} - \boldsymbol{x}_0\| < a\}$$

Definition 9 *A set $S \subset V$ is open if $\forall \boldsymbol{s}_0 \in S,\ \exists \epsilon > 0$ such that $B_\epsilon(\boldsymbol{s}_0) \subset S$.*

Open sets have a boundary which is not included in the set. By convention, we say that the empty set is open.

The opposite of an open set is a closed set.

Definition 10 *A set S is closed if $\sim S$ is open.*

Closed sets have a boundary which is included in the set.

1.3 Convergence

Definition 11 *A sequence of points $\boldsymbol{x}_k$ in normed space $(V, \|\cdot\|)$ converges to a point $\bar{\boldsymbol{x}}$ if*

$$\forall \epsilon > 0,\ \exists N < \infty,\ \text{ such that } \forall k \geq N, \|\boldsymbol{x}_k - \bar{\boldsymbol{x}}\| < \epsilon$$

Convergence means that we can always find a finite time such that after that time, all points in the sequence stay within a specified norm ball.

Definition 12 *A sequence $\boldsymbol{x}_k$ is cauchy if*

$$\forall \epsilon > 0,\ \exists N < \infty \text{ such that } \forall n, m \geq N, \|\boldsymbol{x}_m - \boldsymbol{x}_n\| < \epsilon$$

A Cauchy sequence has a looser type of convergence than a convergent sequence since it only requires all elements to in the sequence to be part of the same norm ball after some time instead of requiring the sequence to get closer and closer to a single point.

Theorem 1 *If $\boldsymbol{x}_n$ is a convergent sequence, then $\boldsymbol{x}_n$ is a also a Cauchy sequence.*

Definition 13 *A normed space $(V, \|\cdot\|)$ is complete if every Cauchy sequence converges to a point in V.*

Because a complete space requires that Cauchy sequences converge, all cauchy sequences are convergent in a complete space. Two important complete spaces are

1. Every finite dimensional vector space
2. $(C[a,b], \|\cdot\|_\infty)$, the set of continuously differentiable functions on the closed interval $[a,b]$ equipped with the infinity norm.

A complete normed space is also called a **Banach Space**.

1.4 Contractions

Definition 14 *A point $\boldsymbol{x}^*$ is a fixed point of a function $P : X \to X$ if $P(\boldsymbol{x}^*) = \boldsymbol{x}^*$.*

Definition 15 *A function $P : X \to X$ is a contraction if $\exists c \in \mathbb{R}, 0 \leq c < 1$ such that*

$$\forall \boldsymbol{x}, \boldsymbol{y} \in X,\ \|P(\boldsymbol{x}) - P(\boldsymbol{y})\| \leq c\|\boldsymbol{x} - \boldsymbol{y}\|$$

Informally, a contraction is a function which makes distances smaller. Suppose we look at a sequence defined by iterates of a function

$$\boldsymbol{x}_{k+1} = P(\boldsymbol{x}_k)$$

where P is a function $P : X \to X$. When does this sequence converge, and to what point will it converge?

Theorem 2 (Contraction Mapping Theorem) *If $P : X \to X$ is a contraction on the Banach space $(X, \|\cdot\|)$, then there is a unique $\boldsymbol{x}^* \in X$ such that $P(\boldsymbol{x}^*) = \boldsymbol{x}^*$ and $\forall \boldsymbol{x}_0 \in X$, the sequence $\boldsymbol{x}_{n+1} = P(\boldsymbol{x}_n)$ converges to $\boldsymbol{x}^*$.*

The contraction mapping theorem proves that contractions have a unique fixed points, and that repeatedly applying the contraction will converge to the fixed point.

1.5 Continuity

Definition 16 *A function $h : V \to W$ on normed spaces $(V, \|\cdot\|_V)$ and $(W, \|\cdot\|_W)$ is continuous at a point $\boldsymbol{x}_0$ if $\forall \epsilon > 0, \exists \delta > 0$ such that*

$$\|\boldsymbol{x} - \boldsymbol{x}_0\|_V < \delta \implies \|h(\boldsymbol{x}) - h(\boldsymbol{x_0})\|_W < \epsilon$$

Continuity essentially means that given an $\epsilon-$ball in W, we can find a $\delta-$ball in V which is mapped to the ball in W. If a function is continuous at all points $\boldsymbol{x}_0$, then we say the function is continuous.

We can make the definition of continuity more restrictive by restraining the rate of growth of the function.

Definition 17 *A function $h : V \to W$ on normed spaces $(V, \|\cdot\|_V)$ and $(W, \|\cdot\|_W)$ is Lipschitz continuous at $\boldsymbol{x}_0 \in V$ if $\exists r > 0$ and $L < \infty$ such that*

$$\forall \boldsymbol{x}, \boldsymbol{y} \in B_r(\boldsymbol{x}_0),\ \|h(\boldsymbol{x}) - h(\boldsymbol{y})\|_W \leq L\|\boldsymbol{x} - \boldsymbol{y}\|_V$$

A good interpretation of Lipschitz Continuity is that given two points in a ball around $\boldsymbol{x}_0$, the slope of the line connecting those two points is less than L. It means that the function is growing slower than linear for some region around $\boldsymbol{x}_0$. Lipschitz continuity implies continuity. If a function is lipschitz continuous with respect to one norm, it is also lipschitz continuous with respect to all equivalent norms.

When the function h is a function on $\mathbb{R}^n$ and is also differentiable, then Lipschitz continuity is easy to determine.

Theorem 3 *For a differentiable function $h : \mathbb{R}^n \to \mathbb{R}^n$,*

$$\exists r > 0, L < \infty, \boldsymbol{x}_0 \in \mathbb{R}^n,\ \forall \boldsymbol{x} \in B_r(\boldsymbol{x}_0), \left|\left|\frac{\partial h}{\partial \boldsymbol{x}}\right|\right|_2 \leq L$$

implies Lipschitz Continuity at $\boldsymbol{x}_0$.

This captures the idea of growing slower than linear in high dimensional space.

Definition 18 *A function $h : \mathbb{R} \to V$ is piecewise continuous if $\forall k \in \mathbb{Z}$, $h : [-k, k] \to V$ is continuous except at a possibly finite number of points, and at the points of discontinuity t_i, $\lim_{s\to 0^+} h(t_i + s)$ and $\lim_{s\to 0^-} h(t_i + s)$ exist and are finite.*

2 Differential Geometry

Definition 19 *$M \subset \mathbb{R}^n$ is a m-dimensional smooth sub-manifold of $\mathbb{R}^n$ if $\forall \boldsymbol{p} \in M$, $\exists r > 0$ and $F : B_r(\boldsymbol{p}) \to \mathbb{R}^{n-m}$ such that*

$$\begin{aligned} M \cap B_r(\boldsymbol{p}) = \{\boldsymbol{x} \in B_r(\boldsymbol{p}) | F(\boldsymbol{x}) = 0\}, \\ F \textit{ is smooth}, \\ \forall \bar{\boldsymbol{x}} \in M \cap B_r(\boldsymbol{p}), \textit{Rank}\left(\left.\frac{\partial F}{\partial \boldsymbol{x}}\right|_{\bar{\boldsymbol{x}}}\right) = n - m \end{aligned}$$

By definition 19, a manifold is essentially defined as the 0-level set of some smooth function F and can be thought of as a surface embedded in a higher dimension.

Definition 20 *The tangent space of a manifold M at $\boldsymbol{p} \in M$ is given by*

$$T_{\boldsymbol{p}}M = \textit{Null}\left(\left.\frac{\partial F}{\partial \boldsymbol{x}}\right|_{\boldsymbol{p}}\right)$$

The tangent space consists of all vectors tangent to the manifold at a particular point $\boldsymbol{p}$.

Definition 21 *The Tangent Bundle of a manifold M is the collection of all tangent spaces*

$$T_M = \bigcup_{\boldsymbol{p} \in M} T_{\boldsymbol{p}}M$$

Definition 22 *A vector field $f : M \to T_M$ on a manifold M is an assignment of each point $\boldsymbol{p} \in M$ to a vector in the tangent space in that point $T_{\boldsymbol{p}}M$.*

Therefore, a vector field can be thought of as a curve through the tangent bundle of a manifold.

Definition 23 *The Lie Derivative of a function V with respect to a vector field f is given by*

$$L_f V = (\nabla_{\boldsymbol{x}} V)^\top f(\boldsymbol{x}).$$

A Lie Derivative is essentially a directional derivative, and it measures how a function changes along a vector field.

