Robot Control - RBE502

Siamak Faal

C-D Semester, 2024

Contents

A Brief Review of Linear Algebra

Vector Space

Definition 6.1 (Vector Space)

V is a *vector space* over the filed $\mathbb{F}$, if for all $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z} \in V$ and α, $\beta \in \mathbb{F}$ we have

1. Vector addition: $\mathbf{x} + \mathbf{y} \in V$,
2. Commutativity of addition: $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$,
3. Associativity of addition: $(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{z} + (\mathbf{y} + \mathbf{z})$,
4. Identity element of addition: $\exists \mathbf{0} \in V$ such that $\mathbf{x} + \mathbf{0} = \mathbf{x}$,
5. Inverse elements of addition: $\exists (-\mathbf{x})$ such that $\mathbf{x} + (-\mathbf{x}) = \mathbf{0}$,
6. Scalar multiplication: $\alpha\,\mathbf{x} \in V$,
7. Compatibility of scalar multiplication with field multiplication: $(\alpha\,\beta)\,\mathbf{x} = \alpha\,(\beta\,\mathbf{x})$,
8. Distributivity of scalar multiplication with respect to vector addition: $\alpha\,(\mathbf{x} + \mathbf{y}) = \alpha\,\mathbf{x} + \alpha\,\mathbf{y}$,
9. Distributivity of scalar multiplication with respect to field addition: $(\alpha + \beta)\,\mathbf{x} = \alpha\,\mathbf{x} + \beta\,\mathbf{x}$,
10. Identity element of scalar multiplication: $\exists 1 \in \mathbb{F}$ such that $1\,\mathbf{x} = \mathbf{x}$.

Examples of Vector Space

The set $\mathbb{F}^n$ of all n-tuples with elements in $\mathbb{F}$ is a vector space.

Proof. First we check vector addition. Let

$$\mathbf{a} = (a_1,\, a_2, \ldots,\, a_n) \quad \text{and} \quad \mathbf{b} = (b_1,\, b_2, \ldots,\, b_n)$$

for a_i and b_i in $\mathbb{F}$ for all i. Clearly $\mathbf{a},\, \mathbf{b} \in \mathbb{F}^n$. Moreover

$$\mathbf{a} + \mathbf{b} = (a_1,\, a_2, \ldots,\, a_n) + (b_1,\, b_2, \ldots,\, b_n) = (a_1 + b_1,\, a_2 + b_2, \ldots,\, a_n + b_n) \in \mathbb{F}^n.$$

Let $\lambda \in \mathbb{F}$, then

$$\lambda \mathbf{a} = \lambda\,(a_1,\, a_2, \ldots,\, a_n) = (\lambda a_1,\, \lambda a_2, \ldots,\, \lambda a_n) \in \mathbb{F}.$$

Since $0 \in \mathbb{F}$, then $\mathbf{0} = (0,\, 0, \ldots,\, 0) \in \mathbb{F}^n$ is the identity element of $\mathbb{F}^n$ since

$$\mathbf{a} + \mathbf{0} = (a_1,\, a_2, \ldots,\, a_n) + (0,\, 0, \ldots,\, 0) = (a_1 + 0,\, a_2 + 0, \ldots,\, a_n + 0) = \mathbf{a}$$

The remaining properties directly follow from the properties of $\mathbb{F}$. $\square$

Examples of Vector Space (contd.)

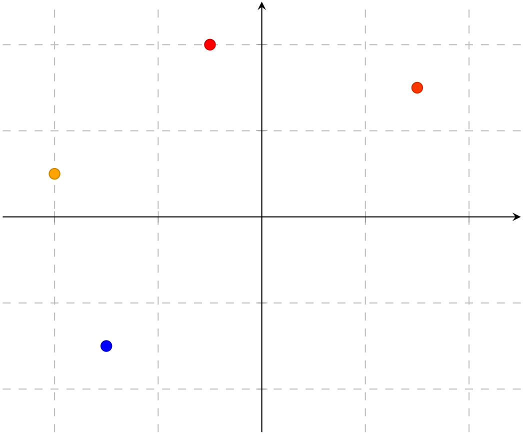
Special examples of $\mathbb{F}^n$ are $\mathbb{R}^n$ and $\mathbb{C}^n$. As an example consider $\mathbb{R}^2$:

Let

$$\mathbf{a} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}, \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} -4 \\ 1 \end{bmatrix},$$

Then

$$\mathbf{a}+\mathbf{b} = \begin{bmatrix} -1 \\ 4 \end{bmatrix}, \quad -1\mathbf{a} = \begin{bmatrix} -3 \\ -3 \end{bmatrix}, \quad \mathbf{a}+\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \mathbf{a}.$$



Examples of Vector Space (contd.)

Let $\mathcal{P}_n(\mathbb{F})$ be the set of all polynomials of degree n with coefficients in $\mathbb{F}$:

$$\mathcal{P}_n(\mathbb{F}) = \left\{a_0x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n \mid a_i \in \mathbb{F},\, \forall i\right\}.$$

We can show that $\mathcal{P}_n(\mathbb{F})$ is a vector space. First, let

$$p_1(x) = \sum_{i=0}^{n} a_i x^{n-i}, \quad \text{and} \quad p_2 = \sum_{i=0}^{n} b_i x^{n-i}$$

in $\mathcal{P}_n(\mathbb{F})$, then

$$(p_1 + p_2)(x) = \sum_{i=0}^{n} (a_i + b_i)\, x^{n-i} \in \mathcal{P}_n(\mathbb{F}).$$

Moreover $\mathbf{0} = \sum_{i=0}^{n} 0\, x^{n-i} \in \mathcal{P}_n(\mathbb{F})$ and for $\lambda \in \mathbb{F}$,

$$\lambda\, p_1(x) = \lambda \sum_{i=0}^{n} a_i x^{n-i} = \sum_{i=0}^{n} \lambda a_i x^{n-i} \in \mathcal{P}_n(\mathbb{F}).$$

Examples of Vector Space (contd.)

Let V be the set of all continuous functions with the domain $[a, b]$ with a zero at a:

$$V = \left\{ f : [a, b] \to \mathbb{R} \mid f \in C^0, \, f(a) = 0 \right\}$$

We can show that V is a vector space. In fact, let $f_1 + f_2$ in V, then

$$f_1(x) + f_2(x) \in C^0, \quad \text{and} \quad f_1(a) + f_2(a) = 0 + 0 = 0 \implies f_1 + f_2 \in V.$$

Moreover $f = 0 \in V$ and for any $\lambda \in \mathbb{R}$, $\lambda f \in V$.

On the contrary, let W be the set of all continuous functions with the domain $[a, b]$ such that $f(a) = 1$. We can easily show that W is **not** a vector space. In fact, let $f_1, f_2 \in W$, then

$$f_1(a) + f_2(a) = 1 + 1 = 2 \implies f_1 + f_2 \notin W.$$

Subspace

Definition 6.2 (Subspace)

A *subspace* of a vector space V is a subset $S \subset V$ that itself is a vector space with vector addition and scalar multiplication defined on V.

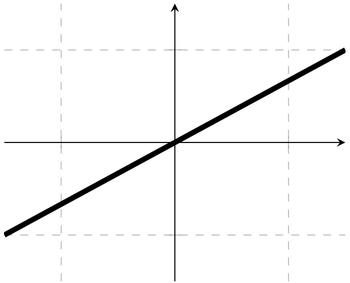
Example

Let $V \subset \mathbb{R}^2$ defined as

$$V = \left\{ \mathbf{x} = (x_1,\, x_2) \in \mathbb{R}^2 \mid 2x_1 - 3x_2 = 0 \right\}$$

Let $\mathbf{x} = (x_1,\, x_2)$, $\mathbf{y} = (y_1,\, y_2) \in V$ and $\lambda \in \mathbb{R}$ then

$$\begin{aligned} 2x_1 - 3x_2 + 2y_1 - 3y_2 &= 0 + 0 = 0, \\ \lambda\,(2x_1 - 3x_2) &= \lambda \cdot 0 = 0, \\ 2 \cdot 0 - 3 \cdot 0 &= 0. \end{aligned}$$



Linear Independence

Definition 6.3 (Linear Independence)

A sequence of vectors $\mathbf{v}_1$, $\mathbf{v}_2$ $\ldots$, $\mathbf{v}_k \in V$ are *linearly independent* if for scalars $\alpha_i \in \mathbb{F}$, $i = 1, \ldots, k$ we have

$$\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \cdots + \alpha_k \mathbf{v}_k = \mathbf{0} \iff \alpha_i = 0, \forall i = 1, \ldots, k.$$

A sequence of vectors $\mathbf{v}_1$, $\mathbf{v}_2$ $\ldots$, $\mathbf{v}_k \in V$ are *linearly dependent* if they are not linearly independent.

Two immediate consequences of Definition 6.3 are

1. Any subset of a linearly independent set is also linearly independent;
2. A set of linearly independent vectors cannot include the zero vector.

Example

Consider $\mathbf{x}_1$, $\mathbf{x}_2$, and $\mathbf{x}_3$ in $\mathbb{R}^3$ defined as

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix}, \quad \mathbf{x}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{x}_3 = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}.$$

We can confirm that $\mathbf{x}_1$, $\mathbf{x}_2$, and $\mathbf{x}_3$ are linearly **independent**. In fact

$$\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + \alpha_3\mathbf{x}_3 = \mathbf{0}$$

leads to

$$\begin{aligned} \alpha_1 - \alpha_2 = 0 &\implies \alpha_1 = \alpha_2, \\ 2\alpha_1 - \alpha_3 = 0 &\implies \alpha_3 = 2\alpha_1, \\ -2\alpha_1 + \alpha_2 + \alpha_3 = 0 &\implies \alpha_1 = 0, \end{aligned}$$

which implies $\alpha_1 = \alpha_2 = \alpha_3 = 0$.

Example (contd.)

As an example of linearly **dependent** vectors in $\mathbb{R}^3$, consider $\mathbf{y}_1$, $\mathbf{y}_2$, and $\mathbf{y}_3$ defined as

$$\mathbf{y}_1 = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}, \quad \mathbf{y}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{y}_3 = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix},$$

Let $\alpha_1 = 1$, $\alpha_2 = -1$ and $\alpha_3 = 2$, then

$$\alpha_1\mathbf{y}_1 + \alpha_2\mathbf{y}_2 + \alpha_3\mathbf{y}_3 \implies \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} + \begin{bmatrix} -1 \\ 0 \\ -1 \end{bmatrix} + \begin{bmatrix} 0 \\ -2 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

Span

Definition 6.4 (Span)

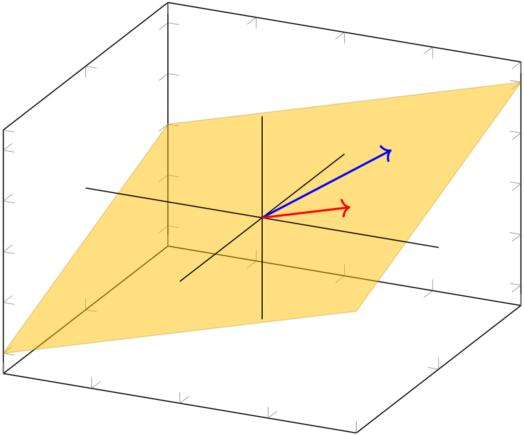
The *span* of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k \in V$ is the set of all linear combinations of $\mathbf{v}_1, \ldots, \mathbf{v}_k$. That is

$$S = \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\} = \sum_{i=1}^{k} \alpha_i \mathbf{v}_i, \quad \forall \alpha_i \in \mathbb{F}.$$

If $S \subset V$ and $S = \text{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ then we say $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is a *spanning set* of S.

Example

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$



Basis

Definition 6.5 (Basis)

The set $\{\mathbf{v}_i\}$, $\mathbf{v}_i \in V$ for $i = 1, \ldots, k$, is a *basis* for V if

1. $\mathbf{v}_i$ are linearly independent.
2. $V = \text{span}\{\mathbf{v}_i\}$.

Example The following are examples of a basis for $\mathbb{R}^2$:

$$\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}, \quad \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix} \right\}, \quad \left\{ \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right\}$$

Definition 6.6 (Finite-Dimensional Vector Space)

If $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ for $n < \infty$ is a basis for V, then V is a *finite-dimensional vector space* and the dimension of V is n; $\dim(V) = n$.

Unique Representation

Theorem 6.1

If $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ is a basis for V, then every $\mathbf{x} \in V$ can be uniquely defined as a linear combination of $\mathbf{v}_i$.

Proof. Since $V = \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$, then $\exists \alpha_i \in \mathbb{F}$, $i = 1, \ldots, k$ such that

$$\mathbf{x} = \sum_{i=1}^{k} \alpha_i \, \mathbf{v}_i.$$

To prove uniqueness, assume $\exists \beta_i \in \mathbb{F}$ such that $\mathbf{x} = \sum_{i=1}^{k} \beta_i \, \mathbf{v}_i$. we have

$$\mathbf{0} = \mathbf{x} - \mathbf{x} = \sum_{i=1}^{k} \alpha_i \, \mathbf{v}_i - \sum_{i=1}^{k} \beta_i \, \mathbf{v}_i = \sum_{i=1}^{k} \underbrace{(\alpha_i - \beta_i)}_{=\gamma_i} \, \mathbf{v}_i = \sum_{i=1}^{k} \gamma_i \, \mathbf{v}_i.$$

Since $\mathbf{v}_i$ are linearly independent, $\sum_{i=1}^{k} \gamma_i \, \mathbf{v}_i = \mathbf{0}$ implies $\gamma_i = 0$. Hence $\alpha_i = \beta_i$, $\forall i$. $\square$

Linear Transformation

Definition 6.7 (Linear Transformation)

A *linear transformation* form vector space V to vector space W, both defined over filed $\mathbb{F}$, is a map $T : V \to W$ that satisfies

$$T(\alpha\,\mathbf{x} + \beta\,\mathbf{y}) = \alpha\,T(\mathbf{x}) + \beta\,T(\mathbf{y}), \quad \forall\,\mathbf{x}, \mathbf{y} \in V, \text{ and } \alpha, \beta \in \mathbb{F}.$$

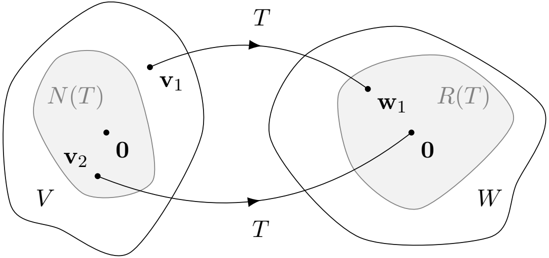
Accordingly, the vector space V is the *domain* and the vector space W is the *co-domain* of the linear transformation T.

Range and Null Space

Definition 6.8 (Range and Null Space)

Let $T : V \to W$ be linear. The *range* or *image* of T, denoted with $R(T)$ or $\mathrm{im}(T)$; And the *null space* or *kernel* of T, denoted with $N(T)$ or $\ker(T)$ are

$$R(T) := \left\{ \mathbf{w} \in W \mid \exists \mathbf{v} \in V \text{ such that } T(\mathbf{v}) = \mathbf{w} \right\} \subset W.$$
$$N(T) := \left\{ \mathbf{v} \in V \mid T(\mathbf{v}) = \mathbf{0} \right\} \subset V.$$



The Rank–Nullity Theorem

Proposition 6.2

Let $T : V \to W$ be a linear transformation, then

1. *Range of T is a subspace of W,*
2. *Null space of T is a subspace of V.*

Theorem 6.3 (The Rank–Nullity Theorem)

Let $T : V \to W$ be linear with $\dim V < \infty$. Then

$$\mathrm{rank}(T) + \mathrm{nullity}(T) = \dim V,$$

where $\mathrm{rank}(T)$ is the dimension of its image. and $\mathrm{nullity}(T)$ is the dimension of its kernel. In other words,

$$\dim(\mathrm{im}\, T) + \dim(\ker T) = \dim(\mathrm{domain}(T)).$$

Proofs are left as an exercise.

Matrix Representation

Theorem 6.4

Let $f : V \to W$ be linear with $\dim(V) = n < \infty$ and $\dim(W) = m < \infty$. Then there exists a matrix $\mathbf{A} \in F^{m \times n}$ such that, for every $\mathbf{v} \in V$, $f(\mathbf{v}) = \mathbf{A}\mathbf{v}$.

Construct.
Let $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be a basis for V. Then every $\mathbf{v} \in V$ is uniquely determined by the coefficients c_1 to c_n in $\mathbb{F}$ as

$$\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n.$$

Since $f : V \to W$ is linear

$$f(\mathbf{v}) = f(c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n) = c_1 f(\mathbf{v}_1) + \cdots + c_n f(\mathbf{v}_n),$$

Matrix Representation (contd.)

which implies that the function f is entirely determined by the vectors $f(\mathbf{v}_1), \ldots, f(\mathbf{v}_n)$. Let $\{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$ be a basis for W. Hence, we can represent each vector $f(\mathbf{v}_j)$ as

$$f(\mathbf{v}_j) = a_{1j}\mathbf{w}_1 + \cdots a_{mj}\mathbf{w}_m.$$

Thus, the function f is entirely determined by the values of a_{ij}. If we put these values into an $m \times n$ matrix $\mathbf{A}$ defined as

$$\mathbf{A} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}$$

we can conveniently use it to compute the vector output of f for any vector in V.

Contents

A Short Review of Numerical Solutions of IVP

Notations

Let $\mathbf{x} \in \mathbb{R}^n$, the n-dimensional point in the space is denoted with

$$\mathbf{x} = (x_1,\ x_2,\ \ldots,\ x_n),$$

and the structure of the vector is denoted with

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

Notations (contd.)

The short hand notation for

$$
\begin{aligned}
z_1 &= f_1(t,\ x_1,\ \ldots, x_n,\ y_1,\ \ldots y_m),\\
z_2 &= f_2(t,\ x_1,\ \ldots, x_n,\ y_1,\ \ldots y_m),\\
&\vdots\\
z_p &= f_p(t,\ x_1,\ \ldots, x_n,\ y_1,\ \ldots y_m),
\end{aligned}
$$

is

$$\mathbf{z} = \mathbf{f}(t,\ \mathbf{x},\ \mathbf{y}).$$

If $t \in \mathbb{R}$, $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{y} \in \mathbb{R}^m$, then

$$\mathbf{f} : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^p.$$

Solution of an IVP

Let $\mathbf{x} \in \mathbb{R}^n$ and consider the following Initial-Value Problem (IVP)

$$\begin{cases} \dot{\mathbf{x}} = \mathbf{f}\left(t,\, \mathbf{x}\right), \quad t > t_0, \\ \mathbf{x}(t_0) = \mathbf{x}_0, \end{cases} \tag{7.1}$$

where $\mathbf{f}$ is a *sufficiently* smooth function. Based on the fundamental theorem of calculus we have

$$\mathbf{x}(t_f) = \mathbf{x}_0 + \int_{t_0}^{t_f} \mathbf{f}\left(t,\, \mathbf{x}\right)\, dt. \tag{7.2}$$

Thus, to find the solution $\mathbf{x}(t_f)$, we need to compute an integral. Hence, the problem of finding the solution $\mathbf{x}(t_f)$ is commonly referred to as an *integration problem*.

Time Discretization

Suppose we divide the time interval $[t_0, t_f]$ into n equally intervals

$$t_0, t_1, \ldots, t_{n-1}, t_n = t_f. \tag{7.3}$$

Using this discretization, we get

$$\mathbf{x}(t_f) = \mathbf{x}(t_0) + \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} \mathbf{f}\left(t, \mathbf{x}(t)\right) dt. \tag{7.4}$$

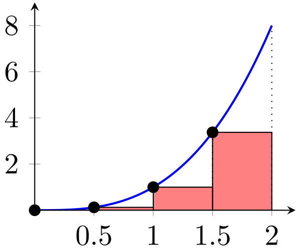
Let $\mathbf{x}_k = \mathbf{x}(t_k)$, the above equality implies

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \int_{t_k}^{t_{k+1}} \mathbf{f}\left(t, \mathbf{x}(t)\right) dt. \tag{7.5}$$

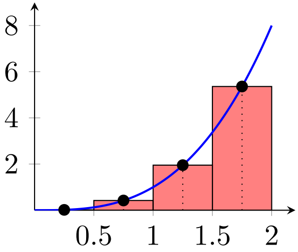
If we can find a way to approximate the integral on the right side of (7.5), we can repeatedly propagate the approximation of $\mathbf{x}(t)$ from time t_k, to time t_{k+1}, thus obtaining an approximation for $\mathbf{x}(t)$ at any desired time t.

Rectangular Integration

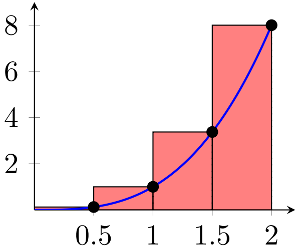
Example of left, middle and right Riemann sums of $f(x) = x^3$ over $x \in [0, 2]$ using 4 equal subintervals:



Left Riemann sum



Middle Riemann sum

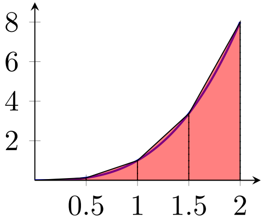


Right Riemann sum

Trapezoidal rule

The trapezoidal rule approximates the integral of the function as the area of a trapezoid

$$\int_a^b f(x)\,dx \approx (b-a)\frac{1}{2}\left(f(a)+f(b)\right).$$



Let x_k be a partition of $[a,b]$ such that

$$a = x_0 < x_1 < \cdots < x_{N-1} < x_N = b$$

and $h_k = x_k - x_{k-1}$ be the length of the k^{th} subinterval, then

$$\int_a^b f(x)\,dx \approx \sum_{k=1}^{N} \frac{f(x_{k-1})+f(x_k)}{2} h_k.$$

The Runge–Kutta method (RK4)

Let (7.1) be the initial value problem. To find an approximation of $\mathbf{x}(t)$ we pick a step size $h > 0$ and for $n \in \{0,\ 1,\ 2,\ 3, \dots\}$ define

$$\mathbf{x}_{n+1} = \mathbf{x}_n + \frac{h}{6}\left(\mathbf{k}_1 + 2\mathbf{k}_2 + 2\mathbf{k}_3 + \mathbf{k}_4\right), \tag{7.6a}$$

$$t_{n+1} = t_n + h, \tag{7.6b}$$

where

$$\mathbf{k}_1 = \mathbf{f}(t_n, \mathbf{x}_n), \tag{7.7a}$$

$$\mathbf{k}_2 = \mathbf{f}\left(t_n + h/2,\ \mathbf{x}_n + h\mathbf{k}_1/2\right), \tag{7.7b}$$

$$\mathbf{k}_3 = \mathbf{f}\left(t_n + h/2,\ \mathbf{x}_n + h\mathbf{k}_2/2\right), \tag{7.7c}$$

$$\mathbf{k}_4 = \mathbf{f}\left(t_n + h,\ \mathbf{x}_n + h\mathbf{k}_3\right). \tag{7.7d}$$

⚠ the above equations have different but equivalent definitions in different texts.

State and State Variables

State Variables

State of a Dynamic System

The *state* of a dynamic system is *the smallest set* of variables (called state variables) such that knowledge of these variables at a specific time t_0, together with knowledge of the input for $t > t_0$, completely determines the behavior of the system for any time $t \geq t_0$.

State Vector. If n state variables x_1, x_2 to x_n are needed to completely describe the behavior of a given system, then these n state variables can be considered the n components of an n-dimensional vector

$$\mathbf{x} = (x_1, x_2, \ldots, x_n)$$

State Space. The n-dimensional space where the state variable x_i defines the coordinates along the i^{th} axis.

Any state $\mathbf{x}$ is a *point* in the state space.

State Space Representation

A state space representation is a mathematical model of a physical system specified as a set of input, output and variables related by **first-order** differential equations or difference equations:

$$\dot{\mathbf{x}} = \mathbf{f}\left(t,\, \mathbf{x},\, \mathbf{u}\right), \tag{7.8a}$$
$$\mathbf{y} = \mathbf{g}\left(t,\, \mathbf{x},\, \mathbf{u}\right). \tag{7.8b}$$

$$\mathbf{x}_{k+1} = \mathbf{f}_k\left(t_k,\, \mathbf{x}_k,\, \mathbf{u}_k\right), \tag{7.9a}$$
$$\mathbf{y}_k = \mathbf{g}_k\left(t_k,\, \mathbf{x}_k,\, \mathbf{u}_k\right). \tag{7.9b}$$

⚠ The choice of a state vector for a dynamic system is not *unique*.

? How can we write system dynamics in state space form?

A possible approach is to formulate the dynamics as a system of first order differential (or difference) equations.

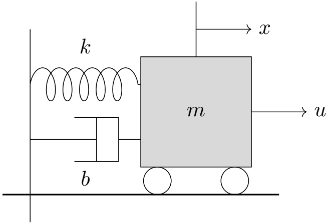
Example 1

The equations of motion for the mass-spring-damper system is

$$m\ddot{x} + b\dot{x} + kx = u.$$

We need a system of first order differential equations to form the state space representation. Let

$$\left.\begin{aligned} x_1 &:= x \\ x_2 &:= \dot{x} \end{aligned}\right\} \implies \begin{cases} \dot{x}_1 = \dot{x} = x_2, \\ \dot{x}_2 = \frac{\mathrm{d}\dot{x}}{\mathrm{d}t} = \ddot{x}. \end{cases}$$



Accordingly, we have

$$\dot{\mathbf{x}} = \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} x_2 \\ \dfrac{1}{m}\left(u - kx_1 - bx_2\right) \end{bmatrix} = \mathbf{f}\left(\mathbf{x}, u\right),$$

$$y = \begin{bmatrix} 1 & 0 \end{bmatrix} \mathbf{x}.$$

Example 2

Consider the system

$$\dddot{y} + ay\ddot{y} + b\dot{y} + cy^3 = u,$$

where $y \in \mathbb{R}$ is the output and $u \in \mathbb{R}$ is the input. In order to convert the dynamics to a system of first order equations, let

$$x_1 := y, \quad x_2 := \dot{y}, \quad x_3 := \ddot{y}.$$

Accordingly,

$$\begin{aligned} \dot{x}_1 &= x_2, \\ \dot{x}_2 &= x_3, \\ \dot{x}_3 &= -ax_1x_3 - bx_2 - cx_1^3 + u, \end{aligned}$$

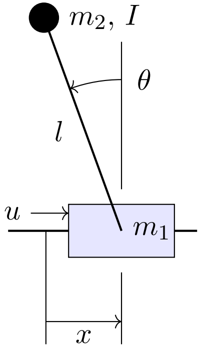
and

$$y = x_1.$$

Example 3 - Cart Pole

Consider a simplified cart-pole system. The Lagrangian of the system is

$$\mathcal{L} = \frac{1}{2}m_1\dot{x}^2 + \frac{1}{2}m_2\left(\dot{x}^2 + l^2\dot{\theta}^2 - 2l\dot{x}\dot{\theta}\cos(\theta)\right) + \frac{1}{2}I\dot{\theta}^2 - m_2 g l\cos(\theta).$$



Evaluating

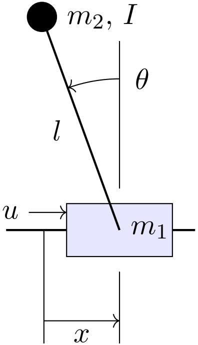
$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{L}}{\partial \dot{x}}\right) - \frac{\partial \mathcal{L}}{\partial x} = u,$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{L}}{\partial \dot{\theta}}\right) - \frac{\partial \mathcal{L}}{\partial \theta} = 0,$$

leads to the equations of motion of the system as

Example 3 - Cart Pole (contd.)

$$\left(m_1 + m_2\right)\ddot{x} - m_2 l\cos(\theta)\ddot{\theta} + m_2 l\dot{\theta}^2\sin(\theta) = u,$$
$$\left(m_2 l^2 + I\right)\ddot{\theta} - m_2 l\cos(\theta)\ddot{x} - m_2 g l\sin(\theta) = 0.$$



Note that, we can write the above equations as

$$\mathbf{M}\left(x,\,\theta\right)\begin{bmatrix}\ddot{x}\\ \ddot{\theta}\end{bmatrix} + \boldsymbol{\phi}\left(x,\,\theta,\,\dot{x},\,\dot{\theta}\right) = \begin{bmatrix}u\\ 0\end{bmatrix}$$

where

$$\mathbf{M}\left(x,\,\theta\right) = \begin{bmatrix} m_1 + m_2 & -m_2 l\cos(\theta) \\ -m_2 l\cos(\theta) & m_2 l^2 + I\end{bmatrix}, \quad \boldsymbol{\phi} = \begin{bmatrix} m_2 l\dot{\theta}^2\sin(\theta) \\ -m_2 g l\sin(\theta)\end{bmatrix}.$$

Example 3 - Cart Pole (contd.)

The determinant of $\mathbf{M}(x, \theta)$ is

$$\det\big(\mathbf{M}(x,\theta)\big) = I(m_1 + m_2) + \left(1 - \cos(\theta)^2 + \frac{m_1}{m_2}\right) m_2^2 l^2 > 0,$$

which implies $\mathbf{M}(x, \theta)$ is invertible. Hence, let

$$\mathbf{z}_1 := \begin{bmatrix} x \\ \theta \end{bmatrix}, \quad \mathbf{z}_2 := \dot{\mathbf{z}}_1 = \begin{bmatrix} \dot{x} \\ \dot{\theta} \end{bmatrix}, \quad \mathbf{u} := \begin{bmatrix} u \\ 0 \end{bmatrix}.$$

The state space representation of the system is

$$\begin{aligned} \dot{\mathbf{z}}_1 &= \mathbf{z}_2, \\ \dot{\mathbf{z}}_2 &= \big[\mathbf{M}(\mathbf{z}_1)\big]^{-1} \big(\mathbf{u} - \boldsymbol{\phi}(\mathbf{z}_1, \mathbf{z}_2)\big). \end{aligned}$$

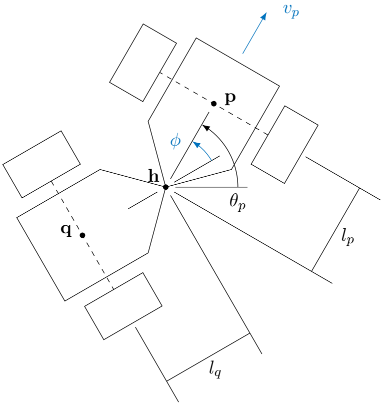
Example 4 - Articulated Vehicle

The kinematics of an articulated vehicle is

$$\dot{p}_1 = v_p \cos\left(\theta_p\right),$$

$$\dot{p}_2 = v_p \sin\left(\theta_p\right),$$

$$\dot{\theta}_p = \frac{v_p \sin(\phi) + l_q \dot{\phi}}{l_p \cos(\phi) + l_q}.$$



Example 4 - Articulated Vehicle (contd.)

Let $(x_1,\ x_2,\ x_3,\ x_4) = \big(p_1,\ p_2,\ \theta_p,\ \phi\big)$, and $(u_1,\ u_2) = (\mathbf{v}_p,\ \dot{\phi})$ then

$$
\begin{aligned}
\dot{x}_1 &= u_1 \cos(x_3),\\
\dot{x}_2 &= u_1 \sin(x_3),\\
\dot{x}_3 &= \frac{u_1 \sin(x_4) + l_q u_2}{l_p\ \cos\left(x_4\right) + l_q},\\
\dot{x}_4 &= u_2.
\end{aligned}
$$

Common State Space Representations

Linear Systems

Definition 7.1 (Linear System)

Let $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{u}$ belong to vector spaces X, Y, and U, respectively. A dynamic system

$$\begin{aligned}\dot{\mathbf{x}} &= \mathbf{f}\left(t,\, \mathbf{x},\, \mathbf{u}\right),\\ \mathbf{y} &= \mathbf{g}\left(t,\, \mathbf{x},\, \mathbf{u}\right),\end{aligned}$$

is linear if $\mathbf{f}$ and $\mathbf{g}$ are linear transformations with respect to $\mathbf{x}$ and $\mathbf{u}$.

If $\mathbf{f}$ and $\mathbf{g}$ are linear with respect to $\mathbf{x}$ and $\mathbf{u}$, then

$$\begin{aligned}\dot{\mathbf{x}} &= \mathbf{f}\left(t,\, \mathbf{x},\, \mathbf{u}\right) = \mathbf{A}(t)\mathbf{x} + \mathbf{B}(t)\mathbf{u},\\ \mathbf{y} &= \mathbf{g}\left(t,\, \mathbf{x},\, \mathbf{u}\right) = \mathbf{C}(t)\mathbf{x} + \mathbf{D}(t)\mathbf{u}.\end{aligned}$$

If $\dim(X) = n$, $\dim(Y) = p$, and $\dim(U) = m$ are finite, then $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ and $\mathbf{D}$ are $n \times n$, $n \times m$, $p \times n$ and $p \times m$ matrices, respectively.

Example

A simplified version of an automobile suspension system:

$$
\begin{aligned}
m_1\ddot{x} + b\dot{x} + (k_1 + k_2)x &= b\dot{y} + k_2 y + k_1 u, \\
m_2\ddot{y} + b\dot{y} + k_2 y &= b\dot{x} + k_2 x.
\end{aligned}
$$

Let $\mathbf{z} = (z_1,\, z_2,\, z_3,\, z_4) := (x,\, y,\, \dot{x},\, \dot{y})$ then

$$
\begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \\ \dot{z}_3 \\ \dot{z}_4 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -\frac{k_1+k_2}{m_1} & \frac{k_2}{m_1} & -\frac{b}{m_1} & \frac{b}{m_1} \\ \frac{k_2}{m_2} & -\frac{k_2}{m_2} & \frac{b}{m_2} & -\frac{b}{m_1} \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \frac{k_1}{m_1} \\ 0 \end{bmatrix} u,
$$

$$
y = \begin{bmatrix} 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{bmatrix}.
$$

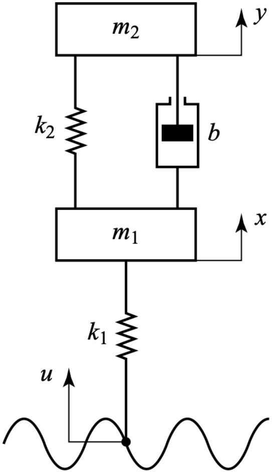


Image src: Katsuhiko, Ogata; Modern control engineering

n$^{\text{th}}$-Order Linear System With No Input Derivatives

Consider the following n$^{\text{th}}$-order linear system defined by

$$y^{(n)} + a_1\, y^{(n-1)} + \cdots + a_{n-1}\, \dot{y} + a_n\, y = u \tag{7.11}$$

Note that the knowledge of

- $y(t_0),\ \dot{y}(t_0),\ \ldots,\ y^{(n-1)}(t_0)$ and
- $u(t)$ for $t \geq t_0$

completely determines the future behavior of the system. Hence we may take

$$\mathbf{x} = (x_1,\, x_2, \ldots,\, x_n) = \left(y,\, \dot{y}, \ldots,\, y^{(n-1)}\right)$$

as the state vector for the system.

n^{th}-Order Linear System With No Input Derivatives (contd.)

Using the state vector $\mathbf{x} = (x_1,\, x_2, \ldots,\, x_n) = \left(y,\, \dot{y}, \ldots,\, y^{(n-1)}\right)$ we can write

$$y^{(n)} + a_1\, y^{(n-1)} + \cdots + a_{n-1}\, \dot{y} + a_n\, y = u$$

the derivative of the state vector as

$$\begin{aligned} \dot{x}_1 &= x_2, \\ \dot{x}_2 &= x_3, \\ &\vdots \\ \dot{x}_n &= -a_n x_1 - a_{n-1} x_2 - \cdots - a_1 x_n + u, \end{aligned}$$

n$^{\text{th}}$-Order Linear System With No Input Derivatives (contd.)

Which could be written in matrix form as

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \vdots \\ \dot{x}_n \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ -a_n & -a_{n-1} & -a_{n-2} & \cdots & -a_1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} u \tag{7.12a}$$

$$y = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \tag{7.12b}$$

n^th-Order Linear System With Input Derivatives

Consider the differential equation that involves derivatives of the input, such as

$$y^{(n)} + a_1\, y^{(n-1)} + \cdots + a_{n-1}\, \dot{y} + a_n\, y = b_0\, u^{(n)} + b_1\, u^{(n-1)} + \cdots + b_{n-1}\, \dot{u} + b_n\, u \qquad (7.13)$$

where u is the input and y is the output.
We need to chose state variables such that they eliminate the derivatives of u in the state equation.
A possible solution is to choose

$$\begin{aligned}
x_1 &= y - \beta_0 u, \\
x_2 &= \dot{y} - \beta_0 \dot{u} - \beta_1 u = \dot{x}_1 - \beta_1 u, \\
x_3 &= \ddot{y} - \beta_0 \ddot{u} - \beta_1 \dot{u} - \beta_2 u = \dot{x}_2 - \beta_2 u, \\
&\ \vdots \\
x_n &= y^{(n-1)} - \beta_0 u^{(n-1)} - \beta_1 u^{(n-2)} - \cdots - \beta_{n-2} \dot{u} - \beta_{n-1} u = \dot{x}_{n-1} - \beta_{n-1} u,
\end{aligned}$$

n^{th}-Order Linear System With Input Derivatives (contd.)

where

$$
\begin{aligned}
\beta_0 &= b_0, \\
\beta_1 &= b_1 - a_1\beta_0, \\
\beta_2 &= b_2 - a_1\beta_1 - a_2\beta_0, \\
\beta_3 &= b_3 - a_1\beta_2 - a_2\beta_1 - a_3\beta_0, \\
&\vdots \\
\beta_{n-1} &= b_{n-1} - a_1\beta_{n-2} - \cdots - a_{n-2}\beta_1 - a_{n-1}\beta_0.
\end{aligned}
$$

n^{th}-Order Linear System With Input Derivatives (contd.)

With this choice of state variables the existence and uniqueness of the solution of the state equation is guaranteed. Taking the derivatives of x_i with respect to time gives

$$\begin{aligned}
\dot{x}_1 &= x_2 + \beta_1 u, \\
\dot{x}_2 &= x_3 + \beta_2 u, \\
&\vdots \\
\dot{x}_{n-1} &= x_n + \beta_{n-1} u, \\
\dot{x}_n &= -a_n x_1 - a_{n-1} x_2 - \cdots - a_1 x_n + \beta_n u.
\end{aligned}$$

n^{th}-Order Linear System With Input Derivatives (contd.)

In the matrix form we have:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \vdots \\ \dot{x}_{n-1} \\ \dot{x}_n \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & \cdots & -a_1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} + \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix} u \tag{7.14a}$$

$$y = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} + \beta_0 u. \tag{7.14b}$$

Autonomous Systems

Definition 7.2

If the function $\mathbf{f}$ in the dynamic system

$$\dot{\mathbf{x}} = \mathbf{f}(t,\ \mathbf{x})$$

does not *explicitly* depend on t, that is

$$\mathbf{f}(t,\ \mathbf{x}) = \mathbf{f}(\mathbf{x})$$

then the dynamic system is said to be *autonomous* or *time invariant*.

The behavior of an autonomous system is invariant to shifts in the time, since changing the t to $t + a$ does not change the state transition function $\mathbf{f}$.

Autonomous Systems (contd.)

Accordingly, a closed loop system

$$\dot{\mathbf{x}} = \mathbf{f}\left(\mathbf{x}, \mathbf{u}(\mathbf{x})\right), \tag{7.15}$$

is *autonomous* if both $\mathbf{f}$ and $\mathbf{u}$ do not *explicitly* depend on t ($\mathbf{u} = \mathbf{u}(x)$ is only a function of the state vector $\mathbf{x}$).

Furthermore, If $\mathbf{f}$ is linear with respect to $\mathbf{x}$ and $\mathbf{u}$ and also autonomous, the system is said to be a *linear time-invariant* (LTI) system, with the representation

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}. \tag{7.16}$$

All physical systems are *non-autonomous*. The concept of an autonomous system is an idealized notion (similar to linear systems). In practice, some system properties change very slowly that we can safely neglect their time variation.

Examples

A simple point-mass pendulum is an example of nonlinear autonomous system. In fact, the equation of motion of the system is

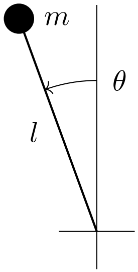
$$ml^2\ddot{\theta} = mgl\sin(\theta),$$

For $x_1 := \theta$ and $x_2 := \dot{\theta}$, the state space representation of the system is

$$\dot{x}_1 = x_2,$$

$$\dot{x}_2 = \frac{g}{l}\sin(x_1),$$

which does not explicitly depend on time.



Control Affine Systems

If the state transition function is an affine function of the input $\mathbf{u}$, then the system is said to be a *control affine* system. The state space form of a control affine system is

$$\dot{\mathbf{x}} = \mathbf{f}_a(t,\ \mathbf{x}) + \mathbf{f}_b(t,\ \mathbf{x})\mathbf{u}. \tag{7.17}$$

State Space Representation vs. Transfer Functions

State Space Representation vs. Transfer Functions

$$\dot{\mathbf{x}} = \mathbf{f}(t, \mathbf{x}, \mathbf{u}),$$
$$\mathbf{y} = \mathbf{g}(t, \mathbf{x}, \mathbf{u}).$$

(+) Suitable for MIMO.

(+) Provides a comprehensive view of the system dynamics.

(+) Suitable for non-linear and time-varying systems.

(−) Not straightforward in frequency domain.

$$\frac{Y(s)}{U(s)} = H(s)$$

(+) Simpler to analyze for SISO.

(+) Easier for frequency response analysis.

(−) Not suitable for non-linear or time-varying systems.

(−) Limited to LTI.

(−) Does not show internal states.

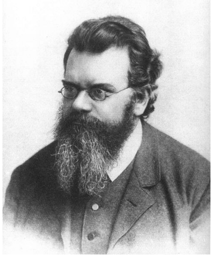
State space representation is versatile and comprehensive, ideal for complex, multi-variable systems, and applicable in both time and frequency domain analysis.

Contents

Phase Portrait Analysis

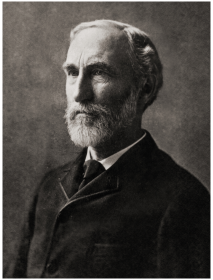
Phase Plane Analysis

- The concept of phase plane (space) was developed in the late 19th century by Ludwig Boltzmann, Henri Poincaré, and Josiah Willard Gibbs.
- Primarily applicable to two-dimensional systems, phase plane analysis is an insightful tool for building intuition about system dynamics.
- The idea is to generate a picture of “motion” trajectories corresponding to different initial conditions.



Ludwig Boltzmann

Henri Poincaré



Henri Josiah Willard Gibbs

Image sources: Wikipedia

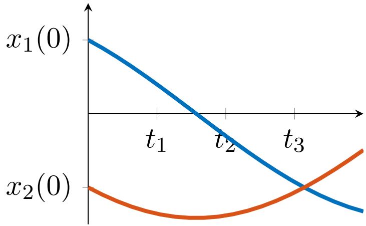
Phase Portrait

Consider the second order system

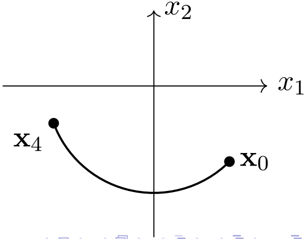
$$\dot{x}_1 = f_1(x_1,\ x_2),$$
$$\dot{x}_2 = f_2(x_1,\ x_2).$$

Let $\mathbf{x}_0 = \big(x_1(0),\, x_2(0)\big)$ be the initial condition; The solution $\mathbf{x}(t)$ may be depicted as

Trajectory in time



Trajectory in the state space



Drawing Phase Portraits

Consider the dynamic system $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ defined as

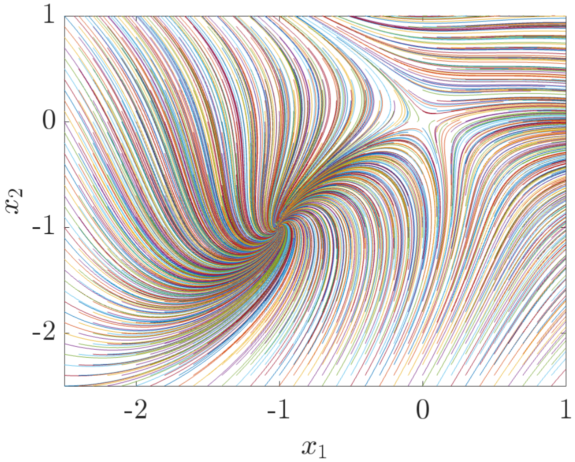
$$\dot{x}_1 = x_1 + x_2 + x_1^2 + x_2^2,$$
$$\dot{x}_2 = x_1 - x_2 + x_1^2 - x_2^2.$$

The portrait on the right is obtained by plotting the solution of

$$\begin{cases} \dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}), \\ \mathbf{x}(0) = \mathbf{x}_0, \end{cases}$$

for all $\mathbf{x}_0$ in

$$\{(x_1,\ x_2) \mid x_1, x_2 \in \{-2.5,\ -2.4, \ldots,\ 1\}\}$$

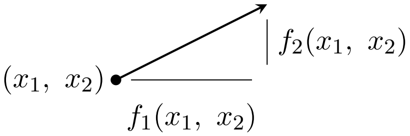


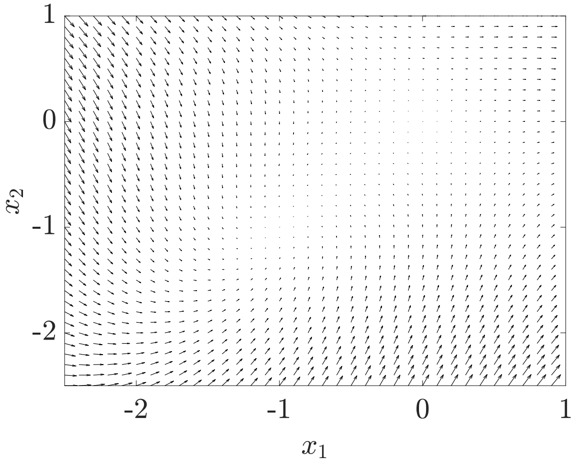
Drawing Phase Portraits (contd.)

For the same system, the portrait on the right is obtained by drawing arrows at every $\mathbf{x}_0$ in

$$\{(x_1,\ x_2) \mid x_1, x_2 \in \{-2.5,\ -2.4, \ldots,\ 1\}\}$$

where the direction of the arrow is defined as



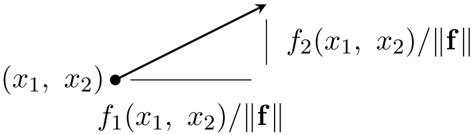


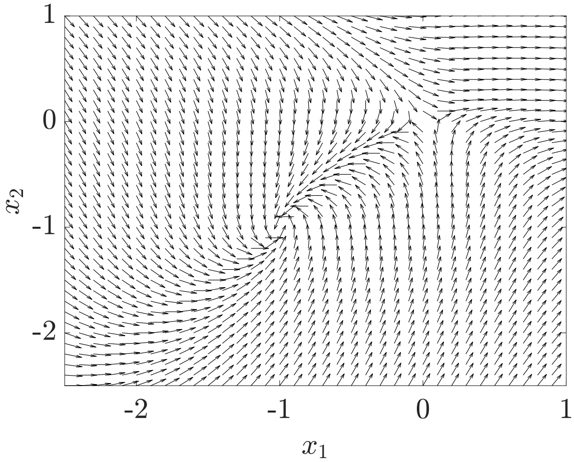
Drawing Phase Portraits (contd.)