Definition 24 *Suppose that $f(\boldsymbol{x})$ and $g(\boldsymbol{x})$ are vector fields. The Lie Bracket of f and g is given by*

$$[f, g] = L_f g - L_g f$$

The Lie Bracket is another vector field, and it essentially measures the difference between moving along vector field f and vector field g across some infinitesimal distance. Another way to think about the Lie Bracket is as a measure of the extent to which f and g commute with each other. The Lie Bracket is also sometimes denoted using the adjoint map

$$\operatorname*{ad}_f g = [f, g].$$

It is helpful when chaining Lie Brackets since we can denote

$$[f, [f, [f, \cdots [f, g]]]] = \operatorname*{ad}_f^i g.$$

Since the Lie Bracket is a vector field, we can look at Lie Derivatives with respect to the Lie Bracket of two vector fields.

Theorem 4 *For a function h and vector fields f and g,*

$$L_{[f,g]}h = L_f L_g h - L_g L_f h$$

We can also use relate repeated Lie Derivatives to doing repeated Lie Brackets.

Theorem 5

$$L_g L_f^i h(\boldsymbol{x}) = 0 \Leftrightarrow L_{\mathrm{ad}_f^i g} h(\boldsymbol{x}) = 0$$

Definition 25 *Suppose $f_1, f_2, \cdots, f_n$ are vector fields. A distribution Δ is the span of the vector fields at each point $\boldsymbol{x}$:*

$$\Delta(\boldsymbol{x}) = \mathit{span}\{f_1(\boldsymbol{x}), f_2(\boldsymbol{x}), \cdots, f_n(\boldsymbol{x})\}.$$

At each point $\boldsymbol{x},\ \Delta(\boldsymbol{x})$ is a subspace of the tangent space at $\boldsymbol{x}$.

Definition 26 *The dimension of a distribution at a point $\boldsymbol{x}$ is given by*

$$\mathit{Dim}\ \Delta(\boldsymbol{x}) = \mathit{Rank}\left(\begin{bmatrix} f_1(\boldsymbol{x}) & \bigg| & f_2(\boldsymbol{x}) & \bigg| & \cdots & \bigg| & f_n(\boldsymbol{x}) \end{bmatrix}\right)$$

Distributions have different properties which are important to look at.

Definition 27 *A distribution Δ is nonsingular, also known as regular, if its dimension is constant.*

Definition 28 *A distribution Δ is involutive if*

$$\forall f, g \in \Delta, \quad [f, g] \in \Delta$$

In involutive distributions, you can never leave the distribution by traveling along vectors inside the distribution.

Definition 29 *A nonsingular K-dimensional distribution $\Delta(\boldsymbol{x}) = span\{f_1(\boldsymbol{x}), \cdots, f_k(\boldsymbol{x})\}$ is completely integrable if $\exists \phi_1, \cdots, \phi_{n-k}$ such that $\forall i, k,\ L_{f_k}\phi_i = 0$ and $\nabla_{\boldsymbol{x}}\phi_i$ are linearly independent.*

It turns out that integrability and involutivity are equivalent to each other.

Theorem 6 (Frobenius Theorem) *A nonsingular Δ is completely integrable if and only if Δ is involutive.*

3 Nonlinear System Dynamics

Consider the nonlinear system

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = f(\boldsymbol{x}, t).$$

f is a vector field which potentially changes with time and governs how the system evolves.

Definition 30 *The system is autonomous if $f(\boldsymbol{x}, t)$ is not explicitly dependent on time t.*

Definition 31 *A point x_0 is an equilibrium point at time t_0 if*

$$\forall t \geq t_0,\ f(\boldsymbol{x}_0, t) = 0$$

Consider a single trajectory $\phi(t, t_0, \boldsymbol{x}_0)$.

Definition 32 *A set S is said to be the $\omega-$limit set of ϕ if*

$$\forall \boldsymbol{y} \in S, \exists t_n \to \infty, \lim_{n\to\infty} \phi(t_n, t_0, \boldsymbol{x}_0) = \boldsymbol{y}$$

Whereas linear systems converge to a single point if they converge at all, nonlinear systems can converge to a set of points. Thus the $\omega-$limit set essentially generalizes the idea of a limit.

Definition 33 *A set $M \subset \mathbb{R}^n$ is said to be invariant if*

$$\forall t \geq t_0,\ \boldsymbol{y} \in M \implies \phi(t, t_0, \boldsymbol{y}) \in M$$

An invariant set is one which a trajectory of the system will never leave once it enters the set. Just like linear systems, non-linear systems can also have periodic solutions.

Definition 34 *A closed orbit γ is a trajectory of the system such that $\gamma(0) = \gamma(T)$ for finite T.*

3.1 Solutions to Nonlinear Systems

Consider the nonlinear system

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = f(\boldsymbol{x}, t), \qquad \boldsymbol{x}(t_0) = \boldsymbol{x}_0 \in \mathbb{R}^n.$$

Definition 35 *A function $\boldsymbol{\Phi}(t)$ is a solution to $\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = f(\boldsymbol{x}, t),\ \boldsymbol{x}(t_0) = \boldsymbol{x}_0$ on the closed interval $[t_0, t]$ if $\boldsymbol{\Phi}(t)$ is defined on the interval $[t_0, t]$, $\frac{\mathrm{d}\boldsymbol{\Phi}}{\mathrm{d}t} = f(\boldsymbol{\Phi}(t), t)$ on the interval $[t_0, t]$, and $\boldsymbol{\Phi}(t_0) = \boldsymbol{x}_0$.*

We say that $\boldsymbol{\Phi}(t)$ is a solution in the sense of Caratheodory if

$$\boldsymbol{\Phi}(t) = \boldsymbol{x}_0 + \int_{t_0}^{t} f(\boldsymbol{\Phi}(\tau), \tau) d\tau.$$

Because the system is nonlinear, it could potentially have no solution, one solution, or many solutions. These solutions could exist locally, or they could exist for all time. We might also want to know when there is a solution which depends continuously on the initial conditions.

Theorem 7 (Local Existence and Uniqueness) *Given* $\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = f(\boldsymbol{x}, t),\ \boldsymbol{x}(t_0) = \boldsymbol{x}_0 \in \mathbb{R}^n$ *where* f *is piecewise continuous in* t *and* $\exists T > t_0$ *such that* $\forall t \in [t_0, T], f$ *is L-Lipschitz Continuous, then* $\exists \delta > 0$ *such that a solution exists and is unique* $\forall t \in [t_0, t_0 + \delta]$.

Theorem 7 can be proved using the Contraction Mapping Theorem (theorem 2) by finding δ such that the function $P : C_n[t_0, t_0 + \delta] \to C_n[t_0, t_0 + \delta]$ given by

$$P(\boldsymbol{\Phi})(t) = \boldsymbol{x}_0 + \int_{t_0}^{t_0+\delta} f(\boldsymbol{\Phi}(\tau), \tau) d\tau$$

is a contraction under the norm $\|\boldsymbol{\Phi}\|_\infty = \sup_{t_0 \leq t \leq t_0+\delta} \|\boldsymbol{\Phi}(t)\|$.

Theorem 8 (Global Existence and Uniqueness) *Suppose* $f(\boldsymbol{x}, t)$ *is piecewise continuous in* t *and* $\forall T \in [t_0, \infty)$, $\exists L_T < \infty$ *such that* f *is* L_T *Lipshitz continuous for all* $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$, *then the nonlinear system has exactly one solution on* $[t_0, T]$.

Once we know that solutions to a nonlinear system exist, we can sometimes bound them.

Theorem 9 (Bellman-Gronwall Lemma) *Suppose* $\lambda \in \mathbb{R}$ *is a constant and* $\mu : [a, b] \to \mathbb{R}$ *is continuous and non-negative, then for a continuous function* $y : [a, b] \to \mathbb{R}$

$$y(t) \leq \lambda + \int_a^t \mu(\tau) y(\tau) d\tau \implies y(t) \leq \lambda exp\left(\int_a^t \mu(\tau) d\tau\right)$$

Another thing we might want to do is understand how the nonlinear system reacts to changes in the initial condition.