To keep the emphasis on the direction, we can normalize the direction vectors. To do so, for every $\mathbf{x}_0$ in

$$\{(x_1,\ x_2) \mid x_1, x_2 \in \{-2.5,\ -2.4, \ldots,\ 1\}\}$$

we define the direction arrow as

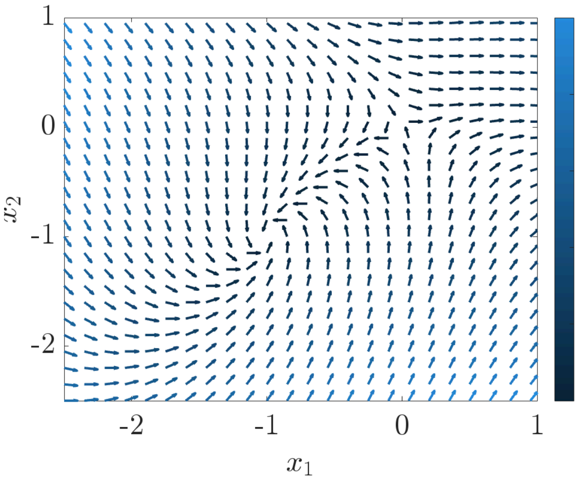




Drawing Phase Portraits (contd.)

To keep the emphasis both on the direction and the magnitude, we can color each *normalized* arrow based on $\|\mathbf{f}(\mathbf{x})\|$, as depicted in the image on the right. Here we have use the grid

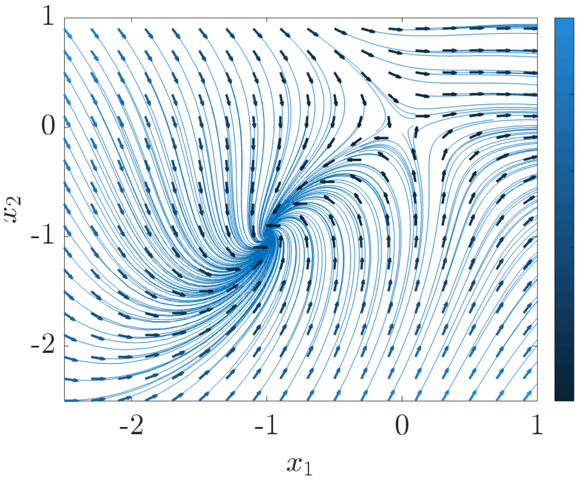
$$\{(x_1,\ x_2) \mid x_1, x_2 \in \{-2.5,\ -2.35, \ldots,\ 0.95\}\}$$



Drawing Phase Portraits (contd.)

The portrait on the right is obtained by drawing both the trajectories of the system starting from $\mathbf{x}_0$ along with the color-coded normalized arrows at every $\mathbf{x}_0$ for all $\mathbf{x}_0$ in

$$\{(x_1,\ x_2)\ |\ x_1, x_2 \in \{-2.5,\ -2.3, \ldots,\ 0.9\}\}$$



Sample MATLAB Code

A sample implementation is available on GitHub at: this link . ⚠ The following syntax is just for demonstration of the concept and it is *not* complete!

```
%...
[x1, x2] = meshgrid(x1_bounds(1):x1_res:x1_bounds(2), ...);
%...
options = odeset('Events', @(t,x)outofbound(t, x, x1_bounds, x2_bounds));
%...
for i = 1:numel(x1)
    x0 = [x1(i); x2(i)];
    [~, x] = ode45(f, tspan, x0, options);
    plot(ax, x(:,1), x(:,2), ...);
    dx = f(0, x0);
    norm_dx(i) = norm(dx);
    dx1(i) = dx(1) / norm_dx(i);
    dx2(i) = dx(2) / norm_dx(i);
end
```

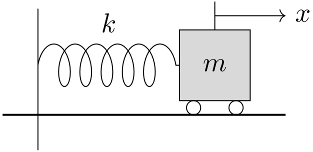
Sample MATLAB Code (contd.)

```
max_norm = max(norm_dx(:));

for i = 1:numel(x1)
    color = interp1(cindex, params.quiver_color_map, norm_dx(i) / max_norm);
    quiver(ax, x1(i), x2(i), dx1(i), dx2(i), ...);
end
end

function [position, isterminal, direction] = outofbound(~, x, x1_bounds, x2_bounds)
    x1_limit = [-0.1, 0.1] + x1_bounds;
    x2_limit = [-0.1, 0.1] + x2_bounds;
    position = (x1_limit(1) < x(1) & x(1) < x1_limit(2) & ...
        x2_limit(1) < x(2) & x(2) < x2_limit(2));
    isterminal = 1;
    direction = 0;
end
```

Phase Portrait of a Mass-Spring System

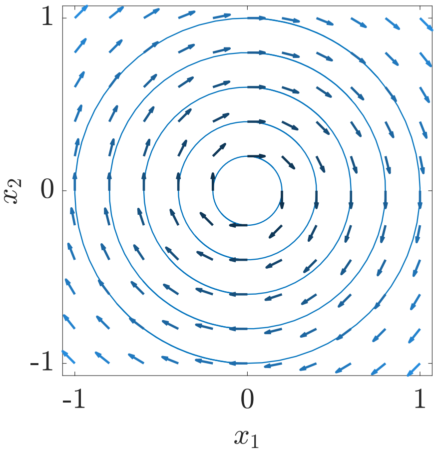


Assuming $m = k = 1$, the equations of motion of the system is

$$\ddot{x} + x = 0,$$

Let $x_1 = x$ and $x_2 = \dot{x}$, then

$$\dot{x}_1 = x_2,$$
$$\dot{x}_2 = -x_1.$$



Phase Portrait of a Mass-Spring System (contd.)

The solution of the IVP

$$\ddot{x} + x = 0, \quad x(0) = x_0, \quad \dot{x}(0) = v_0,$$

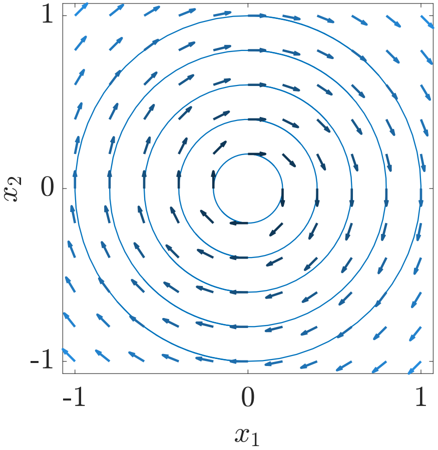
is

$$x(t) = x_1(t) = x_0 \cos(t) + v_0 \sin(t),$$
$$\dot{x}(t) = x_2(t) = -x_0 \sin(t) + v_0 \cos(t),$$

Note that

$$x_1^2(t) + x_2^2(t) = x_0^2 + v_0^2,$$

which indicates the trajectory remains in a constant distance from the origin.

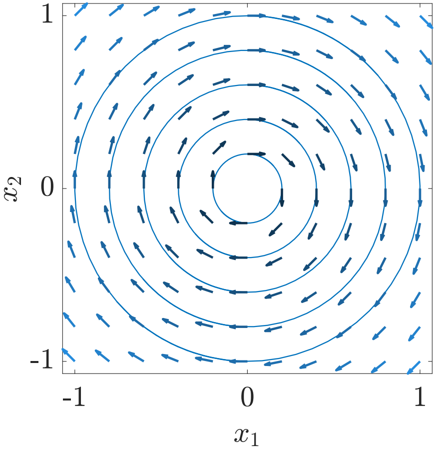


Phase Portrait of a Mass-Spring System (contd.)

Phase portrait of a system illustrates the nature of the system response to various initial conditions.

We can easily see that the trajectories of the mass-spring system does not converge to the origin or diverge to infinity... they remain on a fixed orbit around the origin for all time:

$$x_1^2(t) + x_2^2(t) = x_0^2 + v_0^2 = \text{constant}.$$



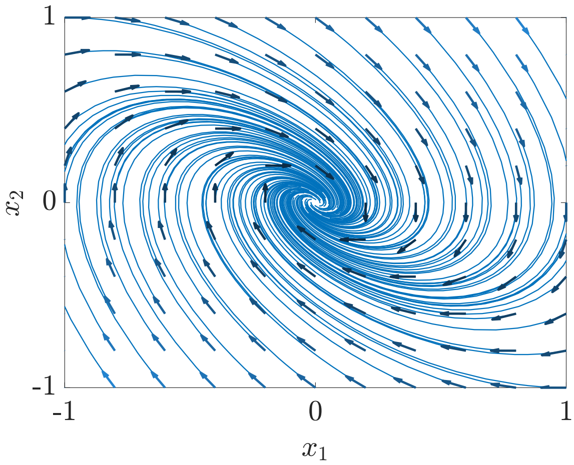
Phase Portrait of a Mass-Spring-Damper System

Addition of a friction changes the equations of motion to

$$\ddot{x} + \dot{x} + x = 0,$$

where we assumed the damping coefficient is 1. Let $x_1 = x$ and $x_2 = \dot{x}$, then

$$\begin{aligned}\dot{x}_1 &= x_2,\\ \dot{x}_2 &= -x_1 - x_2.\end{aligned}$$



Phase Portrait of a Mass-Spring-Damper System (contd.)

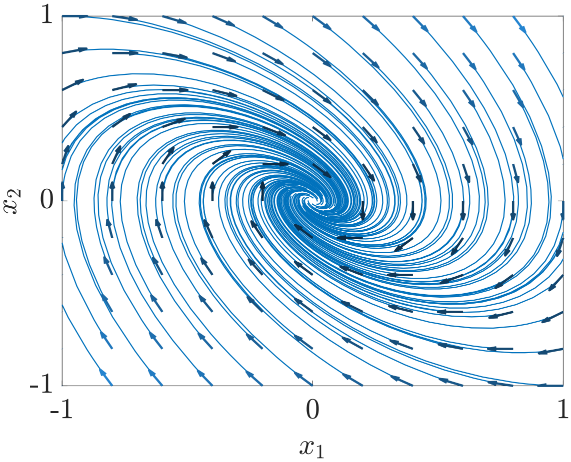
In order to analytically investigate the effect of the damping, let us focus on the total energy of the system

$$E(t) = \frac{1}{2}\dot{x}^2 + \frac{1}{2}x^2.$$

Taking the derivative of $E(t)$ with respect to t yields

$$\begin{aligned}\dot{E} &= \dot{x}\ddot{x} + x\dot{x} \\ &= \dot{x}\left(-\dot{x} - x + x\right) \\ &= -\dot{x}^2,\end{aligned}$$

which implies the system constantly loses energy when $\dot{x} \neq 0$.



Singular Points

A *singular point* is an *equilibrium point* in the phase plane. Precisely, $(x_1^*,\ x_2^*)$ is a singular point if

$$f_1(x_1^*,\ x_2^*) = f_2(x_1^*,\ x_2^*) = 0. \tag{8.18}$$

⚠ We will define *equilibrium points* of a system properly later in the course. For now, it suffices to say: *An equilibrium point is a point where the system can stay forever.*

? Why the point $(x_1^*,\ x_2^*)$ is called a *singular point*?

$$\frac{dx_2}{dx_1} = \frac{dx_2}{dt}\frac{dt}{dx_1} = \frac{f_2(x_1,\ x_2)}{f_1(x_1,\ x_2)}.$$

Thus, at $(x_1^*,\ x_2^*)$ we get the undefined quotient

$$\frac{dx_2}{dx_1}(x_1^*,\ x_2^*) = \frac{f_2(x_1^*,\ x_2^*)}{f_1(x_1^*,\ x_2^*)} = \frac{0}{0},$$

Singular Points (contd.)

Lemma 8.1

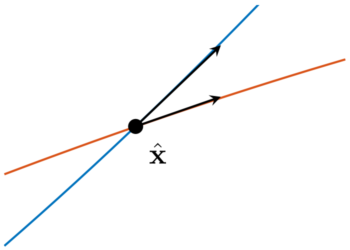
System trajectories do not intersect except possibly at singular points.

Proof. Assume two trajectories intersect at a non-singular point $(\hat{x}_1,\ \hat{x}_2)$. Accordingly, for the two trajectories we have

$$\text{trajectory 1: } \frac{dx_2}{dx_1}(\hat{x}_1,\ \hat{x}_2) = \frac{f_2(\hat{x}_1,\ \hat{x}_2)}{f_1(\hat{x}_1,\ \hat{x}_2)} = a,$$

$$\text{trajectory 2: } \frac{dx_2}{dx_1}(\hat{x}_1,\ \hat{x}_2) = \frac{f_2(\hat{x}_1,\ \hat{x}_2)}{f_1(\hat{x}_1,\ \hat{x}_2)} = b,$$

for $a \neq b$, which is a contradiction, since f_1 and f_2 are single valued functions.



Example: A Nonlinear Second Order System

Consider the system

$$\ddot{x} + 0.6\dot{x} + 3x + x^2 = 0,$$

and the state space form

$$\dot{x}_1 = x_2,$$
$$\dot{x}_2 = -3x_1 - x_1^2 - 0.6x_2.$$

The singular points of the system are

$$x_1^* = 0 \quad \text{and} \quad x_1^* = -3$$
$$x_2^* = 0$$

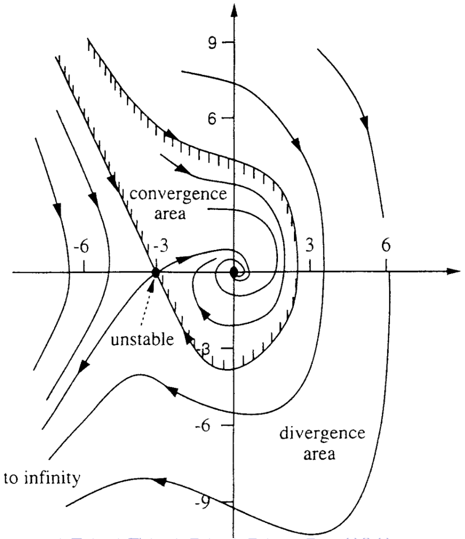
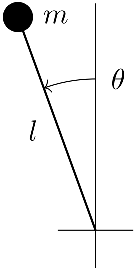


Image and example source: Slotine JJ, Li W. Applied nonlinear control.

A Beginner's Fallacy

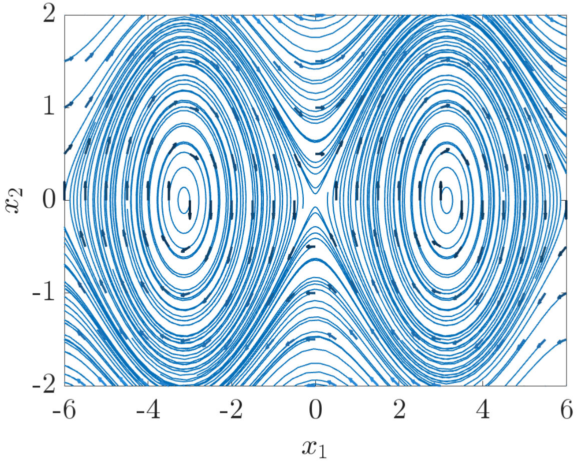
⚠ Shorter paths in state space do *not* necessarily correspond to shorter times.



$$ml^2\ddot{\theta} = mgl\sin(\theta)$$

Let $m = g = l = 1$, $x_1 = \theta$, $x_2 = \dot{\theta}$ then

$$\dot{x}_1 = x_2,$$
$$\dot{x}_2 = \sin(x_1).$$

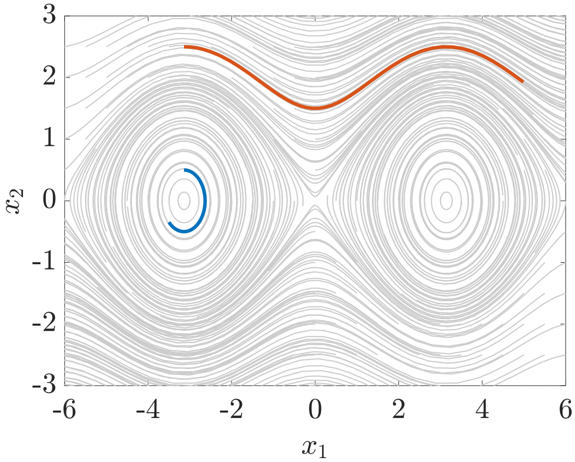


A Beginner's Fallacy (contd.)

Consider two trajectories with the initial conditions

1. $(-\pi, 0.5)$ depicted with blue,
2. $(-\pi, 2.5)$ depicted with orange,

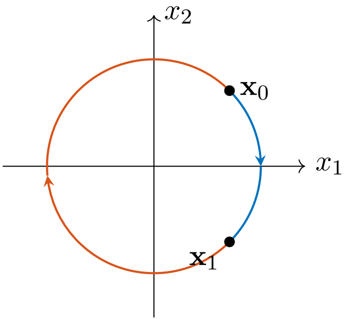
Both highlighted paths correspond to 4 seconds time interval!



A Beginner's Fallacy (contd.)

⚠ A path from point $\mathbf{x}_0$ to $\mathbf{x}_1$ is *not* necessarily also the path from $\mathbf{x}_1$ to $\mathbf{x}_0$.

Consider the mass-spring system with $\ddot{x} + x = 0$. Let $\mathbf{x}_0 = (1, 1)$ and $\mathbf{x}_1 = (1, -1)$



Phase Portrait of Systems with Control Input

? What if the considered system has control inputs?

Consider the second order system with control inputs

$$\dot{x}_1 = f_1(x_1,\, x_2,\, \mathbf{u}),$$
$$\dot{x}_2 = f_2(x_1,\, x_2,\, \mathbf{u}).$$

Substituting $\mathbf{u}$ with its definition, gives us the closed-loop form as

$$\dot{x}_1 = f_1\left(x_1,\, x_2,\, \mathbf{u}\left(x_1,\ x_2\right)\right) = \hat{f}_1(x_1,\, x_2),$$
$$\dot{x}_2 = f_2\left(x_1,\, x_2,\, \mathbf{u}\left(x_1,\ x_2\right)\right) = \hat{f}_2(x_1,\, x_2).$$

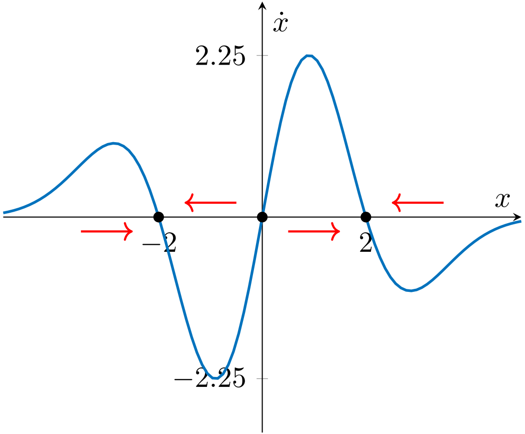
Phase Portraits in First Order Systems

Consider the system

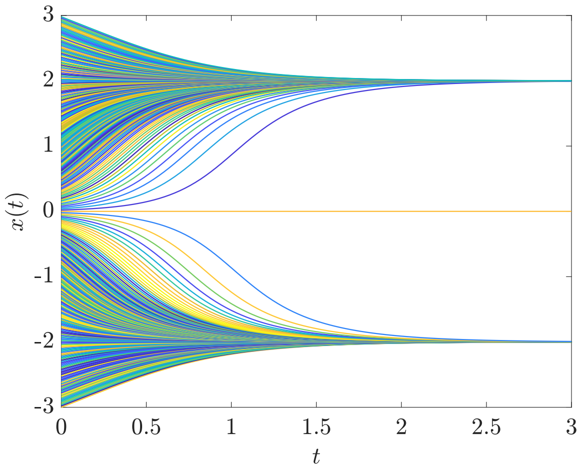
$$\dot{x} = (4x - x^3)e^{-0.3x^2}.$$

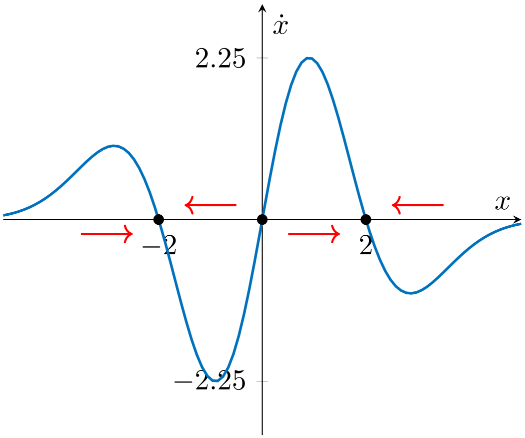
The roots of $\dot{x}$ are located at

$$x \in \{-\infty,\ -2,\ 0,\ 2,\ \infty\}$$



Phase Portraits in First Order Systems (contd.)





Contents

Equilibrium Points

Equilibrium Points

Definition 9.1 (Equilibrium Point)

A point $\bar{\mathbf{x}}$ in the state space of the dynamic system

$$\dot{\mathbf{x}} = \mathbf{f}\,(t,\, \mathbf{x})$$

is an *equilibrium state* or *equilibrium point* if whenever the state of the system starts at $\bar{\mathbf{x}}$, that is $\mathbf{x}(t_0) = \bar{\mathbf{x}}$, it will remain at $\bar{\mathbf{x}}$ for all future time. That is

$$\mathbf{f}(t, \bar{\mathbf{x}}) = \mathbf{0} \quad \forall t > t_0$$

Example - Simple Pendulum

Consider the simple pendulum system

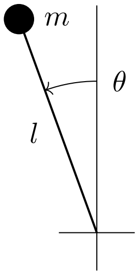
$$ml^2\ddot{\theta} = mgl\sin(\theta).$$

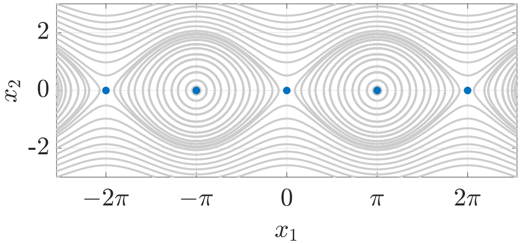
Let $m = g = l = 1$, $x_1 := \theta$ and $x_2 = \dot{\theta}$, then the state space form of the system is

$$\dot{x}_1 = x_2,$$
$$\dot{x}_2 = \sin(x_1).$$

Solving for $\dot{x}_1 = \dot{x}_2 = 0$ gives

$$\bar{x}_1 = \pm 2k\pi, \quad \forall k \in \mathbb{N}_0,$$
$$\bar{x}_2 = 0.$$





Equilibrium Points of Systems With Control Inputs

We can extend the definition of an equilibrium point of a closed loop system to a system with control input as

Definition 9.2

A point $(\bar{\mathbf{x}},\ \bar{\mathbf{u}})$ is an equilibrium point of a system starting at $\big(\mathbf{x}(t_0), \mathbf{u}(t_0)\big) = (\bar{\mathbf{x}},\ \bar{\mathbf{u}})$ if

$$\dot{\mathbf{x}} = \mathbf{f}(t,\ \bar{\mathbf{x}},\ \bar{\mathbf{u}}) = \mathbf{0}, \quad \forall t > t_0.$$

Example - Simple Pendulum With Input

Consider the simple pendulum system

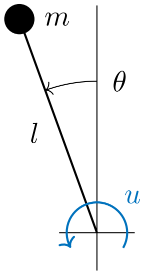
$$ml^2\ddot{\theta} = mgl\sin(\theta) + u.$$

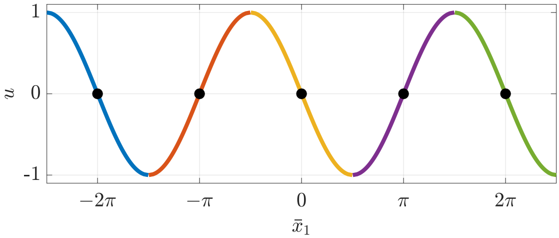
Let $m = g = l = 1$, $x_1 := \theta$ and $x_2 = \dot{\theta}$, then the state space form of the system is

$$\begin{aligned}\dot{x}_1 &= x_2,\\ \dot{x}_2 &= \sin(x_1) + u.\end{aligned}$$

Solving for $\dot{x}_1 = \dot{x}_2 = 0$ gives

$$\begin{aligned}\sin(\bar{x}_1) + u &= 0,\\ \bar{x}_2 &= 0.\end{aligned}$$





Example - Cart Pole System

Recall a state space form for the cart pole system is

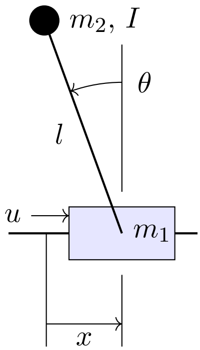
$$\dot{\mathbf{z}}_1 = \mathbf{z}_2,$$
$$\dot{\mathbf{z}}_2 = \left[\mathbf{M}(\mathbf{z}_1)\right]^{-1}\left(\mathbf{u} - \boldsymbol{\phi}(\mathbf{z}_1, \mathbf{z}_2)\right),$$

where $\mathbf{z} = (\mathbf{z}_1, \mathbf{z}_2) = (z_1, z_2, z_3, z_4)$ defined as

$$\mathbf{z}_1 = (z_1, z_2) := (x, \theta), \quad \mathbf{z}_2 = (z_3, z_4) := (\dot{x}, \dot{\theta}), \quad \mathbf{u} := (u, 0),$$

and

$$\mathbf{M} = \begin{bmatrix} m_1 + m_2 & -m_2 l \cos(z_2) \\ -m_2 l \cos(z_2) & m_2 l^2 + I \end{bmatrix}, \quad \boldsymbol{\phi} = \begin{bmatrix} m_2 l z_4^2 \sin(z_2) \\ -m_2 g l \sin(z_2) \end{bmatrix}.$$

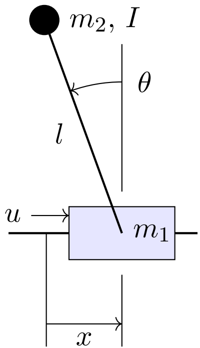


Example - Cart Pole System (contd.)

Since $\mathbf{M}$ is invertible, $\ker(\mathbf{M}) = \{\mathbf{0}\}$. Hence,

$$\dot{\mathbf{z}}_1 = 0 \implies \bar{\mathbf{z}}_2 = \mathbf{0},$$

$$\dot{\mathbf{z}}_2 = 0 \implies \mathbf{u} - \boldsymbol{\phi}(\bar{\mathbf{z}}_1, \bar{\mathbf{z}}_2) = \mathbf{0} \implies \begin{cases} u = 0, \\ m_2 g l \sin(z_2) = 0. \end{cases}$$



Since there is no constraint defining the value of z_1, we get the equilibrium points of the system as

$$\bar{\mathbf{z}} = (z_1, \pm k\pi, 0, 0),$$
$$\bar{u} = 0$$

for all $z_1 \in \mathbb{R}$ and $k \in \mathbb{N}_0$.

Example 3

Example is adopted from: Slotine JJ, Li W. Applied nonlinear control.

? What are the equilibrium points of the system

$$\dot{x} = \frac{a(t)x}{1+x^2} + b(t)$$

for arbitrary piecewise continuous functions $a(t)$ and $b(t)$?

Let x^* be the equilibrium point of the system, then

$$\frac{a(t)x^*}{1+(x^*)^2} + b(t) = 0.$$

Since we assume x^* is an equilibrium point, then

$$\frac{x^*}{1+(x^*)^2} = \xi$$

Example 3 (contd.)

Example is adopted from: Slotine JJ, Li W. Applied nonlinear control.

is constant. Hence, by the assumption, the following equality needs to hold

$$b(t) = -\xi a(t), \quad \forall t,$$

which implies $b(t)$ is an scaled form of $a(t)$ with the scaling factor of $-\xi$. However, this is contradiction to the statement that $a(t)$ and $b(t)$ are arbitrary piecewise continuous functions. Thus, the system does not have an equilibrium point.

If, however, we assume that $b(t) := 0$, then $x^* = 0$ is an the equilibrium point of the system.

Equilibrium Points of Linear Systems

Lets first consider closed-loop LTI systems

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}. \tag{9.19}$$

For any $\bar{\mathbf{x}} \in \ker(\mathbf{A})$ we have $\mathbf{A}\bar{\mathbf{x}} = \mathbf{0}$. Thus:

The null space of $\mathbf{A}$ *contains all the equilibrium points of the system.*

Accordingly, for the closed-loop LTI systems we have

- the zero vector $\bar{\mathbf{x}} = \mathbf{0}$ is an equilibrium point of the system.
- if $\bar{\mathbf{x}} \neq \mathbf{0}$ is an equilibrium point, the are all the points $\lambda\bar{\mathbf{x}}$ for any $\lambda \in \mathbb{R}$.
- if $\bar{\mathbf{x}}_1$ and $\bar{\mathbf{x}}_2$ are equilibrium points, then are all the linear combinations $c_1\bar{\mathbf{x}}_1 + c_2\bar{\mathbf{x}}_2$.

A close-loop LTI system either has a single equilibrium point at $\bar{\mathbf{x}} = \mathbf{0}$ or infinitely many "connected" equilibrium points.

Equilibrium Points of Linear Systems (contd.)

Theorem 9.1

The null space of the linear transformation $T : V \to W$ is a subspace of V.

Proof. Let $v_1, v_2 \in \ker(T) \subset V$, and $\lambda_1, \lambda_2 \in \mathbb{F}$, we have

$$T(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 T(v_1) + \lambda_2 T(v_2) = \lambda_1 \cdot 0 + \lambda_2 \cdot 0 = 0.$$

Since T is a linear transformation, it maps $0 \in V$ to $0 \in W$, hence $0 \in \ker(T)$. In fact,

$$T(0) = T(0 + 0) = T(0) + T(0) = 2T(0),$$

which leads to a contradiction if $T(0) \neq 0 \in W$. $\square$

Equilibrium Points of Linear Systems (contd.)

For an LTI system with an input we have

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}. \tag{9.20}$$

Setting $\dot{\mathbf{x}} = \mathbf{0}$ and solving for $\mathbf{x}$ and $\mathbf{u}$ gives the following equality for all the equilibrium points of the system

$$\mathbf{A}\bar{\mathbf{x}} = -\mathbf{B}\bar{\mathbf{u}}. \tag{9.21}$$

If $\mathbf{A}$ is invertible (if there exists $\mathbf{A}^{-1}$ such that $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$), then we can define $\bar{\mathbf{x}}$ as a function of $\bar{\mathbf{u}}$ as

$$\bar{\mathbf{x}} = -\mathbf{A}^{-1}\mathbf{B}\bar{\mathbf{u}}. \tag{9.22}$$

A particular set of equilibrium points of the system is

$$\left\{(\bar{\mathbf{x}}, \bar{\mathbf{u}}) \mid \bar{\mathbf{x}} \in \ker(\mathbf{A}),\ \bar{\mathbf{u}} \in \ker(\mathbf{B})\right\}. \tag{9.23}$$

Since the zero vector is the smallest null space of a linear operator $(\mathbf{0}, \mathbf{0})$ is an equilibrium point of the system.

Equilibrium Points of Linear Systems (contd.)

Now we can investigate the equilibrium points of non-autonomous closed-loop linear systems

$$\dot{\mathbf{x}} = \mathbf{A}(t)\mathbf{x}. \tag{9.24}$$

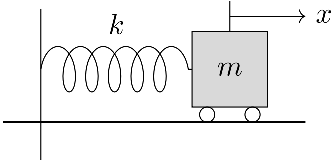
Based on the definition, $\bar{\mathbf{x}}$ is an equilibrium point if

$$\mathbf{x}(t_0) = \bar{\mathbf{x}} \implies \dot{\mathbf{x}}(t) = \mathbf{0}, \quad \forall t \geq t_0 \implies \bar{\mathbf{x}} \in \ker\big(\mathbf{A}(t)\big) \quad \forall t \geq t_0.$$

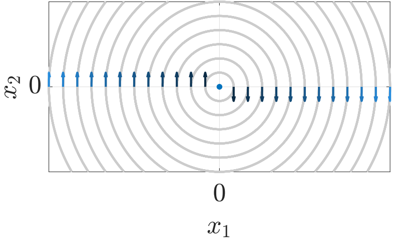
Since $\mathbf{A}(t)$ is a linear map, $\bar{\mathbf{x}} = \mathbf{0} \in \ker(\mathbf{A}(t))$ for all t. Thus, the system has at least one equilibrium point $\bar{\mathbf{x}} = \mathbf{0}$.
Similarly, we can conclude that $(\bar{\mathbf{x}}, \bar{\mathbf{u}}) = (\mathbf{0}, \mathbf{0})$ is an equilibrium point of

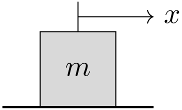
$$\dot{\mathbf{x}} = \mathbf{A}(t)\mathbf{x} + \mathbf{B}(t)\mathbf{u}. \tag{9.25}$$

Examples

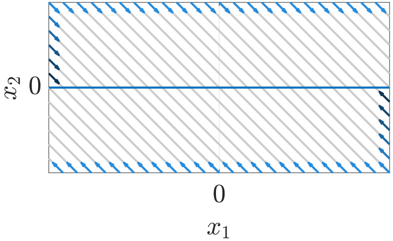


$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -k/m & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$





$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & -b/m \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$



Concept of Stability

Observation

? The system forever remain at an equilibrium point. What happens when the system is disturbed and moves a "small" distance from the equilibrium point?

- How can we define how "small" the disturbance is?
- How can we mathematically characterize system's behaviour when it is moved away from an equilibrium point?

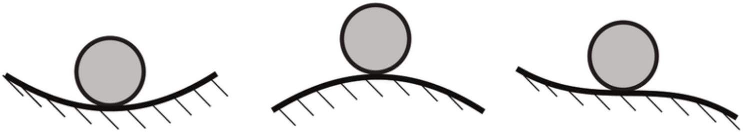
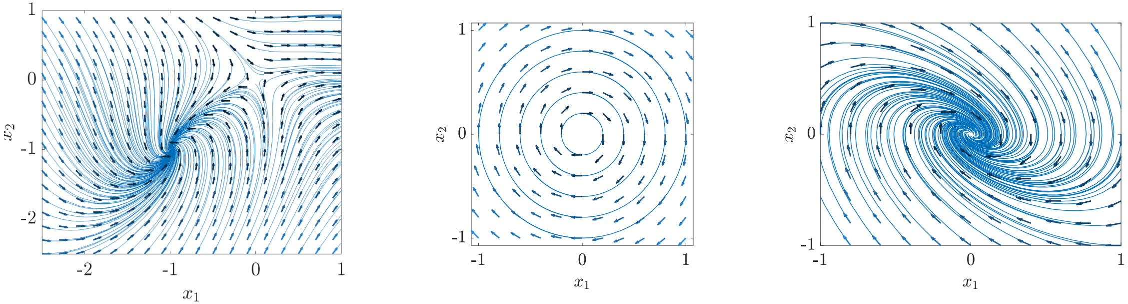


Image source: Ryaboy, Vyacheslav M. "Proper load distribution ensures stable operation." Photonics Spectra 45.12 (2011): 52-54.

Observation (contd.)



Links to some real world examples:

- Tacoma bridge collapse: `https://youtu.be/kZNjbWy6c7c?si=PGl-hEfwSiJAxyZX`
- Trailer tow: `https://youtu.be/6mW_gzdh6to?si=OOaGsC1bMQ-kjhZ7`
- RC jet crash: `https://youtu.be/Zix5HQodYWw?si=NEiNFon4yMorKaJk&t=11`

Roadmap

1. We need to properly define stability of an equilibrium point and *mathematically* define what it means to be close to an equilibrium point;
2. Then, we will proceed to formally introduce Lyapunov's stability theorem:

Let $\mathbf{x} = \mathbf{0}$ *be an equilibrium point of* $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$, *where* $\mathbf{f} : D \to \mathbb{R}^n$ *is locally Lipschitz and* $\mathbf{0} \in D \subset \mathbb{R}^n$. *Let* $V : D \to \mathbb{R}$ *be a continuously differentiable function such that* $V(\mathbf{0}) = 0$ *and such that*

$$V(\mathbf{x}) > 0, \quad \forall \mathbf{x} \in D \setminus \{\mathbf{0}\} \quad and \dot{V}(\mathbf{x}) \leq 0.$$

Then, $\mathbf{x} = \mathbf{0}$ *is stable. Moreover if*

$$\dot{V}(\mathbf{x}) = 0 \iff \mathbf{x} = \mathbf{0},$$

then $\mathbf{x} = \mathbf{0}$ *is asymptotically stable.*

Metric Spaces, Norms and Normed Vector Spaces

Metric Space

Definition 9.3 (Metric Space)

A metric space is an ordered pair $(M,\ d)$ where M is a set and $d : M \times M \to \mathbb{R}$ is a metric on M satisfying:

- $d(x, y) = 0 \iff x = y$;
- $d(x, y) > 0$ if and only if $x \neq y$;
- $d(x, y) = d(y, x)$ for all $x, y \in M$ (*Symmetry*);
- $d(x, z) \leq d(x, y) + d(y, z)$ for all $x, y, z \in M$ (*Triangle inequality*).

The triangle inequality is a natural property of notions of distance: The path from x to z could be divided into segments x to y and y to z. However, this will not make the path any shorter than the shortest path from x to z.

Examples of Metric Spaces

Proofs are left as exercises for interested audience.

- *Real line*: the set of real numbers with the metric $d(x, y) := |x - y|$ is a metric space.
- *Euclidean plane*: $\mathbb{R}^2$ equipped with The Euclidean distance

$$d\left((x_1, y_1), (x_2, y_2)\right) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$$

  is a metric space.
- *Discrete metric space*: Let M be any set and for $\mathbf{x}, \mathbf{y} \in M$ define

$$d(\mathbf{x}, \mathbf{y}) = \begin{cases} 0 & \mathbf{x} = \mathbf{y}, \\ 1 & \mathbf{x} \neq \mathbf{y}, \end{cases}$$

  then (M, d) is a metric space. In this space, all distinct points are 1 unit apart. None of them are close to each other, and none of them are very far away from each other!

More Abstract Examples of Metric Spaces

Proofs are left as exercises for interested audience.

- *Sequence space* ℓ^∞: Let X be the set of all bounded sequences of complex numbers

$$X = \left\{ \mathbf{x} = (x_1,\ x_2\ \ldots) \mid |x_i| \leq c_{\mathbf{x}}, \quad \forall i \in \mathbb{N} \right\},$$

  where $c_{\mathbf{x}}$ is a real number which may depend on $\mathbf{x}$, but not i. We can define $d : X \times X \to \mathbb{R}$ as

$$d(\mathbf{x}, \mathbf{y}) := \sup_{i \in \mathbb{N}} |x_i - y_i|.$$

- *Function space* $C[a, b]$: As a set X we take the set of all real-valued functions that are defined and continuous on a given closed interval $I = [a,\ b]$. We can chose the metric defined as

$$d(f_1, f_2) := \max_{x \in I} |f_1(x) - f_2(x)|.$$

Definition of a Norm

Given a vector space V over a subfield* $\mathbb{F}$ of the complex numbers $\mathbb{C}$, a norm on V is is a real-valued function $\|\cdot\| : V \to \mathbb{R}$ such that

- *Sub-additivity* or *Triangle inequality*: $\|\mathbf{v}_1 + \mathbf{v}_2\| \leq \|\mathbf{v}_1\| + \|\mathbf{v}_2\|$ for all $\mathbf{v}_1,\ \mathbf{v}_2 \in V$;
- *Absolute homogeneity*: $\|\lambda \mathbf{v}\| = |\lambda| \|\mathbf{v}\|$ for all $\mathbf{v} \in V$ and $\lambda \in \mathbb{F}$;
- *Positive definiteness*: For every $\mathbf{v} \in V$, $\|\mathbf{v}\| = 0$ if and only if $\mathbf{v} = \mathbf{0}$;
- *Non-negativity*: $\|\mathbf{v}\| \geq 0$ for all $\mathbf{v} \in V$.

*A subfield K of a field L is a subset $K \subseteq L$ that contains 1, and is closed under addition, subtraction, multiplication, and taking the inverse of a nonzero element of K. In other words, $K \subseteq L$ is a field with respect to the field operations inherited from L.

p-norm

Definition 9.4 (p-norm)

Let $p \geq 1$ in $\mathbb{R}$. The p-norm (also called ℓ^p-norm) of vector $\mathbf{v} = (v_1, \ldots, v_n)$ is

$$\|\mathbf{v}\|_p := \left(\sum_{i=1}^{n} |v_i|^p \right)^{1/p}.$$

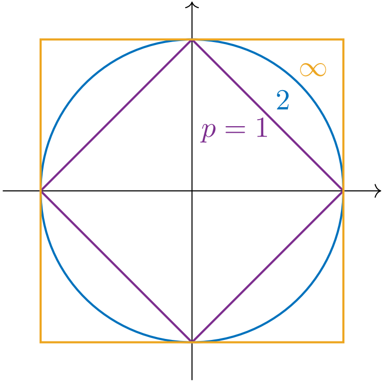
Three special p-norms are

- For $p = 1$ (ℓ^1) we get the Taxicab norm or Manhattan norm $\|\mathbf{v}\|_1 := \sum_{i=1}^{n} |v_i|$.
- For $p = 2$ (ℓ^2) we get the Euclidean norm $\|\mathbf{v}\|_2 := \left(\sum_{i=1}^{n} |x_i|^2\right)^{1/2}$.
- When $p \to \infty$ (ℓ^∞) we get the infinity norm or maximum norm $\|\mathbf{v}\|_\infty := \max_i |v_i|$.

A Unit Circle In p-norms

The following image illustrates the unit circles in $\mathbb{R}^2$ based on ℓ^1, ℓ^2 and ℓ^∞ norms:

$$\|\mathbf{v}\|_p = 1$$



Normed Vector Spaces

Definition 9.5 (Normed Vector Space)

A normed vector space is a vector space equipped with a norm.

Any normed vector space can be equipped with a metric

$$d(x, y) = \|\mathbf{x} - \mathbf{y}\|.$$

The metric d is said to be induced by the norm $\|\cdot\|$.

Conversely, if a metric $d : V \times V \to \mathbb{R}$ is translation invariant and absolutely homogeneous:

$$d(\mathbf{x}, \mathbf{y}) = d(\mathbf{x} + \mathbf{a}, \mathbf{y} + \mathbf{a}), \quad \mathbf{x}, \mathbf{y}, \mathbf{a} \in V,$$
$$d(\alpha\mathbf{x}, \alpha\mathbf{y}) = |\alpha| d(\mathbf{x}, \mathbf{y}), \quad \mathbf{x}, \mathbf{y}, \in V, \quad \alpha \in \mathbb{R},$$

then we can define a norm induced by metric as $\|\mathbf{x}\| = d(\mathbf{x}, \mathbf{0})$.

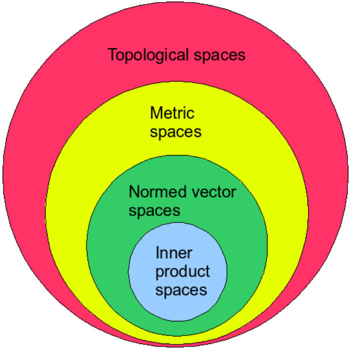


Image source: Wikipedia

Equivalent Norms

Definition 9.6 (Equivalent norms)

A norm $\|\cdot\|$ on a normed vector space X is said to be equivalent to a norm $\|\cdot\|_o$ on X if there are $a, b > 0$ such that for all $x \in X$ we have

$$a\|x\|_o \leq \|x\| \leq b\|x\|_o.$$

Lemma 9.2 (Linear combinations)

Let $\{x_1, \dots, x_n\}$ be a linearly independent set of vectors in a normed space X (of any dimension). Then there is $c > 0$ such that for every choice of scalars $\alpha_1, \dots, \alpha_n$ we have

$$\|\alpha_1 x_1 + \cdots + \alpha_n x_n\| \geq c(|\alpha_1| + \cdots + |\alpha_n|).$$

Proof: Kreyszig, Erwin. Introductory functional analysis with applications.

Equivalent Norms in Finite Dimensions

Theorem 9.3

On a finite dimensional vector space X, any two norms $\|\cdot\|$ and $\|\cdot\|_o$ are equivalent.

Proof. Let dim $X = n$ and $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be any basis for X. Then every $\mathbf{x} \in X$ has a unique representation

$$\mathbf{x} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n.$$

By Lemma 9.2 there exists $\lambda > 0$ such that

$$\|\mathbf{x}\| \geq \lambda \left(|c_1| + \cdots + |c_n|\right).$$

On the other hand, from the triangle inequality we have

$$\|\mathbf{x}\|_o \leq \sum_{i=1}^{n} |c_i| \|v_i\|_o \leq k \sum_{i=1}^{n} |c_i|,$$

Together, $a\|x\|_o \leq \|x\|$ where $a = c/k > 0$. The inequality $\|x\| \leq b\|x\|_o$ is obtained by an interchange of the roles of $\|\cdot\|$ and $\|\cdot\|_o$ in the above argument.

Continuous Mapping

Definition 9.7 (Continuous mapping)

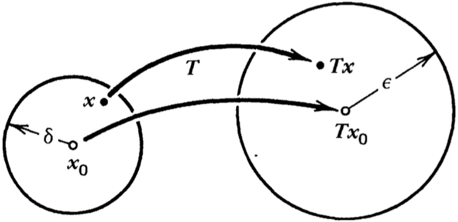
Let $X = (X, d_X)$ and $Y = (Y, d_Y)$ be metric spaces. A mapping $T : X \to Y$ is said to be *continuous* at a point $x_0 \in X$ if for every $\epsilon > 0$ there is a $\delta > 0$ such that

$$d_X(x, x_0) < \delta \implies d_Y\left(T(x), T(x_0)\right) < \epsilon.$$

T is said to be continuous if it is continuous at every point of X.

In a normed vector space, we can say $T : X \to Y$ is continuous at $\mathbf{x}_0$ if for any $\epsilon > 0$ there is a $\delta > 0$ such that

$$\|\mathbf{x} - \mathbf{x}_0\| < \delta \implies \|T(\mathbf{x}) - T(\mathbf{x}_0)\| < \epsilon.$$



Example

Lipschitz Continuous Mapping

Definition 9.8 (Lipschitz Continuous)

Given two metric spaces $X = (X, d_X)$ and $Y = (Y, d_Y)$, a mapping (function) $T : X \to Y$ is called *Lipschitz continuous* if there exists $L \geq 0$ such that, for all x_1 and x_2 in X,

$$d_Y\left(T(x_1), T(x_2)\right) \leq L\, d_X(x_1, x_2).$$

Any such L is referred to as a *Lipschitz constant* for the mapping T.

See that the above condition in normed vector space is

$$\|T(x_1) - T(x_2)\| \leq L\, \|x_1 - x_2\|.$$

Type of Functions in Metric Spaces

The following strict inclusions hold for functions over a compact interval of $\mathbb{R}$:

$$\text{Continuously differentiable} \subset \text{Lipschitz continuous} \subset \alpha\text{-Hölder continuous } (0 < \alpha \leq 1)$$
$$\text{Lipschitz continuous} \subset \text{absolutely continuous} \subset \text{uniformly continuous.}$$

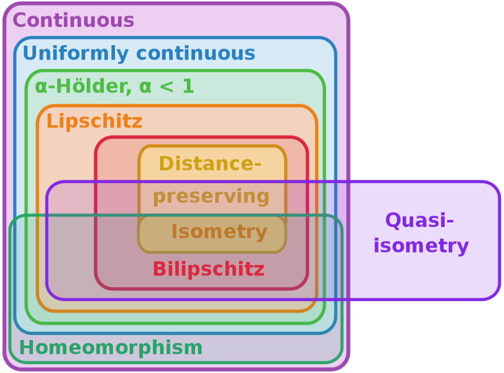


Image source: Wikipedia

A Short Review on Differential Equations

Existence and Uniqueness

For the proofs of the following theorems on existence and uniqueness of the solution of an initial-value problem see: Khalil, H.K., 2002. Nonlinear Systems, third edition (2002).

Theorem 9.4 (Local Existence and Uniqueness)

Let $\mathbf{f}(t, \mathbf{x})$ *be piecewise continuous in* t *and satisfy the Lipschitz condition*

$$\|\mathbf{f}(t, \mathbf{x}) - \mathbf{f}(t, \mathbf{y})\| \leq L\|\mathbf{x} - \mathbf{y}\|.$$

for all $\mathbf{x}$ *and* $\mathbf{y}$ *in* $B(\mathbf{x}_0; r) := \left\{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x} - \mathbf{x}_0\| \leq r\right\}$ *and for all* $t \in [t_0, t_1]$. *Then there exists* $\delta > 0$ *such that the initial value problem*

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{f}(t, \mathbf{x}), \\ \mathbf{x}(t_0) &= \mathbf{x}_0, \end{aligned}$$

has a unique solution over $[t_0, t_0 + \delta)$.

Existence and Uniqueness (contd.)

Theorem 9.5 (Global Existence and Uniqueness)

If $\mathbf{f}(t, \mathbf{x})$ *is piecewise continuous in* t *and satisfies*

$$\|\mathbf{f}(t, \mathbf{x}) - \mathbf{f}(t, \mathbf{y})\| \leq L\|\mathbf{x} - \mathbf{y}\|, \quad \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n, \quad \forall t \in [t_0, t_1],$$

Then, the initial-value problem

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{f}(t, \mathbf{x}), \\ \mathbf{x}(t_0) &= \mathbf{x}_0, \end{aligned}$$

has a unique solution over $[t_0, t_1]$*.*

Existence and Uniqueness (contd.)

Theorem 9.6 (Existence and Uniqueness in Compact Set for Unbounded Time)

Let $\mathbf{f}(t, \mathbf{x})$ *be piecewise continuous in* t *for all* $t \geq t_0$ *and locally Lipschitz in* $\mathbf{x}$ *for all* $\mathbf{x} \in D \subset \mathbb{R}^n$*. If exists a compact set* $W \subset D$ *such that* $\mathbf{x}_0 \in W$ *and every solution of*

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{f}(t, \mathbf{x}), \\ \mathbf{x}(t_0) &= \mathbf{x}_0, \end{aligned}$$

lies entirely in W*. Then, the initial value problem has a unique solution for all* $t \geq t_0$*.*

Stability of equilibrium points

Stability of an Equilibrium Point of an Autonomous System

Consider the autonomous system

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}), \tag{9.26}$$

where $\mathbf{f} : D \to \mathbb{R}^n$ is locally Lipschitz and $D \subset \mathbb{R}^n$. Suppose $\bar{\mathbf{x}} \in D$ is an equilibrium point of (9.26), that is

$$\mathbf{f}(\bar{\mathbf{x}}) = \mathbf{0}.$$

Our goal is to characterize and study the stability of $\bar{\mathbf{x}}$.

Stability of an Equilibrium Point of an Autonomous System (contd.)

Without loss of generality, we will assume $\bar{\mathbf{x}} = \mathbf{0}$, that is

$$\mathbf{f}(\mathbf{0}) = \mathbf{0}.$$

In fact, assume $\bar{\mathbf{x}} \neq \mathbf{0}$. Introducing the change of variable $\boldsymbol{\xi} = \mathbf{x} - \bar{\mathbf{x}}$ gives

$$\dot{\boldsymbol{\xi}} = \dot{\mathbf{x}} - \dot{\bar{\mathbf{x}}} = \mathbf{f}(\mathbf{x}) - \mathbf{f}(\bar{\mathbf{x}}) = \mathbf{f}(\boldsymbol{\xi} + \bar{\mathbf{x}}) =: \mathbf{g}(\boldsymbol{\xi}). \tag{9.27}$$

We can confirm that $\mathbf{g}(\mathbf{0}) = \mathbf{0}$. In particular,

$$\mathbf{g}(\mathbf{0}) = \mathbf{f}(\mathbf{0} + \bar{\mathbf{x}}) = \mathbf{f}(\bar{\mathbf{x}}) = \mathbf{0}. \tag{9.28}$$

Thus, $\bar{\boldsymbol{\xi}} = \mathbf{0}$ is an equilibrium point of

$$\dot{\boldsymbol{\xi}} = \mathbf{g}(\boldsymbol{\xi}). \tag{9.29}$$

Stability of an Equilibrium Point of an Autonomous System (contd.)

Definition 9.9

The equilibrium point $\bar{\mathbf{x}} = \mathbf{0}$ of autonomous system $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$, where $\mathbf{f} : D \to \mathbb{R}^n$ is locally Lipschitz and $D \subset \mathbb{R}^n$, is:

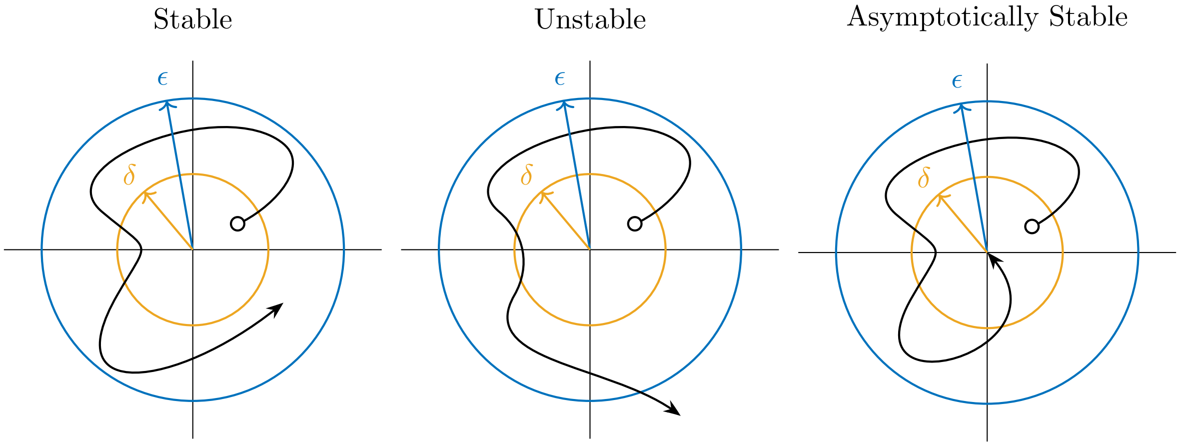
- *stable* if, for any $\epsilon > 0$, there exists $\delta = \delta(\epsilon) > 0$ such that

$$\|\mathbf{x}(0)\| < \delta \implies \|\mathbf{x}(t)\| < \epsilon, \quad \forall t \geq 0.$$

- *unstable* if it is not stable.
- *asymptotically stable* if it is stable and δ can be chosen such that

$$\|\mathbf{x}(0)\| < \delta \implies \lim_{t \to \infty} \|\mathbf{x}(t)\| = 0.$$

Stability of an Equilibrium Point of an Autonomous System (contd.)



Example - Mass Spring System

$$\ddot{x} + x = 0, \quad x(0) = x_0, \quad \dot{x}(0) = v_0,$$

implies

$$\begin{aligned} x(t) &= x_1(t) = x_0 \cos(t) + v_0 \sin(t), \\ \dot{x}(t) &= x_2(t) = -x_0 \sin(t) + v_0 \cos(t), \end{aligned}$$

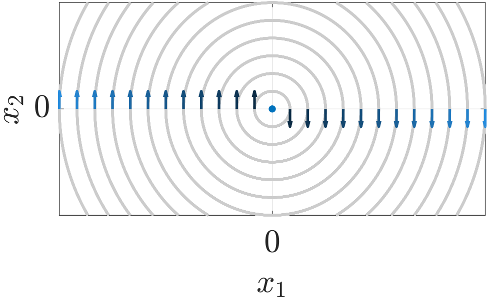
that leads to

$$\|\mathbf{x}(t)\|_2^2 = x_1^2(t) + x_2^2(t) = x_0^2 + v_0^2.$$

Given any $\epsilon > 0$, we can choose $\delta = \epsilon$. Hence,

$$\|\mathbf{x}(0)\| < \delta \implies \|\mathbf{x}(t)\| < \delta = \epsilon.$$

$$\begin{aligned} \dot{x}_1 &= x_2, \\ \dot{x}_2 &= -x_1. \end{aligned}$$



Example - Mass Spring Damper System

$$\ddot{x} + 2\dot{x} + x = 0, \quad x(0) = x_0, \quad \dot{x}(0) = v_0,$$

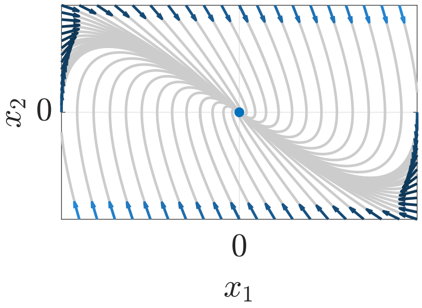
implies

$$x(t) = x_1(t) = e^{-t}\left((1+t)x_0 + tv_0\right),$$
$$\dot{x}(t) = x_2(t) = -e^{-t}\left(tx_0 + (t-1)v_0\right),$$

that leads to

$$\|\mathbf{x}(t)\|_2^2 = e^{-2\,t}\left(at^2 + bt + c\right),$$
$$a = 2\left(x_0 + v_0\right)^2,$$
$$b = 2\left(x_0^2 - v_0^2\right),$$
$$c = x_0^2 + v_0^2.$$

$$\dot{x}_1 = x_2,$$
$$\dot{x}_2 = -2x_2 - x_1.$$



Example - Mass Spring Damper System (contd.)

$b^2 - 4ac = -4(x_0 + v_0)^4 < 0 \implies at^2 + bt + c$ does not have a root in $\mathbb{R}$. Since $c = x_0^2 + v_0^2 > 0$, as expected, $at^2 + bt + c > 0$ for all $t \in \mathbb{R}$. Moreover,

$$\frac{\mathrm{d}}{\mathrm{d}t} \|\mathbf{x}(t)\|_2^2 = -e^{-2t} \left((x_0 + v_0)t - v_0\right)^2 \leq 0, \quad \forall t,$$

meaning $\|\mathbf{x}(t)\|_2^2$ is non-increasing in t. Hence,

$$0 \leq \|\mathbf{x}(t)\|_2^2 \leq \|\mathbf{x}(0)\|_2^2, \quad \forall t \geq 0.$$

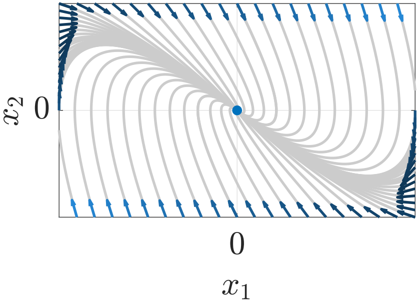
Let $\delta = \epsilon$, since $\frac{\mathrm{d}}{\mathrm{d}t}\|\mathbf{x}(t)\|_2^2 = 0$ only at $t = v_0/(x_0 + v_0)$, we have

$$\|\mathbf{x}(0)\| < \delta \implies \|\mathbf{x}(t)\| < \delta = \epsilon,$$

$$\lim_{t\to\infty} \|\mathbf{x}(t)\| = 0.$$

$$\dot{x}_1 = x_2,$$

$$\dot{x}_2 = -2x_2 - x_1.$$



Example - An Unstable System

$$\ddot{x} - 2\dot{x} + x = 0, \quad x(0) = x_0, \quad \dot{x}(0) = v_0,$$

implies

$$x(t) = x_1(t) = e^t \left(x_0 + \alpha t\right),$$
$$\dot{x}(t) = x_2(t) = e^t \left(v_0 + \alpha t\right),$$

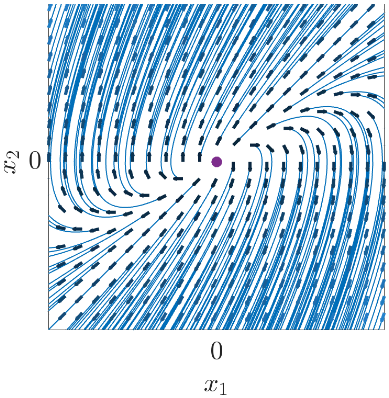
where $\alpha = (v_0 - x_0)$. Accordingly

$$\|\mathbf{x}(t)\|_2^2 = e^{2t} \left((x_0 + \alpha t)^2 + (v_0 + \alpha t)^2 \right).$$

Note that for any $\mathbf{x}(0) \neq \mathbf{0}$,

$$\lim_{t \to \infty} \|\mathbf{x}\|_2 = \infty.$$

$$\dot{x}_1 = x_2,$$
$$\dot{x}_2 = -x_1 + 2x_2.$$



Exponential Stability

Definition 9.10 (Exponential stability of an equilibrium point)

The equilibrium point $\bar{\mathbf{x}} = \mathbf{0}$ of autonomous system $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$, where $\mathbf{f} : D \to \mathbb{R}^n$ is locally Lipschitz and $D \subset \mathbb{R}^n$, is *exponentially stable* if there exists $\alpha, \ \lambda > 0$ such that

$$\|\mathbf{x}(t)\| \leq \alpha \|\mathbf{x}(0)\| e^{-\lambda t},$$

for all $t > t_0$ and $\mathbf{x}(0) \in B_r \subset D$.

In other words, an exponentially stable system converges to origin faster than an exponetial function.

Lyapunov's Stability Theorem

Theorem 9.7 (Lyapunov's Stability Theorem)

Let $\mathbf{x} = \mathbf{0}$ *be an equilibrium point of* $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$, *where* $\mathbf{f} : D \to \mathbb{R}^n$ *is locally Lipschitz and* $\mathbf{0} \in D \subset \mathbb{R}^n$. *Let* $V : D \to \mathbb{R}$ *be a continuously differentiable function such that*

$$\begin{aligned} &V(\mathbf{x}) \geq 0, \quad \textit{and} \quad V(\mathbf{x}) = 0 \iff \mathbf{x} = \mathbf{0}, \\ &\dot{V}(\mathbf{x}) \leq 0. \end{aligned}$$

Then, $\mathbf{x} = \mathbf{0}$ *is stable. Moreover if*

$$\dot{V}(\mathbf{x}) = 0 \iff \mathbf{x} = \mathbf{0},$$

then $\mathbf{x} = \mathbf{0}$ *is asymptotically stable.*

Lyapunov's Stability Theorem (contd.)

Proof. Given $\epsilon > 0$, choose $r \in (0, \epsilon]$ such that

$$\bar{B}_r := \left\{ \mathbf{x} \in D \;\middle|\; \|\mathbf{x}\| \leq r \right\} \subset D.$$

Let

$$\alpha := \min_{\|\mathbf{x}\|=r} V(\mathbf{x}),$$

Since $V(\mathbf{x}) > 0$ for all $\mathbf{x} \in D \setminus \{\mathbf{0}\}$, then $\alpha > 0$. Now choose $\beta \in (0, \alpha)$ and define

$$\Omega_\beta := \left\{ \mathbf{x} \in \bar{B}_r \;\middle|\; V(\mathbf{x}) \leq \beta \right\} \implies \Omega_\beta \subset \mathrm{int}(\bar{B}_r) \quad \text{(why?)}$$

Moreover, if $\mathbf{x}(0) \in \Omega_\beta$, then $\mathbf{x}(t) \in \Omega_\beta$ for all t. In fact

$$\dot{V}\left(\mathbf{x}(t)\right) \leq 0 \implies V\left(\mathbf{x}(t)\right) \leq \left(\mathbf{x}(0)\right) \leq \beta, \quad \forall t.$$

Lyapunov's Stability Theorem (contd.)

Since Ω_β is compact (why?), based on Theorem 9.6 , $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$, with $\mathbf{x}(0) \in \Omega_\beta$ has a unique solution $\forall t \geq 0$. Since $V(\mathbf{x})$ is continuous, and $V(\mathbf{0}) = 0$, there exists $\delta > 0$ such that

$$\|\mathbf{x}\| < \delta \implies V(\mathbf{x}) < \beta \implies B_\delta \subset \Omega_\beta \subset B_r.$$

Hence, for all t, we have

$$\mathbf{x}(0) \in B_\delta \implies \mathbf{x}(0) \in \Omega_\beta \implies \mathbf{x}(t) \in \Omega_\beta \implies \mathbf{x}(t) \in B_r.$$

Equivalently, in the terms of the definition,

$$\|\mathbf{x}(0)\| < \delta \implies \|\mathbf{x}(t)\| < r \leq \epsilon, \quad \forall t.$$

Example - Mass Spring System

Assume the system

$$\dot{x}_1 = x_2,$$
$$\dot{x}_2 = -x_1.$$

Let

$$V = \frac{1}{2}(x_1^2 + x_2^2),$$

We can confirm that $V(\mathbf{x}) \geq 0$ and $V(\mathbf{x}) = 0 \iff \mathbf{x} = \mathbf{0}$. Taking the derivative of V with respect to time gives

$$\dot{V} = x_1\dot{x}_1 + x_2\dot{x}_2 = x_1x_2 - x_2x_1 = 0.$$

Since $\dot{V}(\mathbf{x}) \leq 0$ (in fact $\dot{V}(\mathbf{x}) \leq 0$ for all $\mathbf{x}$), we can conclude that $\mathbf{x} = \mathbf{0}$ is an stable equilibrium point.

Example - Asymptotically Stable System

Consider the system

$$\dot{x}_1 = x_2 - x_1\left(x_1^2 + x_2^2\right),$$
$$\dot{x}_2 = -x_1 - x_2\left(x_1^2 + x_2^2\right).$$

We can confirm $\mathbf{x} = \mathbf{0}$ is an equilibrium point of the system. Similar to the previous example, let

$$V = \frac{1}{2}(x_1^2 + x_2^2).$$

Taking the derivative of V with respect to time gives

$$\dot{V} = x_1\dot{x}_1 + x_2\dot{x}_2 = x_1x_2 - x_1^2\left(x_1^2 + x_2^2\right) - x_1x_2 - x_2^2\left(x_1^2 + x_2^2\right) = -\left(x_1^2 + x_2^2\right)^2.$$

Since $\dot{V}(\mathbf{x}) \leq 0$ and $\dot{V}(\mathbf{x}) = 0 \iff \mathbf{x} = \mathbf{0}$, we can conclude that $\mathbf{x} = \mathbf{0}$ is an asymptotically stable equilibrium point of the system.

Sign Definite Functions (In the Context of Control Theory)

A function $f : D \subset \mathbb{R}^n \to \mathbb{R}$, with $f(\mathbf{0}) = 0$ is

- *Positive definite* if $f(\mathbf{x}) > 0$ for all $\mathbf{x} \in D \setminus \{\mathbf{0}\}$.
- *Positive semidefinite* if $f(\mathbf{x}) \geq 0$ for all $\mathbf{x} \in D$.
- *Negative definite* if $f(\mathbf{x}) < 0$ for all $\mathbf{x} \in D \setminus \{\mathbf{0}\}$.
- *Negative semidefinite* if $f(\mathbf{x}) \leq 0$ for all $\mathbf{x} \in D$.

A class of functions for which sign definiteness can be easily checked is the class of quadratic functions of the form

$$f(\mathbf{x}) := \mathbf{x}^T \mathbf{P} \mathbf{x}, \tag{9.30}$$

where sign definiteness of $\mathbf{P}$ defines the sign definiteness of f.

Rephrased Form of Lyapunov's Stability Theorem

Using the sign definite function definition, we can restate Theorem 9.7 as

Theorem 9.8

The origin of the autonomous system $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ *is stable if there exists a continuously differentiable positive definite function* $V(\mathbf{x})$ *such that* $\dot{V}(\mathbf{x})$ *is negative semidefinite. If* $\dot{V}(\mathbf{x})$ *is negative definite, then the origin is asymptotically stable.*

⚠ The theorems in Lyapunov analysis are all *sufficiency* conditions. If for a specific choice of V, the conditions on $\dot{V}$ are not met, we cannot draw any conclusion on the stability or instability of the system. The only immediate conclusion is that we need to try *another* function.

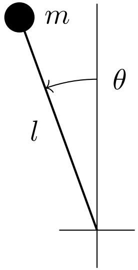
Example - Simple Pendulum

Consider the pendulum system with viscous damping

$$mgl^2\ddot{\theta} = mgl\sin(\theta) - b\dot{\theta}.$$

Let $m = g = l = b = 1$, $x_1 = \theta$, and $x_2 = \dot{\theta}$. Accordingly,

$$\dot{x}_1 = x_2,$$
$$\dot{x}_2 = \sin(x_1) - x_2.$$



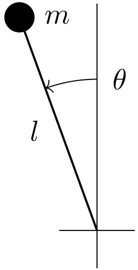
Let us examine the stability of the equilibrium point $\bar{\mathbf{x}} = (\pi, 0)$. Since $\bar{\mathbf{x}} \neq \mathbf{0}$, we can introduce a change of variable

$$\xi_1 = x_1 - \pi,$$
$$\xi_2 = x_2,$$

which leads to

Example - Simple Pendulum (contd.)

$$\dot{\xi}_1 = \frac{\mathrm{d}}{\mathrm{d}t}(x_1 - \pi) = \dot{x}_1 = x_2 = \xi_2,$$
$$\dot{\xi}_2 = \dot{x}_2 = \sin(x_1) - x_2 = \sin(\xi_1 + \pi) - \xi_2 = -\sin(\xi_1) - \xi_2.$$



Now we can confirm that $(\xi_1, \xi_2) = (0, 0)$ is an equilibrium system of

$$\dot{\xi}_1 = \xi_2,$$
$$\dot{\xi}_2 = -\sin(\xi_1) - \xi_2,$$

which coincides with $(x_1, x_2) = (\pi, 0)$ of the original system.

Example - Simple Pendulum (contd.)

To analyze using Lyapunov's theorem, let

$$V(\xi_1, \xi_2) := \big(1 - \cos(\xi_1)\big) + \frac{1}{2}\xi_2^2,$$

which is the total energy of the system. See that $V \geq 0$ and $V = 0 \iff \xi_1 = \xi_2 = 0$.

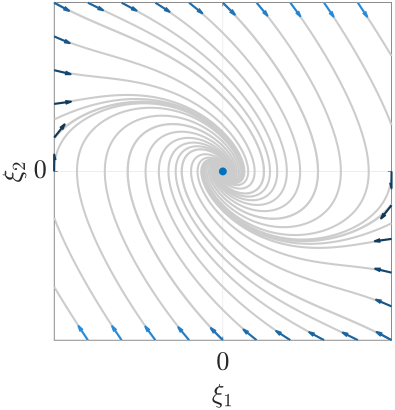
$$\dot{V} = \sin(\xi_1)\dot{\xi}_1 + \xi_2\dot{\xi}_2 = \sin(\xi_1)\dot{\xi}_2 - \xi_2^2 - \sin(\xi_1)\dot{\xi}_2 = -\xi_2^2.$$

Since $\dot{V}$ is a function of ξ_1 and ξ_2, $\dot{V} = -\xi_2^2$ means

$$\dot{V}(\xi_1, \xi_2) = 0 \cdot \xi_1 - \xi_2^2,$$

hence $\dot{V}$ is negative semidefinite. Using current V, we can only conclude the origin is stable (we have nothing on its asymptotic stability)

From the phase portrait, we see $(\xi_1, \xi_2) = (0, 0)$ is an asymptotically stable equilibrium point.



Example - Simple Pendulum (contd.)

As another attempt, let us try

$$W = \frac{1}{2}\xi_2^2 + \frac{1}{2}\left(\xi_1 + \xi_2\right)^2 + 2\left(1 - \cos(\xi_1)\right).$$

See that $W(\xi_1, \xi_2) \geq 0$ and $W = 0 \iff \xi_1 = \xi_2 = 0$. For the derivative of W we have

$$\dot{W} = -\xi_2^2 - \xi_1 \sin(\xi_1)$$

See that for ξ_1 in $(-\pi, 0) \cup (0, \pi)$, we have $\xi_1 \sin(\xi_1) > 0$, which implies $\dot{W}$ is a negative definite function and the origin is asymptotically stable.

Note that $W(\xi_1, \xi_2)$ does not have any obvious physical meaning. Nonetheless, it allows us to prove asymptotic stability of $(\xi_1, \xi_2) = (0, 0)$.

Review of Vector and Matrix Calculus

Vector and Matrix Calculus

Let $\mathbf{A}(t) = [a_{ij}(t)] \in \mathbb{F}^{m \times n}$ where each entry is a function of an independent parameter t. Then

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\mathbf{A}(t) = \left[\frac{\mathrm{d}}{\mathrm{d}t}\,a_{ij}(t)\right], \quad \text{and} \quad \int \mathbf{A}(t)\,dt = \left[\int a_{ij}(t)\ dt\right]. \tag{9.31}$$

Let $\mathbf{A}^{-1}(t)$ be the inverse of $\mathbf{A}(t)$ for every t. In order to find the derivative of $\mathbf{A}^{-1}$ with respect to t we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\mathbf{A}\,\mathbf{A}^{-1}\right) = \left(\frac{\mathrm{d}}{\mathrm{d}t}\,\mathbf{A}\right)\mathbf{A}^{-1} + \mathbf{A}\left(\frac{\mathrm{d}}{\mathrm{d}t}\,\mathbf{A}^{-1}\right). \tag{9.32}$$

Since $\mathbf{A}\,\mathbf{A}^{-1} = \mathbf{I}$, we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\mathbf{A}\,\mathbf{A}^{-1}\right) = \frac{\mathrm{d}}{\mathrm{d}t}\,\mathbf{I} = \mathbf{0} \implies \frac{\mathrm{d}}{\mathrm{d}t}\,\mathbf{A}^{-1} = -\mathbf{A}^{-1}\left(\frac{\mathrm{d}}{\mathrm{d}t}\,\mathbf{A}\right)\mathbf{A}^{-1}. \tag{9.33}$$

Vector and Matrix Calculus (contd.)

Let $\mathbf{x} \in \mathbb{F}^{n\times 1}$ be a column vector and $f : \mathbb{F}^n \to \mathbb{F}$ be a functional. We will use the following convention for the gradient of f with respect to $\mathbf{x}$

$$\frac{\partial f}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial f}{\partial x_1} & \cdots & \frac{\partial f}{\partial x_n} \end{bmatrix}. \tag{9.34}$$

That is:

- if $\mathbf{x}$ being a column vector, we assign the gradient as a row vector,
- if $\mathbf{y}$ is a row vector, then the gradient is a column vector.

Vector and Matrix Calculus (contd.)

Accordingly for $f : \mathbb{F}^n \to \mathbb{F}$, $\mathbf{x} \in \mathbb{F}^{n\times 1}$ and $\mathbf{y} \in \mathbb{F}^{1\times n}$ we have

$$\frac{\mathrm{d}}{\mathrm{d}t} f\left(\mathbf{x}(t)\right) = \frac{\partial f}{\partial \mathbf{x}}\,\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t}\,, \tag{9.35}$$

$$\frac{\mathrm{d}}{\mathrm{d}t} f\left(\mathbf{y}(t)\right) = \frac{\mathrm{d}\mathbf{y}}{\mathrm{d}t}\,\frac{\partial f}{\partial \mathbf{y}}\,. \tag{9.36}$$

Since the inner product of two vector in $\mathbb{F}^n$ is a mapping $\mathbb{F}^n \times \mathbb{F}^n \to \mathbb{F}$, we have

$$\frac{\partial}{\partial \mathbf{x}}(\mathbf{x}^T\mathbf{y}) = \begin{bmatrix} \frac{\partial(\mathbf{x}^T\mathbf{y})}{\partial x_1} & \cdots & \frac{\partial(\mathbf{x}^T\mathbf{y})}{\partial x_n} \end{bmatrix} = \begin{bmatrix} \frac{\partial}{\partial x_1}\left(\sum_{i=1}^{n} x_i y_i\right) & \frac{\partial}{\partial x_n}\left(\sum_{i=1}^{n} x_i\, y_i\right) \end{bmatrix} = \mathbf{y}^T. \tag{9.37}$$

Similarly, we have

$$\frac{\partial}{\partial \mathbf{y}}(\mathbf{x}^T\,\mathbf{y}) = \mathbf{x}^T. \tag{9.38}$$

Vector and Matrix Calculus (contd.)

For matrix $\mathbf{A} = \left[a_{ij}\right] \in \mathbb{F}^{m\times n}$ and functional $f : \mathbb{F}^{m\times n} \to \mathbb{F}$ we have

$$\frac{\partial f}{\partial \mathbf{A}} = \begin{bmatrix} \frac{\partial f}{\partial a_{11}} & \cdots & \frac{\partial f}{\partial a_{1n}} \\ \vdots & & \vdots \\ \frac{\partial f}{\partial a_{m1}} & \cdots & \frac{\partial f}{\partial a_{mn}} \end{bmatrix}. \tag{9.39}$$

Using the above results, we can define the derivative of the quadratic form $\mathbf{x}^T\,\mathbf{A}\,\mathbf{x}$, for $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{A} \in \mathbb{R}^{n\times n}$. To do so let $\mathbf{y}(\mathbf{x}) := \mathbf{A}\,\mathbf{x}$ then

$$\frac{\partial}{\partial \mathbf{x}}\left(\mathbf{x}^T\mathbf{A}\mathbf{x}\right) = \frac{\partial}{\partial \mathbf{x}}\left(\mathbf{x}^T\,\mathbf{y}\right) = \frac{\partial\left(\mathbf{x}^T\,\mathbf{y}(\mathbf{x})\right)}{\partial \mathbf{x}} + \mathbf{x}^T\,\frac{\partial \mathbf{y}}{\partial \mathbf{x}} = \mathbf{y}^T + \mathbf{x}^T\mathbf{A} = \mathbf{x}^T\left(\mathbf{A}^T + \mathbf{A}\right). \tag{9.40}$$

In particular, if $\mathbf{A}$ is symmetric ($\mathbf{A} = \mathbf{A}^T$) then

$$\frac{\partial}{\partial \mathbf{x}}\left(\mathbf{x}^T\,A\,\mathbf{x}\right) = \mathbf{x}^T\left(\mathbf{A}^T + \mathbf{A}\right) = 2\mathbf{x}^T\mathbf{A}. \tag{9.41}$$

Contents

Roadmap

Consider the linear system

$$\dot{\mathbf{x}} = \mathbf{A}(t)\mathbf{x} + \mathbf{B}(t)\mathbf{u}, \tag{10.42}$$

$$\mathbf{y} = \mathbf{C}(t)\mathbf{x} + \mathbf{D}(t)\mathbf{u}. \tag{10.43}$$

our objective is to find a linear map

$$\mathbf{u} = -\mathbf{K}(t)\mathbf{x}, \tag{10.44}$$

such that $\mathbf{x} = \mathbf{0}$ is an exponentially stable equilibrium point of the closed-loop system

$$\dot{\mathbf{x}} = \mathbf{A}(t)\mathbf{x} - \mathbf{B}(t)\mathbf{K}(t)\mathbf{x} = \big(\mathbf{A}(t) - \mathbf{B}(t)\mathbf{K}(t)\big)\,\mathbf{x} = \hat{\mathbf{A}}(t)\mathbf{x},$$

where $\hat{\mathbf{A}} = \mathbf{A}(t) - \mathbf{B}(t)\mathbf{K}(t)$ defines the closed loop model of the system.

Roadmap (contd.)

Recall, we want to find $\mathbf{K}$ such that $\mathbf{x} = \mathbf{0}$ is an exponentially stable equilibrium point of the closed loop system

$$\dot{\mathbf{x}} = \mathbf{A}(t)\mathbf{x} - \mathbf{B}(t)\mathbf{K}(t)\mathbf{x} = \big(\mathbf{A}(t) - \mathbf{B}(t)\mathbf{K}(t)\big)\,\mathbf{x} = \hat{\mathbf{A}}(t)\mathbf{x},$$

To do so, we need to study

- What characteristics of $\hat{\mathbf{A}}$ makes the system exponentially stable?
- What conditions $\mathbf{A}$ and $\mathbf{B}$ should have to let us modify characteristics of $\hat{\mathbf{A}}$ via $\mathbf{u} = -\mathbf{K}\mathbf{x}$?
- How can we compute a matrix $\mathbf{K}$?

Solution of Linear Systems

Solution of Homogeneous State Equations

Before we solve vector-matrix differential equations, let us review the solution of the scalar differential equation

$$\dot{x} = ax \tag{10.45}$$

We may assume a solution $x(t)$ of the form

$$x(t) = b_0 + b_1 t + \cdots + b_k t^k + \cdots \tag{10.46}$$

Taking derivative of $x(t)$ with respect to time and substituting in (10.45) gives

$$b_1 + 2b_2 t + \cdots + k b_k t^{k-1} + \cdots = a\left(b_0 + b_1 t + \cdots + b_k t^k + \cdots\right) \tag{10.47}$$

if (10.46) is the solution to (10.45), then (10.47) must hold true for any t

Solution of Homogeneous State Equations (contd.)

Accordingly, by equating the coefficients of the equal powers of t, we get

$$\begin{aligned} b_1 &= ab_0 \\ b_2 &= \frac{1}{2}ab_1 = \frac{1}{2}a^2b_0 \\ b_3 &= \frac{1}{3}ab_2 = \frac{1}{2}a^3b_0 \\ &\vdots \\ b_k &= \frac{1}{k!}ab_{k-1} = \frac{1}{k!}a^kb_0 \end{aligned}$$

If we set $t = 0$ in (10.46) we get

$$x(0) = b_0, \tag{10.49}$$

Solution of Homogeneous State Equations (contd.)

which implies

$$x(t) = \left(1 + at + \frac{1}{2!}a^2t^2 + \cdots + \frac{1}{k!}a^kt^k + \cdots\right)x(0). \tag{10.50}$$

Recall that for any $z \in \mathbb{R}$,

$$e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!} = 1 + z + \frac{z^2}{2!} + \cdots + \frac{z^k}{k!} + \cdots. \tag{10.51}$$

Hence, using $z = at$, we have

$$x(t) = e^{at}x(0). \tag{10.52}$$

Solution of Homogeneous State Equations (contd.)

We shall now solve the vector-matrix differential equation

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}, \tag{10.53}$$

where $\mathbf{x}(t) \in \mathbb{R}^n$ and $\mathbf{A} \in \mathbb{R}^{n\times n}$. Analogous to the scalar case, we assume that the solution is in the form of a vector power series

$$\mathbf{x}(t) = \mathbf{b}_0 + \mathbf{b}_1 t + \cdots + \mathbf{b}_k t^k + \cdots \tag{10.54}$$

Substituting (10.54) in (10.53) gives

$$\mathbf{b}_1 + 2\mathbf{b}_2 t + \cdots + k\mathbf{b}_k t^{k-1} + \cdots = \mathbf{A}\left(\mathbf{b}_0 + \mathbf{b}_1 t + \cdots + \mathbf{b}_k t^k + \cdots\right) \tag{10.55}$$

if (10.54) is the solution to (10.53), then (10.55) must hold true for any t. Accordingly, by equating the coefficients of the equal powers of t, we get

$$\mathbf{b}_k = \frac{1}{k!}\mathbf{A}^k \mathbf{b}_0 \tag{10.56}$$

Solution of Homogeneous State Equations (contd.)

Assuming $\mathbf{x}(0) = \mathbf{x}_0$ implies $\mathbf{b}_0 = \mathbf{x}_0$, the solution $\mathbf{x}(t)$ can be written as

$$\mathbf{x}(t) = \left(\mathbf{I} + \mathbf{A}t + \frac{1}{2!}\mathbf{A}^2 t^2 + \cdots + \frac{1}{k!}\mathbf{A}^k t^k + \cdots\right)\mathbf{x}_0. \tag{10.57}$$

Matrix Exponential

Because of its similarity to the infinite power series for a scalar exponential, we call the term

$$\mathbf{I} + \mathbf{A}t + \frac{1}{2!}\mathbf{A}^2 t^2 + \cdots + \frac{1}{k!}\mathbf{A}^k t^k + \cdots ,$$

the *matrix exponential* and denote it with $e^{\mathbf{A}t}$, Thus

$$e^{\mathbf{A}t} := \sum_{k=0}^{\infty} \frac{\mathbf{A}^k t^k}{k!}. \tag{10.58}$$

Properties of Matrix Exponential

Convergence of Matrix Exponential

Theorem 10.1

The matrix exponential

$$e^{\mathbf{A}t} := \sum_{k=0}^{\infty} \frac{\mathbf{A}^k t^k}{k!}$$

converges absolutely for all finite t.

Hence, we can use computer calculations to approximate $e^{\mathbf{A}t}$.

What does *absolute convergence* mean and when can we say an infinite power series *converges absolutely*?

To prove the theorem, we need a few definitions

The Radius of Convergence of a Power Series

Definition 10.1 (Radius of Convergence)

Let

$$f(z) := \sum_{n=0}^{\infty} c_n (z-a)^n$$

be a power series centered at a. Then, the *radius of convergence* of $f(z)$ is

$$r := \sup \left\{ |z-a| \mid \sum_{n=0}^{\infty} c_n (z-a)^n < \infty \right\}$$

Note that r belongs to the set of the extended real numbers ($r \in \mathbb{R} \cup \{\pm\infty\}$)

Absolute Convergence

Definition 10.2

Let G be *metric space* and $a_n \in G$ for all n. Then, a G-valued series $\sum_{n=0}^{\infty} a_n$ is *absolutely convergent* if

$$\sum_{n=0}^{\infty} \|a_n\| < \infty.$$

Hence, to prove Theorem 10.1, we need to show that the radius of convergence of

$$e^{\mathbf{A}t} := \sum_{k=0}^{\infty} \left\| \frac{\mathbf{A}^k t^k}{k!} \right\|$$

is in fact infinity.

Proof of Theorem 10.1

Proof. Let $M_{n\times n}(\mathbb{R})$, for $n < \infty$, denote the space of all $n \times n$ real matrices. By the equivalency of norms in finite dimensions, for any $k \in \mathbb{N}_0$ we have

$$0 \leq \left\| \frac{\mathbf{A}^k t^k}{k!} \right\| \leq \frac{\|\mathbf{A}\|^k \, |t|^k}{k!} = \frac{\left(\|\mathbf{A}\| \, |t|\right)^k}{k!}$$

Since for every $x \in \mathbb{R}$ and $x < \infty$ the series

$$\sum_{k=0}^{\infty} \frac{x^k}{k!} = e^x$$

converges, then, for every $\mathbf{A} \in M_{n\times n}(\mathbb{R})$ and $t < \infty$ we have

$$\sum_{k=0}^{\infty} \left\| \frac{\mathbf{A}^k t^k}{k!} \right\| \leq \sum_{k=0}^{\infty} \frac{\left(\|\mathbf{A}\| \, |t|\right)^k}{k!} = e^{\|A\||t|} < \infty,$$

that implies absolute convergence of $e^{\mathbf{A}t}$. $\square$

Properties of the Matrix Exponential

The matrix exponential $e^{\mathbf{A}t}$ has the following properties

$$\frac{\mathrm{d}}{\mathrm{d}t} e^{\mathbf{A}t} = \mathbf{A}e^{\mathbf{A}t} = e^{\mathbf{A}t}\mathbf{A} \tag{10.59a}$$

$$e^{\mathbf{A}(t+s)} = e^{\mathbf{A}t}e^{\mathbf{A}s} \tag{10.59b}$$

$$\left(e^{\mathbf{A}t}\right)^{-1} = e^{-\mathbf{A}t} \tag{10.59c}$$

$$\text{If} \quad AB = BA \quad \text{then} \quad e^{(\mathbf{A}+\mathbf{B})t} = e^{\mathbf{A}t}e^{\mathbf{B}t} \tag{10.59d}$$

Proofs for Properties of the Matrix Exponential

Proof of (10.59a). Owing to the convergence of the infinite series $\sum_{k=0}^{\infty} \frac{\mathbf{A}^k t^k}{k!}$ (Theorem 10.1), we can differentiate each term in the series to get

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} e^{\mathbf{A}t} &= \frac{\mathrm{d}}{\mathrm{d}t} \left(\mathbf{I} + \mathbf{A}t + \frac{1}{2!}\mathbf{A}^2 t^2 + \cdots + \frac{1}{k!}\mathbf{A}^k t^k + \cdots \right) \\
&= \mathbf{A} + \mathbf{A}^2 t + \cdots + \frac{1}{(k-1)!}\mathbf{A}^k t^{k-1} + \cdots \\
&= \mathbf{A} \left(\mathbf{I} + \mathbf{A}t + \frac{1}{2!}\mathbf{A}^2 t^2 + \cdots + \frac{1}{(k-1)!}\mathbf{A}^{k-1} t^{k-1} + \cdots \right) = \mathbf{A} e^{\mathbf{A}t} \\
&= \left(\mathbf{I} + \mathbf{A}t + \frac{1}{2!}\mathbf{A}^2 t^2 + \cdots + \frac{1}{(k-1)!}\mathbf{A}^{k-1} t^{k-1} + \cdots \right) \mathbf{A} = e^{\mathbf{A}t} \mathbf{A}.
\end{aligned}
$$

$\square$

Proofs for Properties of the Matrix Exponential (contd.)

Proof of (10.59b).

$$e^{\mathbf{A}t}e^{\mathbf{A}s} = \left(\sum_{k=0}^{\infty}\frac{\mathbf{A}^k t^k}{k!}\right)\left(\sum_{i=0}^{\infty}\frac{\mathbf{A}^i s^i}{i!}\right) = \sum_{k=0}^{\infty}\mathbf{A}^k\left(\sum_{i=0}^{\infty}\frac{t^i s^{k-i}}{i!(k-i)!}\right)$$

$$= \sum_{k=0}^{\infty}\frac{\mathbf{A}^k (t+s)^k}{k!} = e^{\mathbf{A}(t+s)}.$$

□

Proof of (10.59c). Based on the results of (10.59b), setting $s=-t$ in $e^{\mathbf{A}t}e^{\mathbf{A}s} = e^{\mathbf{A}(t+s)}$ gives

$$e^{\mathbf{A}t}e^{-\mathbf{A}t} = e^{\mathbf{A}(t-t)} = e^{\mathbf{A}0} = \mathbf{I}.$$

Proofs for Properties of the Matrix Exponential (contd.)

Since we can similarly set $t = -s$, we have

$$e^{\mathbf{A}t}e^{-\mathbf{A}t} = e^{-\mathbf{A}t}e^{\mathbf{A}t} = \mathbf{I} \implies \left(e^{\mathbf{A}t}\right)^{-1} = e^{-\mathbf{A}t}.$$

Since the inverse of $e^{\mathbf{A}t}$ always exists, $e^{\mathbf{A}t}$ is nonsingular. $\square$

Proof of (10.59d). Expanding $e^{(\mathbf{A}+\mathbf{B})t}$ gives

$$e^{(\mathbf{A}+\mathbf{B})t} = \mathbf{I} + (\mathbf{A}+\mathbf{B})\,t + \frac{(\mathbf{A}+\mathbf{B})^2\,t^2}{2!} + \cdots + \frac{(\mathbf{A}+\mathbf{B})^k\,t^k}{k!} + \cdots$$

Proofs for Properties of the Matrix Exponential (contd.)

On the other hand, expanding $e^{\mathbf{A}t}e^{\mathbf{B}t}$ gives

$$e^{\mathbf{A}t}e^{\mathbf{B}t} = \left(\mathbf{I} + \mathbf{A}t + \frac{\mathbf{A}^2t^2}{2!} + \frac{\mathbf{A}^3t^3}{3!} + \cdots\right)\left(\mathbf{I} + \mathbf{B}t + \frac{\mathbf{B}^2t^2}{2!} + \frac{\mathbf{B}^3t^3}{3!} + \cdots\right)$$

$$= \mathbf{I} + (\mathbf{A} + \mathbf{B})\,t + \frac{\mathbf{A}^2t^2}{2!} + \mathbf{AB}t^2 + \frac{\mathbf{B}^2t^2}{2!} + \frac{\mathbf{A}^3t^3}{3!} + \frac{\mathbf{A}^2\mathbf{B}t^3}{2!} + \frac{\mathbf{AB}^2t^3}{2!} + \frac{\mathbf{B}^3t^3}{3!} + \cdots$$

Hence, by subtracting $e^{\mathbf{A}t}e^{\mathbf{B}t}$ from $e^{(\mathbf{A}+\mathbf{B})t}$ we get

$$e^{(\mathbf{A}+\mathbf{B})t} - e^{\mathbf{A}t}e^{\mathbf{B}t} = \frac{\mathbf{BA} - \mathbf{AB}}{2!}t^2 + \frac{\mathbf{BA}^2 + \mathbf{ABA} + \mathbf{B}^2\mathbf{A} + \mathbf{BAB} - 2\mathbf{A}2\mathbf{B} - 2\mathbf{AB}^2}{3!}t^3 + \cdots$$

Note that $e^{(\mathbf{A}+\mathbf{B})t} - e^{\mathbf{A}t}e^{\mathbf{B}t} = \mathbf{0}$ only if $\mathbf{A}$ and $\mathbf{B}$ commute (if $\mathbf{AB} = \mathbf{BA}$). $\square$

State Transition Matrix

We can write the solution of homogeneous state equation

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}, \tag{10.60}$$

with the initial condition $\mathbf{x}(0)$ as

$$\mathbf{x}(t) = \mathbf{\Phi}(t)\mathbf{x}(0), \tag{10.61}$$

where $\mathbf{\Phi}(t)$ is an $n \times n$ matrix and is the unique solution of the IVP

$$\dot{\mathbf{\Phi}}(t) = \mathbf{A}\mathbf{\Phi}(t), \quad \mathbf{\Phi}(0) = \mathbf{I}. \tag{10.62}$$

To verify this claim, note that

$$\mathbf{x}(0) = \mathbf{\Phi}(0)\mathbf{x}(0) = \mathbf{I}\mathbf{x}(0) = \mathbf{x}(0), \tag{10.63}$$

$$\dot{\mathbf{x}}(t) = \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{\Phi}(t)\mathbf{x}(0) = \mathbf{A}\mathbf{\Phi}(t)\mathbf{x}(0) = \mathbf{A}\mathbf{x}(t). \tag{10.64}$$

State Transition Matrix (contd.)

Since the solution of $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$, with initial condition $\mathbf{x}(0)$ is $\mathbf{x}(t) = \mathbf{e}^{\mathbf{A}t}\mathbf{x}(0)$, we get:

The State Transition Matrix for LTI Systems

The *state transition matrix* $\mathbf{\Phi}(t)$ for autonomous linear systems is uniquely defined as

$$\mathbf{\Phi}(t) = e^{\mathbf{A}t} \tag{10.65}$$

The state-transition matrix contains all the information about the system's free motions defined by $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$.