Theorem 10 *Suppose the system* $\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = f(\boldsymbol{x}, t),\ \boldsymbol{x}(t_0) = \boldsymbol{x}_0$ *satisfies the conditions of global uniqueness and existence. Fix* $T \in [t_0, \infty]$ *and suppose* $\boldsymbol{x}(\cdot)$ *and* $\boldsymbol{z}(\cdot)$ *are two solutions satisfying* $\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = f(\boldsymbol{x}, t), \boldsymbol{x}(t_0) = \boldsymbol{x}_0$ *and* $\frac{\mathrm{d}\boldsymbol{z}}{\mathrm{d}t} = f(\boldsymbol{z}(t), t),\ \boldsymbol{z}(t_0) = \boldsymbol{z}_0$, *then* $\forall \epsilon > 0, \exists \delta > 0$ *such that*

$$\|\boldsymbol{x}_0 - \boldsymbol{z}_0\| < \delta \implies \|\boldsymbol{x} - \boldsymbol{z}\|_\infty < \epsilon.$$

Theorem 10 is best understood by defining a function $\Psi : \mathbb{R}^n \to C_n[t_0, t]$ where $\Psi(\boldsymbol{x}_0)(t)$ returns the solution to the system given the initial condition. If the conditions of Theorem 10 are satisfied, then the function Ψ will be continuous.

3.2 Planar Dynamical Systems

Planar dynamical systems are those with 2 state variables. Suppose we linearize the autonomous system $\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = f(\boldsymbol{x})$ at an equilibrium point.

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = \left.\frac{\partial f}{\partial \boldsymbol{x}}\right|_{\boldsymbol{x_0}} \boldsymbol{x}$$

Depending on the eigenvalues of $\frac{\partial f}{\partial \boldsymbol{x}}$, the Jacobian, we get several cases for how this linear system behaves. We'll let z_1 and z_2 be the eigenbasis of the *phase space*.

1. The eigenvalues are real, yielding solutions $z_1 = z_1(0)e^{\lambda_1 t}, z_2 = z_2(0)e^{\lambda_2 t}$. If we eliminate the time variable, we can plot the trajectories of the system.

$$\frac{z_1}{z_1(0)} = \left(\frac{z_2}{z_2(0)}\right)^{\frac{\lambda_1}{\lambda_2}}$$

   (a) When $\lambda_1, \lambda_2 < 0$, all trajectories converge to the origin, so we call this a **stable node**.

   (b) When $\lambda_1, \lambda_2 > 0$, all trajectories blow up, so we call this an **unstable node**.

   (c) When $\lambda_1 < 0 < \lambda_2$, the trajectories will converge to the origin along the axis corresponding to λ_1 and diverge along the axis corresponding to λ_2, so we call this a **saddle node**.

2. There is a single repeated eigenvalue with one eigenvector. As before, we can eliminate the time variable and plot the trajectories on the z_1, z_2 axes.

   (a) When $\lambda < 0$, the trajetories will converge to the origin, so we call it an **improper stable node**

   (b) When $\lambda > 0$, the trajetories will diverge from the origin, so we call it an **improper unstable node**

3. When there is a complex pair of eigenvalues, the linear system will have oscillatory behavior. The Real Jordan form of $\frac{\partial f}{\partial \boldsymbol{x}}$ will look like

$$\frac{\partial f}{\partial \boldsymbol{x}} = \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix}.$$

   The parameter β will determine the direction of the trajectories (clockwise if positive).

   (a) When $\alpha < 0$, the trajectories will spiral towards the origin, so we call it a **stable focus**.

(b) When $\alpha = 0$, the trajectories will remain at a constant radius from the origin, so we call it a **center**.

(c) When $\alpha > 0$, the trajectories will spiral away from the origin, so we call it an **unstable focus**.

It turns out that understanding the linear dynamics at equilibrium points can be helpful in understanding the nonlinear dynamics near equilibrium points.

Theorem 11 (Hartman-Grobman Theorem) *If the linearization of a planar dynamical system* $\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = f(\boldsymbol{x})$ *at an equilibrium point* $\boldsymbol{x_0}$ *has no zero or purely imaginary eigenvalues, then there exists a homeomorphism from a neighborhood* U *of* $\boldsymbol{x}_0$ *into* $\mathbb{R}^2$ *which takes trajectories of the nonlinear system and maps them onto the linearization where* $h(\boldsymbol{x_0}) = 0$*, and the homeomorphism can be chosen to preserve the parameterization by time.*

Theorem 11 essentially says that the linear dynamics predict the nonlinear dynamics around equilibria, but only for a neighborhood around the equilibrium point. Outside of this neighborhood, the linearization may be very wrong.

Suppose that we have a simply connected region D (meaning D cannot be contracted to a point) and we want to know if it contains a closed orbit.

Theorem 12 (Bendixon's Theorem) *If* $\mathrm{div}(f)$ *is not identically zero in a sub-region of* D *and does not change sign in* D*, then* D *contains no closed orbits.*

Theorem 12 lets us rule out closed orbits from regions of $\mathbb{R}^2$. If we have a positively invariant region, then we can determine whether it contains closed orbits.

Theorem 13 (Poincare-Bendixson Theorem) *If* M *is a compact, positively invariant set for the flow* $\phi_t(\boldsymbol{x})$*, then if* M *contains no equilibrium points, then* M *has a limit cycle.*

4 Stability of Nonlinear Systems

The equilibria of a system can tell us a great deal about the stability of the system. For nonlinear systems, stability is a property of the equilibrium points, and to be stable is to converge to or stay equilibrium.

Definition 36 *An equilibrium point* $\boldsymbol{x}_e \in \mathbb{R}$ *is a stable equilibrium point in the sense of Lyapunov if and only if* $\forall \epsilon > 0, \exists \delta(t_0, \epsilon)$ *such that*

$$\forall t \geq t_0,\ \|\boldsymbol{x}_0 - \boldsymbol{x}_e\| < \delta(t_0, \epsilon) \implies \|\boldsymbol{x}(t) - \boldsymbol{x}_e\| < \epsilon$$

Lyapunov Stability essentially says that a finite deviation in the initial condition from equilibrium means the resulting trajectory of the system stay close to equilibrium. Notice that this definition is nearly identical to theorem 10. That means stability of an equilibrium point is the same as saying the function which returns the solution to a system given its initial condition is continuous at the equilibrium point.

Definition 37 *An equilibrium point $\boldsymbol{x}_e \in \mathbb{R}$ is an uniformly stable equilibrium point in the sense of Lyapunov if and only if $\forall \epsilon > 0, \exists \delta(\epsilon)$ such that*

$$\forall t \geq t_0,\ \|\boldsymbol{x}_0 - \boldsymbol{x}_e\| < \delta(\epsilon) \implies \|\boldsymbol{x}(t) - \boldsymbol{x}_e\| < \epsilon$$

Uniform stability means that the δ can be chosen independently of the time the system starts at. Both stability and uniform stability do not imply convergence to the equilibrium point. They only guarantee the solution stays within a particular norm ball. Stricter notions of stabilty add this idea in.

Definition 38 *An equilibrium point $\boldsymbol{x}_e$ is attractive if $\forall t_0 > 0,\ \exists c(t_0)$ such that*

$$\boldsymbol{x}(t_0) \in B_c(\boldsymbol{x}_e) \implies \lim_{t\to\infty} \|\boldsymbol{x}(t, t_0, \boldsymbol{x}_0) - \boldsymbol{x}_e\| = 0$$

Attractive equilibria guarantee that trajectories beginning from initial conditions inside of a ball will converge to the equilibrium. However, attractivity does not imply stability since the trajectory could go arbitarily far from the equilibrium so long as it eventually returns.

Definition 39 *An equilibrium point $\boldsymbol{x}_e$ is asymptotically stable if $\boldsymbol{x}_e$ is stable in the sense of Lyapunov and attractive.*

Asymptotic stability fixes the problem of attractivity where trajectories could go far from the equilibrium, and it fixes the problem with stability where the trajectory may not converge to equilibrium. It means that trajectories starting in a ball around equilibrium will converge to equilibrium without leaving that ball. Because the constant for attractivity may depend on time, defining uniform asymptotic stability requires some modifications to the idea of attractivity.

Definition 40 *An equilibrium point is uniformly asympototically stable if $\boldsymbol{x}_e$ is uniformly stable in the sense of Lyapunov, and $\exists c$ and $\gamma : \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}_+$ such that*

$$\forall \boldsymbol{x}_0 \in B_c(\boldsymbol{x}_e),\ \lim_{\tau\to\infty} \gamma(\tau, \boldsymbol{x}_0) = 0, \qquad \forall t \geq t_0,\ \|\boldsymbol{x}(t, t_0, \boldsymbol{x}_0) - \boldsymbol{x}_e\| \leq \gamma(t - t_0, \boldsymbol{x}_0)$$

The existence of the γ function helps guarantee that the rate of converges to equilibrium does not depend on t_0 since the function γ is independent of t_0. Suppose that the γ is an exponential function. Then solutions to the system will converge to the equilibrium exponentially fast.