Properties of State Transition Matrix

Some important properties of the state transition matrix are

$$\mathbf{\Phi}(0) = e^{\mathbf{A}0} = \mathbf{I}, \tag{10.66}$$

$$\mathbf{\Phi}(t_1 + t_2) = e^{\mathbf{A}(t_1+t_2)} = e^{\mathbf{A}t_1}e^{\mathbf{A}t_2} = \mathbf{\Phi}(t_1)\mathbf{\Phi}(t_2) = \mathbf{\Phi}(t_2)\mathbf{\Phi}(t_1), \tag{10.67}$$

$$\mathbf{\Phi}^{-1}(t) = \left[e^{\mathbf{A}t}\right]^{-1} = e^{-\mathbf{A}t} = \mathbf{\Phi}(-t), \tag{10.68}$$

$$\mathbf{\Phi}^{n}(t) = \mathbf{\Phi}(t)\mathbf{\Phi}(t)\mathbf{\Phi}^{n-2}(t) = \mathbf{\Phi}(2t)\mathbf{\Phi}^{n-2}(t) = \cdots = \mathbf{\Phi}(nt), \tag{10.69}$$

$$\mathbf{\Phi}(t_2 - t_1)\mathbf{\Phi}(t_1 - t_0) = e^{\mathbf{A}(t_2-t_1)}e^{\mathbf{A}(t_1-t_0)} = e^{\mathbf{A}t_2}e^{-\mathbf{A}t_0} = \mathbf{\Phi}(t_2 - t_0). \tag{10.70}$$

Solution of Nonhomogeneous State Equations

Similar to the homogeneous case, let us start with the scalar equation

$$\dot{x} = ax + bu. \tag{10.71}$$

Moving ax to the left side and multiplying both sides by e^{-at} gives

$$e^{-at}\left(\dot{x} - ax\right) = \frac{\mathrm{d}}{\mathrm{d}t}\left(e^{-at}x(t)\right) = e^{-at}bu(t). \tag{10.72}$$

Integrating both sides from 0 to t yields

$$x(t) = e^{at}x(0) + e^{at}\int_0^t e^{-a\tau}bu(\tau)\,d\tau. \tag{10.73}$$

Solution of Nonhomogeneous State Equations (contd.)

Let us now consider the nonhomogeneous state equation in vector form described by

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}, \tag{10.74}$$

where $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{A} \in \mathbb{R}^{n\times n}$, $\mathbf{u} \in \mathbb{R}^m$ and $\mathbf{B} \in \mathbb{R}^{n\times m}$. Similar to scalar case, we can move $\mathbf{A}\mathbf{x}$ to the left hand side and premultiply both sides by $e^{-\mathbf{A}t}$. Accordingly

$$e^{-\mathbf{A}t}\left(\dot{\mathbf{x}} - \mathbf{A}\mathbf{x}\right) = \frac{\mathrm{d}}{\mathrm{d}t}\left(e^{-\mathbf{A}t}\mathbf{x}(t)\right) = e^{-\mathbf{A}t}\mathbf{B}\mathbf{u}(t). \tag{10.75}$$

Let $\mathbf{x}_0 = \mathbf{x}(t_0)$. Integrating the above equation between t_0 and t gives

$$\mathbf{x}(t) = e^{\mathbf{A}(t-t_0)}\mathbf{x}_0 + \int_{t_0}^{t} e^{\mathbf{A}(t-\tau)}\mathbf{B}\mathbf{u}(\tau)\ d\tau, \tag{10.76}$$

or, using the state transition matrix, we have

$$\mathbf{x}(t) = \mathbf{\Phi}(t-t_0)\mathbf{x}(0) + \int_{t_0}^{t} \mathbf{\Phi}(t-\tau)\mathbf{B}\mathbf{u}(\tau)\ d\tau. \tag{10.77}$$

A Short Review on Eigenvalues and Eigenvectors

Eigenvalues and Eigenvectors

Definition 10.3 (Eigenvalues and Eigenvectors)

Let $T : V \to V$ be a linear transformation from a vector space V over a field $\mathbb{F}$ into itself and $\mathbf{v} \in V$, $\mathbf{v} \neq \mathbf{0}$, then $\mathbf{v}$ is an *eigenvector* of T if

$$T\mathbf{v} = \lambda\mathbf{v},$$

where λ is a scalar in $\mathbb{F}$, known as the *eigenvalue* associated with $\mathbf{v}$.

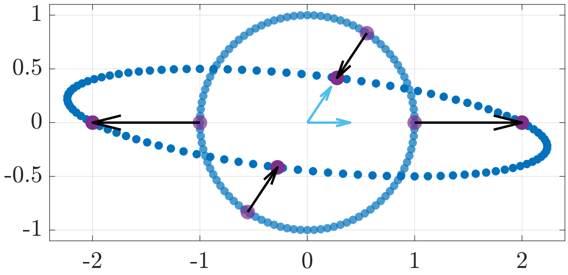
Example

Let $\mathbf{A} : \mathbb{R}^{2\times 2} \to \mathbb{R}^{2\times 2}$ be

$$\mathbf{A} = \begin{bmatrix} 2 & -1 \\ 0 & 1/2 \end{bmatrix},$$

with the eigenvalues and the corresponding eigenvectors:

$$\sigma(\mathbf{A}) = \{\lambda_1 = 2,\ \lambda_2 = 1/2\}, \quad \mathbf{v}_1 = (1, 0), \quad \mathbf{v}_2 = (1, 3/2)$$



Characteristic Equation

Let $\mathbf{A}$ be an $n \times n$ matrix, then $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$ implies

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{v} = \mathbf{0}. \tag{10.78}$$

Since we seek non-trivial solutions[†] to (10.78), we need nullity$\{\mathbf{A} - \lambda\mathbf{I}\} > 0$, which implies

$$\det(\mathbf{A} - \lambda\mathbf{I}) = 0. \tag{10.79}$$

Based on Leibniz formula for determinants, $\det(\mathbf{A} - \lambda\mathbf{I})$ is polynomial of degree n and known as the *characteristic polynomial of* $\mathbf{A}$ and (10.79) is the *characteristic equation of* $\mathbf{A}$. If λ_1 to λ_n are the n eigenvalues of $\mathbf{A}$, based on fundamental theorem of algebra[‡] we have

$$\det(\mathbf{A} - \lambda\mathbf{I}) = (\lambda_1 - \lambda)(\lambda_2 - \lambda)\cdots(\lambda_n - \lambda), \tag{10.80}$$

[†] Since $\mathbf{v} = \mathbf{0}$ is always a solution to (10.78), by a non-trivial solution we mean solutions where $\mathbf{v} \neq \mathbf{0}$.

[‡] Every univariate polynomial of degree $n \in \mathbb{N}$ with coefficients in $\mathbb{C}$ can be factorized as $c(x - r_1)\cdots(x - r_n)$, where $c, r_1, \ldots, r_n \in \mathbb{C}$ and $r_1, \ldots, r_n$ are the roots of the polynomial. A multiple root appears in several factors, and the number of its occurrences is its *multiplicity*.

Algebraic Multiplicity

Definition 10.4 (Algebraic Multiplicity)

Let λ_i be an eigenvalue of an $n \times n$ matrix $\mathbf{A}$, then the *algebraic multiplicity* $\mu_{\mathbf{A}}(\lambda_i)$ of the eigenvalue λ_i is its multiplicity as a root of the characteristic polynomial, that is, the largest integer k such that $(\lambda - \lambda_i)^k$ divides the characteristic polynomial of $\mathbf{A}$ evenly.

As an example, consider the matrix

$$\mathbf{A} = \begin{bmatrix} 4 & 1 \\ 0 & 4 \end{bmatrix}$$

The characteristic polynomial of $\mathbf{A}$ is obtained by calculating

$$\det(\mathbf{A} - \lambda\mathbf{I}) = \det \begin{bmatrix} 4-\lambda & 1 \\ 0 & 4-\lambda \end{bmatrix} = (4-\lambda)^2 = 0,$$

which indicates the eigenvalue $\lambda = 4$ has algebraic multiplicity of 2, that is $\mu_{\mathbf{A}}(4) = 2$.

Geometric Multiplicity

For given eigenvalue λ of the $n \times n$ matrix $\mathbf{A}$, Let

$$E = \left\{ \mathbf{v} \mid (\mathbf{A} - \lambda \mathbf{I})\, \mathbf{v} = \mathbf{0} \right\}. \tag{10.81}$$

See that $E = \text{null}(\mathbf{A} - \lambda \mathbf{I})$ is the set of all eigenvectors of $\mathbf{A}$ associated with λ. The set E is called the *eigenspace* or *characteristic space* of $\mathbf{A}$ associated with λ.

Definition 10.5 (Geometric Multiplicity)

The *geometric multiplicity* of λ, denoted as $\gamma_{\mathbf{A}}(\lambda)$, is the dimension of the nullspace of $\mathbf{A} - \lambda \mathbf{I}$.

Based on Theorem 6.3 (the rank–nullity theorem), we have

$$\gamma_{\mathbf{A}}(\lambda) = n - \text{rank}(\mathbf{A} - \lambda \mathbf{I}) \tag{10.82}$$

Gershgorin Disc

Let $\mathbf{A} = [a_{ij}] \in \mathbb{C}^{n \times n}$. For $i \in \{1, \ldots, n\}$ let r_i be the sum of the absolute values of the non-diagonal entries in the i-th row. That is

$$r_i = \sum_{i \neq j} |a_{ij}| \,. \tag{10.83}$$

Then a closed disc centered at a_{ii} with radius r_i

$$D(a_{ii}, r_i) \subseteq \mathbb{C}, \tag{10.84}$$

is called a *Gershgorin* disc.

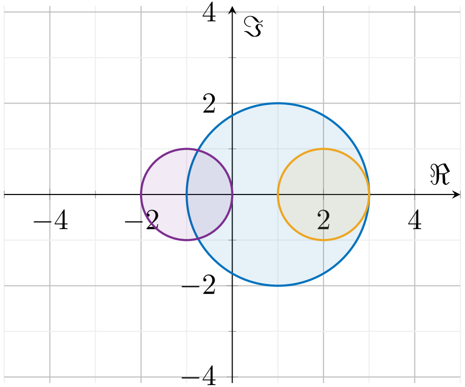
Example of Gershgorin Discs

Consider the following example matrix

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ -1 & 0 & 2 \end{bmatrix}$$

The Gershgorin disc radii are calculated as

$$\begin{aligned} r_1 &= |1| + |1| = 2, \\ r_2 &= |1| + |0| = 1, \\ r_3 &= |-1| + |0| = 1. \end{aligned}$$



Gershgorin Circle Theorem

Theorem 10.2 (Gershgorin Circle Theorem)

Every eigenvalue of $\mathbf{A} \in \mathbb{C}^{n \times n}$ *lies within at least one of the Gershgorin discs.*

Proof. Let λ be an eigenvalue of $\mathbf{A}$ with corresponding eigenvector $\mathbf{x} = (x_1, \ldots, x_n)$. Pick $i \in \{1, \ldots, n\}$ such that $|x_i| \geq |x_j|$. For the i-th row of the equation

$$\mathbf{A}\mathbf{x} = \lambda \mathbf{x},$$

we have

$$\sum_j a_{ij} x_j = \lambda x_i \implies \sum_{j \neq i} a_{ij} x_j = (\lambda - a_{ii}) x_i \implies \sum_{j \neq i} a_{ij} \frac{x_j}{x_i} = \lambda - a_{ii}.$$

Using the triangular inequality and the fact that $|x_j|/|x_i| \leq 1$ we have

$$|\lambda - a_{ii}| = \left| \sum_{j \neq i} a_{ij} \frac{x_j}{x_i} \right| \leq \sum_{j \neq i} |a_{ij}| = r_i.$$

Example of Gershgorin Circle Theorem

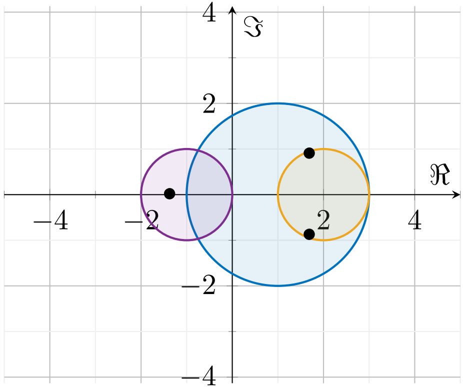
Following the example of Gershgorin disks, for the matrix $\mathbf{A}$

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ -1 & 0 & 2 \end{bmatrix}$$

the set of eigenvalues are

$$\sigma(\mathbf{A}) \approx \{-1.37,\ 1.69 + 0.89i,\ 1.69 - 0.89i\}\,.$$

The eigenvalues of $\mathbf{A}$ are depicted in the figure. Based on the theorem, and as depicted, Every eigenvalue of $\mathbf{A} \in \mathbb{C}^{n\times n}$ lies within at least one of the Gershgorin discs



Roadmap

We evaluated the solution to the closed loop LTI system

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}, \quad \mathbf{x}(0) = \mathbf{x}_0,$$

as

$$\mathbf{x}(t) = e^{\mathbf{A}t}\mathbf{x}_0,$$

where

$$e^{\mathbf{A}t} := \mathbf{I} + \mathbf{A}t + \frac{1}{2!}\mathbf{A}^2 t^2 + \cdots + \frac{1}{k!}\mathbf{A}^k t^k + \cdots .$$

Furthermore, we concluded that $e^{\mathbf{A}t}$ converges absolutely for all finite t. Now we want to answer two questions:

1. Could we evaluate $e^{\mathbf{A}t}$ analytically?
2. How does the matrix $\mathbf{A}$ affect the solution of $\mathbf{x}(t) = e^{\mathbf{A}t}\mathbf{x}_0$?

Matrix Decomposition

Eigendecomposition of a matrix

Let $\mathbf{A}$ be a square $n \times n$ matrix with n *linearly independent eigenvectors* $\mathbf{q}_1$ to $\mathbf{q}_n$. Then $\mathbf{A}$ can be factorized as

$$\mathbf{A} = \mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^{-1}$$

where

$$\mathbf{Q} = \begin{bmatrix} | & | & & | \\ \mathbf{q}_1 & \mathbf{q}_2 & \cdots & \mathbf{q}_n \\ | & | & & | \end{bmatrix}$$

is the square $n \times n$ matrix whose i^{th} column is the eigenvector $\mathbf{q}_i$ of $\mathbf{A}$, and $\boldsymbol{\Lambda}$ is the diagonal matrix of eignevectors

$$\boldsymbol{\Lambda} = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

Eigendecomposition of a matrix (contd.)

⚠ Note that only *diagonalizable* matrices can be factorized in this way.

The decomposition can be derived from the fundamental property of eigenvectors:

$$\mathbf{A}\mathbf{q} = \mathbf{q}\lambda \tag{10.85}$$

since

$$\begin{aligned}
\mathbf{A}\begin{bmatrix}\mathbf{q}_1 & \cdots & \mathbf{q}_n\end{bmatrix} &= \begin{bmatrix}\mathbf{A}\mathbf{q}_1 & \cdots & \mathbf{A}\mathbf{q}_n\end{bmatrix} \\
&= \begin{bmatrix}\mathbf{q}_1\lambda_1 & \cdots & \mathbf{q}_n\lambda_n\end{bmatrix} = \begin{bmatrix}\mathbf{q}_1 & \cdots & \mathbf{q}_n\end{bmatrix}\begin{bmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{bmatrix}
\end{aligned}$$

Jordan Decomposition (Jordan Canonical Form)

In general, $\mathbf{A} \in \mathbb{C}^{n \times n}$ is similar to a block diagonal matrix $\mathbf{J}$ (i.e. there exists invertible $\mathbf{P}$ such that $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{J}$), where $\mathbf{J}$

$$\mathbf{J} = \begin{bmatrix} \mathbf{J}_1 & & \\ & \ddots & \\ & & \mathbf{J}_p \end{bmatrix}$$

where each *Jordan block* $\mathbf{J}_i$ is a square matrix, related to eigenvalue λ_i, of the form

$$\mathbf{J}_i = \begin{bmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{bmatrix}.$$

$\mathbf{J}$ is called the *Jordan normal form* of $\mathbf{A}$.

Jordan Decomposition (Jordan Canonical Form) (contd.)

$$\mathbf{J} = \begin{bmatrix} \begin{array}{ccc|} \lambda_1 & 1 & \\ & \lambda_1 & 1 \\ & & \lambda_1 \\ \hline \end{array} & & & & \\ & \boxed{\begin{array}{cc} \lambda_2 & 1 \\ & \lambda_2 \end{array}} & & & \\ & & \boxed{\lambda_3} & & \\ & & & \ddots & \\ & & & & \begin{array}{|cc} \hline \lambda_n & 1 \\ & \lambda_n \end{array} \end{bmatrix}$$

Jordan Decomposition (Jordan Canonical Form) (contd.)

We can show the following properties for the Jordan form:

- The eigenvalues of matrix $\mathbf{A}$ are the diagonal entries of $\mathbf{J}$, hence are also the eigenvalues of $\mathbf{J}$ with the same algebraic multiplicity.
- The number of Jordan blocks for λ_i equals its geometric multiplicity.
- Algebraic multiplicity of λ_i equals the sum of sizes of all Jordan blocks for λ_i.
- $\mathbf{A}$ is diagonalizable if and only if, for every eigenvalue λ of $\mathbf{A}$, geometric and algebraic multiplicities coincide, implying Jordan blocks are 1×1 matrices.

Jordan Blocks: Case of Equal Geometric and Algebraic Multiplicities

When geometric and algebraic multiplicities of an eigenvalue λ_i are the same, then there are as many blocks as algebraic multiplicity of λ_i. Hence, each block has the size one and the corresponding Jordan blocks are

$$\mathbf{J} = \begin{bmatrix} \lambda_i & 0 & & & \\ & \lambda_i & 0 & & \\ & & \ddots & \ddots & \\ & & & \lambda_i & 0 \\ & & & & \lambda_i \end{bmatrix}.$$

An immediate example is the $n \times n$ identity matrix $\mathbf{I}_n$. It is left as an exercise to show that $\mu_{\mathbf{I}_n}(1) = \gamma_{\mathbf{I}_n}(1) = n$.

Matrix Functions in Jordan Form

Let $f : \mathbb{C} \to \mathbb{C}$ be an analytical function. Applying f on a $n \times n$ Jordan block $\mathbf{J}$ with eigenvalue λ results in an upper triangular matrix:

$$f(\mathbf{J}) = \begin{bmatrix} f(\lambda) & f'(\lambda) & \frac{f''(\lambda)}{2} & \cdots & \frac{f^{(n-1)}(\lambda)}{(n-1)!} \\ & f(\lambda) & f'(\lambda) & \cdots & \frac{f^{(n-2)}(\lambda)}{(n-2)!} \\ & & \ddots & \ddots & \vdots \\ & & & f(\lambda) & f'(\lambda) \\ & & & & f(\lambda) \end{bmatrix},$$

that is, the elements of the k-th superdiagonal of the resulting matrix are

$$\frac{f^{(k)}(\lambda)}{k!}.$$

Matrix Functions in Jordan Form (contd.)

As an example, consider the application of the power function $f(z) = z^n$, for $n \in \mathbb{N}$ to the following Jordan form

$$\begin{bmatrix} \lambda_1 & 1 & 0 & 0 & 0 \\ 0 & \lambda_1 & 1 & 0 & 0 \\ 0 & 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & 0 & \lambda_2 & 1 \\ 0 & 0 & 0 & 0 & \lambda_2 \end{bmatrix}^n = \begin{bmatrix} \lambda_1^n & \binom{n}{1}\lambda_1^{n-1} & \binom{n}{2}\lambda_1^{n-2} & 0 & 0 \\ 0 & \lambda_1^n & \binom{n}{1}\lambda_1^{n-1} & 0 & 0 \\ 0 & 0 & \lambda_1^n & 0 & 0 \\ 0 & 0 & 0 & \lambda_2^n & \binom{n}{1}\lambda_2^{n-1} \\ 0 & 0 & 0 & 0 & \lambda_2^n \end{bmatrix}, \tag{10.86}$$

where the binomial coefficients are defined as

$$\binom{n}{k} = \prod_{i=1}^{k} \frac{n+1-i}{i}.$$

Computation of Matrix Exponential

Using Laplace Transform

Taking the Laplace transform of $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$ gives

$$s\mathbf{X}(s) - \mathbf{x}(0) = \mathbf{A}\mathbf{X}(s), \tag{10.87}$$

where $\mathbf{X}(s) = \mathscr{L}\left[\mathbf{x}(t)\right]$. Solving the above equation for $\mathbf{X}(s)$ gives

$$\mathbf{X}(s) = (s\mathbf{I} - \mathbf{A})^{-1}\mathbf{x}(0). \tag{10.88}$$

To find $\mathbf{x}(t)$ we can take the inverse Laplace transform of $\mathbf{X}(s)$ an obtain

$$\mathbf{x}(t) = \mathscr{L}^{-1}\left[(s\mathbf{I} - \mathbf{A})^{-1}\right]\mathbf{x}(0). \tag{10.89}$$

Knowing that $(s\mathbf{I} - \mathbf{A})^{-1}$ converges, using Neumann series we have

$$(s\mathbf{I} - \mathbf{A})^{-1} = \frac{\mathbf{I}}{s} + \frac{\mathbf{A}}{s^2} + \frac{\mathbf{A}^2}{s^3} + \cdots \tag{10.90}$$

Accordingly we have

$$\mathscr{L}^{-1}\left[(s\mathbf{I} - \mathbf{A})^{-1}\right] = \mathbf{I} + \mathbf{A}t + \frac{\mathbf{A}^2t^2}{2!} + \frac{\mathbf{A}^3t^3}{3!} + \cdots = e^{\mathbf{A}t} \tag{10.91}$$

Using Eigendecomposition of $\mathbf{A}$

If $\mathbf{A}$ is diagonalizable, then there exists matrices $\mathbf{Q}$ and $\mathbf{\Lambda}$ such that

$$\mathbf{A} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^{-1} \implies \mathbf{A}^k = \mathbf{A}\mathbf{A}\cdots\mathbf{A} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^{-1}\mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^{-1}\cdots\mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^{-1} = \mathbf{Q}\mathbf{\Lambda}^k\mathbf{Q}^{-1}. \tag{10.92}$$

Using the decomposition, we can write the series expansion as

$$e^{\mathbf{A}t} = \mathbf{Q}\mathbf{Q}^{-1} + t\mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^{-1} + \frac{t^2}{2!}\mathbf{Q}\mathbf{\Lambda}^2\mathbf{Q}^{-1} + \cdots + \frac{t^k}{k!}\mathbf{Q}\mathbf{\Lambda}^k\mathbf{Q}^{-1} + \cdots \tag{10.93}$$

$$= \mathbf{Q}\left(\mathbf{I} + t\mathbf{\Lambda} + \frac{t^2}{2!}\mathbf{\Lambda}^2 + \cdots + \frac{t^k}{k!}\mathbf{\Lambda}^k + \cdots\right)\mathbf{Q}^{-1} \tag{10.94}$$

$$= \mathbf{Q}\begin{bmatrix} \sum_{k=0}^{\infty}\frac{\lambda_1^k t^k}{k!} & & \\ & \ddots & \\ & & \sum_{k=0}^{\infty}\frac{\lambda_n^k t^k}{k!} \end{bmatrix}\mathbf{Q}^{-1} = \mathbf{Q}\begin{bmatrix} e^{\lambda_1 t} & & \\ & \ddots & \\ & & e^{\lambda_n t} \end{bmatrix}\mathbf{Q}^{-1}. \tag{10.95}$$

Using Jordan form of $\mathbf{A}$

Similar to the eignedecomposition, we can use Jordan decomposition of $\mathbf{A}$

$$\mathbf{A} = \mathbf{P}\mathbf{J}\mathbf{P}^{-1}$$

to write $e^{\mathbf{A}t}$ as

$$e^{\mathbf{A}t} = \mathbf{P}\mathbf{P}^{-1} + t\mathbf{P}\mathbf{J}\mathbf{P}^{-1} + \frac{t^2}{2!}\mathbf{P}\mathbf{J}^2\mathbf{P}^{-1} + \cdots + \frac{t^k}{k!}\mathbf{P}\mathbf{J}^k\mathbf{P}^{-1} + \cdots \tag{10.96}$$

$$= \mathbf{P}\left(\mathbf{I} + t\mathbf{J} + \frac{t^2}{2!}\mathbf{J}^2 + \cdots + \frac{t^k}{k!}\mathbf{J}^k + \cdots\right)\mathbf{P}^{-1}, \tag{10.97}$$

where we can use the results of (10.86) to compute $\mathbf{J}^k$.

Using Jordan form of $\mathbf{A}$ (contd.)

See that for a Jordan block $\mathbf{J}_i$ of size n, associated with the eigenvalue λ_i, we have

$$\sum_{k=0}^{\infty} \frac{t^k \mathbf{J}_i^k}{k!} = \begin{bmatrix} \sum_{k=0}^{\infty} \frac{\lambda_i^k t^k}{k!} & t \sum_{k=0}^{\infty} \frac{\lambda_i^k t^k}{k!} & \frac{t^2}{2} \sum_{k=0}^{\infty} \frac{\lambda_i^k t^k}{k!} & \cdots & \frac{t^{n-1}}{(n-1)!} \sum_{k=0}^{\infty} \frac{\lambda_i^k t^k}{k!} \\ & \sum_{k=0}^{\infty} \frac{\lambda_i^k t^k}{k!} & t \sum_{k=0}^{\infty} \frac{\lambda_i^k t^k}{k!} & \cdots & \frac{t^{n-2}}{(n-2)!} \sum_{k=0}^{\infty} \frac{\lambda_i^k t^k}{k!} \\ & & \ddots & \ddots & \vdots \\ & & & & \sum_{k=0}^{\infty} \frac{\lambda_i^k t^k}{k!} \end{bmatrix} \tag{10.98a}$$

$$= \begin{bmatrix} e^{\lambda_i t} & t e^{\lambda_i t} & \frac{t^2}{2} e^{\lambda_i t} & \cdots & \frac{t^{n-1}}{(n-1)!} e^{\lambda_i t} \\ & e^{\lambda_i t} & t e^{\lambda_i t} & \cdots & \frac{t^{n-2}}{(n-2)!} e^{\lambda_i t} \\ & & \ddots & \ddots & \vdots \\ & & & & e^{\lambda_i t} \end{bmatrix} = e^{\mathbf{J}_i t}. \tag{10.98b}$$

Using Jordan form of $\mathbf{A}$ (contd.)

Using the results of (10.97) and (10.98) we have

$$e^{\mathbf{A}t} = \mathbf{P}\begin{bmatrix} e^{\mathbf{J}_1 t} & & \\ & \ddots & \\ & & e^{\mathbf{J}_p t} \end{bmatrix}\mathbf{P}^{-1} = \mathbf{P}e^{\mathbf{J}t}\mathbf{P}^{-1}, \tag{10.99}$$

where for each Jordan block $\mathbf{J}_i$ of size n associated with the eigenvalue λ_i we have

$$e^{\mathbf{J}_i t} = \begin{bmatrix} e^{\lambda_i t} & te^{\lambda_i t} & \frac{t^2}{2}e^{\lambda_i t} & \cdots & \frac{t^{n-1}}{(n-1)!}e^{\lambda_i t} \\ & e^{\lambda_i t} & te^{\lambda_i t} & \cdots & \frac{t^{n-2}}{(n-2)!}e^{\lambda_i t} \\ & & \ddots & \ddots & \vdots \\ & & & & e^{\lambda_i t} \end{bmatrix}. \tag{10.100}$$

The Putzer Algorithm for Computing $e^{\mathbf{A}t}$

Given $n \times n$ matrix $\mathbf{A}$ with eigenvalues $\lambda_1, \ldots, \lambda_n$, base on Putzer's algorithm[§] we have

$$e^{\mathbf{A}t} = \sum_{j=0}^{n-1} r_{j+1}(t)\mathbf{P}_j, \tag{10.101}$$

where for $j = 0, 1, \ldots, n-1$,

$$\mathbf{P}_0 = \mathbf{I}, \tag{10.102}$$

$$\mathbf{P}_j = \prod_{k=1}^{j}(\mathbf{A} - \lambda_k\mathbf{I}) = \mathbf{P}_{j-1}(\mathbf{A} - \lambda_j\mathbf{I}), \tag{10.103}$$

and for $j = 2, 3, \ldots, n$, r_j is the solution to the initial value problems

$$r_1(0) = 1, \qquad \dot{r}_1 = \lambda_1 r_1, \tag{10.104a}$$

$$r_j(0) = 0, \qquad \dot{r}_j = \lambda_j r_j + r_{j-1}, \tag{10.104b}$$

The Putzer Algorithm for Computing $e^{\mathbf{A}t}$ (contd.)

Note that the Putze algorithm does not require that the matrix $\mathbf{A}$ be diagonalizable and bypasses the complexities of the Jordan canonical form.

One can easily confirm that

$$e^{\mathbf{A}t} = \sum_{j=0}^{n-1} r_{j+1}(t)\mathbf{P}_j, \tag{10.105}$$

is equivalent to

$$e^{\mathbf{A}t} = \sum_{j=0}^{n-1} \alpha_j(t)\mathbf{A}^j = \alpha_0(t)\mathbf{I} + \alpha_1(t)\mathbf{A} + \cdots + \alpha_{n-1}(t)\mathbf{A}^{n-1}, \tag{10.106}$$

where $\alpha_j = \alpha_j(r_1, \ldots, r_j, \lambda_1, \ldots, \lambda_j)$ depends on r_1 to r_j and λ_1 to λ_j.

[§]Reference: Putzer, E. J. "Avoiding the Jordan Canonical Form in the Discussion of Linear Systems with Constant Coefficients." The American Mathematical Monthly, vol. 73, no. 1, 1966, pp. 2–7

Stability in LTI Systems

Theorem 10.3 (Stability of the Origin in Autonomous Linear Systems)

Let $\mathbf{A} \in \mathbb{R}^{n\times n}$ *with eigenvalues* $\{\lambda_1, \ldots, \lambda_n\}$. *The equilibrium point* $\mathbf{x} = \mathbf{0}$ *of* $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$ *is stable if and only if*

- *for all eigenvalues of* $\mathbf{A}$, $\Re(\lambda_i) \leq 0$,
- *for any eigenvalue with* $\Re(\lambda_i) = 0$, *we have* $\mu_{\mathbf{A}}(\lambda_i) = \gamma_{\mathbf{A}}(\lambda_i)$. *In other words, any eigenvalue with zero real part is either a simple or semisimple eigenvalue*[a]. *The equilibrium point* $\mathbf{x} = \mathbf{0}$ *is (globally) asymptotically stable if and only if all eigenvalues of* $\mathbf{A}$ *satisfy* $\Re(\lambda_i) < 0$.

[a] If $\mu_{\mathbf{A}}(\lambda_i) = 1$, then λ_i is said to be a simple eigenvalue. If $\mu_{\mathbf{A}}(\lambda_i)$ equals the geometric multiplicity of $\gamma_{\mathbf{A}}(\lambda_i)$, then (λ_i) is said to be a semisimple eigenvalue.

Stability in LTI Systems (contd.)

Proof.

Hurwitz Matrix

Hurwitz matrix

In engineering and stability theory, a square matrix $\mathbf{A}$ is called a *Hurwitz matrix* if every eigenvalue of $\mathbf{A}$ has strictly negative real part, that is,

$$\Re(\lambda_i) < 0$$

for each eigenvalue λ_i. Since for a Hurwitz matrix $\mathbf{A}$, the differential equation

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x},$$

is asymptotically stable ($\lim_{t\to\infty} \|\mathbf{x}(t)\| = 0$), a Hurwitz matrix is also called a *stable matrix.*

Stability of the Origin in LTI Systems via Lyapunov's Theorem

Asymptotic stability of $\mathbf{x} = \mathbf{0}$ for the system $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$ can also be studied using Lyapunov's method. Consider a quadratic function

$$V(\mathbf{x}) = \mathbf{x}^T \mathbf{P} \mathbf{x}, \tag{10.107}$$

where $\mathbf{P}$ is a real symmetric positive definite matrix. The derivative of V along the trajectories of the system is

$$\dot{V}(\mathbf{x}) = \mathbf{x}^T \mathbf{P} \dot{\mathbf{x}} + \dot{\mathbf{x}}^T \mathbf{P} \mathbf{x} = \mathbf{x}^T \left(\mathbf{P}\mathbf{A} + \mathbf{A}^T \mathbf{P} \right) \mathbf{x} = -\mathbf{x}^T \mathbf{Q} \mathbf{x} \tag{10.108}$$

If $\mathbf{Q}$ is positive definite, we can conclude by Theorem 9.7 that the origin is asymptotically stable; which implies $\Re \lambda_i < 0$ for all eigenvalues of $\mathbf{A}$.

Stability of the Origin in LTI Systems via Lyapunov's Theorem (contd.)

Theorem 10.4

A matrix $\mathbf{A}$ *is Hurwitz; that is,* $\Re(\lambda_i) < 0$ *for all eigenvalues of* $\mathbf{A}$*, if and only if for any given positive definite symmetric matrix* $\mathbf{Q}$ *there exists a positive definite symmetric matrix* $\mathbf{P}$ *that satisfies the Lyapunov equation*

$$\mathbf{P}\mathbf{A} + \mathbf{A}^T\mathbf{P} = -\mathbf{Q} \tag{10.109}$$

Moreover, if $\mathbf{A}$ *is Hurwitz, then* $\mathbf{P}$ *is the unique solution of (10.109).*

In order to prove the theorem, we need the following lemma

Stability of the Origin in LTI Systems via Lyapunov's Theorem (contd.)

Lemma 10.5

Let $\mathbf{A}$ *be Hurwitz and* $\mathbf{Q}$ *be a symmetric positive definite matrix. Define* $\mathbf{P}$ *as*

$$\mathbf{P} := \int_0^\infty \exp(\mathbf{A}^T t)\,\mathbf{Q}\exp(\mathbf{A}t)\,dt.$$

Then $\mathbf{P}$ *is also a symmetric positive matrix.*

Proof of Lemma 10.5.

Stability of the Origin in LTI Systems via Lyapunov's Theorem (contd.)

We can confirm that the integral exists, since the integrand is a sum of terms of the form $t^k e^{\lambda_i t}$, where $\Re(\lambda_i) < 0$. By contradiction, assume $\mathbf{P}$ is not positive definite, then there exists $\mathbf{x} \neq \mathbf{0}$ such that $\mathbf{x}^T\mathbf{P}\mathbf{x} = 0$. However,

$$\begin{aligned}\mathbf{x}^T\mathbf{P}\mathbf{x} = 0 &\implies \int_0^\infty \mathbf{x}^T \exp(\mathbf{A}^T t)\mathbf{Q}\exp(\mathbf{A}t)\mathbf{x}\, dt = 0 \\ &\implies \exp(\mathbf{A}t)\mathbf{x} = 0,\ \forall t \geq 0 \implies \mathbf{x} = \mathbf{0}\end{aligned}$$

since $\exp(\mathbf{A}t)$ is nonsingular for all t. This contradiction shows that $\mathbf{P}$ is positive definite. $\square$

Stability of the Origin in LTI Systems via Lyapunov's Theorem (contd.)

Proof of Theorem 10.4.

Proof is adopted from "Nonlinear Systems", 3rd Edition by Hassan Khalil

Using the Lyapunov function $V(\mathbf{x}) = \mathbf{x}^T \mathbf{P} \mathbf{x}$, we immediately get the sufficiency from Theorem 9.7. To prove necessity, assume that all eigenvalues of $\mathbf{A}$ satisfy $\Re(\lambda_i) < 0$ and consider the matrix $\mathbf{P}$, defined as

$$\mathbf{P} = \int_0^\infty \exp(\mathbf{A}^T t)\, \mathbf{Q} \exp(\mathbf{A} t)\, dt.$$

Stability of the Origin in LTI Systems via Lyapunov's Theorem (contd.)

Based on Lemma 10.5, $\mathbf{P}$ is symmetric and positive definite. By substituting $\mathbf{P}$, as defined above, in the left-hand side of $\mathbf{PA} + \mathbf{A}^T\mathbf{P} = -\mathbf{Q}$ gives

$$\begin{aligned}\mathbf{PA} + \mathbf{A}^T\mathbf{P} &= \int_0^\infty \exp(\mathbf{A}^T t)\mathbf{Q}\exp(\mathbf{A}t)\mathbf{A}\,dt + \int_0^\infty \mathbf{A}^T\exp(\mathbf{A}^T t)\mathbf{Q}\exp(\mathbf{A}t)\,dt \\ &= \int_0^\infty \frac{d}{dt}\exp(\mathbf{A}^T t)\mathbf{Q}\exp(\mathbf{A}t)\,dt = \exp(\mathbf{A}^T t)\mathbf{Q}\exp(\mathbf{A}t)\Big|_0^\infty = -\mathbf{Q}\end{aligned}$$

which shows that P is indeed a solution of $\mathbf{PA} + \mathbf{A}^T\mathbf{P} = -\mathbf{Q}$. To show that it is the unique solution, suppose there is another solution $\bar{\mathbf{P}} \neq \mathbf{P}$. Then,

$$(\mathbf{P} - \bar{\mathbf{P}})\mathbf{A} + \mathbf{A}^T(\mathbf{P} - \bar{\mathbf{P}}) = \mathbf{0}$$

Stability of the Origin in LTI Systems via Lyapunov's Theorem (contd.)

Premultiplying by $\exp(\mathbf{A}^T t)$ and postmultiplying by $\exp(\mathbf{A}t)$, we obtain

$$\begin{aligned}
\mathbf{0} &= \exp(\mathbf{A}^T t)\left[(\mathbf{P}-\bar{\mathbf{P}})\mathbf{A} + \mathbf{A}^T(\mathbf{P}-\bar{\mathbf{P}})\right]\exp(\mathbf{A}t) \\
&= \frac{d}{dt}\{\exp(\mathbf{A}^T t)(\mathbf{P}-\bar{\mathbf{P}})\exp(\mathbf{A}t)\}
\end{aligned}$$

Hence,

$$\exp(\mathbf{A}^T t)(\mathbf{P}-\bar{\mathbf{P}})\exp(\mathbf{A}t) = \text{a constant } \forall t$$

In particular, since $\exp(\mathbf{A}0) = \mathbf{I}$, we have

$$(\mathbf{P}-\bar{\mathbf{P}}) = \exp(\mathbf{A}^T t)(\mathbf{P}-\bar{\mathbf{P}})\exp(\mathbf{A}t) \to 0 \text{ as } t \to \infty$$

Therefore, $\bar{\mathbf{P}} = \mathbf{P}$. $\square$

Notes on Finding $\mathbf{Q}$ and $\mathbf{P}$ Matrices

To show stability of the origin for $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$, it suffices to find one $\mathbf{P}$ and $\mathbf{Q}$ pair that solve

$$\mathbf{P}\mathbf{A} + \mathbf{A}^T\mathbf{P} = -\mathbf{Q}.$$

An easier approach is to assume a symmetric positive definite $\mathbf{Q}$ and find the corresponding $\mathbf{P}$.

$\mathbf{P}\mathbf{A} + \mathbf{A}^T\mathbf{P} = -\mathbf{Q}$ is a linear algebraic equation that can be solved by rearranging it as

$$\mathbf{M}\mathbf{z} = \mathbf{b}, \tag{10.110}$$

where $\mathbf{z}$ and $\mathbf{b}$ are defined by stacking the elements of $\mathbf{P}$ and $\mathbf{Q}$ in vectors.

Lyapunov equation is an example of Sylvester equation

$$\mathbf{C}\mathbf{X} + \mathbf{X}\mathbf{B} = \mathbf{Y},$$

where $\mathbf{X} = \mathbf{P}$, $\mathbf{Y} = \mathbf{Q}$ and $\mathbf{C} = \mathbf{B} = \mathbf{A}$. The solution to Sylvester equation is well studied and there are numerically efficient methods for solving the problem.

Example

adopted from "Nonlinear Systems", 3rd Edition by Hassan Khalil

Let

$$\mathbf{A} = \begin{bmatrix} 0 & -1 \\ 1 & -1 \end{bmatrix}, \quad \mathbf{Q} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \text{and} \quad \mathbf{P} = \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix},$$

where, due to symmetry, $p_{12} = p_{21}$. The Lyapunov equation $\mathbf{PA} + \mathbf{A}^T\mathbf{P} = -\mathbf{Q}$ can be rewritten as

$$\begin{bmatrix} 0 & 2 & 0 \\ -1 & -1 & 1 \\ 0 & -2 & -2 \end{bmatrix} \begin{bmatrix} p_{11} \\ p_{12} \\ p_{22} \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ -1 \end{bmatrix}.$$

The unique solution of this equation is given by

$$\begin{bmatrix} p_{11} \\ p_{12} \\ p_{22} \end{bmatrix} = \begin{bmatrix} 1.5 \\ -0.5 \\ 1.0 \end{bmatrix} \quad \Longrightarrow \quad \mathbf{P} = \begin{bmatrix} 1.5 & -0.5 \\ -0.5 & 1.0 \end{bmatrix}.$$

One can easily confirm that the matrix $\mathbf{P}$ is positive definite.

Two Dimensional Autonomous Linear Closed Loop Systems

2D Autonomous Linear Closed Loop Systems

As a general form of the 2D autonomous linear closed loop system we have

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad \Longrightarrow \quad \begin{aligned} \dot{x}_1 &= ax_1 + bx_2, && (10.111a) \\ \dot{x}_2 &= cx_1 + dx_2. && (10.111b) \end{aligned}$$

Multiplying (10.111b) by b and substituting bx_2 from (10.111a) in (10.111a) gives

$$b\dot{x}_2 = bcx_1 + d(\dot{x}_1 - ax_1) \tag{10.112}$$

Taking the derivative of (10.111a) with respect to time and substituting $b\dot{x}_2$ from (10.112) gives

$$\begin{aligned} \ddot{x}_1 &= a\dot{x}_1 + b\dot{x}_2 && (10.113a) \\ &= a\dot{x}_1 + bcx_1 + d(\dot{x}_1 - ax_1) && (10.113b) \\ &= (a + d)\dot{x}_1 + (bc - da)x_1. && (10.113c) \end{aligned}$$

2D Autonomous Linear Closed Loop Systems (contd.)

Let $x := x_1$, $p := -(a+d)$ and $q := (ad - bc)$. We can write (10.113c) as

$$\ddot{x} + p\dot{x} + qx = 0, \tag{10.114}$$

which implies, for any matrix $\mathbf{A} \in \mathbb{R}^{2\times 2}$, the system $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$ has an equivalent linear second order system in the form of (10.114). The generic solution of (10.114) is

$$x(t) = \begin{cases} k_1 e^{\lambda_1 t} + k_2 e^{\lambda_2 t} & \lambda_1, \lambda_2 \in \mathbb{R},\ \lambda_1 \neq \lambda_2, \\ k_1 e^{\lambda_1 t} + k_2 t e^{\lambda_1 t} & \lambda_1 = \lambda_2 \in \mathbb{R}, \\ e^{\alpha t}\left(k_1 \cos(\beta t) + k_2 \sin(\beta t)\right) & \lambda_1, \lambda_2 \in \mathbb{C},\ \alpha = \Re(\lambda_1) = \Re(\lambda_2),\ \beta = \Im(\alpha_1) = \Im(\alpha_2), \end{cases}$$

where k_1 and k_2 depend on the initial (or boundary) conditions. λ_1 and λ_2 are the roots of the characteristic polynomial

$$\lambda^2 + p\lambda + q = 0. \tag{10.115}$$

2D Autonomous Linear Closed Loop Systems (contd.)

Since $p := -(a+d)$ and $q := (ad - bc)$, the characteristic polynomial of (10.114) is in fact

$$\lambda^2 + p\lambda + q = \lambda^2 - (a+d)\lambda + (ad - bc) = 0. \tag{10.116}$$

Note that, for matrix $\mathbf{A}$, as defined in (10.111), we have

$$\det\left(\mathbf{A} - \lambda\mathbf{I}\right) = \det\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}\right) = \lambda^2 - (a+d)\lambda + (ad - bc), \tag{10.117}$$

which implies that

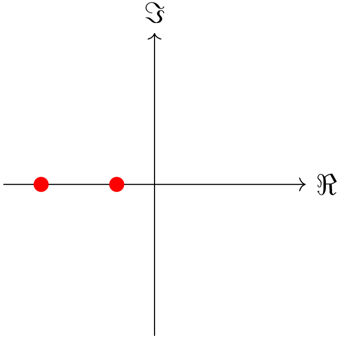
The eigenvalues of $\mathbf{A}$ are the roots of the characteristic polynomial of the equivalent second order system (10.114).

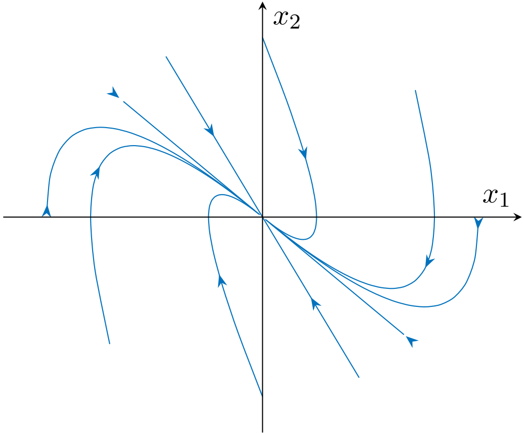
2D Autonomous Linear Closed Loop Systems (contd.)