Definition 41 *An equilibrium point $\boldsymbol{x}_e$ is locally exponentially stable if $\exists h, m, \alpha$ such that*

$$\forall \boldsymbol{x}_0 \in B_h(\boldsymbol{x}_e),\ \|\boldsymbol{x}(t, t_0, \boldsymbol{x}_0) - \boldsymbol{x}_e\| \leq m e^{-\alpha(t-t_0)} \|\boldsymbol{x}(t) - \boldsymbol{x}_e\|$$

Definitions 36, 37 and 39 to 41 are all local definitions because the only need to hold for $\boldsymbol{x}_0$ inside a ball around the equilibrium. If they hold $\forall \boldsymbol{x}_0 \in \mathbb{R}^n$, then they become global properties.

Just as we can define stability, we can also define instability.

Definition 42 *An equilibrium point $\boldsymbol{x}_e$ is unstable in the sense of Lyapunov if $\exists \epsilon > 0, \forall delta > 0$ such that*

$$\exists \boldsymbol{x}_0 \in B_\delta(\boldsymbol{x}_e) \implies \exists T \geq t_0, x(T, t_0, \boldsymbol{x}_0) \notin B_\epsilon(\boldsymbol{x}_e)$$

Instability means that for any $\delta-$ball, we can find an $\epsilon-$ball for which there is at least one initial condition whose corresponding trajectory leaves the $\epsilon-$ball.

4.1 Lyapunov Functions

In order to prove different types of stability, we will construct functions which have particular properties around equilibrium points of the system. The properties of these functions will help determine what type of stable the equilibrium point is.

Definition 43 *A class $\mathcal{K}$ function is a function $\alpha : \mathbb{R}_+ \to \mathbb{R}_+$ such that $\alpha(0) = 0$ and $\alpha(s)$ is strictly monotonically increasing in s.*

A subset of the class $\mathcal{K}$ functions grow unbounded as the argument approaches infinity.

Definition 44 *A class $\mathcal{KR}$ function is a class $\mathcal{K}$ function α where $\lim_{s\to\infty} \alpha(s) = s$.*

Class $\mathcal{KR}$ functions are "radially unbounded". We can use class $\mathcal{K}$ and class $\mathcal{KR}$ to bound "energy-like" functions called **Lyapunov Functions**.

Definition 45 *A function $V(\boldsymbol{x}, t) : \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R}$ is locally positive definite (LPDF) on a set $G \subset \mathbb{R}^n$ containing $\boldsymbol{x}_e$ if $\exists \alpha \in \mathcal{K}$ such that*

$$V(\boldsymbol{x}, t) \geq \alpha(\|\boldsymbol{x} - \boldsymbol{x}_e\|)$$

LPDF functions are locally "energy-like" in the sense that the equilibrium point is assigned the lowest "energy" value, and the larger the deviation from the equilibrium, the higher the value of the "energy".

Definition 46 *A function $V(\boldsymbol{x}, t) : \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R}$ is positive definite (PDF) if $\exists \alpha \in \mathcal{KR}$ such that*

$$\forall \boldsymbol{x} \in \mathbb{R}^n,\ V(\boldsymbol{x}, t) \geq \alpha(\|\boldsymbol{x} - \boldsymbol{x}_e\|)$$

Positive definite functions act like "energy functions" everywhere in $\mathbb{R}^n$.

Definition 47 *A function $V(\boldsymbol{x}, t) : \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R}$ is decrescent if $\exists \alpha \in \mathcal{K}$ such that*

$$\forall \boldsymbol{x} \in B_h(\boldsymbol{x}_e),\ V(\boldsymbol{x}, t) \leq \beta(\|\boldsymbol{x} - \boldsymbol{x}_e\|)$$

Descresence means that for a ball around the equilibrium, we can upper bound the growth of the energy.

Note that we can assume $\boldsymbol{x}_e = 0$ without loss of generality for definitions 45 to 47 since for a given system, we can always define a linear change of variables that shifts the equilibrium point to the origin.

4.1.1 Quadratic Lyapunov Functions

Definition 48 *A Quadratic Lypunov function is of the form*

$$V(\boldsymbol{x}) = \boldsymbol{x}^\top P \boldsymbol{x}, \quad P \succ 0$$

Quadratic Lyapunov Functions are one of the simplest types of Lyapunov Functions. Their level sets are ellipses where the major axis is the eigenvector corresponding to $\lambda_{min}(P)$, and the minor axis is the eigenvecctor corresponding to $\lambda_{max}(P)$.

Theorem 14 *Consider the sublevel set $\Omega_c = \{\boldsymbol{x} | V(\boldsymbol{x}) \leq c\}$. Then r_* is the radius of the largest circle contained inside Ω_c, and r^* is the radius of the largest circle containing Ω_c.*

$$r_* = \sqrt{\frac{c}{\lambda_{max}(P)}} \qquad r^* = \sqrt{\frac{c}{\lambda_{min}(P)}}$$

4.1.2 Sum-of-Squares Lyapunov Functions

Definition 49 *A polynomial $p(\boldsymbol{x})$ is sum-of-squares (SOS) if $\exists g_1, \cdots, g_r$ such that*

$$p(\boldsymbol{x}) = \sum_{i=1}^{r} g_i^2(\boldsymbol{x})$$

SOS polynomials have the nice property that they are always non-negative due to being a sum of squared numbers. Since any polynomial can be written in a quadratic form $P(\boldsymbol{x}) = z^\top(\boldsymbol{x}) Q z(\boldsymbol{x})$ where z is a vector of monomials, the properties of Q can tell us if P is SOS or not.

Theorem 15 *A polynomial is SOS if and only if it can be written as*

$$p(\boldsymbol{x}) = z^\top(\boldsymbol{x}) Q z(\boldsymbol{x}), \quad Q \succeq 0$$

Note that Q is not necessarily unique, and if we construct a linear operator which maps Q to P, then this linear operator will have a Null Space. Mathematically, consider

$$\mathcal{L}(Q)(\boldsymbol{x}) = z^\top(\boldsymbol{x}) Q z(\boldsymbol{x}).$$

This linear operator has a null space spanned by the polynomials N_j. Given a matrix $Q_0 \succeq 0$ such that $p(\boldsymbol{x}) = z^\top(\boldsymbol{x}) Q_0 z(\boldsymbol{x})$ (i.e p is SOS), it is also true that

$$p(\boldsymbol{x}) = z^\top(\boldsymbol{x}) \left(Q_0 + \sum_j \lambda_j N_j(\boldsymbol{x}) \right) z(\boldsymbol{x}).$$

SOS polynomials are helpful in finding Lyapunov functions because we can use SOS Programming to find SOS polynomials which satisfy desired properties. For example, if we want $V(\boldsymbol{x})$ to be PDF, then one constraint in our SOS program will be that

$$V(\boldsymbol{x}) - \epsilon \boldsymbol{x}^\top \boldsymbol{x}, \quad \epsilon > 0$$

is SOS.

4.2 Proving Stability

To prove the stability of an equilibrium point for a given nonlinear system, we will construct a Lyapunov function and determine stability from the properties of the Lyapunov functions which we can find. Given properties of V and $\frac{\mathrm{d}V}{\mathrm{d}t}$, we can use the **Lyapunov Stability Theorems** to prove the stability of equilibria.

Theorem 16 *If* $\exists V(\boldsymbol{x}, t)$ *such that* V *is LPDF and the Lie Derivative* $-\dfrac{\mathrm{d}V}{\mathrm{d}t} \geq 0$ *locally, then* $\boldsymbol{x}_e$ *is stable in the sense of Lyapunov.*

Theorem 17 *If* $\exists V(\boldsymbol{x}, t)$ *such that* V *is LPDF and decrescent, and the Lie Derivative* $-\dfrac{\mathrm{d}V}{\mathrm{d}t} \geq 0$ *locally, then* $\boldsymbol{x}_e$ *is uniformly stable in the sense of Lyapunov.*

Theorem 18 *If* $\exists V(\boldsymbol{x}, t)$ *such that* V *is LPDF and decrescent, and the Lie Derivative* $-\dfrac{\mathrm{d}V}{\mathrm{d}t}$ *is LPDF, then* $\boldsymbol{x}_e$ *is uniformly asymptotically stable in the sense of Lyapunov.*

Theorem 19 *If* $\exists V(\boldsymbol{x}, t)$ *such that* V *is PDF and decrescent, and the Lie Derivative* $-\dfrac{\mathrm{d}V}{\mathrm{d}t}$ *is LPDF, then* $\boldsymbol{x}_e$ *is globally uniformly asymptotically stable in the sense of Lyapunov.*

Theorem 20 *If* $\exists V(\boldsymbol{x}, t)$ *and* $h, \alpha > 0$ *such that* V *is LPDF is decrescent, The Lie derivative* $-\dfrac{\mathrm{d}V}{\mathrm{d}t}$ *is LDPF, and*

$$\forall \boldsymbol{x} \in B_h(\boldsymbol{x}_e), \ \left\| \frac{\partial V}{\partial t} \right\| \leq \alpha \|\boldsymbol{x} - \boldsymbol{x}_e\|$$

The results of theorems 16 to 20 are summarized in table 1. Going down the rows

Conditions on V	**Conditions on** $-\dfrac{\mathrm{d}\mathbf{V}}{\mathrm{d}\mathbf{t}}$	**Conclusion**
LPDF	≥ 0 locally	Stable
LPDF, Decrescent	≥ 0 locally	Uniformly Stable
LPDF, Decrescent	LPDF	Uniformly, Asymptotically Stable
LPDF, Decrescent	LDPF, $\exists \alpha > 0$ such that $\left\| \dfrac{\mathrm{d}V}{\mathrm{d}t} \right\| \leq \alpha \|\boldsymbol{x} - \boldsymbol{x}_e\|$	Exponentially Stable
PDF, Decrescent	PDF	Globally, Uniformly, Asymptotically Stable

Table 1: Summary of Lyapunov Stability Theorems

of table 1 lead to increasingly stricter forms of stability. Descresence appears to add uniformity to the stability, while $-\dfrac{\mathrm{d}V}{\mathrm{d}t}$ being LPDF adds asymptotic convergence.

However, these conditions are only sufficient, meaning if we cannot find a suitable V, that does not mean that an equilibrium point is not stable.

One very common case where it can be difficult to find appropriate Lyapunov functions is in proving asymptotic stability since it can be hard to find V such that $-\frac{\mathrm{d}V}{\mathrm{d}t}$ is LPDF. In the case of autonomous systems, we can still prove asymptotic stability without such a V.

Theorem 21 (LaSalle's Invariance Principle) *Consider a smooth function $V : \mathbb{R}^n \to \mathbb{R}$ with bounded sub-level sets $\Omega_c = \{\boldsymbol{x} | V(\boldsymbol{x}) \leq c\}$ and $\forall \boldsymbol{x} \in \Omega_c$, the Lie derivative $\frac{\mathrm{d}V}{\mathrm{d}t} \leq 0$. Define $S = \left\{\boldsymbol{x} | \frac{\mathrm{d}V}{\mathrm{d}t} = 0\right\}$ and let M be the largest invariant set in S, then*

$$\forall \boldsymbol{x}_0 \in \Omega_c,\ \boldsymbol{x}(t, t_0, x_0) \to M \text{ as } t \to \infty.$$

LaSalle's theorem helps prove general convergence to an invariant set. Since V is always decreasing in the sub-level set Ω_c, trajectories starting in Ω_c must eventually reach S. At some point, they will reach the set M in S, and then they will stay there. Thus if the set M is only the equilibrium point, or a set of equilibrium points, then we can show that the system trajectories asymptotically converges to this equilibrium or set of equilibria. Moreover, if $V(\boldsymbol{x})$ is PDF, and $\forall \boldsymbol{x} \in \mathbb{R}^n$, $\frac{\mathrm{d}V}{\mathrm{d}t} \leq 0$, then we can show global asymptotic stability as well.

LaSalle's theorem can be generalized to non-autonomous systems as well, but it is slightly more complicated since the set S may change over time.

4.2.1 Indirect Method of Lyapunov

It turns out that we can also prove the stability of systems by looking at the linearization around the equilibrium. Without loss of generality, suppose $\boldsymbol{x}_e = 0$. The linearization at the equilibrium is given by

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = f(x, t) = f(0, t) + \frac{\partial f}{\partial \boldsymbol{x}}|_{\boldsymbol{x}=0}\boldsymbol{x} + f_1(\boldsymbol{x}, t) \approx A(t)\boldsymbol{x}.$$

The function $f_1(\boldsymbol{x}, t)$ is the higher-order terms of the linearization. The linearization is a time-varying system. Consider the time-varying linear system

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = A(t)\boldsymbol{x},\ \boldsymbol{x}(t_0) = \boldsymbol{x}_0.$$

Definition 50 *The state transition matrix $\Phi(t, t_0)$ of a time-varying linear system is a matrix satisfying*

$$\boldsymbol{x}(t) = \Phi(t, t_0)\boldsymbol{x}_0, \ \frac{\mathrm{d}\Phi}{\mathrm{d}t} = A(t)\Phi(t, t_0), \ \Phi(t_0, t_0) = I$$

The state transition matrix is useful in determining properties of the system.

1. $\sup_{t \geq t_0} \|\Phi(t, t_0)\| = m(t_0) < \infty \implies$ the system is stable at the origin at t_0.
2. $\sup_{t_0 \geq 0} \sup_{t \geq t_0} \|\Phi(t, t_0)\| = m < \infty \implies$ the system is uniformly stable at the origin at t_0.
3. $\lim_{t \to \infty} \|\Phi(t, t_0)\| = 0 \implies$ the system is asymptotically stable.
4. $\forall t_0, \epsilon > 0, \exists T$ such that $\forall t \geq t_0 + T, \ \|\Phi(t, t_0)\| < \epsilon \implies$ the system is uniformly asymptotically stable.
5. $\|\Phi(t, t_0)\| \leq Me^{-\lambda(t - t_0)} \implies$ exponential stability.

If the system was Time-Invariant, then the system would be stable so long as the eigenvalues of A were in the open left-half of the complex plane. In fact, we could use A to construct positive definite matrices.

Theorem 22 (Lyapunov Lemma) *For a matrix $A \in \mathbb{R}^{n \times n}$, its eigenvalues λ_i satisfy $\Re(\lambda_i) < 0$ if and only if $\forall Q \succ 0$, there exists a solution $P \succ 0$ to the equation*

$$A^T P + PA = -Q.$$

In general, we can use the **Lyapunov Equation** to count how many eigenvalues of A are stable.

Theorem 23 (Tausskey Lemma) *For $A \in \mathbb{R}^{n \times n}$ and given $Q \succ 0$, if there are no eigenvalues on the $j\omega$ axis, then the solution P to $A^T P + PA = -Q$ has as many positive eigenvalues as A has eigenvalues in the complex left half plane.*

The Lyapunov Lemma has extensions to the time-varying case.

Theorem 24 (Time-Varying Lyapunov Lemma) *If $A(\cdot)$ is bounded and for some $Q(t) \succeq \alpha I$, the solution $P(t)$ to $A(t)^T P(t) + P(t)A(t) = -Q(t)$ is bounded, then the origin is a asymptotically stable equilibrium point.*

It turns out that uniform asymptotic stability of the linearization of a system corresponds to uniform, asymptotic stability of the nonlinear system.

Theorem 25 (Indirect Theorem of Lyapunov) *For a nonlinear system whose higher-order terms of the linearization are given by $f(\boldsymbol{x}, t)$, if*

$$\lim_{\|\boldsymbol{x}\| \to 0} \sup_{t \geq 0} \frac{\|f_1(\boldsymbol{x}, t)\|}{\|\boldsymbol{x}\|} = 0$$

and if $\boldsymbol{x}_e$ is a uniformly asymptotic stable equilibrium point of $\frac{\mathrm{d}\boldsymbol{z}}{\mathrm{d}t} = A(t)\boldsymbol{z}$ where $A(t)$ is the Jacobian at the $\boldsymbol{x}_e$, then $\boldsymbol{x}_e$ is a uniformly asymptotic stable equilibrium point of $\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = f(\boldsymbol{x}, t)$

4.3 Proving Instability

Theorem 26 *An equilibrium point $\boldsymbol{x}_e$ is unstable in the sense of Lyapunov if $\exists V(\boldsymbol{x}, t)$ which is decrescent, the Lie derivative $\frac{\mathrm{d}V}{\mathrm{d}t}$ is LPDF, $V(\boldsymbol{x}_e, t)$, and $\exists \boldsymbol{x}$ in the neighborhood of $\boldsymbol{x}_e$ such that $V(\boldsymbol{x}_0, t) > 0$.*

4.4 Region of Attraction

For asymptotically stable and exponential stable equilibria, it makes sense to wonder which initial conditions will cause trajectories to converge to the equilibrium.