Now we can study the behavior of the system in the vicinity of origin ($x_1 = x_2 = 0$) as a function of the eigenvalues of $\mathbf{A}$ (or the roots of the characteristic polynomial of the equivalent second order system)

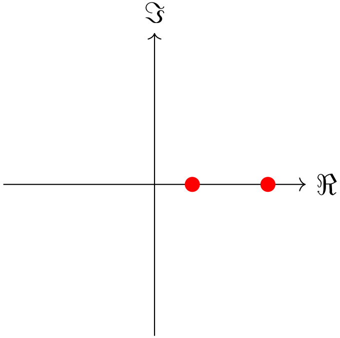
- Stable and unstable Nodes
- Saddle point
- Stable or unstable focus
- Center point

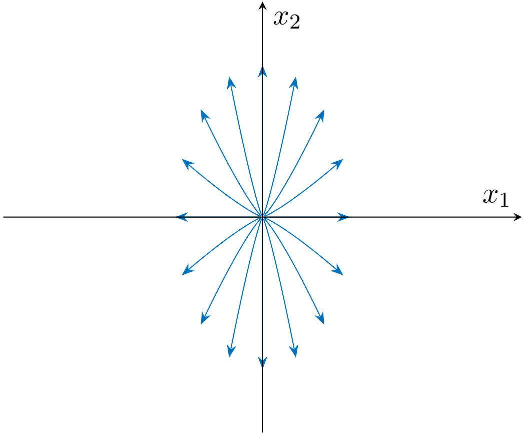
Stable Node



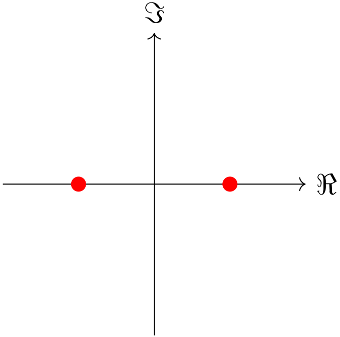


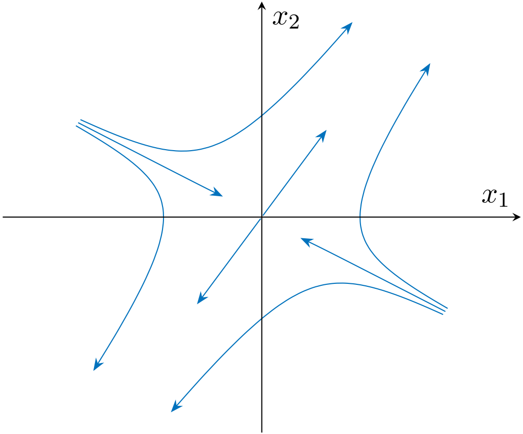
Unstable Node



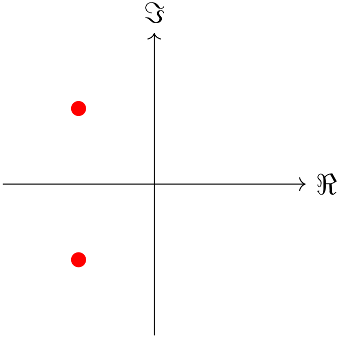


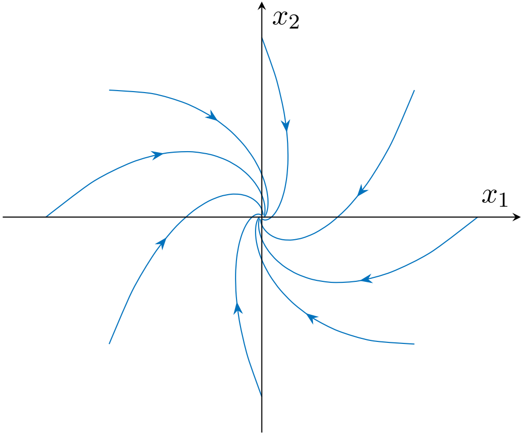
Saddle Point



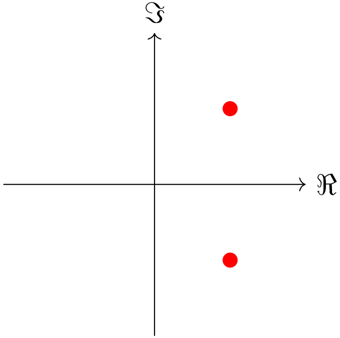


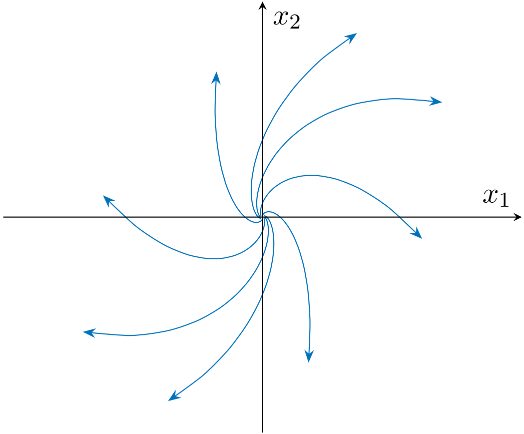
Stable Focus



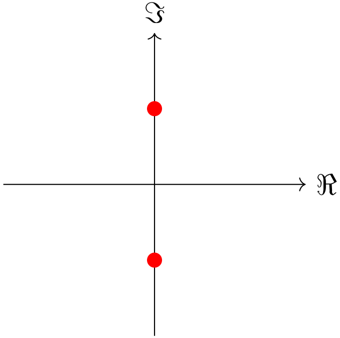


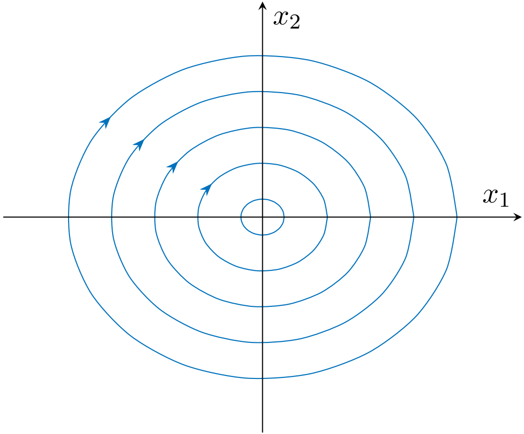
Unstable Focus



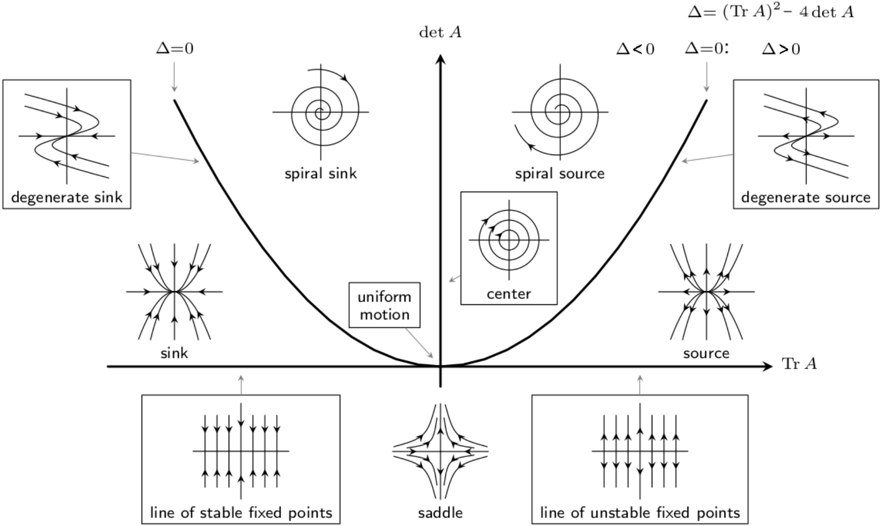


Center Point





Poincaré Diagram: Classification of Phase Portaits in the $(\det A, \mathrm{Tr}\, A)$-plane



Illustrative Examples

Case Study 1

Consider the system $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}u$, defined as

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -3 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u. \tag{10.118}$$

First, let us examine the open loop system with $u = 0$.

$$u = 0 \longrightarrow \boxed{\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}u} \longrightarrow \boxed{\int} \longrightarrow \mathbf{x}(t)$$

Since $u = 0$, the system dynamics simplifies to

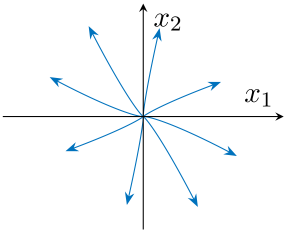
$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -3 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}. \tag{10.119}$$

Case Study 1 (contd.)

The eigenvalues of $\mathbf{A}$ are

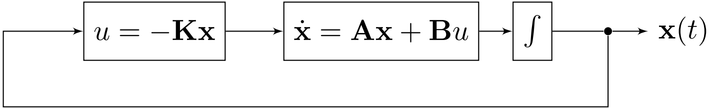
$$\lambda_1 = \frac{1 + i\sqrt{11}}{2}, \quad \lambda_2 = \frac{1 - i\sqrt{11}}{2}, \tag{10.120}$$

which implies $\mathbf{A}$ is not singular. Hence $\mathbf{x} = \mathbf{0}$ is the only equilibrium point of the system and, in fact, it is an unstable focus.



Case Study 1 (contd.)

We would like to use a closed loop control system to convert the origin into a stable node (recall that for a stable node, both eigenvalues are real and negative)



Let us pick our *desired* closed loop eigenvalues as $\xi_1, \xi_2 \in \mathbb{R}$, such that $\xi_2 \leq \xi_1 < 0$. Since the state-space is $\mathbb{R}^2$, and the input space is $\mathbb{R}$. As our control function, we seek a linear transformation

$$u : \mathbb{R}^2 \to \mathbb{R}.$$

Based on Theorem 6.4, such a linear transformation is represented by a 1×2 matrix. Accordingly let

$$u = -\mathbf{K}\mathbf{x} = -\begin{bmatrix} k_1 & k_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = -k_1 x_1 - k_2 x_2. \tag{10.121}$$

Case Study 1 (contd.)

Our objective is to find k_1 and k_2 such that the eigenvalues of the closed loop system coincide with ξ_1 and ξ_2.

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}u = \mathbf{A}\mathbf{x} - \mathbf{B}\mathbf{K}\mathbf{x} = (\mathbf{A} - \mathbf{B}\mathbf{K})\,\mathbf{x} \tag{10.122}$$

$$= \left(\begin{bmatrix} 0 & 1 \\ -3 & 1 \end{bmatrix} - \begin{bmatrix} 0 \\ 1 \end{bmatrix} \begin{bmatrix} k_1 & k_2 \end{bmatrix} \right) \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -3 - k_1 & 1 - k_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \hat{\mathbf{A}}\mathbf{x}. \tag{10.123}$$

To find k_1 and k_2, we can proceed to find the eigenvalues of $\hat{\mathbf{A}}$. Accordingly we have

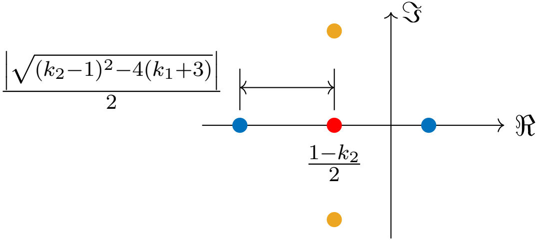
$$\det(\hat{\mathbf{A}} - \lambda \mathbf{I}) = \det \left(\begin{bmatrix} -\lambda & 1 \\ -3 - k_1 & 1 - k_2 - \lambda \end{bmatrix} \right) = \lambda^2 + (k_2 - 1)\lambda + k_1 + 3 \tag{10.124}$$

Case Study 1 (contd.)

The eigenvalues of $\hat{\mathbf{A}}$ are the solutions of $\lambda^2 + (k_2 - 1)\lambda + k_1 + 3 = 0,$, which are

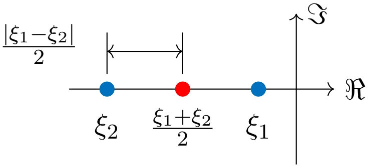
$$\lambda_1 = \frac{1 - k_2}{2} + \frac{\sqrt{(k_2 - 1)^2 - 4(k_1 + 3)}}{2} \tag{10.125}$$

$$\lambda_2 = \frac{1 - k_2}{2} - \frac{\sqrt{(k_2 - 1)^2 - 4(k_1 + 3)}}{2}. \tag{10.126}$$



Case Study 1 (contd.)

As for ξ_1 and ξ_2 we have



Accordingly, to find k_1 and k_2 we can match the centers and radii as

$$k_2 = 1 - \xi_1 - \xi_2, \tag{10.127}$$

and

$$(k_2 - 1)^2 - 4(k_1 + 3) = (\xi_1 - \xi_2)^2 \implies k_1 = \xi_1\xi_2 - 3. \tag{10.128}$$

Case Study 1 (contd.)

Thus, given *desired* closed loop eigenvalues ξ_1 and ξ_2, we can construct *the* control function

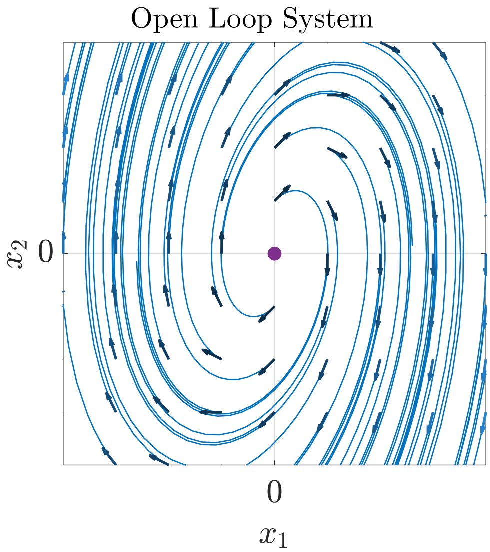
$$u = -\begin{bmatrix} \xi_1\xi_2 - 3 & 1 - \xi_1 - \xi_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \tag{10.129}$$

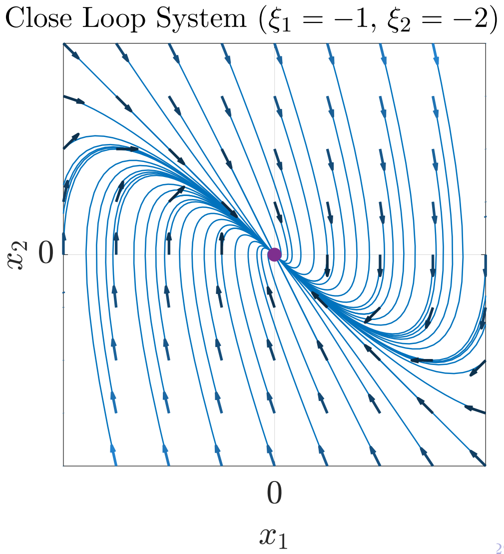
which causes the eignevalues of the closed loop system $\dot{\mathbf{x}} = \mathbf{Ax} - \mathbf{BKx}$ to coincide with ξ_1 and ξ_2. In fact, for λ_1 and λ_2 we have

$$\lambda_1 = \frac{1 - k_2}{2} + \frac{\sqrt{(k_2 - 1)^2 - 4(k_1 + 3)}}{2} = \frac{\xi_1 + \xi_2}{2} + \frac{\sqrt{(\xi_1 - \xi_2)^2}}{2} = \xi_1 \tag{10.130}$$

$$\lambda_2 = \frac{1 - k_2}{2} - \frac{\sqrt{(k_2 - 1)^2 - 4(k_1 + 3)}}{2} = \frac{\xi_1 + \xi_2}{2} - \frac{\sqrt{(\xi_1 - \xi_2)^2}}{2} = \xi_2 \tag{10.131}$$

Case Study 1 (contd.)





Key Insights from Case Study 1

- The structure of matrices $\mathbf{A}$ and $\mathbf{B}$ allows us to use a linear controller $u = -\mathbf{K}\mathbf{x}$ to dictate the eigenvalues of the closed loop system (despite the fact that the origin of the open loop system $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$ is an unstable focus.)
- Using this approach, we have converted the origin of the closed loop system

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} - \mathbf{B}\mathbf{K}\mathbf{x} = (\mathbf{A} - \mathbf{B}\mathbf{K})\mathbf{x} = \hat{\mathbf{A}}\mathbf{x},$$

  to a *stable node.*
- The desired eigenvalues ξ_1 and ξ_2 *uniquely* define k_1 and k_2 and, accordingly, the control matrix $\mathbf{K} = [k_1,\, k_2]$,

$$\begin{aligned} k_1 &= \xi_1 \xi_2 - 3, \\ k_2 &= 1 - \xi_1 + \xi_2. \end{aligned}$$

Case Study 2

Let us now consider a two dimensional state space with a two dimensional input

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}. \tag{10.132}$$

We can see that for the open loop system with $\mathbf{u} = \mathbf{0}$, the origin is a saddle point. In fact,

$$\mathbf{u} = \mathbf{0} \implies \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \tag{10.133}$$

and for matrix $\mathbf{A}$, as defined, we have

$$\det\left(\begin{bmatrix} -\lambda & 1 \\ 1 & 2-\lambda \end{bmatrix}\right) = \lambda^2 - 2\lambda - 1 = 0 \implies \begin{cases} \lambda_1 = 1 + \sqrt{2} > 0 \\ \lambda_2 = 1 - \sqrt{2} < 0. \end{cases} \tag{10.134}$$

Case Study 2 (contd.)

Our objective, once more, is to find a linear control function

$$\mathbf{u} : \mathbb{R}^2 \to \mathbb{R}^2, \tag{10.135}$$

so that the eigenvalues of the closed loop system coincide with a set of desired eigenvalues. Similar to the case study 1, based on Theorem 6.4, such a linear transformation is represented by a 2×2 matrix $\mathbf{K}$, that is

$$\mathbf{u} = \mathbf{K}\mathbf{x} = \begin{bmatrix} k_{11} & k_{12} \\ k_{21} & k_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{10.136}$$

Following the same steps as before, we can compute the eigenvalues of the closed loop system

$$\dot{\mathbf{x}} = (\mathbf{A} - \mathbf{B}\mathbf{K})\,\mathbf{x} = \hat{\mathbf{A}}\mathbf{x},$$

and find eigenvalues of $\hat{\mathbf{A}}$, namely λ_1 and λ_2, as functions of k_{ij}.

Case Study 2 (contd.)

Accordingly, for $\hat{\mathbf{A}}$ we have

$$\hat{\mathbf{A}} = \begin{bmatrix} -k_{11} & 1 - k_{12} \\ 1 - k_{22} & 2 - k_{22} \end{bmatrix}. \tag{10.137}$$

Hence, the characteristic polynomial of $\hat{\mathbf{A}}$ is

$$\det\left(\begin{bmatrix} -k_{11} - \lambda & 1 - k_{12} \\ 1 - k_{22} & 2 - k_{22} - \lambda \end{bmatrix}\right) = \lambda^2 + (k_{11} + k_{22} - 2)\,\lambda + k_{11}(k_{22} - 2) + (k_{12} - 1)(1 - k_{21}) \tag{10.138}$$

$$\lambda_1 = \frac{2 - k_{11} - k_{22}}{2} + \frac{\sqrt{(k_{11} + k_{22} - 2)^2 - 4k_{11}(k_{22} - 2) - 4(k_{12} - 1)(1 - k_{21})}}{2}, \tag{10.139a}$$

$$\lambda_2 = \frac{2 - k_{11} - k_{22}}{2} - \frac{\sqrt{(k_{11} + k_{22} - 2)^2 - 4k_{11}(k_{22} - 2) - 4(k_{12} - 1)(1 - k_{21})}}{2}. \tag{10.139b}$$

Case Study 2 (contd.)

Following the idea as in case study 1, given *desired* closed loop eigenvalues ξ_1 and ξ_2, we can match the center and radii of λ_1 and λ_2 with ξ_1 and ξ_2. Accordingly, for k_{ij} we have

$$k_{11} + k_{22} = 2 - \xi_1 - \xi_2, \tag{10.140a}$$

$$k_{11}(k_{22} - 2) + (k_{12} - 1)(1 - k_{21}) = \xi_1 \xi_2. \tag{10.140b}$$

See that (10.140) is a system of two nonlinear equations with four unknowns and does not have a unique solution. To find "a" solution, we can either assume more constraints on $\mathbf{K}$ or find it as a solution to an optimization problem.

Case Study 2 (contd.)

To follow with our example, assume

$$k_{11} = k_{22} = \kappa_1, \quad \text{and} \quad k_{12} = k_{21} = \kappa_2, \tag{10.141}$$

Using the above assumptions, we can solve (10.140) for κ_1 and κ_2 and get

$$\kappa_1 = 1 - (\xi_1 + \xi_2)/2 \tag{10.142}$$

$$\kappa_2 = 1 \pm \left(\sqrt{(\xi_1 - \xi_2)^2 - 4}\right)/2 \tag{10.143}$$

Since the control output is usually (always in the context of robot control) is a real physical signal (e.g. force, torque, pump set point, joint kinematics), the output of the controller should be in $\mathbb{R}$, which implies, the assumption (10.141) restricts the choice of desired eigenvalues to

$$(\xi_1 - \xi_2)^2 \geq 4. \tag{10.144}$$

Case Study 2 (contd.)

For example, if we pick

$$\xi_1 = -1, \quad \xi_2 = -3, \tag{10.145}$$

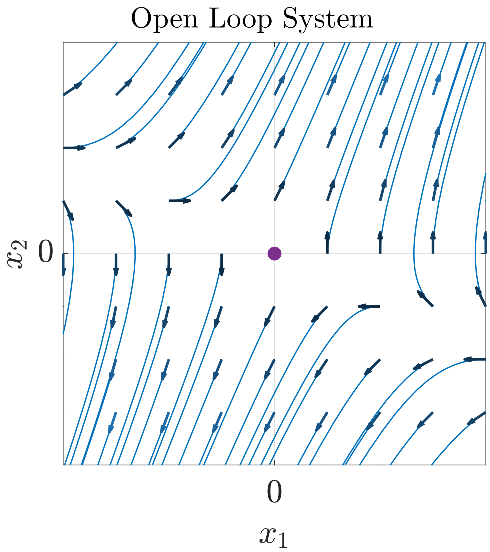
we get

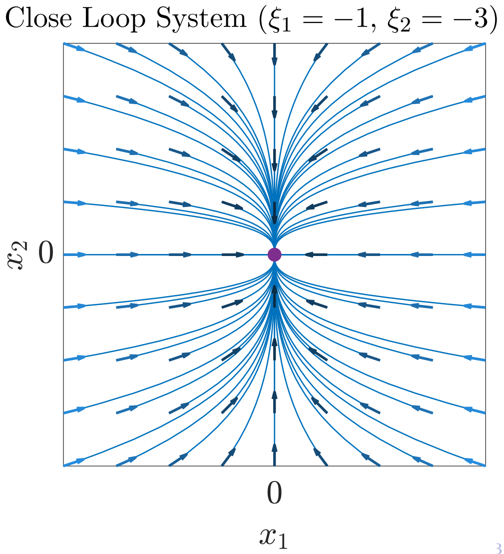
$$(\xi_1 - \xi_2)^2 = (-1+3)^2 = 4 \quad \begin{cases} \kappa_1 = 1 - \frac{\xi_1+\xi_2}{2} = 3, \\ \kappa_2 = 1 \pm \left(\sqrt{(\xi_1 - \xi_2)^2 - 4}\right)/2 = 1. \end{cases} \tag{10.146}$$

Hence, the corresponding control matrix is

$$\mathbf{K} = \begin{bmatrix} k_{11} & k_{12} \\ k_{21} & k_{22} \end{bmatrix} = \begin{bmatrix} \kappa_1 & \kappa_2 \\ \kappa_2 & \kappa_1 \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}. \tag{10.147}$$

Case Study 2 (contd.)





Key Insights from Case Study 2

- Similar to case study 1, the structure of matrices $\mathbf{A}$ and $\mathbf{B}$ allows us to use a linear controller $u = -\mathbf{K}\mathbf{x}$ to dictate the eigenvalues of the closed loop system.
- For the given desired eigenvalues ξ_1 and ξ_2, the choice of $\mathbf{K}$ is *not unique*. In fact, to find *a* solution, we considered a specific structure for $\mathbf{K}$,

$$\mathbf{K} = \begin{bmatrix} \kappa_1 & \kappa_2 \\ \kappa_2 & \kappa_1 \end{bmatrix}.$$

- The assumed structure on $\mathbf{K}$ restricts the choice of desired eigenvalues to

$$(\xi_1 - \xi_2)^2 \geq 4, \quad \xi_1,\, \xi_2 \in \mathbb{R}$$

In fact, for any complex conjugate pair $\xi_1, \xi_2 \in \mathbb{C}$, we have

$$(\xi_1 - \xi_2)^2 - 4 = \big(a + ib - (a - ib)\big)^2 - 4 = (i2b)^2 - 4 = -4b^2 - 4 < 0.$$

Case Study 3

As another example, let us consider the system

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} u, \tag{10.148}$$

where $\mathbf{x} = (x_1, x_2, x_3) \in \mathbb{R}^3$ and $u \in \mathbb{R}$. Clearly, $\mathbf{x} = \mathbf{0}$ of the open loop system with $u = 0$ is an unstable node¶. Let us follow our strategy of stabilizing the origin by manipulating the eigenvalues of the closed loop system. Since as a closed loop controller, we seek a function

$$\mathbf{x} \mapsto u(\mathbf{x}) \in \mathbb{R}, \tag{10.149}$$

based on Theorem 6.4, if u is a linear function, then it has a matrix representation

$$u = -\mathbf{K}\mathbf{x} = -\begin{bmatrix} k_1 & k_2 & k_3 \end{bmatrix} \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}^T = -\sum_{i=1}^{3} k_i x_i. \tag{10.150}$$

Case Study 3 (contd.)

For the closed loop system we have

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}u = \mathbf{A}\mathbf{x} - \mathbf{B}\mathbf{K}\mathbf{x} = (\mathbf{A} - \mathbf{B}\mathbf{K})\,\mathbf{x} \tag{10.151}$$

$$= \left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} k_1 & k_2 & k_3 \end{bmatrix} \right) \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -k_1 & -k_2 & 1 - k_3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \hat{\mathbf{A}}\mathbf{x}. \tag{10.152}$$

For the eigenvalues of $\hat{\mathbf{A}}$, we have

$$\det\left(\hat{\mathbf{A}} - \lambda\mathbf{I}\right) = -(\lambda - 1)^2(\lambda + k_{13} - 1) = 0 \implies \begin{cases} \lambda_1 = \lambda_2 = 1, \\ \lambda_3 = 1 - k_{13}, \end{cases} \tag{10.153}$$

which indicates *only one of the three* eigenvalues is adjustable via the input $u = -\mathbf{K}\mathbf{x}$.

¶since the only eigenvalue of $\mathbf{A}$ is $\lambda = 1$ (with algebraic multiplicity of 3)

Remarks on Case Studies

Consider the state-space equation of a single input, single output (SISO) LTI system

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}u \tag{10.154}$$

$$y = \mathbf{C}\mathbf{x}. \tag{10.155}$$

Taking the Laplace transform of the system and solving for $Y(s)/U(s)$ gives

$$\frac{Y(s)}{U(s)} = G(s) = \mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}. \tag{10.156}$$

where for $Y(s) = \mathscr{L}\left[y(t)\right]$. We can write $(s\mathbf{I} - \mathbf{A})^{-1}$ using adjugate and determinant as

$$(s\mathbf{I} - \mathbf{A})^{-1} = \frac{\mathrm{adj}(s\mathbf{I} - \mathbf{A})}{\det(s\mathbf{I} - \mathbf{A})}. \tag{10.157}$$

Remarks on Case Studies (contd.)

Note that the poles of $G(s)$ are the values that satisfy

$$\det(s\mathbf{I} - \mathbf{A}) = 0. \tag{10.158}$$

On the other hand, (10.157) is the characteristic equation of matrix $\mathbf{A}$, whose solutions are the eigenvalues of A.
Accordingly, with respect to classical control, the eigenvalues of the system matrix $\mathbf{A}$ are refereed to as *poles* of the system.

Pole Placement

In linear systems, pole placement involves assigning specified locations to the closed-loop poles via a feedback loop. Hence, controlling the system's closed loop behavior.

Remarks on Case Studies (contd.)

- Not all Linear Time-Invariant (LTI) systems permit pole placement.
  - How can we identify which systems allow for pole placement?
- For systems where pole placement is feasible:
  - The presence of multiple inputs might lead to non-unique solutions for pole placement.
  - Computing the controller matrix $\mathbf{K}$ can become complex, even for small-sized systems. It's essential to employ appropriate tools and algorithms for this task. Given the Abel-Ruffini theorem (also known as the impossibility theorem), we understand that there is no general solution in radicals for polynomial equations of degree five or higher. This complicates finding explicit solutions for pole placement in higher-dimensional systems.
  - Determining an optimal set of poles is crucial for achieving desired system performance and stability.

Inverse Mapping

Injective Maps and Null Space

Proposition 10.6

Let $T : V \to W$ be a linear transformation, then the following statements are equivalent.

1. *T is injective (one–to-one);*
2. *The null space of T only consists of the zero vector:* $\text{null}(T) = \{\mathbf{0}\}$*;*
3. *T maps linearly independent vectors in V to linearly independent vectors in W.*

Proof of (1) $\implies$ (2). Since T is injective, for every $\mathbf{w} \in \text{range}(T)$, there exists a unique $\mathbf{v} \in V$ such that $T(\mathbf{v}) = \mathbf{w}$. Let $\mathbf{v} = \mathbf{0}$, since T is linear, $T(\mathbf{0}) = \mathbf{0} \in W$. Thus, by uniqueness of $\mathbf{0} \in V$ we get $\text{null}(T) = \{\mathbf{0}\}$. □

Injective Maps and Null Space (contd.)

Proof of (2) $\implies$ (1). Let $\mathbf{v}_1, \mathbf{v}_2 \in V$ such that $T(\mathbf{v}_1) = T(\mathbf{v}_2) = \mathbf{w}$. We have

$$T(\mathbf{v}_1) - T(\mathbf{v}_2) = \mathbf{w} - \mathbf{w} = \mathbf{0} \implies T(\mathbf{v}_1 - \mathbf{v}_2) = \mathbf{0}.$$

Since $\text{null}(T) = \{\mathbf{0}\}$, we get $\mathbf{v}_1 - \mathbf{v}_2 = \mathbf{0} \implies \mathbf{v}_1 = \mathbf{v}_2$. $\square$

Proof of (2) $\implies$ (3). Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a set of linearly independent vectors. We need

$$\sum_{i=1}^{n} \alpha_i T(\mathbf{v}_i) = \mathbf{0} \implies \alpha_i = 0 \text{ for every } i.$$

We have

$$\sum_{i=1}^{n} \alpha_i T(\mathbf{v}_i) = \mathbf{0} \implies T\Big(\sum_{i=1}^{n} \alpha_i \mathbf{v}_i\Big) = \mathbf{0}.$$

Injective Maps and Null Space (contd.)

Since $\operatorname{null}(T) = \{\mathbf{0}\}$, then

$$\sum_{i=1}^{n} \alpha_i \, \mathbf{v}_i = \mathbf{0}.$$

Since $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ are linearly independent, $\sum_{i=1}^{n} \alpha_i \, \mathbf{v}_i = \mathbf{0} \implies \alpha_i = 0$ for every i. $\square$

Proof of (3) $\implies$ *(2).* Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a basis for V and $\mathbf{v} \in \operatorname{null}(T)$. We have

$$\mathbf{0} = T(\mathbf{v}) = T\Big(\sum_{i=1}^{n} \alpha_i \mathbf{v}_i\Big) = \sum_{i=1}^{n} \alpha_i \, T(\mathbf{v}_i).$$

Since $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ are linearly independent, $\sum_{i=1}^{n} \alpha_i \, T(\mathbf{v}_i) = \mathbf{0} \implies \alpha_i = 0$ for every i, that implies $\mathbf{v} = \mathbf{0}$. Thus, $\operatorname{null}(T) = \{\mathbf{0}\}$. $\square$

Rank

Definition 10.6 (Rank)

The rank of a linear map Φ is defined as the dimension of its image. That is

$$\operatorname{rank}(\Phi) = \dim(\operatorname{img}(\Phi))$$

The *column rank* of $\mathbf{A}$ is the dimension of the column space of $\mathbf{A}$, while the *row rank* of $\mathbf{A}$ is the dimension of the row space of $\mathbf{A}$.

A matrix is said to have full rank if its rank equals the largest possible for a matrix of the same dimensions, which is the lesser of the number of rows and columns. A matrix is said to be rank-deficient if it does not have full rank.

Theorem 10.7

For any matrix $\mathbf{A}$, the row rank is the same as the column rank and is simply refereed to as the rank of the matrix. In other words, $\operatorname{rank}(\mathbf{A}) = \operatorname{rank}(\mathbf{A}^T)$.

Example

The matrix

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

has rank 2: the first two columns are linearly independent, but the third column is a linear combination of the first two. As another example, the matrix

$$\mathbf{B} = \begin{bmatrix} 1 & 1 & 0 & 2 \\ -1 & -1 & 0 & -2 \end{bmatrix}$$

has rank 1: any pair of columns is linearly dependent. Hence, $\mathrm{rank}(\mathbf{B}^T) = 1$:

$$\mathbf{B}^{\mathrm{T}} = \begin{bmatrix} 1 & -1 \\ 1 & -1 \\ 0 & 0 \\ 2 & -2 \end{bmatrix}$$

Important Equivalent Statements

Theorem 10.8

For any $\mathbf{A} \in \mathbb{C}^{n\times n}$, *the following statements are equivalent:*

1. $\mathbf{A}$ *has an inverse* $\mathbf{A}^{-1}$ *such that* $\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$,
2. $\operatorname{rank}(\mathbf{A}) = n$,
3. $\operatorname{range}(\mathbf{A}) = \mathbb{C}^n$,
4. $\operatorname{null}(\mathbf{A}) = \{\mathbf{0}\}$,
5. 0 *is not an eigenvalue of* $\mathbf{A}$,
6. 0 *is not a singularvalue of* $\mathbf{A}$,
7. $\det(\mathbf{A}) \neq 0$.

Proof is left as an exercise for interested audience.

Controllability

Controllability Definition

Definition 10.7 (Complete (Full) State Controllability)

A dynamic system is said to be *controllable* at time t_0 if there exists an unconstrained control vector that transfer the system from any initial state $\mathbf{x}(t_0)$ to any other state in a finite interval of time.

⚠ Note that:

- Controllability does not mean that a reached state can be maintained, only that any state can be reached.
- Controllability does not mean that arbitrary paths can be made through state space, only that there exists a path within the prescribed finite time interval.

Controllability in Discrete Linear Autonomous Systems

Consider the a discrete-time linear system

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k, \tag{10.159}$$

where $k \in \mathbb{Z}$ represents the time variable; Accordingly, $\mathbf{x}_k = \mathbf{x}(t_k)$ and $\mathbf{u}_k = \mathbf{u}(t_k)$. Given the state $\mathbf{x}_0$ and control $\mathbf{u}_0$ at an initial time t_0, based on (10.159), we have

$$\mathbf{x}_1 = \mathbf{A}\mathbf{x}_0 + \mathbf{B}\mathbf{u}_0, \tag{10.160}$$

$$\mathbf{x}_2 = \mathbf{A}\mathbf{x}_1 + \mathbf{B}\mathbf{u}_1 = \mathbf{B}\mathbf{u}_1 + \mathbf{A}\mathbf{B}\mathbf{u}_0 + \mathbf{A}^2\mathbf{x}_0, \tag{10.161}$$

$$\mathbf{x}_3 = \mathbf{A}\mathbf{x}_2 + \mathbf{B}\mathbf{u}_2 = \mathbf{B}\mathbf{u}_2 + \mathbf{A}\mathbf{B}\mathbf{u}_1 + \mathbf{A}^2\mathbf{B}\mathbf{u}_0 + \mathbf{A}^3\mathbf{x}_0, \tag{10.162}$$

$$\vdots \tag{10.163}$$

$$\mathbf{x}_n = \mathbf{B}\mathbf{u}_{n-1} + \mathbf{A}\mathbf{B}\mathbf{u}_{n-2} + \cdots + \mathbf{A}^{n-1}\mathbf{B}\mathbf{u}_0 + \mathbf{A}^n\mathbf{x}_0 \tag{10.164}$$

Controllability in Discrete Linear Autonomous Systems (contd.)

Equivalently, we can write (10.164) as

$$\mathbf{x}_n - \mathbf{A}^n\,\mathbf{x}_0 = \begin{bmatrix} \mathbf{B} & \mathbf{AB} & \cdots & \mathbf{A}^{n-1}\mathbf{B} \end{bmatrix} \begin{bmatrix} \mathbf{u}_{n-1} \\ \mathbf{u}_{n-2} \\ \vdots \\ \mathbf{u}_0 \end{bmatrix}. \tag{10.165}$$

Thus, if $\mathbf{x}_n - \mathbf{A}^n\,\mathbf{x}_0$ is in the range of the $n \times (n \cdot m)$ matrix

$$\mathcal{C} := \begin{bmatrix} \mathbf{B} & \mathbf{AB} & \cdots & \mathbf{A}^{n-1}\mathbf{B} \end{bmatrix}, \tag{10.166}$$

then there exists a control vector $(\mathbf{u}_0, \mathbf{u}_1, \ldots, \mathbf{u}_{n-1})$ that takes the system from $\mathbf{x}_0$ to $\mathbf{x}_n$ in a finite k. Matrix $\mathcal{C} = \mathcal{C}(\mathbf{A}, \mathbf{B})$ as defined in (10.166) is called the *controllability matrix* of the pair $(\mathbf{A}, \mathbf{B})$.

Controllability in Discrete Linear Autonomous Systems (contd.)

Since based on the definition of full state controllability, we need to traverse from any initial state $\mathbf{x}_0$ to any final state $\mathbf{x}_f$ in finite time, we want range($\mathcal{C}$) to span the n dimensional state space. Hence:

Theorem 10.9 (Full State Controllability of Discrete LTI Systems)

A linear time invariant discrete system

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k,$$

with n dimensional state is full state controllable if

$$\operatorname{rank}(\mathcal{C}) = \operatorname{rank}\left(\begin{bmatrix}\mathbf{B} & \mathbf{A}\mathbf{B} & \cdots & \mathbf{A}^{n-1}\mathbf{B}\end{bmatrix}\right) = n.$$

Slight Abuse of Notation!

Note that the $(\mathbf{A}, \mathbf{B})$ pair that defines a continuous system $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t)$ is *not* the same $(\mathbf{A}, \mathbf{B})$ pair that defines a *discretized* system $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k$. In fact, let $h = t_{k+1} - t_k$. Assuming $\mathbf{u}(t_k)$ is a constant vector in the interval $[t_k,\, t_{k+1})$, then

$$\mathbf{x}(t_{k+1}) = \mathbf{\Phi}(h)\mathbf{x}(t_k) + \int_0^h \mathbf{\Phi}(h-\tau)\, d\tau \mathbf{B}\mathbf{u}(t_k). \tag{10.167}$$

Alternatively, an approximation of $\mathbf{x}(t_{k+1})$ could achieved based on Taylor expansion of $\dot{\mathbf{x}}$. Since $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t)$ we have

$$\begin{aligned} \mathbf{x}(t_{k+1}) &= \mathbf{x}(t_k + h) = \mathbf{x}(t_k) + \dot{\mathbf{x}}(t_k)(t_{k+1} - t_k) + \mathcal{O}(h^2) && (10.168a)\\ &= \mathbf{x}(t_k) + h\left(\mathbf{A}\mathbf{x}(t_k) + \mathbf{B}\mathbf{u}(t_k)\right) + \mathcal{O}(h^2) && (10.168b)\\ &\approx (\mathbf{I} + h\mathbf{A})\mathbf{x}(t_k) + h\mathbf{B}\mathbf{u}(t_k). && (10.168c) \end{aligned}$$

We can also use better approximations, such as RK4, to calculate $(\mathbf{A}, \mathbf{B})$ pair for the *discretized* system.

Controllability in Continuous Linear Autonomous Systems

We are now ready to derive the condition for complete state controllability of continuous LTI systems.

Without loss of generality, we assume

- The final state $\mathbf{x}_f$ is the origin ($\mathbf{x}_f = \mathbf{0}$):

  Traversing from any $\mathbf{x}_0$ to $\mathbf{x}_f$ could be translated to traversing from any initial state $\mathbf{x}_0$ to origin (i.e. $\mathbf{x}_f = \mathbf{0}$) through a proper shift of the state-space via a change of variable (similar to our discussion on stability of any equilibrium point, could be translated to stability of the origin of the shifted system).

- The initial time $t_0 = 0$:

  Since the system is autonomous, the system dynamics is not an explicit function of time. Hence, to simplify, we can assume $t_0 = 0$.

Controllability in Continuous Linear Autonomous Systems (contd.)

For $\mathbf{x}(0) = \mathbf{x}_0$, based on the solution of the nonhomogeneous linear system we have

$$\mathbf{x}(t_f) = \mathbf{0} = e^{\mathbf{A}t_f}\mathbf{x}_0 + \int_0^{t_f} e^{\mathbf{A}(t_f-\tau)}\mathbf{B}\mathbf{u}(\tau)\, d\tau, \tag{10.169}$$

that implies

$$\mathbf{x}_0 = -\int_0^{t_f} e^{-\mathbf{A}\tau}\mathbf{B}\mathbf{u}(\tau)\, d\tau. \tag{10.170}$$

Using the results of the Putze algorithm we have

$$e^{\mathbf{A}t} = \sum_{k=0}^{n-1} \alpha_k(t)\mathbf{A}^k \tag{10.171}$$

Controllability in Continuous Linear Autonomous Systems (contd.)

Substituting (10.171) in (10.170) gives

$$\mathbf{x}_0 = -\int_0^{t_f} \left(\sum_{k=0}^{n-1} \alpha_k(\tau)\mathbf{A}^k \right) \mathbf{B}\mathbf{u}(\tau)\ d\tau = -\sum_{k=0}^{n-1} \mathbf{A}^k\mathbf{B} \int_0^{t_f} \alpha_k(\tau)\mathbf{u}(\tau)\ d\tau. \tag{10.172}$$

Let us define $\mathbf{v}_k$ as

$$\mathbf{v}_k := \int_0^{t_f} \alpha_k(\tau)\mathbf{u}(\tau)\ d\tau, \tag{10.173}$$

then we have

$$\mathbf{x}_0 = -\sum_{k=0}^{n-1} \mathbf{A}^k\mathbf{B}\mathbf{v}_k = \begin{bmatrix} \mathbf{B} & \mathbf{A}\mathbf{B} & \cdots & \mathbf{A}^{n-1}\mathbf{B} \end{bmatrix} \begin{bmatrix} \mathbf{v}_0 \\ \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_{n-1} \end{bmatrix} \tag{10.174}$$

Controllability in Continuous Linear Autonomous Systems (contd.)

Similar to discrete case, if the continuous system is full state controllable, there should exists a control vector $(\mathbf{v}_0, \dots, \mathbf{v}_{n-1})$ to take system from any initial state $\mathbf{x}_0$ to origin. Hence, let

$$\mathcal{C} := \begin{bmatrix} \mathbf{B} & \mathbf{AB} & \cdots & \mathbf{A}^{n-1}\mathbf{B} \end{bmatrix} \tag{10.175}$$

be *controllability matrix*, we can conclude:

Theorem 10.10 (Full State Controllability of Continuous LTI Systems)

A linear time invariant continuous system

$$\dot{\mathbf{x}} = \mathbf{Ax} + \mathbf{Bu},$$

with n dimensional state is full state controllable if

$$\mathrm{rank}(\mathcal{C}) = \mathrm{rank}\left(\begin{bmatrix} \mathbf{B} & \mathbf{AB} & \cdots & \mathbf{A}^{n-1}\mathbf{B} \end{bmatrix}\right) = n.$$

Examples

Examples are adopted from: Ogata, Katsuhiko. "Modern Control Engineering". fifth edition. 2010.

Consider the system given by

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u. \tag{10.176}$$

The corresponding controllability matrix is

$$\mathcal{C} = \begin{bmatrix} \mathbf{B} & \mathbf{AB} \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}. \tag{10.177}$$

Since $\mathcal{C}$ is singular (also we can check that $\text{rank}(\mathcal{C}) < 2$) the system is not completely state controllable.

Examples (contd.)

Examples are adopted from: Ogata, Katsuhiko. “Modern Control Engineering”. fifth edition. 2010.

Consider the system given by

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u$$

In this case the controllability matrix is

$$\mathcal{C} = \begin{bmatrix} \mathbf{B} & \mathbf{AB} \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 1 & 1 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix}.$$

$\mathcal{C}$ has two linearly independent columns, which implies $\text{rank}(\mathcal{C}) = 2$. Hence, the system is completely state controllable.

Examples (contd.)

Examples are adopted from: Ogata, Katsuhiko. “Modern Control Engineering”. fifth edition. 2010.

Consider the system given by

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 4 \\ 3 \end{bmatrix} u.$$

We can confirm the controllability of the system by the rank of the controllability matrix:

$$\mathcal{C} = \begin{bmatrix} \mathbf{B} & \mathbf{AB} & \mathbf{A}^2\mathbf{B} \end{bmatrix}$$

$$= \begin{bmatrix} \begin{bmatrix} 0 \\ 4 \\ 3 \end{bmatrix} & \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -2 \end{bmatrix} \begin{bmatrix} 0 \\ 4 \\ 3 \end{bmatrix} & \begin{bmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} 0 \\ 4 \\ 3 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 0 & 4 & -8 \\ 4 & -4 & 4 \\ 3 & -6 & 12 \end{bmatrix} \Longrightarrow \text{rank}(\mathcal{C}) = 3.$$

Examples (contd.)

Examples are adopted from: Ogata, Katsuhiko. "Modern Control Engineering". fifth edition. 2010.

Consider the system given by

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 4 & 2 \\ 0 & 0 \\ 3 & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}.$$

Note that this system is not controllable, since for the rank of the controllability matrix we have

$$\mathcal{C} = \begin{bmatrix} \mathbf{B} & \mathbf{AB} & \mathbf{A}^2\mathbf{B} \end{bmatrix} = \begin{bmatrix} 4 & 2 & -4 & -2 & 4 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 3 & 0 & -6 & 0 & 12 & 0 \end{bmatrix} \implies \mathrm{rank}(\mathcal{C}) = 2 < 3.$$

Matrix Similarity

Definition 10.8 (Matrix Similarity)

two $n \times n$ matrices $\mathbf{A}$ and $\mathbf{B}$ are *similar* if there exists an invertible $n \times n$ matrix $\mathbf{P}$ such that

$$\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}.$$

A transformation $\mathbf{A} \mapsto \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ is called a *similarity transformation* or *conjugation* of the matrix $\mathbf{A}$.

Similar matrices represent *the same linear map under* two different bases, with $\mathbf{P}$ being the change of basis matrix.

Spectrum of Similar Matrices

Theorem 10.11

Two similar matrices have the same eigenvalues with the same multiplicities.

Proof. Let $\mathbf{A}$ and $\mathbf{B}$ be similar, then exists $\mathbf{P}$ invertible such that $\mathbf{A} = \mathbf{P}^{-1}\mathbf{B}\mathbf{P}$ and

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v} \implies \mathbf{B}\mathbf{P}\mathbf{v} = \lambda\mathbf{P}\mathbf{v},$$

implying any eigenvalue of $\mathbf{A}$ is an eigenvalue of $\mathbf{B}$ with eigenvector $\mathbf{P}\mathbf{v}$. Let $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ be a basis for $\mathrm{null}(\mathbf{A} - \lambda\mathbf{I})$. Then, based on Proposition 10.6, $\{\mathbf{P}\mathbf{v}_1, \dots, \mathbf{P}\mathbf{v}_k\} \subset \mathrm{null}(\mathbf{B} - \lambda\mathbf{I})$ are linearly independent, implying $\mathrm{nullity}(\mathbf{B} - \lambda\mathbf{I}) \geq k$. Furthermore, we can show that $\mathbf{A}$ and $\mathbf{B}$ have the same characteristic equation. In fact

$$\det(\mathbf{A} - \lambda\mathbf{I}) = \det\left(\mathbf{P}^{-1}(\mathbf{B} - \lambda\mathbf{I})\mathbf{P}\right) = \det(\mathbf{P}^{-1})\det(\mathbf{B} - \lambda\mathbf{I})\det(\mathbf{P}) = \det(\mathbf{B} - \lambda\mathbf{I}), \qquad (10.178)$$

Repeating the argument for $\mathbf{B} = \mathbf{P}\mathbf{A}\mathbf{P}^{-1}$, we get any eigenvalue of $\mathbf{B}$ is an eigenvalue of $\mathbf{A}$ with the same algebraic and geometric multiplicities. $\square$

Similarity Transformation of LTI Systems

Consider the autonomous linear system

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}, \tag{10.179a}$$

$$\mathbf{y} = \mathbf{C}\mathbf{x} + \mathbf{D}\mathbf{u}. \tag{10.179b}$$

let $\mathbf{T}$ be an invertible matrix and consider the change of variable

$$\mathbf{z} = \mathbf{T}\mathbf{x}. \tag{10.180}$$

Accordingly we have

$$\dot{\mathbf{z}} = \mathbf{T}\dot{\mathbf{x}} = \mathbf{T}\mathbf{A}\mathbf{x} + \mathbf{T}\mathbf{B}\mathbf{u} = \mathbf{T}\mathbf{A}\mathbf{T}^{-1}\mathbf{z} + \mathbf{T}\mathbf{B}\mathbf{u}, \tag{10.181a}$$

$$\mathbf{y} = \mathbf{C}\mathbf{T}^{-1}\mathbf{z} + \mathbf{D}\mathbf{u}. \tag{10.181b}$$

Similarity Transformation of LTI Systems (contd.)