Definition 51 *If $\boldsymbol{x}_e$ is an equilibrium point of a time-invariant system $\frac{\mathrm{d}x}{\mathrm{d}t} = f(x)$, then the Region of Attraction of $\boldsymbol{x}_e$ is*

$$\mathcal{R}_A(\boldsymbol{x}_e) = \{\boldsymbol{x}_0 \in \mathbb{R}^n | \lim_{t \to \infty} \boldsymbol{x}(t, t_0) = \boldsymbol{x}_e\}$$

Suppose that we have a Lyapunov function $V(\boldsymbol{x})$ and a region D such that $V(\boldsymbol{x}) > 0$ and $\frac{\mathrm{d}V}{\mathrm{d}t} < 0$ in D. Define a sublevel set of the Lypunov function Ω_c which is a subset of D. We know that if $\boldsymbol{x}_0 \in \Omega_c$, then the trajectory will stay inside Ω_c and converge to the equilibrium point. Thus we can use the largest Ω_c that is compact and contained in D as an estimate of the region of attraction.

When we have a Quadratic Lyapunov Function, we can set D to be the largest circle which satisfies the conditions on V, and the corresponding Ω_c contained inside D will be the estimate of the Region of Attraction.

We can find even better approximations of the region of attraction using SOS programming. Suppose we have a V which we used to prove asymptotic stability. Then if there exists an s which satisfies the following SOS program, then the sublevel set Ω_c is an estimate of the Region of Attraction.

$$\begin{aligned} &\max_{c,s} c \\ \text{such that } &s(\boldsymbol{x}) \text{ is SOS}, \\ &-\left(\frac{\mathrm{d}V}{\mathrm{d}t} + \epsilon \boldsymbol{x}^\top \boldsymbol{x}\right) + s(\boldsymbol{x})(c - V(\boldsymbol{x})) \text{ is SOS.} \end{aligned}$$

5 Nonlinear Feedback Control

In nonlinear control problems, we have a system

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = f(\boldsymbol{x}, \boldsymbol{u}).$$

$\boldsymbol{x}$ is the state of the system, and $\boldsymbol{u}$ is the input to the system. Note that for simplicity, the system is time-invariant. Further assume, without loss of generality, that $f(0,0) = 0$. The goal of nonlinear feedback control is to find a state feedback law $\alpha(\boldsymbol{x})$ such that the equilibrium point $\boldsymbol{x}_e = 0$ is globally asymptotically stable for the closed loop system

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = f(\boldsymbol{x}, \alpha(\boldsymbol{x}))$$

Sometimes, the control impacts the state evolution in an affine manner.

Definition 52 *A control affine system is given by the differential equation*

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = f(\boldsymbol{x}) + G(\boldsymbol{x})\boldsymbol{u}$$

where $G(\boldsymbol{x})$ is a matrix dependent on the state vector $\boldsymbol{x}$.

When designing controllers, there is a wide variety of techniques we can use. Some simple techniques involve canceling out various types of nonlinearities in the system using the input. Here are some examples.

1. Set $\boldsymbol{u}$ such that it cancels out nonlinear terms and adds a stable linear term, effectively making the nonlinear system behave linear in the closed loop.

2. Set $\boldsymbol{u}$ to cancel destabilitizing nonlinear terms and add a stable linear term, so the stable nonlinearities help the input drive the system to equilibrium.

3. Set $\boldsymbol{u}$ to cancel destabiliizing nonlinear terms, so the nonlinear system dynamics drive the system to equilibrium.

4. Set $\boldsymbol{u}$ to dominate destabilizing terms so they have a minimal impact on the overall system behavior.

While these techniques can work, there are also more principled ways of designing controllers to satisfy different criteria, particularly for the case of control affine systems.

5.1 Control Lyapunov Functions

If we can find an $\alpha(\boldsymbol{x})$ that makes the origin globally asymptotically stable, then the converse Lyapunov theorem says that we can find a corresponding Lyapunov function for the system.

$$\begin{aligned}
\forall \boldsymbol{x} \neq 0, \frac{\mathrm{d}V}{\mathrm{d}t} < 0 &\implies \frac{\partial V}{\partial \boldsymbol{x}} f(\boldsymbol{x}, \alpha(\boldsymbol{x})) < 0 \\
&\implies \exists \boldsymbol{u} \text{ s.t } \frac{\partial V}{\partial \boldsymbol{x}} f(\boldsymbol{x}, \boldsymbol{u}) < 0 \\
&\Leftrightarrow \inf_{\boldsymbol{u}} \frac{\partial V}{\partial \boldsymbol{x}} f(\boldsymbol{x}, \boldsymbol{u}) < 0
\end{aligned}$$

This result motivates the following definition.

Definition 53 *A continuously differentiable, PDF, radially unbounded $V : \mathbb{R}^n \to \mathbb{R}$ is a Control Lyapunov Function for the system $\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = f(\boldsymbol{x}, \boldsymbol{u})$ if*

$$\forall \boldsymbol{x} \neq 0, \inf_{\boldsymbol{u}} \frac{\partial V}{\partial \boldsymbol{x}} f(\boldsymbol{x}, \boldsymbol{u}) < 0$$

Once we have a control lyapunov function, we can prove that it is possible to find a state feedback law that will make the origin globally asymptotically stable.

Theorem 27 *Suppose f is Lipschitz and V is a control Lyapunov function, then there exists a smooth function α such that the origin is a globally asympototically stable equilibrium point of $\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = f(\boldsymbol{x}, \alpha(\boldsymbol{x}))$.*

Suppose that we have a control affine system, and we want to construct a control lyapunov function for the system.

$$\inf_{\boldsymbol{u}} \frac{\partial V}{\partial \boldsymbol{x}} f(\boldsymbol{x}, \boldsymbol{u}) = \inf_{\boldsymbol{u}} L_f V + \sum_i L_{g_i} V \boldsymbol{u}_i < 0$$

If $\forall i,\ L_{g_i} V = 0$, then definition 53 is satisfied so long as $L_f V < 0$.

Theorem 28 *A function V is a control lyapunov function for a control affine system*

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = f(\boldsymbol{x}) + \sum_i g_i(\boldsymbol{x})\boldsymbol{u}$$

if

$$L_{g_i}V\,\forall i \implies L_fV \leq 0$$

Notice that the condition in theorem 28 is essentially saying that the l_2-norm of the vector composed of the $L_{g_i}V$ is equal to 0. The choice of CLF is important because different CLFs have different properties when used to derive controllers.

Definition 54 *A CLF $V(\boldsymbol{x})$ satisfies the small control property if $\forall\epsilon > 0$, $\exists\delta > 0$ such that $\boldsymbol{x} \in B_\delta(0)$, then if $\boldsymbol{x} \neq 0, \exists\boldsymbol{u} \in B_\epsilon(0)$ satisfying*

$$\frac{\mathrm{d}V}{\mathrm{d}t} = L_fV + L_gV^T\boldsymbol{u} < 0.$$

The small control property means that CLF will lead to a controller which has a small value that does not get too large when close to the equilibrium.

Given a control lyapunov function for a control affine system $V(\boldsymbol{x}, \boldsymbol{u})$, we can devise a controller which stabilizes the system. In particular, we need

$$\frac{dV}{dt}(x, u) = L_fV(\boldsymbol{x}) + L_gV(\boldsymbol{x})^\top\boldsymbol{u} \leq 0.$$

Hence, let

$$\boldsymbol{u} = \begin{cases} 0, & \text{if } L_fV < 0, \\ (L_gVL_gV^\top)^{-1}(-L_fVL_gV^\top), & \text{if } L_fV > 0. \end{cases}$$

When the plant dynamics are naturally stabilizing, this controller exerts no control effort. When the plant dynamics are not naturally stabilizing, then the controller applies some control to stabilize the system. We can show that this is a minimum norm controller as it solves the optimization problem

$$\begin{aligned} \min \quad & \boldsymbol{u}^\top\boldsymbol{u} \\ \text{s.t} \quad & L_fV + L_gV^\top\boldsymbol{u} \leq 0. \end{aligned}$$

Another type of controller is known as the Sontag controller.

Theorem 29 *Suppose $V : \mathbb{R}^n \to \mathbb{R}$ is a CLF for SISO control affine system $\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = f(\boldsymbol{x}) + g(\boldsymbol{x})u$ where $f(0) = 0$ where f and g are Lipschitz. Then the Sontag feedback control law is given by*

$$\alpha_S(\boldsymbol{x}) = \begin{cases} \frac{-L_fV - \sqrt{(L_fV)^2 + (L_gV)^4}}{L_gV}, & \textit{if } L_gV \neq 0, \\ 0, & \textit{else.} \end{cases}$$

makes the origin globally asymptotically stable. Moreover, $\alpha_S(\boldsymbol{x})$ is continuous everywhere except $\boldsymbol{x} = 0$, is continuous at $\boldsymbol{x} = 0$ if V satisfies the small control property, and if $V(\boldsymbol{x})$ is $K + 1$ times continuously differentiable and $f(\boldsymbol{x}), g(\boldsymbol{x})$ are K times continously differentiable $\forall \boldsymbol{x} \neq 0$, then α_S is K times continuously differentiable.

5.2 Feedback Linearization

Since we have a large number of tools which allow us to control linear systems, it would be ideal if we could somehow leverage those tools for nonlinear control. Feedback linearization is the process of finding a feedback control law $\boldsymbol{u} = \alpha(\boldsymbol{x})$ such that under a nonlinear change of coordinates $z = \Phi(\boldsymbol{x})$, the system $\frac{\mathrm{d}x}{\mathrm{d}t} = f(\boldsymbol{x}, \alpha(\boldsymbol{x}))$ behaves like a linear system $\frac{\mathrm{d}\boldsymbol{z}}{\mathrm{d}t} = A\boldsymbol{z}$. When the system is control-affine, there are well-established results which help us do this.

5.2.1 SISO Case

Suppose we have a SISO control-affine system

$$\begin{aligned} \frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} &= f(\boldsymbol{x}) + g(\boldsymbol{x})u \\ y &= h(\boldsymbol{x}) \end{aligned}$$

Definition 55 *A SISO control affine system with an equilibrium point $\boldsymbol{x}_e$ has strict relative degree γ if in a neighborhood U around the equilibirum point, $L_gL_f^{\gamma-1}h(\boldsymbol{x})$ is bounded away from 0 and*

$$\forall i = 0, \cdots, \gamma - 2,\ L_gL_f^{\gamma-1}h(\boldsymbol{x}) = 0$$

To understand relative degree, suppose we differentiate y once

$$\frac{\mathrm{d}y}{\mathrm{d}t} = L_fh + L_ghu.$$

If $\forall \boldsymbol{x} \in U,\ L_gh(\boldsymbol{x}) = 0$ where U is some region around the equilibrium, then

$$\forall \boldsymbol{x} \in U,\ \frac{\mathrm{d}y}{\mathrm{d}t} = L_fh(\boldsymbol{x}).$$

If we differentiate again, then

$$\forall \boldsymbol{x} \in U,\ \frac{\mathrm{d}^2y}{\mathrm{d}t^2} = L_f^2h(\boldsymbol{x}) + L_gL_f(\boldsymbol{x}).$$

Suppose that $\forall \boldsymbol{x} \in U,\ L_g L_f(\boldsymbol{x}) = 0$, then we can differentiate again. At some point, after γ differentiations, we will get

$$\forall \boldsymbol{x} \in U,\ \frac{\mathrm{d}\gamma}{\mathrm{d}y} t = L_f^{\gamma} h(\boldsymbol{x}) + L_g L_f^{\gamma-1} h(\boldsymbol{x}) u.$$

Therefore, the relative degree of the system is essentially telling us which derivative of the output that we can control. By sequentially taking derivatives, we are essentially looking at the system

$$\begin{aligned}
y &= h(\boldsymbol{x}) \\
\frac{\mathrm{d}y}{\mathrm{d}t} &= L_f h(\boldsymbol{x}) \\
\frac{\mathrm{d}^2 y}{\mathrm{d}t^2} &= L_f^2 h(\boldsymbol{x}) \\
&\vdots \\
\frac{\mathrm{d}^{\gamma} y}{\mathrm{d}t^{\gamma}} &= L_f^{\gamma} h(\boldsymbol{x}) + L_g L_f^{\gamma-1} h(\boldsymbol{x})
\end{aligned}$$

Suppose $\forall i = 0, \cdots, \gamma - 1$, we let $\xi_i(\boldsymbol{x}) = \dfrac{\mathrm{d}^i y}{\mathrm{d}t^i}$. These are γ linearly independent coordinates. Since the distribution

$$\Delta(\boldsymbol{x}) = \text{span}\{g(\boldsymbol{x})\}$$

is involutive, it is integrable, and so there must be $n - 1$ functions η_i such that

$$\forall \boldsymbol{x} \in U,\ (\nabla_{\boldsymbol{x}} \boldsymbol{\eta})^{\top} g(\boldsymbol{x}) = 0.$$

We can now choose $n - \gamma$ of them which are linearly independent of the ξ_i and linearly independent with each other, and this forms a change of coordinates

$$\begin{bmatrix} \xi_1 \\ \vdots \\ \xi_{\gamma} \\ \eta_1 \\ \vdots \\ \eta_{n-\gamma} \end{bmatrix} = \Phi(\boldsymbol{x}) = \begin{bmatrix} h(\boldsymbol{x}) \\ L_f h(\boldsymbol{x}) \\ \vdots \\ L_f^{\gamma-1} \\ \eta_1 \\ \vdots \\ \eta_{n-\gamma} \end{bmatrix}$$

This change of coordinates allows us to put the system into a canonical form.

Definition 56 *The normal form of a SISO control affine system is given by*

$$
\begin{aligned}
\frac{\mathrm{d}\xi_1}{\mathrm{d}t} &= \xi_2 \\
\frac{\mathrm{d}\xi_2}{\mathrm{d}t} &= \xi_3 \\
&\vdots \\
\frac{\mathrm{d}\xi_\gamma}{\mathrm{d}t} &= b(\boldsymbol{\xi}, \boldsymbol{\eta}) + a(\boldsymbol{\eta}, \boldsymbol{\xi})u \\
\frac{\mathrm{d}\boldsymbol{\eta}}{\mathrm{d}t} &= q(\boldsymbol{\xi}, \boldsymbol{\eta}), \\
y &= \eta_1
\end{aligned}
$$

When the original system is given by $\frac{\mathrm{d}x}{\mathrm{d}t} = f(\boldsymbol{x}) + g(\boldsymbol{x})u$, then

$$
b(\boldsymbol{\xi}, \boldsymbol{\eta}) = L_f^\gamma h(\Phi^{-1}(\boldsymbol{\xi}, \boldsymbol{\eta})) \qquad a(\boldsymbol{\xi}, \boldsymbol{\eta}) = L_g L_f^{\gamma-1} h(\Phi^{-1}(\boldsymbol{\xi}, \boldsymbol{\eta}))
$$

With this parameterization, it is quite easy to see how we can make our system behave linearly. In particular, choose

$$
u = \frac{1}{a(\boldsymbol{\xi}, \boldsymbol{\eta})} \left(-b(\boldsymbol{\xi}, \boldsymbol{\eta}) + \boldsymbol{v}\right)
$$

where $\boldsymbol{v}$ is some control input. Then the system becomes

$$
\begin{aligned}
\frac{\mathrm{d}\xi_1}{\mathrm{d}t} &= \xi_2 \\
\frac{\mathrm{d}\xi_2}{\mathrm{d}t} &= \xi_3 \\
&\vdots \\
\frac{\mathrm{d}\xi_\gamma}{\mathrm{d}t} &= \boldsymbol{v} \\
\frac{\mathrm{d}\boldsymbol{\eta}}{\mathrm{d}t} &= q(\boldsymbol{\xi}, \boldsymbol{\eta}) \\
y &= \eta_1
\end{aligned}
$$

, which is a linear system. Therefore, we can design a linear controller input $\boldsymbol{v}$ where we have all of the tools of linear control at our disposal. However, notice that the η_i cannot be impacted by the control effort. These are known as the **internal dynamics** of the system. When $\boldsymbol{\xi} = 0$, then

$$
\frac{\mathrm{d}\boldsymbol{\eta}}{\mathrm{d}t} = q(0, \boldsymbol{\eta})
$$

are known as the **Zero Dynamics** of the system. Zero dynamics for a system can be dangerous because if they are unstable, then the system could be blowing up.

When $\gamma = n$, then there are no zero dynamics. When this happens, we say the system is **Full State Linearizable**. Fortunately, there are necessary and sufficient conditions which guarantee full state linearization.