Let $\bar{\mathbf{x}}$ be an equilibrium point of $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$. Then, for $\bar{\mathbf{z}} = \mathbf{T}\bar{\mathbf{x}}$ we have

$$\mathbf{T}\mathbf{A}\mathbf{T}^{-1}\mathbf{T}\bar{\mathbf{z}} = \mathbf{T}\mathbf{A}\bar{\mathbf{x}} = \mathbf{0} \implies \bar{\mathbf{z}} \text{ is an equilibrium point of } \dot{\mathbf{z}} = \mathbf{T}\mathbf{A}\mathbf{T}^{-1}\mathbf{z}. \tag{10.182}$$

The controllability matrix for $\dot{\mathbf{z}} = \mathbf{T}\mathbf{A}\mathbf{T}^{-1}\mathbf{z} + \mathbf{T}\mathbf{B}\mathbf{u}$ is

$$\mathcal{C}(\mathbf{T}\mathbf{A}\mathbf{T}^{-1}, \mathbf{T}\mathbf{B}) = \begin{bmatrix} \mathbf{T}\mathbf{B} & \mathbf{T}\mathbf{A}\mathbf{T}^{-1}\mathbf{T}\mathbf{B} & \left(\mathbf{T}\mathbf{A}\mathbf{T}^{-1}\right)^2\mathbf{T}\mathbf{B} & \cdots & \left(\mathbf{T}\mathbf{A}\mathbf{T}^{-1}\right)^{n-1}\mathbf{T}\mathbf{B} \end{bmatrix} \tag{10.183}$$

$$= \begin{bmatrix} \mathbf{T}\mathbf{B} & \mathbf{T}\mathbf{A}\mathbf{B} & \mathbf{T}\mathbf{A}^2\mathbf{B} & \cdots & \mathbf{T}\mathbf{A}^{n-1}\mathbf{B} \end{bmatrix} \tag{10.184}$$

$$= \mathbf{T}\begin{bmatrix} \mathbf{B} & \mathbf{A}\mathbf{B} & \mathbf{A}^2\mathbf{B} & \cdots & \mathbf{A}^{n-1}\mathbf{B} \end{bmatrix} = \mathbf{T}\,\mathcal{C}(\mathbf{A}, \mathbf{B}). \tag{10.185}$$

Since $\mathbf{T}$ is invertible, based on Proposition 10.6, $\mathbf{T}$ preserves linear independence of the columns of $\mathcal{C}(\mathbf{A}, \mathbf{B})$. Hence, we can conclude

Lemma 10.12

Controllability of $(\mathbf{A},\ \mathbf{B})$ *implies controllability of* $(\mathbf{T}\mathbf{A}\mathbf{T}^{-1}, \mathbf{T}\mathbf{B})$ *and vice versa.*

Alternative Form of the Complete State Controllability

Assume $\mathbf{A}$ is diagonalizable. Therefore, exists $\mathbf{V}$, invertible, such that

$$\mathbf{A}\mathbf{V} = \mathbf{V}\boldsymbol{\Lambda}, \tag{10.186}$$

where $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ is the diagonal matrix composed of eigenvalues of $\mathbf{A}$ and $\mathbf{V} = [\mathbf{v}_1, \ldots, \mathbf{v}_n]$ is the matrix of the eigenvectors of $\mathbf{A}$. Now consider the change of variable

$$\mathbf{x} = \mathbf{V}\mathbf{z} \implies \dot{\mathbf{x}} = \mathbf{V}\dot{\mathbf{z}} \implies \mathbf{V}\dot{\mathbf{z}} = \mathbf{A}\mathbf{V}\mathbf{z} + \mathbf{B}\mathbf{u}. \tag{10.187}$$

Since $\mathbf{V}$ is invertible, we get

$$\dot{\mathbf{z}} = \mathbf{V}^{-1}\mathbf{A}\mathbf{V}\mathbf{z} + \mathbf{V}^{-1}\mathbf{B}\mathbf{u} = \boldsymbol{\Lambda}\mathbf{z} + \mathbf{P}\mathbf{u}, \tag{10.188}$$

where $\mathbf{P} := \mathbf{V}^{-1}\mathbf{B}$.

Alternative Form of the Complete State Controllability (contd.)

See that (10.188) is

$$
\begin{aligned}
\dot{z}_1 &= \lambda_1 z_1 + p_{11}u_1 + p_{12}u_2 + \cdots + p_{1m}u_m \\
&\vdots \\
\dot{z}_n &= \lambda_n z_n + p_{n1}u_1 + p_{n2}u_2 + \cdots + p_{nm}u_m.
\end{aligned}
$$

If the elements of any one row of the $n \times m$ matrix P are all zero, then the corresponding state variable cannot be controlled by any of the u_j. Hence:

Lemma 10.13

if the eigenvectors of $\mathbf{A}$ *are distinct, then the system is completely state controllable if and only if no row of* $\mathbf{V}^{-1}\mathbf{B}$ *has all zero elements.*

For the general case, we can use the Jordan form of $\mathbf{A}$. For details, see: Ogata, Katsuhiko. "Modern Control Engineering". fifth edition. 2010.

Stabilizability for LTI Systems

Consider an uncontrollable pair $(\mathbf{A}, \mathbf{B})$. There exists an invertible $\mathbf{T}$ such that

$$\begin{bmatrix} \dot{\mathbf{x}}_c \\ \dot{\mathbf{x}}_u \end{bmatrix} = \begin{bmatrix} \mathbf{A}_c & \mathbf{A}_{12} \\ \mathbf{0} & \mathbf{A}_u \end{bmatrix} \begin{bmatrix} \mathbf{x}_c \\ \mathbf{x}_u \end{bmatrix} + \begin{bmatrix} \mathbf{B}_c \\ \mathbf{0} \end{bmatrix} \mathbf{u} = \bar{\mathbf{A}} \begin{bmatrix} \mathbf{x}_c \\ \mathbf{x}_u \end{bmatrix} + \bar{\mathbf{B}}\mathbf{u}, \tag{10.189}$$

where

$$\bar{\mathbf{A}} = \mathbf{T}^{-1}\mathbf{A}\mathbf{T}, \quad \bar{\mathbf{B}} = \mathbf{T}^{-1}\mathbf{B}. \tag{10.190}$$

Expanding (10.189) gives

$$\dot{\mathbf{x}}_c = \mathbf{A}_c\mathbf{x}_c + \mathbf{A}_{12}\mathbf{x}_u + \mathbf{B}_c\mathbf{u}, \tag{10.191a}$$

$$\dot{\mathbf{x}}_u = \mathbf{A}_u\mathbf{x}_u, \tag{10.191b}$$

Indicating that:

An uncontrollable system has a subsystem that is physically disconnected from the input.

Stabilizability for LTI Systems (contd.)

If $\mathbf{A}_u$ is Hurwitz, then $\lim_{t\to\infty} \|\mathbf{x}_u\| = 0$, and the control design problem simplifies to finding $\mathbf{u}$ so that $\lim_{t\to\infty} \|\mathbf{x}_c\| = 0$. Accordingly we have:

Definition 10.9 (Stabilizable LTI System)

The pair $(\mathbf{A}, \mathbf{B})$ is stabilizable if either $\mathbf{A}_u$ does not exist or $\mathbf{A}_u$ is Hurwitz.

Definition 10.10 (Stabilizable LTI System – Alternative Definition)

The pair $(\mathbf{A}, \mathbf{B})$ is stabilizable if there exists a state feedback gain matrix $\mathbf{K}$ for which all the eigenvalues of $\mathbf{A} - \mathbf{BK}$ have strictly negative real part.

Stabilizability is a weaker condition than controllability.

A Method to Find $\mathbf{T}$

Let $r < n$ be the rank of $\mathcal{C}(\mathbf{A}, \mathbf{B})$. It means, $\mathcal{C}$ has r linearly independent columns.

- Find $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ as a basis for range($\mathcal{C}$) and form $\mathbf{T}_c = \begin{bmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_r \end{bmatrix}$.
- Find $\{\mathbf{n}_{r+1}, \ldots, \mathbf{n}_n\}$ as a basis for null($\mathcal{C}$) and form $\mathbf{T}_n = \begin{bmatrix} \mathbf{n}_{r+1} & \cdots & \mathbf{n}_n \end{bmatrix}$.
- Form $\mathbf{T} = \begin{bmatrix} \mathbf{T}_c & \mathbf{T}_u \end{bmatrix}$, $\bar{\mathbf{A}} = \mathbf{T}^{-1}\mathbf{A}\mathbf{T}$ and $\bar{\mathbf{B}} = \mathbf{T}^{-1}\mathbf{B}$.

```
Ctr = ctrb(A,B);
display(rank(Ctr));
Tc = orth(Ctr);
Tu = null(Ctr);
T = [Tc, Tu];
A_ = T\(A*T);
B_ = T\B;
```

Output Controllability

In the practical design of a control system, we may want to control the output rather than the state. Hence, it is desirable to define separately complete output controllability. Consider

Definition 10.11 (Complete Output Controllability)

The system the system described by

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u},$$
$$\mathbf{y} = \mathbf{C}\mathbf{x} + \mathbf{D}\mathbf{u},$$

is said to be completely output controllable if it is possible to construct an unconstrained control vector $\mathbf{u}(t)$ that will transfer any given initial output $\mathbf{y}(t_0)$ to any final output $\mathbf{y}(t_f)$ in a finite time interval $t_0 \leq t \leq t_f$.

⚠ Complete state controllability is neither necessary nor sufficient for controlling the output of the system.

Output Controllability (contd.)

Theorem 10.14 (Complete Output Controllability)

The system

$$\begin{aligned}\dot{\mathbf{x}} &= \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u},\\ \mathbf{y} &= \mathbf{C}\mathbf{x},\end{aligned}$$

with n dimensional state, m dimensional input and r dimensional output, is completely output controllable if and only if $r \times (nm)$ matrix

$$\mathcal{P} = \begin{bmatrix}\mathbf{CB} & \mathbf{CAB} & \mathbf{CA}^2\mathbf{B} & \cdots & \mathbf{CA}^{n-1}\mathbf{B}\end{bmatrix}$$

is of rank r.

Popov-Belevitch-Hautus Test for Controllability and Stablizability

The following lemmas provide an alternative test for controllability and stablizability of LTI system

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}.$$

Lemma 10.15 (Hautus Lemma for Controllability)

given a square matrix $\mathbf{A} \in \mathbb{M}_n(\mathbb{R})$ *and a matrix* $\mathbf{B} \in \mathbb{M}_{n\times m}(\mathbb{R})$, *the following are equivalent:*

1. *The pair* $(\mathbf{A}, \mathbf{B})$ *is controllable.*
2. *For all* $\lambda \in \mathbb{C}$, *it holds that* $\operatorname{rank}\left(\begin{bmatrix} \lambda\mathbf{I} - \mathbf{A} & \mathbf{B} \end{bmatrix}\right) = n.$
3. *For all* $\lambda \in \mathbb{C}$ *that are eigenvalues of* $\mathbf{A}$, *it holds that* $\operatorname{rank}\left(\begin{bmatrix} \lambda\mathbf{I} - \mathbf{A} & \mathbf{B} \end{bmatrix}\right) = n.$

Popov-Belevitch-Hautus Test for Controllability and Stablizability (contd.)

Lemma 10.16 (Hautus Lemma for Stabilizability)

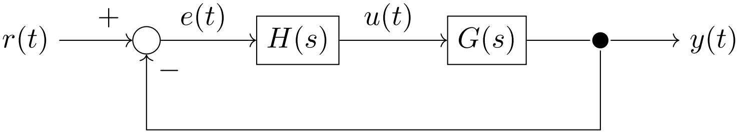
given a square matrix $\mathbf{A} \in \mathbb{M}_n(\mathbb{R})$ *and a matrix* $\mathbf{B} \in \mathbb{M}_{n\times m}(\mathbb{R})$, *the following are equivalent:*

1. *The pair* $(\mathbf{A}, \mathbf{B})$ *is stabilizable.*
2. *For all* $\lambda \in \mathbb{C}$ *that are eigenvalues of* $\mathbf{A}$ *and for which* $\Re(\lambda) \geq 0$, *it holds that* $\mathrm{rank}[\lambda\mathbf{I} - \mathbf{A}, \mathbf{B}] = n$.

Pole Placement

Pole Placement Objective

Classical Control Approach of Compensator Design: Consider a closed loop SISO system



$$\frac{Y(s)}{R(s)} = \frac{H(s)G(s)}{1 + H(s)G(s)} = \frac{\sum_{j=1}^{m}(z - z_j)}{\sum_{i=1}^{n}(p - p_i)}. \tag{10.193}$$

In the classical approach to controller design of SISO systems, we design $H(s)$ such that the dominant closed-loop poles have a desired damping ratio ξ and a desired undamped natural frequency ω_n (assuming the effects of nondominant poles are negligible). Doing so, the order of the system may be raised by 1 or 2, unless pole–zero cancellation takes place.

Pole Placement Objective (contd.)

Regulator Design via Pole Placement: Consider an LTI system

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}, \tag{10.194}$$
$$\mathbf{y} = \mathbf{C}\mathbf{x} + \mathbf{D}\mathbf{u}, \tag{10.195}$$

where $\mathbf{x}$ is n-dimensional state, $\mathbf{u}$ is m-dimensional input and $\mathbf{y}$ is r-dimensional output vectors. Our objective is: for a *state feedback* control function

$$\mathbf{u} = -\mathbf{K}\mathbf{x}, \tag{10.196}$$

chose an appropriate *gain matrix* $\mathbf{K}$ such that the eigenvalues of the closed loop system

$$\dot{\mathbf{x}} = (\mathbf{A} - \mathbf{B}\mathbf{K})\mathbf{x}, \tag{10.197}$$

coincide with a set of given *arbitrary* poles. If we choose the *arbitrary* eigenvalues such that $\mathbf{A} - \mathbf{B}\mathbf{K}$ is Hurwitz, then the output will drive to zero from any initial state.

Pole Placement Objective (contd.)

The closed loop system

$$\dot{\mathbf{x}} = (\mathbf{A} - \mathbf{BK})\mathbf{x}, \tag{10.198}$$

has no input.

Its objective is to maintain the zero output.

Due to the disturbances, the output will deviate from zero. The nonzero output will be returned to the *zero reference input* owing to $\mathbf{A} - \mathbf{BK}$ being Hurwits.

Regulator system

Such a system where the reference input is a always a constant (which could be zero) is called a regulator system.

Pole Placement Objective (contd.)

- Pole placement approach specifies all closed loop poles (in contrast to specifying only dominant closed loop poles in the conventional design approach).
- placing all closed-loop poles requires successful measurements of all state variables:
  - We may not have access to all the states of the system
  - We may need *observers* to estimate states (observers are discussed later in the course)
- We need controllability to be able to place the closed loop poles at arbitrarily chosen locations.
- It is suitable for both SISO and MIMO systems,

$(\mathbf{A},\mathbf{B})$ Not Controllable $\implies \nexists$ a Solution to Pole Placement

If $(\mathbf{A},\mathbf{B})$ is not controllable, then $\operatorname{rank}\left(\mathcal{C}(\mathbf{A},\mathbf{B})\right) = q < n$ and $\exists$ $\mathbf{T}$, invertible, such that

$$\mathbf{T}^{-1}\mathbf{A}\mathbf{T} = \bar{\mathbf{A}} = \begin{bmatrix} \mathbf{A}_c & \mathbf{A}_{12} \\ \mathbf{0} & \mathbf{A}_u \end{bmatrix}, \quad \mathbf{T}^{-1}\mathbf{B} = \bar{\mathbf{B}} = \begin{bmatrix} \mathbf{B}_c \\ \mathbf{0} \end{bmatrix}. \tag{10.199}$$

Let $\bar{\mathbf{K}} := \mathbf{K}\mathbf{T} = \begin{bmatrix} \mathbf{K}_1 & \mathbf{K}_2 \end{bmatrix}$. The characteristic equation of the closed loop system is

$$|\mathbf{A} - \mathbf{B}\mathbf{K} - \lambda\mathbf{I}| = \left|\mathbf{T}^{-1}\left(\mathbf{A} - \mathbf{B}\mathbf{K} - \lambda\mathbf{I}\right)\mathbf{T}\right| = \left|\mathbf{T}^{-1}\mathbf{A}\mathbf{T} - \mathbf{T}^{-1}\mathbf{B}\mathbf{K}\mathbf{T} - \lambda\mathbf{I}\right| \tag{10.200}$$

$$= \left|\bar{\mathbf{A}} - \bar{\mathbf{B}}\bar{\mathbf{K}} - \lambda\mathbf{I}\right| = \left|\begin{bmatrix} \mathbf{A}_c - \mathbf{B}_c\mathbf{K}_1 - \lambda\mathbf{I}_q & \mathbf{A}_{12} - \mathbf{B}_c\mathbf{K}_2 \\ \mathbf{0} & \mathbf{A}_u - \lambda\mathbf{I}_{n-q} \end{bmatrix}\right| \tag{10.201}$$

$$= \left|\mathbf{A}_c - \mathbf{B}_c\mathbf{K}_1 - \lambda\mathbf{I}_q\right| \cdot \left|\mathbf{A}_u - \lambda\mathbf{I}_{n-q}\right| = 0, \tag{10.202}$$

Since eigenvalues of $\mathbf{A}_u$ does not depend on $\mathbf{K}$, we conclude, if $\operatorname{rank}(\mathbf{A},\mathbf{B}) < n$, there are eigenvalues of matrix $\mathbf{A}$ that cannot be arbitrarily placed.

Controllability and Pole Placement

Theorem 10.17

Let $\mathbf{A} \in \mathbb{R}^{n\times n}$ *and* $\mathbf{B} \in \mathbb{R}^{n\times m}$*. The pair* $(\mathbf{A}, \mathbf{B})$ *is controllable if and only if for every symmetric set* Λ *of* n *complex numbers, there exists* $\mathbf{K}$ *such that*

$$\sigma(\mathbf{A} - \mathbf{BK}) = \Lambda.$$

For proof see: W. Murray Wonham, "Linear Multivariable Control: A Geometric Approach". second edition. 1979

Corollary 10.17.1

Necessary and sufficient condition for arbitrary pole placement is that the system be completely state controllable.

Pole Placement For Single Input Systems

What Makes Single Input Systems Special!

Consider a **single input** linear time invariant system with n dimensional state

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}u, \tag{10.203}$$

$$\mathbf{y} = \mathbf{C}\mathbf{x} + \mathbf{D}u, \tag{10.204}$$

The controllability matrix of the $(\mathbf{A}, \mathbf{B})$ pair is

$$\mathcal{C}(\mathbf{A}, \mathbf{B}) = \begin{bmatrix} \mathbf{B} & \mathbf{A}\mathbf{B} & \mathbf{A}^2\mathbf{B} & \cdots & \mathbf{A}^{n-1}\mathbf{B} \end{bmatrix}. \tag{10.205}$$

In general, the controllability matrix $\mathcal{C}(\mathbf{A}, \mathbf{B})$ is a matrix of size $n \times (nm)$, where m is the dimension of input. Hence,

If the input is one dimensional ($m = 1$), then $\mathcal{C}$ is a square matrix of size n.

If a single input system is controllable, then $\text{rank}(\mathcal{C}) = n$ (it is a full-rank matrix). Hence, there exists $\mathcal{C}^{-1}$ such that

$$\mathcal{C}\mathcal{C}^{-1} = \mathcal{C}^{-1}\mathcal{C} = \mathbf{I}.$$

Determination of $\mathbf{K}$ Using Direct Substitution Method

If the system is of low order ($n < 3$), direct substitution $\mathbf{K}$ into

$$\det(\mathbf{A} - \mathbf{BK} - \lambda\mathbf{I}) \tag{10.206}$$

may be simple. For example, if $n = 3$, we have

$$\mathbf{K} = \begin{bmatrix} k_1 & k_2 & k_3 \end{bmatrix}. \tag{10.207}$$

Therefore, we can find k_1 to k_3 by equating the coefficients of (10.206) with the polynomial obtained from the set of desired eigenvalues μ_1 to μ_3. That is

$$\det(\mathbf{A} - \mathbf{BK} - \lambda\mathbf{I}) = (\mu_1 - \lambda)(\mu_2 - \lambda)(\mu_3 - \lambda). \tag{10.208}$$

This approach is convenient if $n < 4$ otherwise it may become very tedious. If the system is not completely controllable, matrix $\mathbf{K}$ cannot be determined (No solution exists).

Example

adopted from: Ogata, Katsuhiko. "Modern Control Engineering". fifth edition. 2010.

Consider the system $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}$ with

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & -5 & -6 \end{bmatrix}, \quad \text{and} \quad \mathbf{B} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}. \tag{10.209}$$

We want to find the gain matrix $\mathbf{K}$ for the state feedback control $\mathbf{u} = -\mathbf{K}\mathbf{x}$, such that the eigenvalues of the closed loop system match the desired eigenvalues

$$\mu_1 = -2 + 4i, \quad \mu_2 = -2 - 4i, \quad \mu_3 = -10. \tag{10.210}$$

We make such a choice because we know from experience that such a set of closed-loop poles will result in a reasonable or acceptable transient response.

Example (contd.)

adopted from: Ogata, Katsuhiko. "Modern Control Engineering". fifth edition. 2010.

Let $\mathbf{K} = \begin{bmatrix} k_1 & k_2 & k_3 \end{bmatrix}$, for the characteristic polynomial of the closed loop system we have

$$|\mathbf{A} - \mathbf{BK} - \lambda \mathbf{I}| = \left| \begin{bmatrix} -\lambda & 1 & 0 \\ 0 & -\lambda & 1 \\ -1 - k_1 & -5 - k_2 & -6 - k_3 - \lambda \end{bmatrix} \right| \tag{10.211}$$

$$= -\lambda^3 - (6 + k_3)\lambda^2 - (5 + k_2)\lambda - 1 - k_1. \tag{10.212}$$

For the desired polynomial we have

$$(\mu_1 - \lambda)(\mu_2 - \lambda)(\mu_3 - \lambda) = (-2 + 4i - \lambda)(-2 - 4i - \lambda)(-10 - \lambda) = -\lambda^3 - 14\lambda^2 - 60\lambda - 200. \tag{10.213}$$

Example (contd.)

adopted from: Ogata, Katsuhiko. “Modern Control Engineering”. fifth edition. 2010.

Matching the coefficients of (10.212) with (10.213) gives

$$6 + k_3 = 14, \implies k_3 = 8, \tag{10.214}$$

$$5 + k_2 = 60, \implies k_2 = 55, \tag{10.215}$$

$$1 + k_1 = 200 \implies k_1 = 199. \tag{10.216}$$

Hence, the gain matrix $\mathbf{K}$ is

$$\mathbf{K} = \begin{bmatrix} 199 & 55 & 8 \end{bmatrix}. \tag{10.217}$$

Example (contd.)

adopted from: Ogata, Katsuhiko. “Modern Control Engineering”. fifth edition. 2010.

```
syms k1 k2 k3 'real'; syms s

A = [0, 1, 0; 0, 0, 1; -1, -5, -6];
B = [0; 0; 1];
K = [k1, k2, k3];
mu = [-2-4i, -2+4i, -10];

charpoly = det(A - B*K - s*eye(3));
despoly = expand(prod(mu - s));

sol = solve(coeffs(charpoly,s) == coeffs(despoly,s), [k1, k2, k3]);
K_num = [sol.k1, sol.k2, sol.k3];
disp(eig(A - B*K_num));
```

Controllable Canonical Form

For the transfer function

$$\frac{Y(s)}{U(s)} = \frac{b_0 s^n + b_1 s^{n-1} + \cdots + b_{n-1}s + b_n}{s^n + a_1 s^{n-1} + \cdots + a_{n-1}s + a_n}, \tag{10.218}$$

the following state-space representation is called a *controllable canonical form*:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \vdots \\ \dot{x}_{n-1} \\ \dot{x}_n \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & \cdots & -a_1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} u \tag{10.219}$$

The output equation is given by

$$y = \begin{bmatrix} b_n - a_n b_0 & b_{n-1} - a_{n-1} b_0 & \cdots & b_1 - a_1 b_0 \end{bmatrix} \mathbf{x} + b_0 u \tag{10.220}$$

Controllable Canonical Form (contd.)

Given a **single input** and **controllable** LTI system with n dimensional state

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}u, \qquad \mathbf{y} = \mathbf{C}\mathbf{x} + \mathbf{D}u, \tag{10.221}$$

we can convert it to controllable canonical form via the change of variable

$$\mathbf{x} = \mathbf{T}\mathbf{z}, \qquad \mathbf{T} := \mathcal{C}(\mathbf{A}, \mathbf{B})\,\mathbf{W}, \tag{10.222}$$

for $\mathbf{W}$ defined as

$$\mathbf{W} := \begin{bmatrix} c_{n-1} & c_{n-2} & \cdots & c_1 & 1 \\ c_{n-2} & c_{n-3} & \cdots & 1 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ c_1 & 1 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \end{bmatrix}, \tag{10.223}$$

where the c_i's are coefficients of the characteristic polynomial

$$\det(\lambda \mathbf{I} - \mathbf{A}) = \lambda^n + c_1 \lambda^{n-1} + \cdots + c_{n-1}\lambda + c_n. \tag{10.224}$$

Controllable Canonical Form (contd.)

Based on the change of variable

$$\mathbf{x} = \mathbf{T}\mathbf{z} = \mathcal{C}\mathbf{W}\mathbf{z}, \tag{10.225}$$

Since $\text{rank}(\mathbf{W}) = \text{rank}(\mathcal{C}) = n$, then $\mathbf{T}^{-1} = \mathbf{W}^{-1}\mathcal{C}^{-1}$ exists and accordingly:

$$\mathbf{x} = \mathbf{T}\mathbf{z} = \mathcal{C}\mathbf{W}\mathbf{z}, \qquad \Longrightarrow \qquad \mathbf{z} = \mathbf{T}^{-1}\mathbf{x}, \tag{10.226}$$

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}u = \mathbf{A}\mathbf{T}\mathbf{z} + \mathbf{B}u, \qquad \Longrightarrow \qquad \dot{\mathbf{z}} = \mathbf{T}^{-1}\mathbf{A}\mathbf{T}\mathbf{z} + \mathbf{T}^{-1}\mathbf{B}u = \hat{\mathbf{A}}\mathbf{x} + \hat{\mathbf{B}}u, \tag{10.227}$$

$$\mathbf{y} = \mathbf{C}\mathbf{x}, \qquad \Longrightarrow \qquad \mathbf{y} = \mathbf{C}\mathbf{T}\mathbf{z}, \tag{10.228}$$

where

$$\hat{\mathbf{A}} = \mathbf{T}^{-1}\mathbf{A}\mathbf{T} = \mathbf{W}^{-1}\ \mathcal{C}^{-1}\ \mathbf{A}\ \mathcal{C}\ \mathbf{W}, \tag{10.229}$$

$$\hat{\mathbf{B}} = \mathbf{T}^{-1}\mathbf{B} = \mathbf{W}^{-1}\ \mathcal{C}^{-1}\ \mathbf{B}. \tag{10.230}$$

For derivation details see: Ogata, Katsuhiko. “Modern Control Engineering”. fifth edition. 2010.

Example

Consider the system $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}u$, where

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ -3 & -4 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} u. \tag{10.231}$$

The controllability matrix for the $(\mathbf{A}, \mathbf{B})$ pair is

$$\mathcal{C}(\mathbf{A}, \mathbf{B}) = \begin{bmatrix} \mathbf{B} & \mathbf{AB} & \mathbf{A}^2\mathbf{B} \end{bmatrix} = \begin{bmatrix} 0 & 2 & -2 \\ 1 & 1 & 1 \\ 2 & -4 & -10 \end{bmatrix}, \tag{10.232}$$

which has rank of 3. Thus the system is controllable.

Example (contd.)

In order to find the transformation $\mathbf{T} = \mathcal{C}\mathbf{W}$, we need to find the coefficients of the characteristic polynomial of $\mathbf{A}$. we have

$$\det(\lambda\mathbf{I} - \mathbf{A}) = \left| \begin{bmatrix} \lambda - 1 & 0 & -1 \\ 0 & \lambda - 1 & 0 \\ 3 & 4 & \lambda \end{bmatrix} \right| = \lambda^3 - 2\lambda^2 + 4\lambda - 3, \tag{10.233}$$

which implies

$$c_0 = 1, \quad c_1 = -2, \quad c_2 = 4, \quad c_3 = -3. \tag{10.234}$$

Thus, we can form $\mathbf{W}$ as

$$\mathbf{W} := \begin{bmatrix} c_2 & c_1 & 1 \\ c_1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 4 & -2 & 1 \\ -2 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \tag{10.235}$$

Example (contd.)

Accordingly, setting $\mathbf{T} = \mathcal{C}\mathbf{W}$, we get

$$\hat{\mathbf{A}} = \mathbf{T}^{-1}\mathbf{A}\mathbf{T} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 3 & -4 & 2 \end{bmatrix}, \qquad \hat{\mathbf{B}} = \mathbf{T}^{-1}\mathbf{B} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \tag{10.236}$$

And the controllable canonical form of the system is

$$\begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \\ \dot{z}_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 3 & -4 & 2 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} u, \quad \text{and} \quad \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} = \begin{bmatrix} -1/6 & 1/9 & -1/18 \\ 0 & 1/3 & -1/6 \\ 1/2 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}. \tag{10.237}$$

Example (contd.)

```
A = [1, 0, 1; 0, 1, 0; -3, -4, 0];
B = [0; 1; 2];
n = length(A);
Ctr = ctrb(A,B);
if rank(Ctr) < n
    error("(A,B) pair is not controllable.");
end
c = charpoly(A);
W = zeros(n);
for row =1:n
    W(row,1:n+1-row) = c(end-row:-1:1);
end
T = Ctr*W;
A_ = T\(A*T);
B_ = T\B;
```

Determination of $\mathbf{K}$ Using Controllable Canonical Form

Consider the pole placement problem for LTI system

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}u, \tag{10.238}$$

with n dimensional state and one dimensional input. Let

$$\Lambda = \{\mu_1, \ldots, \mu_n\} \subset \mathbb{C} \tag{10.239}$$

be a self conjugate set of desired closed loop poles. Our goal is to determine the gain matrix $\mathbf{K}$ for the full state control function $u = -\mathbf{K}\mathbf{x}$, such that

$$\sigma(\mathbf{A} - \mathbf{B}\mathbf{K}) = \Lambda. \tag{10.240}$$

A *self-conjugate set*, informally, is a set of complex numbers that is closed under conjugation, meaning, if $z = a + bi$ in the set, then its conjugate $\bar{z} = a - bi$ is also in the set. A self-conjugate set is symmetric with respect to the real axis in the complex plane.

Determination of **K** Using Controllable Canonical Form (contd.)

Step 0. Check the controllability of the system. If the system is full state controllable, then continue with the following steps.
Step 1. From the characteristic polynomial of **A** and find c_i:

$$\det(\lambda \mathbf{I} - \mathbf{A}) = \lambda^n + c_1 \lambda^{n-1} + \cdots + c_{n-1}\lambda + c_n. \tag{10.241}$$

Step 2. Determine **T** that transforms the system into the controllable canonical form. If the system is already in the controllable canonical form, then $\mathbf{T} = \mathbf{I}$.
Step 3. Using the set Λ, find the coefficients of the desired characteristic polynomial α_i:

$$(\lambda - \mu_1)(\lambda - \mu_2)\cdots 1(\lambda - \mu_n) = \lambda^n + \alpha_1 \lambda^{n-1} + \cdots + \alpha_{n-1}\lambda + \alpha_n. \tag{10.242}$$

Step 4. The gain matrix **K** can be determined from

$$\mathbf{K} = \begin{bmatrix} \alpha_n - c_n & \alpha_{n-1} - c_{n-1} & \cdots & \alpha_1 - c_1 \end{bmatrix} \mathbf{T}^{-1}. \tag{10.243}$$

Example

Let us consider the same system $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}$ as defined in (10.209) with the same Λ:

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & -5 & -6 \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \qquad \Lambda = \{-2+4i, -2-4i, -10\}. \tag{10.244}$$

From the previous example, we know the system is controllable. Also, $(\mathbf{A}, \mathbf{B})$ is already in the Controllable form, thus $\mathbf{T} = \mathbf{I}$. All left to do is to find c_i and α_i. We have:

$$\det(\lambda \mathbf{I} - \mathbf{A}) = c_0\lambda^3 + c_1\lambda^2 + c_2\lambda + c_3 = \lambda^3 + 6\lambda^2 + 5\lambda + 1, \tag{10.245}$$

$$\prod_{\mu \in \Lambda} (\lambda - \mu) = \alpha_0\lambda^3 + \alpha_1\lambda^2 + \alpha_2\lambda + \alpha_3 = \lambda^3 + 14\lambda^2 + 60\lambda + 200. \tag{10.246}$$

Accordingly, the gain matrix $\mathbf{K}$ is

$$\mathbf{K} = \begin{bmatrix} \alpha_3 - c_3 & \alpha_2 - c_2 & \alpha_1 - c_1 \end{bmatrix} \mathbf{T}^{-1} = \begin{bmatrix} 199 & 55 & 8 \end{bmatrix} \mathbf{I} = \begin{bmatrix} 199 & 55 & 8 \end{bmatrix}. \tag{10.247}$$

Example (contd.)

```
A = [0, 1, 0; 0, 0, 1; -1, -5, -6];
B = [0; 0; 1];
mu = [-2-4i, -2+4i, -10];
c = charpoly(A);
alpha = poly(mu);

n = length(A);
Ctr = ctrb(A,B);
if rank(Ctr) < n; return; end

W = zeros(n);
for row =1:n
    W(row,1:n+1-row) = c(end-row:-1:1)
end
T = Ctr*W;
K = (alpha(end:-1:2) - c(end:-1:2))/T;
disp(eig(A - B*K));
```

Derivation Details

In what follows, we discuss detailed derivation of (10.243). Let $\mathbf{T} = \mathcal{C}\mathbf{W}$, and introduce the change of variable

$$\mathbf{x} = \mathbf{T}\mathbf{z} \qquad \Longrightarrow \qquad \dot{\mathbf{z}} = \mathbf{T}^{-1}\mathbf{A}\mathbf{T}\mathbf{z} + \mathbf{T}^{-1}\mathbf{B}u = \hat{\mathbf{A}}\mathbf{z} + \hat{\mathbf{B}}u, \tag{10.248}$$

where

$$\hat{\mathbf{A}} = \mathbf{T}^{-1}\mathbf{A}\mathbf{T} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -c_n & -c_{n-1} & -c_{n-2} & \cdots & -c_1 \end{bmatrix}, \quad \text{and} \quad \hat{\mathbf{B}} = \mathbf{T}^{-1}\mathbf{B} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}. \tag{10.249}$$

Let

$$\hat{\mathbf{K}} := \mathbf{K}\mathbf{T} = \begin{bmatrix} \delta_n & \delta_{n-1} & \cdots & \delta_1 \end{bmatrix}. \tag{10.250}$$

Derivation Details (contd.)

Using the gain matrix $\hat{\mathbf{K}} = \mathbf{K}\mathbf{T}$, the characteristic equation of the closed loop system

$$\dot{\mathbf{z}} = \mathbf{T}^{-1}\mathbf{A}\mathbf{T}\mathbf{z} - \mathbf{T}^{-1}\mathbf{B}\mathbf{K}\mathbf{T}\mathbf{z} = \mathbf{T}^{-1}\left(\mathbf{A} - \mathbf{B}\mathbf{K}\right)\mathbf{T}\mathbf{z}, \tag{10.251}$$

is

$$\left|\lambda\mathbf{I} - \mathbf{T}^{-1}\left(\mathbf{A} - \mathbf{B}\mathbf{K}\right)\mathbf{T}\right| = \left|\mathbf{T}^{-1}\right|\left|\lambda\mathbf{I} - \mathbf{A} + \mathbf{B}\mathbf{K}\right|\left|\mathbf{T}\right| = \left|\lambda\mathbf{I} - \mathbf{A} + \mathbf{B}\mathbf{K}\right| = 0, \tag{10.252}$$

which is equivalent to the characteristic equation of the original system.

Derivation Details (contd.)

For the characteristic polynomial of the closed loop system we have

$$\left|\lambda\mathbf{I} - \mathbf{T}^{-1}\left(\mathbf{A} - \mathbf{B}\mathbf{K}\right)\mathbf{T}\right| = \left|\lambda\mathbf{I} - \mathbf{T}^{-1}\mathbf{A}\mathbf{T} - \mathbf{T}^{-1}\mathbf{B}\mathbf{K}\mathbf{T}\right| = \left|\lambda\mathbf{I} - \hat{\mathbf{A}} - \hat{\mathbf{B}}\hat{\mathbf{K}}\right| \tag{10.253}$$

$$= \left|\begin{bmatrix} \lambda & -1 & \cdots & 0 & 0 \\ 0 & \lambda & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda & -1 \\ c_n + \delta_n & c_{n-1} + \delta_{n-1} & \cdots & \lambda + c_1 + \delta_1 & \lambda \end{bmatrix}\right| \tag{10.254}$$

$$= \lambda^n + (c_1 + \delta_1)\lambda^{n-1} + \cdots + (c_{n-1} + \delta_{n-1})\lambda + (c_n + \delta_n) \tag{10.255}$$

Using the set Λ, we have the coefficients of the desired characteristic polynomial as

$$\prod_{\mu \in \Lambda} (\lambda - \mu) = \lambda^n + \alpha_1 \lambda^{n-1} + \cdots + \alpha_{n-1}\lambda + \alpha_n. \tag{10.256}$$

Derivation Details (contd.)

By equating the coefficients of the two polynomials we get

$$\begin{aligned} c_1 + \delta_1 &= \alpha_1 & \implies && \delta_1 &= \alpha_1 - c_1, \\ c_2 + \delta_2 &= \alpha_2 & \implies && \delta_2 &= \alpha_2 - c_2, \\ &\vdots &&& &\vdots \\ c_n + \delta_n &= \alpha_n & \implies && \delta_n &= \alpha_n - c_n. \end{aligned}$$

Since $\hat{\mathbf{K}}$, based on the definition is

$$\hat{\mathbf{K}} = \mathbf{K}\mathbf{T} = \begin{bmatrix} \delta_n & \delta_{n-1} & \cdots & \delta_1 \end{bmatrix} = \begin{bmatrix} \alpha_n - c_n & \alpha_{n-1} - c_{n-1} & \cdots & \alpha_1 - c_1 \end{bmatrix}. \tag{10.257}$$

Accordingly, the gain matrix $\mathbf{K}$ is

$$\mathbf{K} = \hat{\mathbf{K}}\mathbf{T}^{-1} = \begin{bmatrix} \alpha_n - c_n & \alpha_{n-1} - c_{n-1} & \cdots & \alpha_1 - c_1 \end{bmatrix} \mathbf{T}^{-1}. \tag{10.258}$$

Determination of $\mathbf{K}$ Using Ackermann's Formula

Consider a **single input** and **controllable** LTI system with n dimensional state

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}u. \tag{10.259}$$

Given a self conjugate set $\Lambda = \{\mu_1, \ldots, \mu_n\} \subset \mathbb{C}$ as the set of desired closed loop poles, let $\phi(\lambda)$ denote the desired polynomial:

$$\phi(\lambda) := \prod_{\mu \in \Lambda} (\lambda - \mu) = \lambda^n + \alpha_1 \lambda^{n-1} + \cdots + \alpha_{n-1}\lambda + \alpha_n. \tag{10.260}$$

Ackermann's formula

The gain matrix $\mathbf{K}$ for the control $u = -\mathbf{K}\mathbf{x}$, as the solution to the pole placement problem, is defined as

$$\mathbf{K} = \begin{bmatrix} 0 & \cdots & 0 & 1 \end{bmatrix} \mathcal{C}^{-1} \phi(\mathbf{A}) \tag{10.261}$$

where $\mathcal{C} = \mathcal{C}(\mathbf{A}, \mathbf{B})$ is the controllability matrix and $\phi(\mathbf{A})$ is the desired characteristic polynomial evaluated at the matrix $\mathbf{A}$.

Example

Let us find the gain matrix $\mathbf{K}$ for the same problem as before, where, as defined in (10.209),

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & -5 & -6 \end{bmatrix}, \qquad \mathbf{B} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \qquad \Lambda = \{-2+4i, -2-4i, -10\}. \tag{10.262}$$

The controllability matrix for the system is

$$\mathcal{C} = \begin{bmatrix} \mathbf{B} & \mathbf{AB} & \mathbf{A}^2\mathbf{B} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & -6 \\ 1 & -6 & 31 \end{bmatrix} \implies \mathcal{C}^{-1} = \begin{bmatrix} 5 & 6 & 1 \\ 6 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \tag{10.263}$$

For the desired characteristic polynomial we have

$$\phi(\lambda) = \prod_{\mu \in \Lambda} (\lambda - \mu) = \alpha_0 \lambda^3 + \alpha_1 \lambda^2 + \alpha_2 \lambda + \alpha_3 = \lambda^3 + 14\lambda^2 + 60\lambda + 200. \tag{10.264}$$

Example (contd.)

Thus, for $\phi(\mathbf{A})$ we get

$$\phi(\mathbf{A}) = \mathbf{A}^3 + 14\mathbf{A}^2 + 60\mathbf{A} + 200\mathbf{I} = \begin{bmatrix} 199 & 55 & 8 \\ -8 & 159 & 7 \\ -7 & -43 & 117 \end{bmatrix}. \tag{10.265}$$

Finally, based on Ackermann's formula we have

$$\mathbf{K} = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \mathcal{C}^{-1}\phi(\mathbf{A}) = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 5 & 6 & 1 \\ 6 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 199 & 55 & 8 \\ -8 & 159 & 7 \\ -7 & -43 & 117 \end{bmatrix} = \begin{bmatrix} 199 & 55 & 8 \end{bmatrix}.$$

MATLAB function `acker` calculates the gain matrix **K** using the Akermann's Formula.

Example (contd.)

Using MATLAB function `acker`

```
A = [0, 1, 0; 0, 0, 1; -1, -5, -6];
B = [0; 0; 1];
mu = [-2-4i, -2+4i, -10];
K = acker(A,B,mu);
```

Direct calculation

```
A = [0, 1, 0; 0, 0, 1; -1, -5, -6];
B = [0; 0; 1];
mu = [-2-4i, -2+4i, -10];
alpha = poly(mu);
phi = polyvalm(alpha,A);
K = Ctr\phi;
K = K(end,:);
```

Cayley–Hamilton Theorem

Theorem 10.18 (Cayley–Hamilton)

Every square matrix over a commutative ring satisfies its own characteristic equation. In other words, if $\mathbf{A}$ *is an* $n \times n$ *matrix and* $\phi(\lambda) = \det(\lambda\mathbf{I} - \mathbf{A})$ *is its characteristic polynomial, then* $\phi(\mathbf{A}) = \mathbf{0}$*.*

Consider a square matrix $\mathbf{A}$ and let

$$\phi(\lambda) = \lambda^n + c_1\lambda^{n-1} + \cdots + c_{n-1}\lambda + c_n, \tag{10.266}$$

be the characteristic polynomial of the matrix $\mathbf{A}$. The theorem claims that:

$$\mathbf{A}^n + c_1\mathbf{A}^{n-1} + \cdots + c_{n-1}\mathbf{A} + c_n\mathbf{I} = 0 \tag{10.267}$$

Cayley–Hamilton Theorem (contd.)

For a direct algebraic proof see: Ogata, Katsuhiko. "Modern Control Engineering". fifth edition. 2010.

If $\mathbf{v}$ is an eigenvector of $\mathbf{A}$ with eigenvalue λ, then

$$\begin{aligned}\phi(\mathbf{A})\ \mathbf{v} &= \mathbf{A}^n \cdot \mathbf{v} + c_1 \mathbf{A}^{n-1} \cdot \mathbf{v} + \cdots + c_{n-1} \mathbf{A} \cdot \mathbf{v} + c_n \mathbf{I}_n \cdot \mathbf{v} && (10.268)\\ &= \lambda^n \mathbf{v} + c_1 \lambda^{n-1} \mathbf{v} + \cdots + c_{n-1} \lambda \mathbf{v} + c_n \mathbf{v} = \phi(\lambda)\mathbf{v} = \mathbf{0}, && (10.269)\end{aligned}$$

since the eigenvalues of $\mathbf{A}$ are precisely the roots of $\phi(t)$.

If $\mathbf{A}$ admits a basis of eigenvectors, in other words, if $\mathbf{A}$ is diagonalizable, then

$$\mathbf{A} = \mathbf{V}\mathbf{D}\mathbf{V}^{-1}, \tag{10.270}$$

Cayley–Hamilton Theorem (contd.)

where $\mathbf{V}$ is the matrix of eigenvectors, and $\mathbf{D}$ is the diagonal matrix of eigenvalues. The characteristic polynomial of $\mathbf{A}$ is given by

$$\phi_{\mathbf{A}}(\lambda) = |\lambda\mathbf{I} - \mathbf{A}| = \prod_{i=1}^{n}(\lambda - \lambda_i) \equiv \sum_{k=0}^{n} c_k \lambda^k \tag{10.271}$$

Applying the characteristic polynomial to $\mathbf{A}$ gives us

$$\phi_{\mathbf{A}}(\mathbf{A}) = \sum c_k \mathbf{A}^k = \mathbf{V}\phi_{\mathbf{A}}(\mathbf{D})\mathbf{V}^{-1} = \mathbf{V}\mathbf{P}\mathbf{V}^{-1} \tag{10.272}$$

where $\mathbf{P} = \mathrm{diag}(p_{11},\ p_{22},\ \ldots,\ p_{nn})$ is a diagonal matrix such that

$$p_{ii} = \sum_{k=0}^{n} c_k \lambda_i^k = \prod_{j=1}^{n}(\lambda_i - \lambda_j) = 0, \tag{10.273}$$

Therefore,

$$\phi_{\mathbf{A}}(\mathbf{A}) = \mathbf{V}\mathbf{P}\mathbf{V}^{-1} = \mathbf{0}. \tag{10.274}$$

Derivation Details of Ackermann's Formula

Let $\hat{\mathbf{A}} = \mathbf{A} - \mathbf{BK}$, and

$$|\lambda\mathbf{I} - \mathbf{A} + \mathbf{BK}| = \left|\lambda\mathbf{I} - \hat{\mathbf{A}}\right| = \prod_{\mu\in\Lambda}(\lambda - \mu) = \lambda^n + \alpha_1\lambda^{n-1} + \cdots \alpha_{n-1}\lambda + \alpha_n =: \phi_{\hat{\mathbf{A}}}(\lambda). \quad (10.275)$$

Based on Cayley–Hamilton Theorem we have

$$\phi_{\hat{\mathbf{A}}}(\hat{\mathbf{A}}) = \hat{\mathbf{A}}^n + \alpha_1\hat{\mathbf{A}}^{n-1} + \cdots \alpha_{n-1}\hat{\mathbf{A}} + \alpha_n\mathbf{I} = \mathbf{0}. \quad (10.276)$$

To simplify the derivation, let us consider the case where $n = 3$. The derivation is extendable to any $n \in \mathbb{N}$. Consider the following identities

$$\hat{\mathbf{A}}^0 = \mathbf{I}, \quad (10.277)$$

$$\hat{\mathbf{A}}^1 = (\mathbf{A} - \mathbf{BK})^1 = \mathbf{A} - \mathbf{BK}, \quad (10.278)$$

$$\hat{\mathbf{A}}^2 = (\mathbf{A} - \mathbf{BK})^2 = \mathbf{A}^2 - \mathbf{ABK} - \mathbf{BK}\hat{\mathbf{A}} \quad (10.279)$$

$$\hat{\mathbf{A}}^3 = (\mathbf{A} - \mathbf{BK})^3 = \mathbf{A}^3 - \mathbf{A}^2\mathbf{BK} - \mathbf{ABK}\hat{\mathbf{A}} - \mathbf{BK}\hat{\mathbf{A}}^2 \quad (10.280)$$

Derivation Details of Ackermann's Formula (contd.)