Theorem 30 *There exists a function h such that a control affine system $\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = f(\boldsymbol{x}) + g(\boldsymbol{x})u$ has relative degree n at $\boldsymbol{x}_0$ if and only if*

$$\begin{bmatrix} g(\boldsymbol{x}) & \cdots & \mathrm{ad}_f^{n-2}\, g(\boldsymbol{x}) & \mathrm{ad}_f^{n-1}\, g(\boldsymbol{x}) \end{bmatrix}$$

has rank n and

$$\begin{bmatrix} g(\boldsymbol{x}) & \cdots & \mathrm{ad}_f^{n-3}\, g(\boldsymbol{x}) & \mathrm{ad}_f^{n-2}\, g(\boldsymbol{x}) \end{bmatrix}$$

has rank $n-1$ and is involutive in the neighborhood of $\boldsymbol{x}_0$. The h is chosen to satisfy

$$(\nabla_{\boldsymbol{x}} h)^\top \begin{bmatrix} g(\boldsymbol{x}) & \cdots & \mathrm{ad}_f^{n-3}\, g(\boldsymbol{x}) & \mathrm{ad}_f^{n-2}\, g(\boldsymbol{x}) \end{bmatrix} = 0$$

5.2.2 MIMO Case

Suppose instead we have a MIMO control affine system where

$$\begin{aligned} \frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} &= f(\boldsymbol{x}) + G(\boldsymbol{x})\boldsymbol{u} \\ \boldsymbol{y} &= h(\boldsymbol{x}) \end{aligned}$$

We will assume that the number of outputs is equal to the number of inputs (i.e $\boldsymbol{y}, \boldsymbol{u} \in \mathbb{R}^m$). To linearize the system, we can take the same idea of relative degree from the SISO case and apply it to the MIMO case. Define γ_j to be the lowest derivative of y_j which is impacted by at least one input.

$$\begin{bmatrix} \frac{\mathrm{d}^{\gamma_1} y_1}{\mathrm{d}t^{\gamma_1}} \\ \vdots \\ \frac{\mathrm{d}^{\gamma_m} y_m}{\mathrm{d}t^{\gamma_m}} \end{bmatrix} = \begin{bmatrix} L_f^{\gamma_1} h_1(\boldsymbol{x}) \\ \vdots \\ L_f^{\gamma_m} h_m(\boldsymbol{x}) \end{bmatrix} + A(\boldsymbol{x})\boldsymbol{u}, \qquad A(\boldsymbol{x}) = \begin{bmatrix} L_{g_1} L_f^{\gamma_1 - 1} h_1(\boldsymbol{x}) & \cdots & L_{g_m} L_f^{\gamma_1 - 1} h_m(\boldsymbol{x}) \\ \vdots & \ddots & \vdots \\ L_{g_1} L_f^{\gamma_m - 1} h_1(\boldsymbol{x}) & \cdots & L_{g_m} L_f^{\gamma_m - 1} h_m(\boldsymbol{x}) \end{bmatrix}$$

Definition 57 *A square control affine system has a vector relative degree $(\gamma_1, \cdots, \gamma_m)$ at $\boldsymbol{x}_0 \in U$ if $A(\boldsymbol{x}_0)$ is nonsingular and*

$$\forall 1 \leq i \leq m,\ 1 \leq j \leq m,\ 0 \leq k \leq \gamma_j - 2,\ \forall \boldsymbol{x} \in U,\ L_{g_i} L_f^k h_j(\boldsymbol{x}) = 0$$

As before, we can assign $\dfrac{\mathrm{d}^i y_j}{\mathrm{d}t^i} = \xi_i^j$ as a partial change of coordinates and then choose linearly independent $\boldsymbol{\eta}$.

Definition 58 *The normal form of a square MIMO system is given by*

$$\begin{aligned}\frac{\mathrm{d}\boldsymbol{\eta}}{\mathrm{d}t} &= q(\boldsymbol{\xi},\boldsymbol{\eta}) + p(\boldsymbol{\xi},\boldsymbol{\eta})\boldsymbol{u} \\ \frac{\mathrm{d}\xi_i^j}{\mathrm{d}t} &= \xi_{i+1}^j, && \forall j,\ \forall i < \gamma_j - 1 \\ \frac{\mathrm{d}\xi_{\gamma_j-1}^j}{\mathrm{d}t} &= b^j(\boldsymbol{\xi},\boldsymbol{\eta}) + \boldsymbol{a}^j(\boldsymbol{\xi},\boldsymbol{\eta})^\top \boldsymbol{u}\end{aligned}$$

As before the $\dfrac{\mathrm{d}\boldsymbol{\eta}}{\mathrm{d}t}$ represent the internal dynamics of the system that are not impacted by the control. As with the linear case, we can design a controller

$$\boldsymbol{u} = A^{-1}(\boldsymbol{x})\left(\begin{bmatrix} L_f^{\gamma_1} h_1(\boldsymbol{x}) \\ \vdots \\ L_f^{\gamma_m} h_m(\boldsymbol{x})\end{bmatrix} + \boldsymbol{v}\right)$$

which renders the system linear. We can now choose $\boldsymbol{v}$ where each entry of $\boldsymbol{v}$ controls a different output. For this reason, we call $A(\boldsymbol{x})$ the decoupling matrix. As in the SISO case, unless $\sum_j \gamma_j = n$, there are zero dynamics to the system.

Theorem 31 *A control affine square system* $\dfrac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = f(\boldsymbol{x}) + g(\boldsymbol{x})\boldsymbol{u}$ *has vector relative degree* $\sum_j \gamma_j = n$ *if and only if* Δ_i *is involutive for all* $i \leq n-2$, Δ_i *has constant rank for all* $1 \leq i \leq n-1$ *and* Δ_{n-1} *has rank* n *where*

$$\begin{aligned}\Delta_0(\boldsymbol{x}) &= span\{g_1(\boldsymbol{x}), \cdots, g_m(\boldsymbol{x})\} \\ \Delta_i(\boldsymbol{x}) &= span\{\operatorname*{ad}_f^k g_i(\boldsymbol{x}) | \forall 0 \leq k \leq i, 1 \leq j \leq m\}, \qquad \forall 1 \leq i \leq n-1\end{aligned}$$

5.2.3 Dynamic Extension

Sometimes, we can use full-state linearization even if h does not satisfy the conditions in theorem 30. We do this by adding additional states to the system and corresponding pseudo-control inputs which help control these states. Sometimes, this can be done in a way which makes the extended system full-state linearizable.

5.2.4 Sliding Mode Control

In sliding mode control, we design a controller

$$u = \beta(\boldsymbol{x})\text{sgn}(s)$$

where $s(\boldsymbol{x})$ describes a manifold called the "sliding manifold". Sliding mode controllers have two states

1. Reaching Mode
2. Sliding Mode

During the reaching mode, the controller drives the state towards the sliding manifold $s(\boldsymbol{x}) = 0$. We choose s such that on the manifold, when the system is in sliding mode, the system naturally converges asymptotically to the equilibrium. If $s(\boldsymbol{x}) = 0$ is an invariant manifold, then the system will smoothly travel along the manifold to equilibrium. If the sliding manifold is not invariant, then the state will chatter around the manifold towards equilibrium as the controller continuously drives it back to the manifold once it leaves. To choose s, we need to find a CLF $V(s)$ which converges to 0 in finite time when applying the sliding mode controller.

5.2.5 Backstepping

Definition 59 *A system expressed in strict feedback form is given by*

$$\begin{aligned}
\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} &= f_0(\boldsymbol{x}) + g_0(\boldsymbol{x})\xi_1 \\
\frac{\mathrm{d}\xi_1}{\mathrm{d}t} &= f_1(\boldsymbol{x}, \xi_1) + g_1(\boldsymbol{x}, \xi_1)\xi_2 \\
&\vdots \\
\frac{\mathrm{d}\xi_k}{\mathrm{d}t} &= f_k(\boldsymbol{x}, \xi_1, \cdots, \xi_k) + g_k(\boldsymbol{x}, \xi_1, \cdots, \xi_k)u
\end{aligned}$$

When systems are expressed in this way, we have a convenient method of designing controllers.

Theorem 32 (Backstepping Lemma) *Suppose there is a continuously differentiable $u = \alpha(\boldsymbol{x})$ and a CLF $V(\boldsymbol{x})$ such that*

$$L_f V + L_g V \alpha \leq -W$$

where W is a positive semi-definite function for the system $\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = f(\boldsymbol{x}) + g(\boldsymbol{x})u$. Then for the system

$$\begin{aligned}\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} &= f(\boldsymbol{x}) + g(\boldsymbol{x})\xi \\ \frac{\mathrm{d}\xi}{\mathrm{d}t} &= u\end{aligned}$$

the function

$$V_a(\boldsymbol{x}, \xi) = V(\boldsymbol{x}) + \frac{1}{2}(\xi - \alpha(\boldsymbol{x}))^2$$

is a valid CLF and the control input

$$u = -c(\xi - \alpha(\boldsymbol{x})) + (\nabla_{\boldsymbol{x}}\alpha)^\top (f(\boldsymbol{x}) + g(\boldsymbol{x})\xi) - (\nabla_{\boldsymbol{x}} V)^\top g(\boldsymbol{x}), \qquad c > 0$$

is a stabilizing controller.

If we apply theorem 32 to a system expressed in strict feedback form, then we can recursively define controllers until we arrive at a controller for the full system.