Noting that $\alpha_0 = 1$, we have

$$\phi_{\hat{\mathbf{A}}}(\hat{\mathbf{A}}) = \alpha_3 \mathbf{I} + \alpha_2 \hat{\mathbf{A}} + \alpha_1 \hat{\mathbf{A}}^2 + \alpha_0 \hat{\mathbf{A}}^3 \tag{10.281}$$

$$= \alpha_3 \mathbf{I} + \alpha_2 (\mathbf{A} - \mathbf{BK}) + \alpha_1 \left(\mathbf{A}^2 - \mathbf{ABK} - \mathbf{BK}\hat{\mathbf{A}} \right) + \mathbf{A}^3 - \mathbf{A}^2\mathbf{BK} - \mathbf{ABK}\hat{\mathbf{A}} - \mathbf{BK}\hat{\mathbf{A}}^2 \tag{10.282}$$

$$= \alpha_3 \mathbf{I} + \alpha_2 \mathbf{A} + \alpha_1 \mathbf{A}^2 + \mathbf{A}^3 - \alpha_2 \mathbf{BK} - \alpha_1 \left(\mathbf{ABK} + \mathbf{BK}\hat{\mathbf{A}} \right) - \mathbf{A}^2\mathbf{BK} - \mathbf{ABK}\hat{\mathbf{A}} - \mathbf{BK}\hat{\mathbf{A}}^2 \tag{10.283}$$

Noting that $\phi_{\hat{\mathbf{A}}}(\mathbf{A}) \neq \mathbf{0}$ ($\phi_{\hat{\mathbf{A}}}$ is the characteristic polynomial of $\hat{\mathbf{A}}$ not $\mathbf{A}$), we have

$$\phi_{\hat{\mathbf{A}}}(\hat{\mathbf{A}}) = \mathbf{0} = \phi_{\hat{\mathbf{A}}}(\mathbf{A}) - \mathbf{B} \left(\alpha_2 \mathbf{K} + \alpha_1 \mathbf{K}\hat{\mathbf{A}} + \mathbf{K}\hat{\mathbf{A}}^2 \right) - \mathbf{AB} \left(\alpha_1 \mathbf{K} + \mathbf{K}\hat{\mathbf{A}} \right) - \mathbf{A}^2\mathbf{BK} \tag{10.284}$$

Solving the above equation for $\phi_{\hat{\mathbf{A}}}(\mathbf{A})$ gives

$$\phi_{\hat{\mathbf{A}}}(\mathbf{A}) = \begin{bmatrix} \mathbf{B} & \mathbf{AB} & \mathbf{A}^2\mathbf{B} \end{bmatrix} \begin{bmatrix} \alpha_2 \mathbf{K} + \alpha_1 \mathbf{K}\hat{\mathbf{A}} + \mathbf{K}\hat{\mathbf{A}}^2 \\ \alpha_1 \mathbf{K} + \mathbf{K}\hat{\mathbf{A}} \\ \mathbf{K} \end{bmatrix} \implies \begin{bmatrix} * \\ * \\ \mathbf{K} \end{bmatrix} = \mathcal{C}^{-1}(\mathbf{A}, \mathbf{B})\, \phi_{\hat{\mathbf{A}}}(\mathbf{A}). \tag{10.285}$$

Pole Placement For Multi Input Systems

Robust Pole Assignment

Source: Kautsky, J., Nichols, N. K., & Van Dooren, P. (1985). Robust pole assignment in linear state feedback. International Journal of control, 41(5), 1129-1155.

Theorem 10.19

Given $\mathbf{\Lambda} = \mathrm{diag}\{\lambda_i, \lambda_2, \ldots, \lambda_j\}$ *and* $\mathbf{X}$ *non-singular, then there exists* $\mathbf{F}$*, a solution to* $(\mathbf{A} + \mathbf{BF})\,\mathbf{X} = \mathbf{X\Lambda}$*, if and only if*

$$\mathbf{U}_1^T\left(\mathbf{AX} - \mathbf{X\Lambda}\right) = \mathbf{0}, \tag{10.286}$$

where

$$\mathbf{B} = \begin{bmatrix}\mathbf{U}_0 & \mathbf{U}_1\end{bmatrix}\begin{bmatrix}\mathbf{Z}\\ \mathbf{0}\end{bmatrix} \tag{10.287}$$

with $\mathbf{U} = \begin{bmatrix}\mathbf{U}_0 & \mathbf{U}_1\end{bmatrix}$ *orthogonal and* $\mathbf{Z}$ *non-singular. Then* $\mathbf{F}$ *is given explicitly by*

$$\mathbf{F} = \mathbf{Z}^{-1}\mathbf{U}_0^T\left(\mathbf{X\Lambda X}^{-1} - \mathbf{A}\right) \tag{10.288}$$

Robust Pole Assignment (contd.)

Proof. The assumption that $\mathbf{B}$ is of full rank implies the existence of decomposition (10.287). From $(\mathbf{A} + \mathbf{B}\mathbf{F})\,\mathbf{X} = \mathbf{X}\boldsymbol{\Lambda}$, $\mathbf{F}$ must satisfy

$$\mathbf{B}\mathbf{F} = \mathbf{X}\boldsymbol{\Lambda}\mathbf{X}^{-1} - \mathbf{A}$$

and pre-multiplication by $\mathbf{U}^T$ then gives the two equations

$$\mathbf{Z}\mathbf{F} = \mathbf{U}_0^T\left(\mathbf{X}\boldsymbol{\Lambda}\mathbf{X}^{-1} - \mathbf{A}\right),$$
$$\mathbf{0} = \mathbf{U}_1^T\left(\mathbf{X}\boldsymbol{\Lambda}\mathbf{X}^{-1} - \mathbf{A}\right).$$

□

$\mathbf{U}_1$, $\mathbf{U}_0$ and $\mathbf{Z}$ could be obtained by an orthogonal decomposition method, such as QR or SVD. For example, based on QR decomposition of $\mathbf{B}$, we have $\mathbf{B} = \mathbf{Q}\mathbf{R}$. The matrix $\mathbf{Q}$ would then be split into $\mathbf{U}_0$ and $\mathbf{U}_1$, the components that correspond to the non-zero and zero parts of $\mathbf{R}$, respectively. The nonsingular part of $\mathbf{R}$ is $\mathbf{Z}$.

Robust Pole Assignment (contd.)

In MATLAB, the command

```
K = place(A,B,p)
[K, prec] = place(A,B,p)
```

is an implementation of the robust pole assignment algorithm. The function returns the gain matrix $\mathbf{K}$ as the solution to

$$(\mathbf{A} - \mathbf{B}\mathbf{K})\mathbf{X} = \mathbf{X}\boldsymbol{\Lambda},$$

where $\mathbf{X}$ is chosen such that the sensitivity of the closed-loop poles to perturbations in $\mathbf{A}$ or $\mathbf{B}$ is minimized.

`place` also returns `prec`, an accuracy estimate of how closely the eigenvalues of $\mathbf{A} - \mathbf{B}\mathbf{K}$ match the specified locations defined by `p`.

Linear Quadratic Regulator

Consider a linear autonomous multi input, multi output system

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}. \tag{10.289}$$

The Linear Quadratic Regulator (LQR) problem is to find the linear control function

$$\mathbf{u} = -\mathbf{K}\mathbf{x}, \tag{10.290}$$

that minimizes the cost function

$$J = \int_0^\infty \mathbf{x}^T\mathbf{Q}\mathbf{x} + \mathbf{u}^T\mathbf{R}\mathbf{u}\; dt, \tag{10.291}$$

where $\mathbf{Q}$ is a symmetric positive definite or positive semidefinite matrix and $\mathbf{R}$ is a symmetric positive definite matrix. The matrices $\mathbf{Q}$ and $\mathbf{R}$ determine the relative importance of the error and the expenditure of this energy. In this problem, we assume that the control vector $\mathbf{u}(t)$ is unconstrained.

Linear Quadratic Regulator (contd.)

We will revisit LQR derivation using tools of optimal control. Here, we follow a derivation that does not require techniques from calculus of variations.

For the closed loop system we have

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} - \mathbf{B}\mathbf{K}\mathbf{x} = (\mathbf{A} - \mathbf{B}\mathbf{K})\mathbf{x}. \tag{10.292}$$

In the following derivations, we assume that $\mathbf{A} - \mathbf{B}\mathbf{K}$ is stable (Hurwitz). Substituting $\mathbf{u} = -\mathbf{K}\mathbf{x}$ in the definition of the cost function gives

$$J = \int_0^\infty \mathbf{x}^T\mathbf{Q}\mathbf{x} + \mathbf{u}^T\mathbf{R}\mathbf{u}\; dt = \int_0^\infty \mathbf{x}^T\mathbf{Q}\mathbf{x} + \mathbf{x}^T\mathbf{K}^T\mathbf{R}\mathbf{K}\mathbf{x}\; dt = \int_0^\infty \mathbf{x}^T\left(\mathbf{Q} + \mathbf{K}^T\mathbf{R}\mathbf{K}\right)\mathbf{x}\; dt \tag{10.293}$$

Let $\mathbf{P}$ be a symmetric positive definite matrix, set

$$-\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x}^T\mathbf{P}\mathbf{x} = \mathbf{x}^T\left(\mathbf{Q} + \mathbf{K}^T\mathbf{R}\mathbf{K}\right)\mathbf{x}. \tag{10.294}$$

Linear Quadratic Regulator (contd.)

From (10.294) we get

$$\mathbf{x}^T\left(\mathbf{Q}+\mathbf{K}^T\mathbf{R}\mathbf{K}\right)\mathbf{x} = -\dot{\mathbf{x}}^T\mathbf{P}\mathbf{x} - \mathbf{x}^T\mathbf{P}\dot{\mathbf{x}} = \mathbf{x}^T\left((\mathbf{A}-\mathbf{B}\mathbf{K})^T\mathbf{P}+\mathbf{P}(\mathbf{A}-\mathbf{B}\mathbf{K})\right)\mathbf{x}. \quad (10.295)$$

Since the above equation needs to hold for any $\mathbf{x}$, we get

$$(\mathbf{A}-\mathbf{B}\mathbf{K})^T\mathbf{P}+\mathbf{P}(\mathbf{A}-\mathbf{B}\mathbf{K}) = -\left(\mathbf{Q}+\mathbf{K}^T\mathbf{R}\mathbf{K}\right). \quad (10.296)$$

The existence of $\mathbf{P}$ is a corollary of Theorem 10.4, since $\mathbf{A}-\mathbf{B}\mathbf{K}$ is stable (by assumption) and $\mathbf{Q}+\mathbf{K}^T\mathbf{R}\mathbf{K}$ is symmetric positive definite.

Recall that, based on Theorem 10.4, for any Hurwitz $\mathbf{A}$ and any positive definite symmetric matrix $\mathbf{Q}$ there exists a positive definite symmetric matrix $\mathbf{P}$ such that $\mathbf{P}\mathbf{A}+\mathbf{A}^T\mathbf{P}=-\mathbf{Q}$.

Linear Quadratic Regulator (contd.)

Since we know $\mathbf{P}$ exists, we can proceed to find its value. Using (10.294), and $\mathbf{A} - \mathbf{BK}$ being Hurwitz, we can evaluate the performance index J as

$$J = \int_0^\infty \mathbf{x}^T \left(\mathbf{Q} + \mathbf{K}^T \mathbf{R} \mathbf{K} \right) \mathbf{x} \; dt = -\mathbf{x}^T \mathbf{P} \mathbf{x} \Big|_0^\infty = \mathbf{x}(0)^T \mathbf{P} \mathbf{x}(0). \tag{10.297}$$

Since $\mathbf{R}$ is a symmetric positive definite matrix, we can write $\mathbf{R} = \mathbf{T}^T\mathbf{T}$, for some nonsingular $\mathbf{T}$. Accordingly, (10.296) gives

$$\begin{aligned}
\mathbf{0} &= \left(\mathbf{A}^T - \mathbf{K}^T \mathbf{B}^T \right) \mathbf{P} + \mathbf{P} \left(\mathbf{A} - \mathbf{BK} \right) + \mathbf{Q} + \mathbf{K}^T \mathbf{T}^T \mathbf{T} \mathbf{K} && (10.298) \\
&= \mathbf{A}^T \mathbf{P} + \mathbf{P} \mathbf{A} + \mathbf{Q} - \mathbf{K}^T \mathbf{B}^T \mathbf{P} - \mathbf{P} \mathbf{B} \mathbf{K} + \mathbf{K}^T \mathbf{T}^T \mathbf{T} \mathbf{K} && (10.299) \\
&= \mathbf{A}^T \mathbf{P} + \mathbf{P} \mathbf{A} + \mathbf{Q} + (\mathbf{T} \mathbf{K} - \mathbf{T}^{-T} \mathbf{B}^T \mathbf{P})^T (\mathbf{T} \mathbf{K} - \mathbf{T}^{-T} \mathbf{B}^T \mathbf{P}) - \mathbf{P} \mathbf{B} \mathbf{R}^{-1} \mathbf{B}^T \mathbf{P} && (10.300)
\end{aligned}$$

Linear Quadratic Regulator (contd.)

The minimization of J with respect to $\mathbf{K}$ requires minimization of

$$\mathbf{x}^T(\mathbf{TK} - \mathbf{T}^{-T}\mathbf{B}^T\mathbf{P})^T(\mathbf{TK} - \mathbf{T}^{-T}\mathbf{B}^T\mathbf{P})\mathbf{x}, \tag{10.301}$$

with respect to $\mathbf{K}$ (for details see: Ogata, Katsuhiko. "Modern Control Engineering". fifth edition. 2010.). The minimum of the above expression with respect to $\mathbf{K}$ happens at

$$\mathbf{TK} - \mathbf{T}^{-T}\mathbf{B}^T\mathbf{P} = \mathbf{0} \implies \mathbf{K} = \mathbf{T}^{-1}\mathbf{T}^{-T}\mathbf{B}^T\mathbf{P} = \mathbf{R}^{-1}\mathbf{B}^T\mathbf{P}. \tag{10.302}$$

Linear Quadratic Regulator (contd.)

Summary Given a linear autonomous multi input, multi output system

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}. \tag{10.303}$$

The linear control function that minimizes the cost function

$$J = \int_0^\infty \mathbf{x}^T\mathbf{Q}\mathbf{x} + \mathbf{u}^T\mathbf{R}\mathbf{u}\, dt, \tag{10.304}$$

for a given symmetric $\mathbf{Q}$ and $\mathbf{R}$ matrices, where $\mathbf{Q}$ is positive definite (or positive semidefinite) and $\mathbf{R}$ is positive definite, is

$$\mathbf{u} = -\mathbf{K}\mathbf{x} = -\mathbf{R}^{-1}\mathbf{B}^T\mathbf{P}\mathbf{x}(t), \tag{10.305}$$

where the symmetric positive definite matrix $\mathbf{P}$ is the solution to the reduced-matrix Riccati equation

$$\mathbf{A}^T\mathbf{P} + \mathbf{P}\mathbf{A} - \mathbf{P}\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^T\mathbf{P} + \mathbf{Q} = \mathbf{0}. \tag{10.306}$$

Example

Adopted from: Ogata, Katsuhiko. “Modern Control Engineering”. fifth edition. 2010.

Consider second order system

$$\ddot{y} = u, \tag{10.307}$$

known as *double integrator*. Let $x_1 = y$ and $x_2 = \dot{y}$ we have

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u. \tag{10.308}$$

Our goal is to find linear control $u = -\mathbf{K}\mathbf{x}$ that minimizes the cost function

$$J = \int_0^\infty \mathbf{x}^T \mathbf{Q} \mathbf{x} + u^2 \; dt, \tag{10.309}$$

where $\mathbf{R} = [1]$ and for $\mu > 0$,

$$\mathbf{Q} = \begin{bmatrix} 1 & 0 \\ 0 & \mu \end{bmatrix}. \tag{10.310}$$

Example (contd.)

Adopted from: Ogata, Katsuhiko. "Modern Control Engineering". fifth edition. 2010.

Based on LQR derivation, first we need to find symmetric positive definite $\mathbf{P}$ as the solution of the reduced-matrix Riccati equation

$$\mathbf{A}^T\mathbf{P} + \mathbf{PA} - \mathbf{PBR}^{-1}\mathbf{B}^T\mathbf{P} + \mathbf{Q} = \mathbf{0}. \tag{10.311}$$

Since $\mathbf{A}$, $\mathbf{B}$, $\mathbf{Q}$ and $\mathbf{R}$ are real, then $\mathbf{P}$ must be a real matrix. Since $\mathbf{P}$ is symmetric, let

$$\mathbf{P} = \begin{bmatrix} p_1 & p_3 \\ p_3 & p_2 \end{bmatrix}. \tag{10.312}$$

Accordingly, for the reduced-matrix Riccati equation we get

$$\underbrace{\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}}_{\mathbf{A}^T} \underbrace{\begin{bmatrix} p_1 & p_3 \\ p_3 & p_2 \end{bmatrix}}_{\mathbf{P}} + \underbrace{\begin{bmatrix} p_1 & p_3 \\ p_3 & p_2 \end{bmatrix}}_{\mathbf{P}} \underbrace{\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}}_{\mathbf{A}} - \underbrace{\begin{bmatrix} p_1 & p_3 \\ p_3 & p_2 \end{bmatrix}}_{\mathbf{P}} \underbrace{\begin{bmatrix} 0 \\ 1 \end{bmatrix}}_{\mathbf{B}} \underbrace{\begin{bmatrix} 1 \end{bmatrix}}_{\mathbf{R}^{-1}} \underbrace{\begin{bmatrix} 0 & 1 \end{bmatrix}}_{\mathbf{B}^T} \underbrace{\begin{bmatrix} p_1 & p_3 \\ p_3 & p_2 \end{bmatrix}}_{\mathbf{P}} + \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & \mu \end{bmatrix}}_{\mathbf{Q}} = \mathbf{0} \tag{10.313}$$

Example (contd.)

Adopted from: Ogata, Katsuhiko. "Modern Control Engineering". fifth edition. 2010.

Simplifying the equation yields

$$\begin{bmatrix} 0 & 0 \\ p_1 & p_3 \end{bmatrix} + \begin{bmatrix} 0 & p_1 \\ 0 & p_3 \end{bmatrix} - \begin{bmatrix} p_3^2 & p_2 p_3 \\ p_2 p_3 & p_2^2 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & \mu \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \tag{10.314}$$

which is equivalent to the following system of equations

$$1 - p_3^2 = 0, \tag{10.315a}$$

$$p_1 - p_2 p_3 = 0, \tag{10.315b}$$

$$2p_3 - p_2^2 + \mu = 0. \tag{10.315c}$$

which gives

$$p_1 = p_2 = \sqrt{\mu + 2}, \quad \text{and} \quad p_3 = 1. \tag{10.316}$$

Example (contd.)

Adopted from: Ogata, Katsuhiko. "Modern Control Engineering". fifth edition. 2010.

Given the values of p_1, p_2 and p_3, we can easily check that $\mathbf{P}$ is a positive definite matrix (how?). Using the matrix $\mathbf{P}$, we can define the gain matrix and the control function as

$$u = -\mathbf{K}\mathbf{x} = -\mathbf{R}^{-1}\mathbf{B}^T\mathbf{P}\mathbf{x} = -\begin{bmatrix}1\end{bmatrix}\begin{bmatrix}0 & 1\end{bmatrix}\begin{bmatrix}\sqrt{\mu+2} & 1 \\ 1 & \sqrt{\mu+2}\end{bmatrix}\begin{bmatrix}x_1 \\ x_2\end{bmatrix} = -x_1 - x_2\sqrt{\mu+2}, \tag{10.317}$$

and the optimal cost from any initial state $\mathbf{x}(0)$ is

$$J = \mathbf{x}(0)^T\mathbf{P}\mathbf{x}(0) = \begin{bmatrix}x_1(0) & x_2(0)\end{bmatrix}\begin{bmatrix}\sqrt{\mu+2} & 1 \\ 1 & \sqrt{\mu+2}\end{bmatrix}\begin{bmatrix}x_1(0) \\ x_2(0)\end{bmatrix} \tag{10.318}$$

$$= \sqrt{\mu+2}\left(x_1^2(0) + x_2^2(0)\right) + 2x_1(0)x_2(0). \tag{10.319}$$

Solving LQR Problems with MATLAB

In MATLAB, the command

```
K = lqr(A,B,Q,R)
[K,P,E] = lqr(A,B,Q,R)
```

solves the continuous-time LQR problem and returns the gain matrix $\mathbf{K}$ such that

$$\mathbf{u} = -\mathbf{K}\mathbf{x}$$

minimizes the performance index

$$J = \int_0^\infty \left(\mathbf{x}^T \mathbf{Q}\mathbf{x} + \mathbf{u}^T \mathbf{R}\mathbf{u}\right) dt$$

subject to the constraint equation

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}.$$

`lqr` also returns optional outputs: $\mathbf{P}$ as the solution of the associated algebraic Riccati equation and E as poles of the closed-loop system.

Contents

Linearization and Nominal Motion

Linearization of Autonomous Closed Loop System

Consider autonomous closed loop system

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}), \tag{11.320}$$

where $\mathbf{f} : D \subset \mathbb{R}^n \to \mathbb{R}^n$ is pointwise differentiable in D and $\mathbf{f}(\mathbf{0} \in D) = \mathbf{0}$ (i.e. the system has an equilibrium at the origin). Based on the mean value theorem, for every f_i of $\mathbf{f}(\mathbf{x}) = \big(f_1(\mathbf{x}), \ldots, f_n(\mathbf{x})\big)$, we have

$$f_i(\mathbf{x}) = f_i(\mathbf{0}) + \frac{\partial f_i}{\partial \mathbf{x}}(\mathbf{z}_i)\mathbf{x}, \tag{11.321}$$

where $\mathbf{z}_i$ is a point on the line segment connecting $\mathbf{x}$ to the origin.

Mean Value Theorem

Let $f : [a, b] \to \mathbb{R}$ be a continuous function on the closed interval $[a, b]$, and differentiable on the open interval (a, b). Then, $\exists\ c \in (a, b)$ such that

$$f'(c) = \big(f(b) - f(a)\big)\,(b - a)\,.$$

Linearization of Autonomous Closed Loop System (contd.)

Since $\mathbf{f}(\mathbf{0}) = \mathbf{0}$, $f_i(\mathbf{0}) = 0$ for all i. Hence

$$f_i(\mathbf{x}) = \frac{\partial f_i}{\partial \mathbf{x}}(\mathbf{z}_i)\mathbf{x} = \frac{\partial f_i}{\partial \mathbf{x}}(\mathbf{0})\mathbf{x} + \frac{\partial f_i}{\partial \mathbf{x}}(\mathbf{z}_i)\mathbf{x} - \frac{\partial f_i}{\partial \mathbf{x}}(\mathbf{0})\mathbf{x} \tag{11.322}$$

$$= \frac{\partial f_i}{\partial \mathbf{x}}(\mathbf{0})\mathbf{x} + \left(\frac{\partial f_i}{\partial \mathbf{x}}(\mathbf{z}_i) - \frac{\partial f_i}{\partial \mathbf{x}}(\mathbf{0})\right)\mathbf{x} \tag{11.323}$$

$$= \begin{bmatrix} a_{i1} & \cdots & a_{in} \end{bmatrix}\mathbf{x} + g_i(\mathbf{x}). \tag{11.324}$$

Accordingly, we get

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{g}(\mathbf{x}), \tag{11.325}$$

where

$$\mathbf{A} = \begin{bmatrix} a_{ij} \end{bmatrix} = \begin{bmatrix} \frac{\partial f_i}{\partial x_j} \end{bmatrix}, \qquad \mathbf{g}(\mathbf{x}) = \begin{bmatrix} g_i(\mathbf{x}) \end{bmatrix} = \begin{bmatrix} \left(\frac{\partial f_i}{\partial \mathbf{x}}(\mathbf{z}_i) - \frac{\partial f_i}{\partial \mathbf{x}}(\mathbf{0})\right)\mathbf{x} \end{bmatrix}. \tag{11.326}$$

Linearization of Autonomous Closed Loop System (contd.)

For every $g_i(\mathbf{x})$ we have

$$\left|g_i(\mathbf{x})\right| \leq \left\|\frac{\partial f_i}{\partial \mathbf{x}}(\mathbf{z}_i) - \frac{\partial f_i}{\partial \mathbf{x}}(\mathbf{0})\right\| \|\mathbf{x}\| . \tag{11.327}$$

By continuity of $\frac{\partial f_i}{\partial \mathbf{x}}$ we get

$$\frac{\|\mathbf{g}(\mathbf{x})\|}{\|\mathbf{x}\|} \to 0 \quad \text{as} \quad \|\mathbf{x}\| \to \mathbf{0}. \tag{11.328}$$

This suggests that in a *small neighborhood of the origin* we can approximate the nonlinear system $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ by its linearization about the origin

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}, \quad \text{where} \quad \mathbf{A} = \frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{0}). \tag{11.329}$$

Example

Adopted from Slotine JJ, Li W. "Applied Nonlinear Control". 1991

Consider the system

$$\dot{x}_1 = f_1(x_1, x_2) = x_1 \cos(x_2) + x_2^2, \tag{11.330}$$

$$\dot{x}_2 = f_2(x_1, x_2) = x_1 (x_1 + 1) + x_1 \sin(x_2) + x_2. \tag{11.331}$$

Clearly, $x_1 = x_2 = 0$ is an equilibrium point of the system. To obtain linear approximation of the system, we can compute the evaluation of the Jacobin of $\mathbf{f} = (f_1, f_2)$ with respect to $\mathbf{x} = (x_1, x_2)$ at the origin. Accordingly, we have

$$\mathbf{A} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(0,0) & \frac{\partial f_1}{\partial x_2}(0,0) \\ \frac{\partial f_2}{\partial x_1}(0,0) & \frac{\partial f_2}{\partial x_2}(0,0) \end{bmatrix} = \begin{bmatrix} \cos(x_2) & -x_1 \sin(x_2) + 2x_2 \\ 2x_1 + 1 + \sin(x_2) & x_1 \cos(x_2) + 1 \end{bmatrix} \Bigg|_{(0,0)} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}. \tag{11.332}$$

Example (contd.)

Adopted from Slotine JJ, Li W. “Applied Nonlinear Control”. 1991

Thus, the linear approximation of

$$\dot{x}_1 = f_1(x_1, x_2) = x_1 \cos(x_2) + x_2^2, \tag{11.333}$$

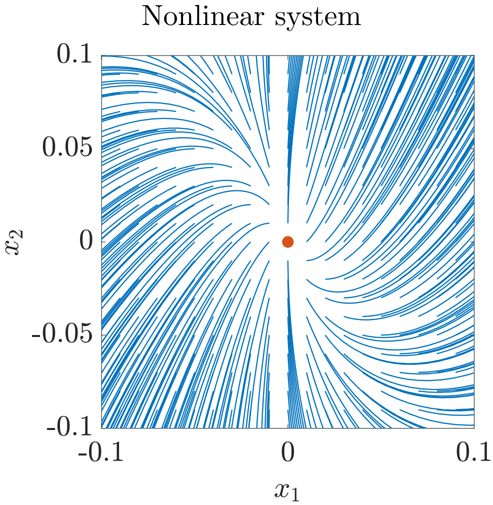
$$\dot{x}_2 = f_2(x_1, x_2) = x_1 (x_1 + 1) + x_1 \sin(x_2) + x_2. \tag{11.334}$$

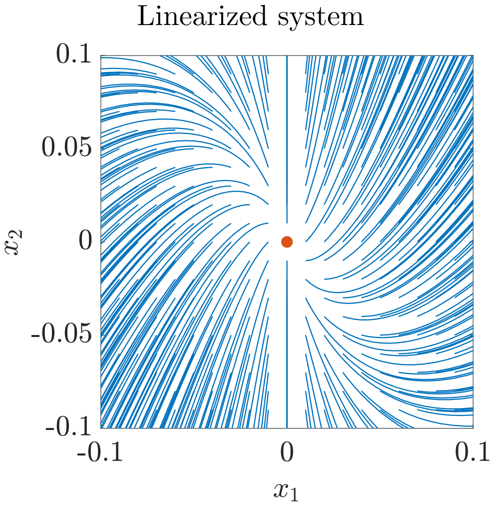
around the origin is

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \implies \begin{cases} \dot{x}_1 = x_1, \\ \dot{x}_2 = x_1 + x_2. \end{cases} \tag{11.335}$$

Example (contd.)

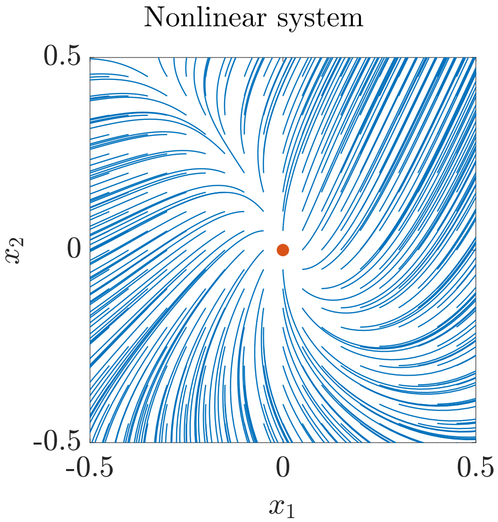
Adopted from Slotine JJ, Li W. "Applied Nonlinear Control". 1991

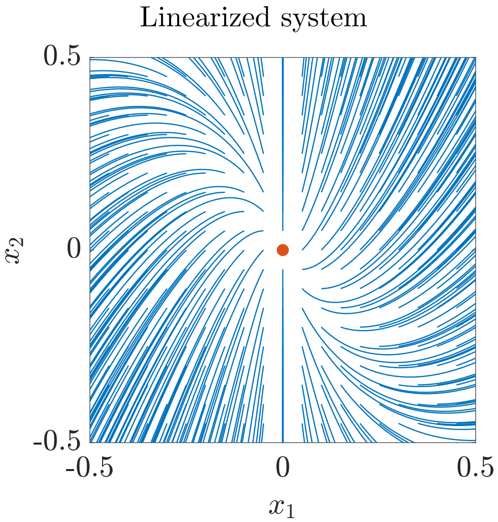




Example (contd.)

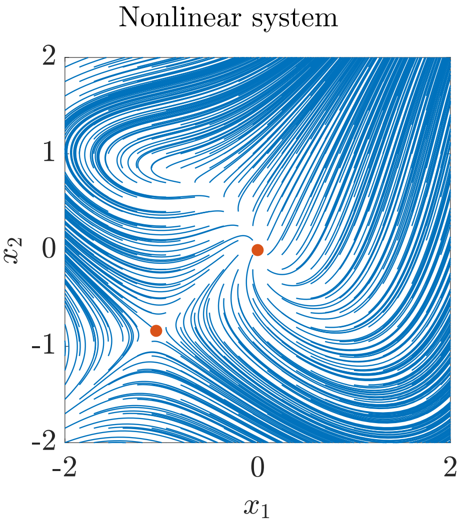
Adopted from Slotine JJ, Li W. "Applied Nonlinear Control". 1991

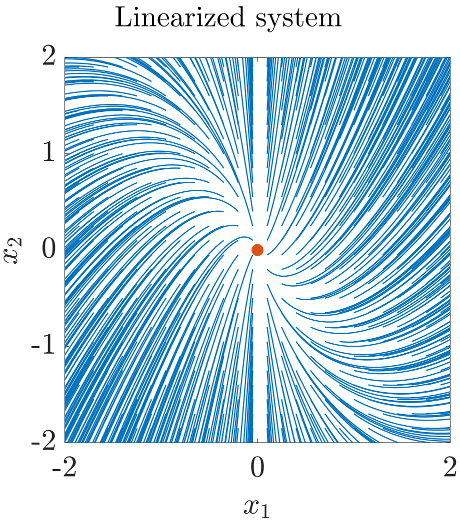




Example (contd.)

Adopted from Slotine JJ, Li W. "Applied Nonlinear Control". 1991





Lyapunov's Indirect Method: Condition for Stability of Nonlinear Systems

Theorem 11.1 (Lyapunov's Linearization Method)

Let $\mathbf{x} = \mathbf{0}$ *be an equilibrium point for the nonlinear system*

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}),$$

where $\mathbf{f} : D \to \mathbb{R}^n$ *is continuously differentiable and* D *is a neighborhood of the origin. Let*

$$\mathbf{A} = \frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{0}),$$

with eigenvalues λ_1 *to* λ_n. *Then*

1. *The origin is asymptotically stable if* $\Re(\lambda_i) < 0$ *for all* λ_i.
2. *The origin is unstable if* $\Re(\lambda_i) > 0$ *for one or more of* λ_i.

Lyapunov's Indirect Method: Condition for Stability of Nonlinear Systems (contd.)

Theorem 11.1 provides a simple procedure for determining the stability of an equilibrium point at the origin. We calculate the Jacobian matrix

$$\mathbf{A} = \frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{0}),$$

and test its eigenvalues λ_i:

1. If $\Re(\lambda_i) < 0$ *for all* λ_i then the origin is asymptotically stable,
2. If for *at least one* of $\Re(\lambda_i) > 0$, then the origin is unstable.

Theorem 11.1 does not say anything about the case when $\Re(\lambda_i) \leq 0$ for all i. In this case, linearization fails to determine the stability of the equilibrium point. *If the linearized system is marginally stable ($\Re(\lambda_i) = 0$ for some i), then we cannot conclude anything about stability of the equilibrium point of the nonlinear system.*

Lyapunov's Indirect Method: Condition for Stability of Nonlinear Systems (contd.)

Proof of 1.

adopted from Khalil, H.K. "Nonlinear Systems". third edition. 2002

Let $\mathbf{A}$ be a Hurwitz. Then, based on Lyapunov's theorem for LTI systems, for any positive definite symmetric matrix $\mathbf{Q}$, the solution $\mathbf{P}$ of the Lyapunov equation $\mathbf{PA} + \mathbf{A}^T\mathbf{P} = -\mathbf{Q}$ is symmetric positive definite. Let $V = \mathbf{x}^T\mathbf{Px}$. For the nonlinear system we have

$$\begin{aligned}
\dot{V}(\mathbf{x}) &= \mathbf{x}^T\mathbf{Pf}(\mathbf{x}) + \mathbf{f}^T(\mathbf{x})\mathbf{Px} && (11.336)\\
&= \mathbf{x}^T\mathbf{P}\left(\mathbf{Ax} + \mathbf{g}(\mathbf{x})\right) + \left(\mathbf{x}^T\mathbf{A}^T + \mathbf{g}^T(\mathbf{x})\right)\mathbf{Px} && (11.337)\\
&= \mathbf{x}^T\left(\mathbf{PA} + \mathbf{A}^T\mathbf{P}\right)\mathbf{x} + \mathbf{x}^T\mathbf{Pg}(\mathbf{x}) + \mathbf{g}^T(\mathbf{x})\mathbf{Px} && (11.338)\\
&= -\mathbf{x}^T\mathbf{Qx} + 2\mathbf{x}^T\mathbf{Pg}(\mathbf{x}). && (11.339)
\end{aligned}$$

The last equality is obtained by noting that $\mathbf{x}^T\mathbf{Pg}(\mathbf{x}) \in \mathbb{R} \implies \mathbf{x}^T\mathbf{Pg}(\mathbf{x}) = \mathbf{g}^T(\mathbf{x})\mathbf{Px}$.

Lyapunov's Indirect Method: Condition for Stability of Nonlinear Systems (contd.)

The first term on the right-hand side is negative definite, while the second term is indefinite. However, for the function $\mathbf{g}(\mathbf{x})$ we have

$$\frac{\|\mathbf{g}(\mathbf{x})\|_2}{\|\mathbf{x}\|_2} \to 0 \quad \text{as} \quad \|\mathbf{x}\|_2 \to \mathbf{0}, \tag{11.340}$$

which implies for any $\gamma > 0$, there exists $r > 0$ such that, for all $\mathbf{x}$ with $\|\mathbf{x}\|_2 < r$ we have

$$\|\mathbf{g}(\mathbf{x})\|_2 < \gamma \|\mathbf{x}\|_2. \tag{11.341}$$

Accordingly, for all $\mathbf{x}$ with $\|\mathbf{x}\|_2 < r$ we get

$$\dot{V}(\mathbf{x}) < -\mathbf{x}^T \mathbf{Q} \mathbf{x} + 2\gamma \|\mathbf{P}\|_2 \|\mathbf{x}\|_2^2. \tag{11.342}$$

Lyapunov's Indirect Method: Condition for Stability of Nonlinear Systems (contd.)

We have

$$\mathbf{x}^T\mathbf{Q}\mathbf{x} \geq \lambda_{\min}(\mathbf{Q})\|\mathbf{x}\|_2^2, \tag{11.343}$$

where $\lambda_{\min}(\mathbf{Q})$ is the smallest eigenvalues of $\mathbf{Q}$. Since $\mathbf{Q}$ is a symmetric positive definite matrix, $0 < \lambda_{\min}(\mathbf{Q})$. Hence, for all $\mathbf{x}$ with $\|\mathbf{x}\|_2 < r$ we get

$$\dot{V}(\mathbf{x}) < -\left(\lambda_{\min}(\mathbf{Q}) - 2\gamma\|\mathbf{P}\|_2\right)\|\mathbf{x}\|_2^2. \tag{11.344}$$

Choosing

$$\gamma < \frac{\lambda_{\min}(\mathbf{Q})}{2\|P\|_2}, \tag{11.345}$$

ensures that $V(\mathbf{x})$ is negative definite. Accordingly, based on Lyapunov's direct method, we conclude that the origin is an asymptotically stable equilibrium point. $\square$

For the proof of (2) see Khalil, H.K. "Nonlinear Systems". third edition. 2002.

Example

Adopted from Khalil, H.K. "Nonlinear Systems". third edition. 2002

Consider the general pendulum equation

$$\dot{x}_1 = x_2, \tag{11.346}$$

$$\dot{x}_2 = -a \sin x_1 - b x_2. \tag{11.347}$$

The equilibrium points of the system are $\{(k\pi, 0) \mid k \in \mathbb{Z}\}$. Let us investigate stability of two equilibrium points $(0,0)$ and $(\pi, 0)$ by using linearization.
The Jacobian matrix is given by

$$\frac{\partial \mathbf{f}}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -a \cos x_1 & -b \end{bmatrix}. \tag{11.348}$$

Example (contd.)

Adopted from Khalil, H.K. “Nonlinear Systems”. third edition. 2002

To determine the stability of the origin, we evaluate the Jacobian at $\mathbf{x} = \mathbf{0}$:

$$\mathbf{A} = \left.\frac{\partial \mathbf{f}}{\partial \mathbf{x}}\right|_{\mathbf{x}=\mathbf{0}} = \begin{bmatrix} 0 & 1 \\ -a & -b \end{bmatrix} \tag{11.349}$$

The eigenvalues of $\mathbf{A}$ are

$$\lambda_1 = -\frac{b}{2} - \frac{1}{2}\sqrt{b^2 - 4a}, \qquad \lambda_2 = -\frac{b}{2} + \frac{1}{2}\sqrt{b^2 - 4a}. \tag{11.350}$$

For all $a, b > 0$, the eigenvalues satisfy $\Re(\lambda_i) < 0$, which implies, the origin is an asymptotically stable equilibrium point.

In the absence of friction ($b = 0$), both eigenvalues are on the imaginary axis. Thus, we cannot determine the stability of the origin through linearization.

Example (contd.)

Adopted from Khalil, H.K. "Nonlinear Systems". third edition. 2002

In order to study the stability of the equilibrium point at $(\pi, 0)$, we introduce the change of variable

$$z_1 = x_1 - \pi, \tag{11.351a}$$
$$z_2 = x_2. \tag{11.351b}$$

The corresponding state space form of $\mathbf{z}$ is

$$\dot{z}_1 = \dot{x}_1 - 0 = x_2 = z_2, \tag{11.352a}$$
$$\dot{z}_2 = \dot{x}_2 = -a\sin(z_1 + \pi) - bz_2 = a\sin z_1 - bz_2. \tag{11.352b}$$

The Jacobian matrix is given by

$$\frac{\partial \mathbf{f}_\pi}{\partial \mathbf{z}} = \begin{bmatrix} 0 & 1 \\ a\cos z_1 & -b \end{bmatrix}. \tag{11.353}$$

Example (contd.)

Adopted from Khalil, H.K. “Nonlinear Systems”. third edition. 2002

To determine the stability of the origin, we evaluate the Jacobian at $\mathbf{z} = \mathbf{0}$:

$$\mathbf{A}_\pi = \left.\frac{\partial \mathbf{f}}{\partial \mathbf{z}}\right|_{\mathbf{z}=\mathbf{0}} = \begin{bmatrix} 0 & 1 \\ a & -b \end{bmatrix} \tag{11.354}$$

The eigenvalues of $\mathbf{A}_\pi$ are

$$\mu_1 = -\frac{b}{2} - \frac{1}{2}\sqrt{b^2 + 4a}, \qquad \mu_2 = -\frac{b}{2} + \frac{1}{2}\sqrt{b^2 + 4a}. \tag{11.355}$$

For all $a > 0$ and $b \geq 0$, there is one eigenvalue in the open right-half plane. Hence, the equilibrium point at $(z_1 = 0, z_2 = 0) = (x_1 = \pi, x_2 = 0)$ is unstable.

See that $\left.\frac{\partial \mathbf{f}_\pi}{\partial \mathbf{z}}\right|_{\mathbf{z}=(0,0)} = \left.\frac{\partial \mathbf{f}}{\partial \mathbf{x}}\right|_{\mathbf{x}=(\pi,0)}$.

Linearization Via Taylor Series

A function $f : D \subset \mathbb{R} \to \mathbb{R}$, D open, is *real analytic* if for any $x_0 \in D$ we have

$$f(x) = \sum_{n=0}^{\infty} a_n (x - x_0)^n = a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + \cdots \tag{11.356}$$

for some $a_0, a_1, \cdots \in \mathbb{R}$ and the series is convergent to $f(x)$ for x in a neighborhood of x_0. Alternatively, a real analytic function is an *infinitely differentiable* function such that the *Taylor series* at any point x_0 in its domain

$$T(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n = f(x_0) + f'(x_0) \cdot (x - x_0) + \cdots \tag{11.357}$$

converges to $f(x)$ for x in a neighborhood of x_0 pointwise.

Linearization Via Taylor Series (contd.)

Let $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^n$ be a continuously differentiable and consider the autonomous system

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}), \tag{11.358}$$

Based on Taylor series, for every $i \in \{1, \dots, n\}$, we have

$$\dot{x}_i = f_i(\mathbf{0}) + \sum_{j=1}^{n} \frac{\partial f_i}{\partial x_j}(\mathbf{0}) \cdot x_j + \epsilon_i(\mathbf{x}) = \begin{bmatrix} a_{i1} = \frac{\partial f_i}{\partial x_1}(\mathbf{0}) & \cdots & a_{in} = \frac{\partial f_i}{\partial x_n}(\mathbf{0}) \end{bmatrix} \mathbf{x} + \epsilon_i(\mathbf{x}) \tag{11.359a}$$

where ϵ_i captures the effect of higher-order terms. Thus for $\mathbf{f} = (f_1, \dots, f_n)$ we get

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{0}) + \mathbf{A}\mathbf{x} + \boldsymbol{\epsilon}, \tag{11.360}$$

Linearization Via Taylor Series (contd.)

where $\mathbf{A}$ is $n \times n$ Jacobian matrix of $\mathbf{f}$ with respect to $\mathbf{x}$ evaluated at $\mathbf{x} = \mathbf{0}$. That is

$$\mathbf{A} = \frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{0}) = \left[\frac{\partial f_i}{\partial x_j}(\mathbf{0}) \right]. \tag{11.361}$$

If $\mathbf{f}(\mathbf{0}) = \mathbf{0}$, i.e. the system has an equilibrium point at the origin, we can obtain the *linearization* (or *linear approximation*) of the original nonlinear system at origin as

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}. \tag{11.362}$$

More generally, for a non-autonomous nonlinear system with a control input $\mathbf{u}$

$$\dot{\mathbf{x}} = \mathbf{f}(t, \mathbf{x}, \mathbf{u}) \tag{11.363}$$

Linearization Via Taylor Series (contd.)

with continuously differentiable $\mathbf{f} : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ such that $\mathbf{f}(t, \mathbf{0}, \mathbf{0}) = \mathbf{0}$, we can obtain *linear approximation* of the system as

$$\dot{\mathbf{x}} = \left[\frac{\partial \mathbf{f}}{\partial \mathbf{x}}(t, \mathbf{x} = \mathbf{0}, \mathbf{u} = \mathbf{0})\right] \mathbf{x} + \left[\frac{\partial \mathbf{f}}{\partial \mathbf{u}}(t, \mathbf{x} = \mathbf{0}, \mathbf{u} = \mathbf{0})\right] \mathbf{u} = \mathbf{A}(t)\mathbf{x} + \mathbf{B}(t)\mathbf{u}, \tag{11.364}$$

where $\mathbf{A} : \mathbb{R} \to \mathbb{R}^{n \times n}$ and $\mathbf{B} : \mathbb{R} \to \mathbb{R}^{n \times m}$ are explicit time dependent Jacobian matrices of $\mathbf{f}$ with respect to $\mathbf{x}$ and $\mathbf{u}$ evaluated at $(\mathbf{x} = \mathbf{0}, \mathbf{u} = \mathbf{0})$, respectively.

Linearization of Regulatory Systems

Consider the nonlinear control system

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}). \tag{11.365}$$

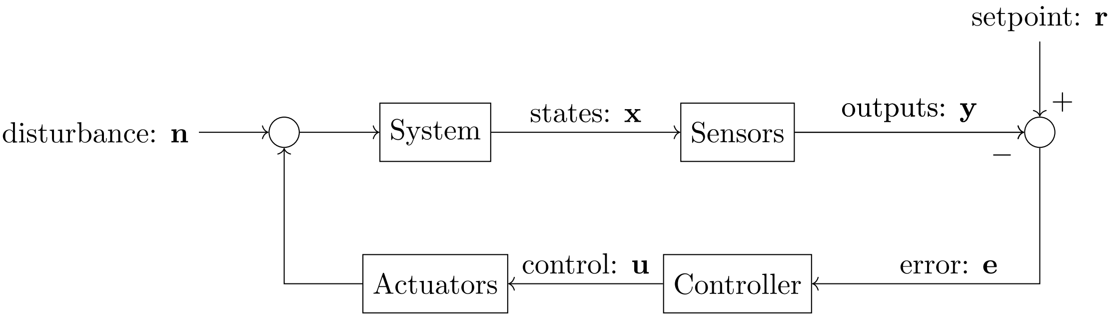
If $\mathbf{u} = \mathbf{u}(\mathbf{x})$ we get $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}) = \mathbf{f}\left(\mathbf{x}, \mathbf{u}(\mathbf{x})\right)$. Accordingly,

$$\begin{aligned}
\mathbf{f}\left(\mathbf{x}, \mathbf{u}(\mathbf{x})\right) &= \mathbf{f}\left(\mathbf{0}, \mathbf{u}(\mathbf{0})\right) + \left.\frac{\partial \mathbf{f}}{\partial \mathbf{x}}\right|_{\mathbf{x}=\mathbf{0}} \mathbf{x} + \boldsymbol{\epsilon} && (11.366) \\
&= \mathbf{f}\left(\mathbf{0}, \mathbf{u}(\mathbf{0})\right) + \left(\left.\frac{\partial \mathbf{f}}{\partial \mathbf{x}}\right|_{\mathbf{x}=\mathbf{0}} + \left.\left(\frac{\partial \mathbf{f}}{\partial \mathbf{u}} \frac{\partial \mathbf{u}}{\partial \mathbf{x}}\right)\right|_{\mathbf{x}=\mathbf{0}} \right) \mathbf{x} + \boldsymbol{\epsilon} && (11.367) \\
&= \mathbf{f}\left(\mathbf{0}, \mathbf{u}(\mathbf{0})\right) + \left(\mathbf{A} + \mathbf{B}\mathbf{G}\right) \mathbf{x} + \boldsymbol{\epsilon}. && (11.368)
\end{aligned}$$

Hence, if $\mathbf{f}(\mathbf{0}, \mathbf{u}(\mathbf{0})) = \mathbf{0}$, we get $\dot{\mathbf{x}} \approx \left(\mathbf{A} + \mathbf{B}\mathbf{G}\right) \mathbf{x}$, where $\mathbf{G} = \left.\frac{\partial \mathbf{u}}{\partial \mathbf{x}}\right|_{\mathbf{x}=\mathbf{0}}$.

Nominal Motion

Setpoint


setpoint: **r**
disturbance: **n**
System
states: **x**
Sensors
outputs: **y**
+
−
Actuators
control: **u**
Controller
error: **e**

Nominal Motion

What if we are concerned with the stability of a *motion* rather than a fixed point? Consider the system

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}), \tag{11.369}$$

and let $\mathbf{x}^*$ be *nominal motion* of the system, which is the solution of

$$\begin{cases} \dot{\mathbf{x}}^* = \mathbf{f}(\mathbf{x}^*), \\ \mathbf{x}^*(0) = \mathbf{x}_0. \end{cases} \tag{11.370}$$

Consider a perturbation $\delta\mathbf{x}_0$ from the initial condition $\mathbf{x}_0$. We are interested in studying the solution of

$$\begin{cases} \dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}), \\ \mathbf{x}(0) = \mathbf{x}_0 + \delta\mathbf{x}_0. \end{cases} \tag{11.371}$$

Nominal Motion (contd.)

Let the error $\mathbf{e}(t)$ be

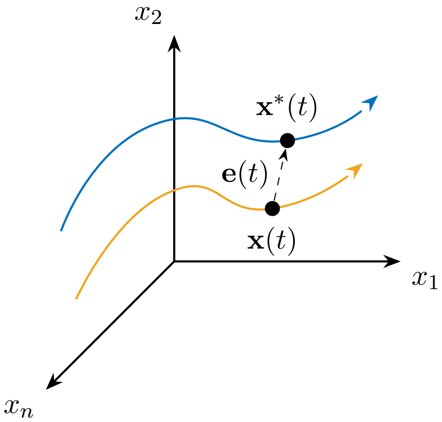
$$\mathbf{e}(t) := \mathbf{x}^*(t) - \mathbf{x}(t). \tag{11.372}$$

Since $\mathbf{x}^*(t)$ and $\mathbf{x}(t)$ are solutions of $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$, we get

$$\begin{aligned} \dot{\mathbf{e}} &= \dot{\mathbf{x}}^*(t) - \dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}^*) - \mathbf{f}(\mathbf{x}) && (11.373a) \\ &= \mathbf{f}(\mathbf{x}^*) - \mathbf{f}\left(\mathbf{x}^*(t) - \mathbf{e}\right) = \mathbf{g}(t, \mathbf{e}), && (11.373b) \end{aligned}$$

Hence, $\mathbf{e}(t)$ is the solution of *non-autonomous* IVP

$$\dot{\mathbf{e}} = \mathbf{g}(t, \mathbf{e}), \qquad \mathbf{e}(0) = \delta\mathbf{x}_0. \tag{11.374}$$



Example

Adopted from: Slotine JJ, Li W. "Applied Nonlinear Control". 1991

Consider a nonlinear mass-spring-damper system

$$m\ddot{x} + k_1 x + k_2 x^3 = 0. \tag{11.375}$$

Let $x^*(t)$ denote the nominal motion and set $e(t) = x^*(t) - x(t)$, which implies

$$\dot{e} = \dot{x}^* - \dot{x} \implies \dot{x} = \dot{x}^* - \dot{e}, \qquad \ddot{e} = \ddot{x}^* - \ddot{x} \implies \ddot{x} = \ddot{x}^* - \ddot{e}. \tag{11.376}$$

Substituting $x = x^* - e$, $\dot{x} = \dot{x}^* - \dot{e}$ and $\ddot{x} = \ddot{x}^* - \ddot{e}$ in (11.375) gives

$$m\left(\ddot{x}^* - \ddot{e}\right) + k_1\left(x^* - e\right) + k_2\left(x^* - e\right)^3 = 0 \tag{11.377}$$

$$\implies m\ddot{x}^* + k_1 x^* + k_2 x^{*3} - \left(m\ddot{e} + k_1 e + k_2\left(e^3 - 3x^* e^2 + 3x^{*2} e\right)\right) = 0 \tag{11.378}$$

$$\implies m\ddot{e} + k_1 e + k_2\left(e^2 - 3x^*(t)\, e + 3x^{*2}(t)\right) e = 0, \tag{11.379}$$

which is clearly a non-autonomous system.

Error Dynamics

Consider the system

$$\dot{\mathbf{x}} = \mathbf{f}(t, \mathbf{x}, \mathbf{u}). \tag{11.380}$$

Let $\mathbf{r}(t)$ be the desired reference input and set the error as

$$\mathbf{e}(t) := \mathbf{r}(t) - \mathbf{x}(t). \tag{11.381}$$

Taking derivative of the error with respect to time gives

$$\dot{\mathbf{e}} = \dot{\mathbf{r}} - \dot{\mathbf{x}} = \dot{\mathbf{r}} - \mathbf{f}(t, \mathbf{x}, \mathbf{u}) \tag{11.382}$$

$$= \dot{\mathbf{r}} - \mathbf{f}(t, \mathbf{r} - \mathbf{e}, \mathbf{u}) \tag{11.383}$$

See that (11.383) is a non-autonomous system. Let us define

$$\mathbf{f}_e := \dot{\mathbf{r}} - \mathbf{f}(t, \mathbf{r} - \mathbf{e}, \mathbf{u}), \tag{11.384}$$

Accordingly, we obtain a representation for *error dynamics* as

$$\dot{\mathbf{e}} = \mathbf{f}_e(t, \mathbf{e}, \mathbf{u}). \tag{11.385}$$

Error Dynamics (contd.)

Now let us proceed to find a linear approximation of the error. Since our goal is to minimize the error, let us find an approximation for the point $\mathbf{e} = \mathbf{0}$ and $\mathbf{u} = \mathbf{0}$. We have

$$\mathbf{f}_e(t, \mathbf{e}, \mathbf{u}) = \mathbf{f}_e(t, \mathbf{0}, \mathbf{0}) + \left.\frac{\partial \mathbf{f}_e}{\partial \mathbf{e}}\right|_{\substack{\mathbf{e}\,=\,\mathbf{0}\\ \mathbf{u}\,=\,\mathbf{0}}} \mathbf{e} + \left.\frac{\partial \mathbf{f}_e}{\partial \mathbf{u}}\right|_{\substack{\mathbf{e}\,=\,\mathbf{0}\\ \mathbf{u}\,=\,\mathbf{0}}} \mathbf{u} + \boldsymbol{\epsilon} \tag{11.386}$$

Recall that $\mathbf{f}_e := \dot{\mathbf{r}} - \mathbf{f}(t, \mathbf{r} - \mathbf{e}, \mathbf{u})$. Hence,

$$\mathbf{f}_e(t, \mathbf{0}, \mathbf{0}) = \dot{\mathbf{r}} - \mathbf{f}(t, \mathbf{r} - \mathbf{0}, \mathbf{0}) = \dot{\mathbf{r}} - \mathbf{f}(t, \mathbf{r}, \mathbf{0}), \tag{11.387}$$

which is <u>*not*</u> necessarily zero! To address this, let us define base input $\mathbf{u}_0(t)$ such that

$$\mathbf{f}\left(t,\ \mathbf{r}(t),\ \mathbf{u}_0(t)\right) = \dot{\mathbf{r}}(t). \tag{11.388}$$

Error Dynamics (contd.)

Using $\mathbf{u}_0(t)$, we can find linear approximation of the error at the point $(\mathbf{e} = \mathbf{0},\ \mathbf{u} = \mathbf{u}_0)$, which leads to

$$\mathbf{f}_e(t, \mathbf{e}, \mathbf{u}) = \mathbf{f}_e(t, \mathbf{0}, \mathbf{u}_0) + \left.\frac{\partial \mathbf{f}_e}{\partial \mathbf{e}}\right|_{\substack{\mathbf{e}\,=\,\mathbf{0} \\ \mathbf{u}\,=\,\mathbf{u}_0}} \mathbf{e} + \left.\frac{\partial \mathbf{f}_e}{\partial \mathbf{u}}\right|_{\substack{\mathbf{e}\,=\,\mathbf{0} \\ \mathbf{u}\,=\,\mathbf{u}_0}} (\mathbf{u} - \mathbf{u}_0) + \boldsymbol{\epsilon} \tag{11.389}$$

$$= \dot{\mathbf{r}} - \mathbf{f}(t, \mathbf{r}, \mathbf{u}_0) + \mathbf{A}(t)\mathbf{e} + \mathbf{B}(t)\,(\mathbf{u} - \mathbf{u}_0) \tag{11.390}$$

$$= \mathbf{A}(t)\mathbf{e} + \mathbf{B}(t)\mathbf{u} + \mathbf{w}(t), \tag{11.391}$$

which again leads to an *affine* approximation.

As an approach to systematically circumvent the presence of nonzero term, we can introduce variations in state and control simultaneously, as explained next.

Error Dynamics (contd.)

As before, let the error $\mathbf{e}$ be the variation of the states $\mathbf{x}$ from a desired reference $\mathbf{r}$. Similarly, let us define control $\mathbf{v}$ as a variation of the input $\mathbf{u}$ from a base input $\mathbf{u}_0$. Hence,

$$\mathbf{e}(t) := \mathbf{r}(t) - \mathbf{x}(t), \tag{11.392}$$

$$\mathbf{v}(t) := \mathbf{u}_0(t) - \mathbf{u}(t), \tag{11.393}$$

where the base input $\mathbf{u}_0$ is defined such that

$$\mathbf{f}\left(t,\ \mathbf{r}(t),\ \mathbf{u}_0(t)\right) = \dot{\mathbf{r}}(t). \tag{11.394}$$

Now for the error dynamics we get

$$\dot{\mathbf{e}} = \mathbf{f}_e(t, \mathbf{e}, \mathbf{v}) := \dot{\mathbf{r}} - \mathbf{f}(t,\ \mathbf{r} - \mathbf{e},\ \mathbf{u}_0 - \mathbf{v}). \tag{11.395}$$

Error Dynamics (contd.)

Note that $\mathbf{e}$ and $\mathbf{v}$ are variations from our desired states and inputs, respectively. Now let us proceed to find a linear approximation of the system at $(\mathbf{e} = \mathbf{0}, \mathbf{v} = \mathbf{0})$. We have

$$\mathbf{f}_e(t, \mathbf{e}, \mathbf{v}) = \mathbf{f}_e(t, \mathbf{0}, \mathbf{0}) + \left.\frac{\partial \mathbf{f}_e}{\partial \mathbf{e}}\right|_{\substack{\mathbf{e}=\mathbf{0}\\ \mathbf{v}=\mathbf{0}}} \mathbf{e} + \left.\frac{\partial \mathbf{f}_e}{\partial \mathbf{v}}\right|_{\substack{\mathbf{e}=\mathbf{0}\\ \mathbf{v}=\mathbf{0}}} \mathbf{v} + \boldsymbol{\epsilon} \tag{11.396}$$

$$= \dot{\mathbf{r}} - \mathbf{f}(t, \mathbf{r}, \mathbf{u}_0) + \left.\frac{\partial \mathbf{f}_e}{\partial \mathbf{e}}\right|_{\substack{\mathbf{e}=\mathbf{0}\\ \mathbf{v}=\mathbf{0}}} \mathbf{e} + \left.\frac{\partial \mathbf{f}_e}{\partial \mathbf{v}}\right|_{\substack{\mathbf{e}=\mathbf{0}\\ \mathbf{v}=\mathbf{0}}} \mathbf{v} + \boldsymbol{\epsilon} \tag{11.397}$$

$$= \mathbf{A}_e(t)\ \mathbf{e} + \mathbf{B}_e(t)\ \mathbf{v} + \boldsymbol{\epsilon}, \tag{11.398}$$

Since, based on the definition of $\mathbf{u}_0$, we have $\mathbf{f}_e(t, \mathbf{0}, \mathbf{0}) = \dot{\mathbf{r}} - \mathbf{f}(t, \mathbf{r}, \mathbf{u}_0) = \mathbf{0}$.

Now let us explore the relation between $\mathbf{A}_e(t)$ and $\mathbf{B}_e(t)$ with the $\mathbf{A}$ and $\mathbf{B}$ matrices as Jacobians of the system dynamics $\mathbf{f}(t, \mathbf{x}, \mathbf{u})$.

Error Dynamics (contd.)

Noting that $\partial \mathbf{x}/\partial \mathbf{e} = -\mathbf{I}$ and $\partial \mathbf{u}/\partial \mathbf{v} = -\mathbf{I}$, for the Jacobians of $\mathbf{f}_e$ we have

$$\left.\frac{\partial \mathbf{f}_e}{\partial \mathbf{e}}\right|_{\substack{\mathbf{e}=\mathbf{0}\\ \mathbf{v}=\mathbf{0}}} = \left.\left(\left[\frac{\partial \dot{\mathbf{r}}}{\partial \mathbf{x}} - \frac{\partial \mathbf{f}}{\partial \mathbf{x}}\right]\frac{\partial \mathbf{x}}{\partial \mathbf{e}}\right)\right|_{\substack{\mathbf{e}=\mathbf{0}\\ \mathbf{v}=\mathbf{0}}} = -\left.\left(\frac{\partial}{\partial \mathbf{x}}\mathbf{f}(t,\mathbf{x},\mathbf{u})\right)\right|_{\substack{\mathbf{x}=\mathbf{r}(t)\\ \mathbf{u}=\mathbf{u}_0(t)}} \frac{\partial \mathbf{x}}{\partial \mathbf{e}} = \left.\frac{\partial \mathbf{f}}{\partial \mathbf{x}}\right|_{\substack{\mathbf{x}=\mathbf{r}(t)\\ \mathbf{u}=\mathbf{u}_0(t)}} \tag{11.399}$$

$$\left.\frac{\partial \mathbf{f}_e}{\partial \mathbf{v}}\right|_{\substack{\mathbf{e}=\mathbf{0}\\ \mathbf{v}=\mathbf{0}}} = \left.\left(\left[\frac{\partial \dot{\mathbf{r}}}{\partial \mathbf{u}} - \frac{\partial \mathbf{f}}{\partial \mathbf{u}}\right]\frac{\partial \mathbf{u}}{\partial \mathbf{v}}\right)\right|_{\substack{\mathbf{e}=\mathbf{0}\\ \mathbf{v}=\mathbf{0}}} = -\left.\left(\frac{\partial}{\partial \mathbf{u}}\mathbf{f}(t,\mathbf{x},\mathbf{u})\right)\right|_{\substack{\mathbf{x}=\mathbf{r}(t)\\ \mathbf{u}=\mathbf{u}_0(t)}} \frac{\partial \mathbf{u}}{\partial \mathbf{v}} = \left.\frac{\partial \mathbf{f}}{\partial \mathbf{u}}\right|_{\substack{\mathbf{x}=\mathbf{r}(t)\\ \mathbf{u}=\mathbf{u}_0(t)}} \tag{11.400}$$

Error Dynamics (contd.)

In summary, let $\mathbf{e} := \mathbf{r} - \mathbf{x}$ and $\mathbf{v} := \mathbf{u}_0 - \mathbf{u}$, then

$$\dot{\mathbf{x}} = \mathbf{f}(t, \mathbf{x}, \mathbf{u}) \qquad \dot{\mathbf{e}} = \dot{\mathbf{r}} - \mathbf{f}(t,\ \mathbf{r} - \mathbf{e},\ \mathbf{u}_0 - \mathbf{v}).$$

and for the linear approximation of the error dynamics we have

$$\dot{\mathbf{e}} \approx \mathbf{A}_e\ \mathbf{e} + \mathbf{B}_e\ \mathbf{v},$$

where

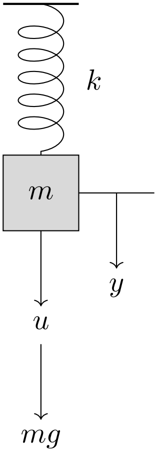
$$\mathbf{A}_e = \frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{r}, \mathbf{u}_0), \qquad \mathbf{B}_e = \frac{\partial \mathbf{f}}{\partial \mathbf{v}}(\mathbf{r}, \mathbf{u}_0).$$

Note that the control input to the system is $\mathbf{u} = \mathbf{u}_0 - \mathbf{v}$.

Example

Consider the hanging weight problem with the dynamics

$$m\ddot{y} = -ky + mg + u. \tag{11.401}$$



Let $x_1 = y$ and $x_2 = \dot{y}$, and for simplicity assume $m = 1$. Accordingly we have

$$\dot{x}_1 = x_2, \tag{11.402}$$
$$\dot{x}_2 = -kx_1 + u + g. \tag{11.403}$$

Let $\mathbf{r}$ be a reference state and define $\mathbf{e} = \mathbf{r} - \mathbf{x}$. Accordingly, we can form the error dynamics as

$$\dot{e}_1 = \dot{r}_1 - f_1(\mathbf{r} - \mathbf{e}, u) = \dot{r}_1 - r_2 + e_2 \tag{11.404}$$
$$\dot{e}_2 = \dot{r}_2 - f_2(\mathbf{r} - \mathbf{e}, u) = \dot{r}_2 + k(r_1 - e_1) - g - u. \tag{11.405}$$

Contents

Some Common Nonlinear Behaviors

Multiple Equilibria

Limit Cycles

Limit cycles are oscillations of fixed amplitude and fixed period that are self-excited or caused without external excitation.

An example of a limit cycle is the following famous oscillator dynamics, first studied in the 1920's by the Dutch electrical engineer Balthasar Van der Pol.

$$\ddot{x} + \mu(x^2 - 1)\dot{x} + kx = 0, \tag{12.406}$$

where $\mu, k > 0$, is known as *Van der Pol equation.* It can be considered as a mass-spring-damper system with a nonlinear damping (or an RLC electrical circuit with a nonlinear resistor).

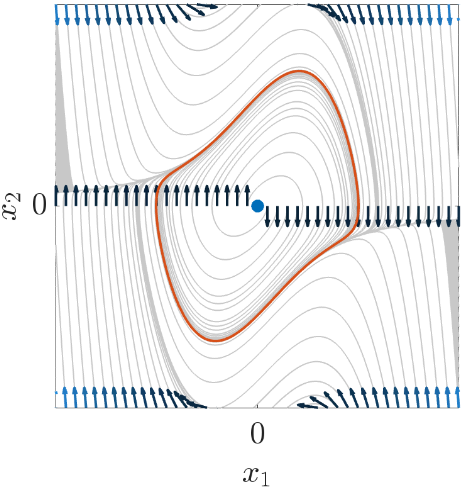
- $|x| > 1 \implies 2c(x^2 - 1) > 0$ and the system loses energy,
- $|x| < 1 \implies 2c(x^2 - 1) < 0$ and system gains energy.

Limit Cycles (contd.)

Limit cycles are sustained by periodically absorbing and releasing energy from and to the environment(in contrast with the conservative mass-spring systems where the rate of change of energy is zero).

Sustained oscillations can also be found in linear systems. However, limit cycles are different from linear oscillations:

- the amplitude of the self-sustained excitation is independent of the initial condition,
- marginally stable linear systems are sensitive to changes in system parameters, while limit cycles are not easily affected by parameter changes.



Invariant Set

Definition 12.1

A set M is an *invariant set* with respect to $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ if

$$\mathbf{x}(0) \in M \implies \mathbf{x}(t) \in M \quad \forall t \in \mathbb{R}.$$

In other words, if a set M is an invariant set, then if a solution belongs to M at some time instant, then it belongs to M for all future and past time. Some examples of invariant sets are:

- Any equilibrium point of the system,
- Limit cycles
- Trivially, the whole state space

Invariant Set (contd.)

Definition 12.2 (Positive Limit Set)

A point $\mathbf{p}$ is said to be a *positive limit point* of $\mathbf{x}(t)$ if there is a sequence $\{t_n\}$, such that

- $t_n \to \infty$ as $n \to \infty$,
- $\mathbf{x}(t_n) \to \mathbf{p}$

The set of all positive limit points of $\mathbf{x}(t)$ is called the *positive limit set* of $\mathbf{x}(t)$.

Definition 12.3 (Positively Invariant Set)

A set M is said to be a positively invariant set with respect to $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ if

$$\mathbf{x}(0) \in M \implies \mathbf{x}(t) \in M, \quad \forall t \geq 0.$$

Invariant Set Theorem

Theorem 12.1 (LaSalle's Invariant Set)

Let $\Omega \subset D$ be a compact set that is positively invariant with respect to $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$. Let $V : D \to \mathbb{R}$ be a continuously differentiable function such that

$$\dot{V}(\mathbf{x}) \leq 0, \quad \forall \mathbf{x} \in \Omega,$$

and let E be the set of all points in Ω where $\dot{V}(\mathbf{x}) = 0$. That is

$$E = \left\{ \mathbf{x} \in \Omega \mid \dot{V}(\mathbf{x}) = 0 \right\}.$$

Let M be the largest invariant set in E. Then every solution starting in Ω approaches M as $t \to \infty$.

For proof see: Khalil, H.K. "Nonlinear Systems". third edition. 2002.

Invariant Set Theorem (contd.)

Unlike Lyapunov's theorem, LaSalle's Invariant Set theorem does not require the function $V(\mathbf{x})$ to be positive definite.

The construction of the set Ω is not necessarily tied to the construction of the function $V(\mathbf{x})$. However, in many applications we can use $V(\mathbf{x})$ itself to construct Ω.

In particular, let

$$\Omega_c := \left\{\mathbf{x} \in \mathbb{R}^n \mid V(\mathbf{x}) < c\right\} \tag{12.407}$$

If Ω_c is bounded and $\dot{V}(\mathbf{x}) \leq 0$ in Ω_c, then we can take $\Omega = \Omega_c$

When $V(\mathbf{x})$ is positive definite, Ω_c is bounded for sufficiently small $c > 0$. This is not necessarily true when $V(\mathbf{x})$ is not positive definite.

Invariant Set Theorem (contd.)

Alternatively, we can state the theorem as:

see Slotine JJ, Li W. "Applied Nonlinear Control". 1991

Consider an autonomous system $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$, with $\mathbf{f}$ continuous. let $V(\mathbf{x})$ be a scalar function with continuous first partial derivatives. Assume

- for some $c > 0$, the set $\Omega_c = \left\{\mathbf{x} \in \mathbb{R}^n \mid V(\mathbf{x}) < c\right\}$ is bounded
- $\dot{V}(\mathbf{x}) \leq 0$ for all $\mathbf{x} \in \Omega_c$

Let $E \subset \Omega_c$ be the set of all points within $\mathbf{x} \in \Omega_c$ where $\dot{V}(\mathbf{x}) = 0$, and M be the largest invariant set in E. Then, every solution $\mathbf{x}(t)$ originating in Ω_c tends to M as $t \to \infty$.

Barbashin and Krasovskii Theorems

The following two corollaries of Theorem 12.1 are known as the theorems of Barbashin and Krasovskii, who proved them before the introduction of LaSalle's invariance principle.

Corollary 12.1.1

Let $\mathbf{x} = \mathbf{0}$ be an equilibrium point for $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$. Let $V : D \to \mathbb{R}$ be a continuously differentiable positive definite function on $D \subset \mathbb{R}^n$, with $\mathbf{0} \in D$, such that $\dot{V}(x) \leq 0$ in D. Let $S = \left\{ \mathbf{x} \in D \mid \dot{V}(\mathbf{x}) = 0 \right\}$ where no solution can stay identically in S, other than the trivial solution $\mathbf{x}(t) = 0$. Then, the origin is asymptotically stable.

Corollary 12.1.2

Let $\mathbf{x} = \mathbf{0}$ be an equilibrium point for $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$. Let $V : \mathbb{R}^n \to \mathbb{R}$ be a continuously differentiable, radially unbounded, positive definite function such that $\dot{V}(\mathbf{x}) \leq 0$ for all $\mathbf{x} \in \mathbb{R}^n$. Let $S = \left\{ x \in \mathbb{R}^n \mid \dot{V}(\mathbf{x}) = 0 \right\}$ where no solution can stay identically in S, other than the trivial solution $\mathbf{x}(t) = \mathbf{0}$. Then, the origin is globally asymptotically stable.

Region of Attraction

also known as: region of asymptotic stability, domain of attraction, or basin

Region of Attraction

Definition 12.4

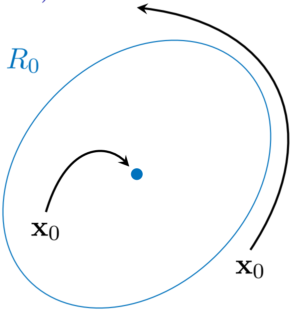
Let the origin $\mathbf{x} = \mathbf{0}$ be an asymptotically stable equilibrium point for the system

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}), \tag{12.408}$$

where $\mathbf{f} : D \to \mathbb{R}^n$ is locally Lipschitz and $\mathbf{0} \in D \subseteq \mathbb{R}^n$. Let $\mathbf{\Phi}(t; \mathbf{x}_0)$ be the solution of (12.408) that starts at initial state $\mathbf{x}_0$ at time $t = 0$. The *region of attraction of the origin*, denoted by R_A, is defined by

$$R_A = \left\{ \mathbf{x}_0 \in D \;\middle|\; \mathbf{\Phi}(t; \mathbf{x}_0) \text{ is defined } \forall t \geq 0 \text{ and } \lim_{t \to \infty} \|\mathbf{\Phi}(t; \mathbf{x}_0)\| = 0 \right\}.$$

Region of Attraction (contd.)



Lemma 12.2

If $\mathbf{x} = \mathbf{0}$ is an asymptotically stable equilibrium point for $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$, then its region of attraction R_A is an open, connected, invariant set where its boundary is formed by trajectories.

For proof of the lemma see: Khalil, H.K. "Nonlinear Systems". third edition. 2002

Example 1 - Van der Pol Equation in Reverse Time

Adopted from Khalil, H.K. “Nonlinear Systems”. third edition. 2002

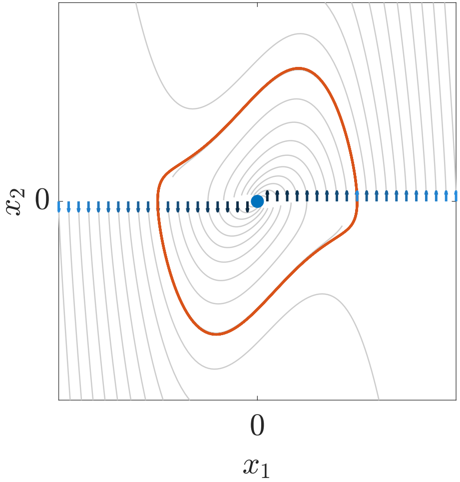
Consider the system

$$\dot{x}_1 = -x_2,$$
$$\dot{x}_2 = x_1 + \left(x_1^2 - 1\right) x_2,$$

which is a Van der Pol equation in reverse time (t replaced by $-t$). The system has one equilibrium point at $\mathbf{x} = \mathbf{0}$ and one unstable limit cycle. The origin is a stable focus, since

$$\left.\frac{\partial \mathbf{f}}{\partial \mathbf{x}}\right|_{\mathbf{x}=\mathbf{0}} = \begin{bmatrix} 0 & -1 \\ 1 & -1 \end{bmatrix}$$

with eigenvalues $1/2 \pm i\sqrt{3}/2$.



Example 2

Adopted from Khalil, H.K. “Nonlinear Systems”. third edition. 2002

Consider the second order system

$$\dot{x}_1 = x_2, \tag{12.409a}$$

$$\dot{x}_2 = -x_1 + \frac{1}{3}x_1^3 - x_2. \tag{12.409b}$$

Solving for $\dot{x}_1 = \dot{x}_2 = 0$ gives the equilibrium points of the system as

$$\bar{\mathbf{x}} \in \left\{ (0,0)\,, \left(\sqrt{3},0\right), \left(-\sqrt{3},0\right) \right\}. \tag{12.410}$$

Example 2 (contd.)

Adopted from Khalil, H.K. "Nonlinear Systems". third edition. 2002

For each of the three equilibrium points we have

$$\left.\frac{\partial \mathbf{f}}{\partial \mathbf{x}}\right|_{\mathbf{x}=(0,0)} = \begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix} \implies \lambda_1, \lambda_2 = -\frac{1}{2} \pm i\frac{\sqrt{3}}{2} \implies (0,0) \text{ is a stable focus,} \tag{12.411}$$

$$\left.\frac{\partial \mathbf{f}}{\partial \mathbf{x}}\right|_{\mathbf{x}=(-\sqrt{3},0)} = \begin{bmatrix} 0 & 1 \\ 2 & -1 \end{bmatrix} \implies \lambda_1 = 1,\ \lambda_2 = -2 \implies \left(-\sqrt{3}, 0\right) \text{ is a saddle point,} \tag{12.412}$$

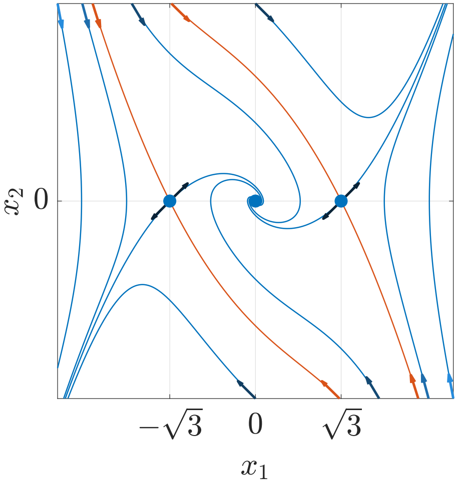
$$\left.\frac{\partial \mathbf{f}}{\partial \mathbf{x}}\right|_{\mathbf{x}=(\sqrt{3},0)} = \begin{bmatrix} 0 & 1 \\ 2 & -1 \end{bmatrix} \implies \lambda_1 = 1,\ \lambda_2 = -2 \implies \left(\sqrt{3}, 0\right) \text{ is a saddle point.} \tag{12.413}$$

Example 2 (contd.)

Adopted from Khalil, H.K. "Nonlinear Systems". third edition. 2002

From the phase portrait, we see that the stable trajectories of the saddle points form two separatrices[‖] that are the boundaries of the region of attraction.

In this example, R_A is unbounded and its boundary is formed of stable trajectories of saddle points.



---

*a *separatrix* is the boundary separating two modes of behaviour in a differential equation.

Example 3

Adopted from Khalil, H.K. “Nonlinear Systems”. third edition. 2002

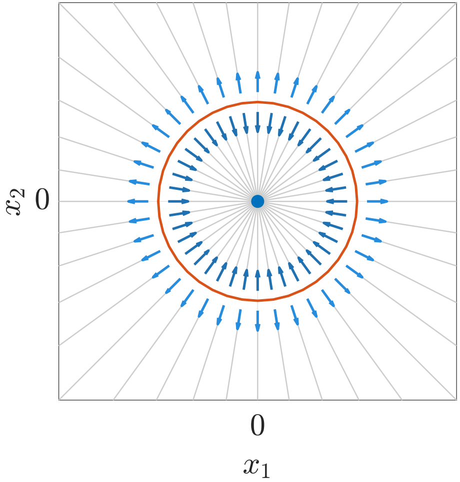
Consider the system

$$\dot{x}_1 = -x_1(1 - x_1^2 - x_2^2), \tag{12.414a}$$

$$\dot{x}_2 = -x_2(1 - x_1^2 - x_2^2), \tag{12.414b}$$

which has an isolated equilibrium point at the origin and a continuum of equilibrium points on the unit circle;

Note that in the case, the boundary of R_A is a closed curve of equilibrium points.



Example 3 (contd.)

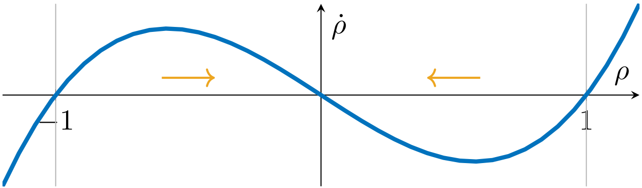
Adopted from Khalil, H.K. “Nonlinear Systems”. third edition. 2002

Alternatively, we can study the system using the change of variable

$$x_1 = \rho\cos(\theta), \qquad x_2 = \rho\sin(\theta), \tag{12.415a}$$

which leads to

$$\dot{\rho} = -\rho\left(1-\rho^2\right), \qquad \dot{\theta} = 0. \tag{12.416a}$$



Boundary Determination for the Region of Attraction

Determination of R_A via Lyapunov's Theory

We can find estimates of R_A by using Lyapunov's method. That is, we want to find $\Omega \subset R_A$ such that

$$\mathbf{x}_0 \in \Omega \implies \boldsymbol{\Phi}(t; \mathbf{x}_0) \to \mathbf{0} \quad \text{as} \quad t \to \infty.$$

Recall, based on Lyapunov's theorem for autonomous systems we have:

if D is a domain that contains the origin, and if we can find a Lyapunov function $V(\mathbf{x})$ that is positive definite in D and $\dot{V}(\mathbf{x})$ is negative definite in D [a], then the origin is asymptotically stable.

[a] or negative semidefinite, but no solution can stay identically in the set $\{V(\mathbf{x}) = 0\}$ except for $\mathbf{x} = \mathbf{0}$

⚠ We may incorrectly conclude that D is an estimate of R_A. However, this conjecture is not true!

Let us explore this in an example.

Example

Adopted from Khalil, H.K. "Nonlinear Systems". third edition. 2002

Consider the system in the previous example:

$$\dot{x}_1 = x_2, \tag{12.417a}$$

$$\dot{x}_2 = -x_1 + \frac{1}{3}x_1^3 - x_2. \tag{12.417b}$$

let us proceed with Lyapunov's theory. Let

$$V(\mathbf{x}) = \frac{3}{4}x_1^2 - \frac{1}{12}x_1^4 + \frac{1}{2}x_1x_2 + \frac{1}{2}x_2^2, \tag{12.418}$$

with the time derivative defined as

$$\dot{V}(\mathbf{x}) = -\frac{1}{2}x_1^2\left(1 - \frac{1}{3}x_1^2\right) - \frac{1}{2}x_2^2. \tag{12.419}$$

Example (contd.)

Adopted from Khalil, H.K. "Nonlinear Systems". third edition. 2002

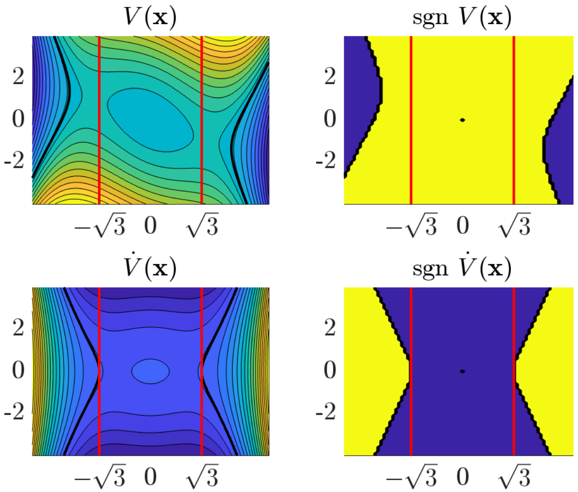
$$V(\mathbf{x}) = \frac{3}{4}x_1^2 - \frac{1}{12}x_1^4 + \frac{1}{2}x_1x_2 + \frac{1}{2}x_2^2,$$
$$\dot{V}(\mathbf{x}) = -\frac{1}{2}x_1^2\left(1 - \frac{1}{3}x_1^2\right) - \frac{1}{2}x_2^2.$$

Let D defined as

$$D := \left\{(x_1, x_2) \mid -\sqrt{3} \le x_1 \le \sqrt{3}\right\}.$$

Then, for all $\mathbf{x} \in D \setminus \{\mathbf{0}\}$ we have

$$V(\mathbf{x}) > 0,$$
$$\dot{V}(\mathbf{x}) < 0.$$



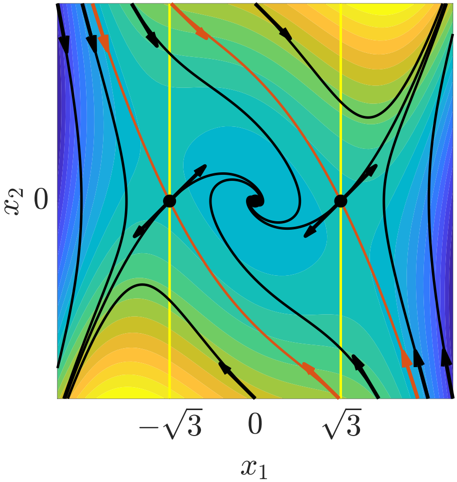
Example (contd.)

Adopted from Khalil, H.K. "Nonlinear Systems". third edition. 2002

Clearly, $D \not\subset R_A$

> Even though a trajectory starting in D will move from one Lyapunov surface $V(\mathbf{x}) = c_1$ to an inner Lyapunov surface $V(\mathbf{x}) = c_2 < c_1$, there is no guarantee that the trajectory will remain forever in D.

This problem does not arise when R_A is estimated by a compact positively invariant subset of D where $\Omega \subset D$ such that every trajectory starting in Ω stays in Ω for all future time.



A Conservative Approach to Estimate R_A

Clearly Ω in Theorem 12.1 is a subset of R_A. Accordingly, the simplest estimate is the set

$$\Omega_c := \left\{\mathbf{x} \in \mathbb{R}^n \mid V(\mathbf{x}) \leq c\right\} \quad \text{when } \Omega_c \subset D \text{ is bounded.} \tag{12.420}$$

Quadratic Lyapunov Function on Hyperspherical Domain

For function $V(\mathbf{x}) = \mathbf{x}^T\mathbf{P}\mathbf{x}$ and $D = \left\{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\|_2 < r\right\}$, choosing c as

$$c < \min_{\|\mathbf{x}\|_2 = r} \mathbf{x}^T\mathbf{P}\mathbf{x} = \lambda_{\min}(\mathbf{P})r^2, \text{ guarantees that } \Omega_c \subset D. \tag{12.421}$$

To obtain the results in (12.421) we use Rayleigh quotient:

Rayleigh Quotient

For a given complex Hermitian matrix $\mathbf{M}$ and nonzero vector $\mathbf{x}$,

$$\lambda_{\min}(\mathbf{M}) \leq \frac{\mathbf{x}^T\mathbf{M}\mathbf{x}}{\mathbf{x}^T\mathbf{x}} \leq \lambda_{\max}(\mathbf{M})$$

A Conservative Approach to Estimate R_A (contd.)

Quadratic Lyapunov Function on Half-Space

For $D = \left\{\mathbf{x} \in \mathbb{R}^n \mid |\mathbf{b}^T\mathbf{x}| < r\right\}$, where $\mathbf{b} \in \mathbb{R}^n$,

$$\min_{|\mathbf{b}^T\mathbf{x}|=r} \mathbf{x}^T\mathbf{P}\mathbf{x} = \frac{r^2}{\mathbf{b}^T\mathbf{P}^{-1}\mathbf{b}} \tag{12.422}$$

Therefore, $\left\{\mathbf{x} \mid \mathbf{x}^T\mathbf{P}\mathbf{x} \leq c\right\}$ will be a subset of $D = \left\{\mathbf{x} \mid |\mathbf{b}_i^T\mathbf{x}| < r_i, i = 1, \ldots, p\right\}$, if

$$c < \min_i \frac{r_i^2}{\mathbf{b}_i^T\mathbf{P}^{-1}\mathbf{b}_i} \tag{12.423}$$

A Conservative Approach to Estimate R_A (contd.)

From the discussions on linearization we obtained: if the Jacobian matrix

$$\mathbf{A} = \left.\frac{\partial \mathbf{f}}{\partial \mathbf{x}}\right|_{\mathbf{x}=\mathbf{0}} \tag{12.424}$$

is Hurwitz, then we can always find a quadratic Lyapunov function $V(\mathbf{x}) = \mathbf{x}^T\mathbf{P}\mathbf{x}$ by solving the Lyapunov equation $\mathbf{P}\mathbf{A} + \mathbf{A}^T\mathbf{P} = -\mathbf{Q}$ for any positive definite matrix $\mathbf{Q}$. Hence, if $\mathbf{A}$ is Hurwitz, we can estimate the region of attraction of the origin via

$$\Omega_c = \left\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x}^T\mathbf{P}\mathbf{x} \leq c\right\}$$

Example

Global Asymptotic Stability

As we discussed Ω_c is an estimate of the region of attraction. However, Ω_c may be a conservative estimate; Meaning, Ω_c may be much smaller than the actual region of attraction of the system.

? Under what conditions will the region of attraction be the whole space $\mathbb{R}^n$?

Definition 12.5 (Globally Asymptotically Stable Equilibrium Point)

Let $\bar{\mathbf{x}}$ be *the* equilibrium point of the system $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$. We say $\bar{\mathbf{x}}$ is *globally asymptotically stable* equilibrium point of the system if for any initial state $\mathbf{x}_0$, the trajectory $\boldsymbol{\Phi}(t, \mathbf{x}_0)$ approaches $\bar{\mathbf{x}}$ as $t \to \infty$.

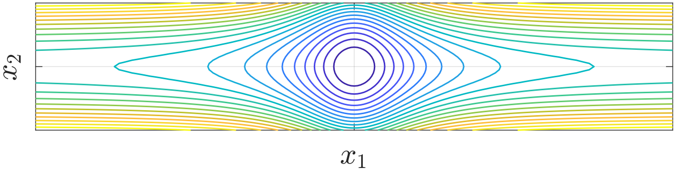
we can see that global asymptotic stability can be established if any point $\mathbf{x} \in \mathbb{R}^n$ can be included in the interior of a bounded set Ω_c. Obviously, this requires that the conditions for Lyapunov's stability theorem hold globally ($D = \mathbb{R}^n$). However, if $V(\mathbf{x})$ is not carefully chosen, then for large c, the set Ω_c is not necessarily bounded.

Global Asymptotic Stability (contd.)

As an example consider,

$$V(\mathbf{x}) = \frac{x_1^2}{1 + x_1^2} + x_2^2. \tag{12.425}$$

As depicted in the level sets, owing to the continuity and positive definiteness of $V(\mathbf{x})$, for small c, the set $V(\mathbf{x}) = c$ is closed and $\Omega_c \subset B_r$ for some $0 < r < \infty$ is bounded.



For Ω_c to be in the interior of a ball B_r, then

$$c < \inf_{\|\mathbf{x}\| > r} V(\mathbf{x}). \tag{12.426}$$

Global Asymptotic Stability (contd.)

Let l defined as

$$l := \lim_{r\to\infty} \inf_{\|x\|\geq r} V(x), \tag{12.427}$$

If $l < \infty$, then Ω_c will be bounded if $c < l$. In this example

$$l = \lim_{r\to\infty} \min_{\|x\|=r} \left[\frac{x_1^2}{1+x_1^2} + x_2^2\right] = \lim_{|x_1|\to\infty} \frac{x_1^2}{1+x_1^2} = 1 \tag{12.428}$$

Thus, Ω_c is bounded only for $c < 1$. An extra condition that ensures that Ω_c is bounded for all values of $c > 0$ is

$$V(x) \to \infty \quad \text{as} \quad \|x\| \to \infty \tag{12.429}$$

A function satisfying this condition is said to be *radially unbounded*.

Global Asymptotic Stability (contd.)

Theorem 12.3 (Barbashin-Krasovskii Theorem)

Let $\mathbf{x} = \mathbf{0}$ *be an equilibrium point for*

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}). \tag{12.430}$$

Let $V : \mathbb{R}^n \to \mathbb{R}$ *be continuously differentiable function such that*

$$V(\mathbf{0}) = 0, \quad \textit{and} \quad V(\mathbf{x}) > 0 \quad \forall \mathbf{x} \neq \mathbf{0}, \tag{12.431}$$

$$\|\mathbf{x}\| \to \infty \implies V(\mathbf{x}) \to \infty, \tag{12.432}$$

$$\dot{V}(\mathbf{x}) < 0, \quad \forall \mathbf{x} \neq \mathbf{0}, \tag{12.433}$$

then $\mathbf{x} = \mathbf{0}$ *is globally asymptotically stable.*

Contents

Position Control of Manipulators via Gravity Compensation

The general form of the manipulator dynamics is

$$\mathbf{M}(\mathbf{q})\ddot{\mathbf{q}} + \mathbf{b}(\mathbf{q}, \dot{\mathbf{q}}) + \mathbf{g}(\mathbf{q}) = \boldsymbol{\tau}, \qquad (13.434)$$

where $\mathbf{q} \in \mathbb{R}^n$ denotes the generalized coordinates of the arm (vector of n joint variables), $\boldsymbol{\tau} \in \mathbb{R}^n$ is the input forces and torques applied to the joints, $\mathbf{M}$ is an $n \times n$ positive definite inertia matrix. $\mathbf{b}$ includes Coriolis and centripetal forces and $\mathbf{g}$ includes torques due to the gravity.

Image

Position Control of Manipulators via Gravity Compensation (contd.)

Let $\mathbf{r}$ denote the vector of desired joint angles of the robot. Assume a simple PD control function

$$\boldsymbol{\tau} = \mathbf{K}_p(\mathbf{r} - \mathbf{q}) - \mathbf{K}_d\dot{\mathbf{q}}, \tag{13.435}$$

Where $\mathbf{K}_p$ and $\mathbf{K}_d$ are symmetric positive definite matrices.

Let us explore the efficacy of this controller.
To obtain the equilibrium points of the closed loop system, we set $\ddot{\mathbf{q}} = \dot{\mathbf{q}} = \mathbf{0}$ in (13.434).
Accordingly, the equilibrium points of the closed loop system are the solutions of

$$\mathbf{q} = \mathbf{r} - \mathbf{K}_p^{-1}\mathbf{g}(\mathbf{q}), \tag{13.436}$$

since, by assumption $\mathbf{K}_p$ is non-singular. Note that the term $\mathbf{K}_p^{-1}\mathbf{g}(\mathbf{q})$ defines the steady state error of the closed loop system.

Position Control of Manipulators via Gravity Compensation (contd.)

In order to eliminate this steady state error, we can reformulate the control function as

$$\boldsymbol{\tau} = \mathbf{K}_p(\mathbf{r} - \mathbf{q}) - \mathbf{K}_d\dot{\mathbf{q}} + \mathbf{g}(\mathbf{q}). \tag{13.437}$$

Substituting the above controller in (13.434), and setting $\ddot{\mathbf{q}} = \dot{\mathbf{q}} = \mathbf{0}$, we find the closed loop equilibrium points as

$$\mathbf{K}_p(\mathbf{r} - \mathbf{q}) = \mathbf{0} \implies \mathbf{q} = \mathbf{r}, \tag{13.438}$$

owing to $\mathbf{K}_p$ not being singular. Hence, the reference position is an equilibrium point of the closed loop system with gravity compensation.

Before we proceed to investigate the asymptotic stability of the equilibrium point $(\mathbf{q} = \mathbf{r}, \dot{\mathbf{q}} = \mathbf{0})$, let us study the rate of change of energy for the manipulator system.

The Rate of Change of Energy of Rigid Body Systems

Consider a mechanical system with the dynamics

$$\mathbf{M}(\mathbf{q})\ddot{\mathbf{q}} + \mathbf{b}(\mathbf{q}, \dot{\mathbf{q}}) + \mathbf{g}(\mathbf{q}) = \boldsymbol{\tau}, \tag{13.439}$$

where

$$\mathbf{b}(\mathbf{q}, \dot{\mathbf{q}}) = \mathbf{C}(\mathbf{q}, \dot{\mathbf{q}})\dot{\mathbf{q}} \qquad \text{and} \qquad \mathbf{g}(\mathbf{q}) = \frac{\partial U(\mathbf{q})}{\partial \mathbf{q}}, \tag{13.440}$$

where $U(\mathbf{q})$ is the potential energy of the system. Hence, the total energy of the system is

$$E = \frac{1}{2}\dot{\mathbf{q}}^T \mathbf{M}\dot{\mathbf{q}} + U(\mathbf{q}) \tag{13.441}$$

The Rate of Change of Energy of Rigid Body Systems (contd.)

Taking the derivative of E with respect to time yields

$$\dot{E} = \dot{\mathbf{q}}^T \mathbf{M}\ddot{\mathbf{q}} + \frac{1}{2}\dot{\mathbf{q}}^T \dot{\mathbf{M}}\dot{\mathbf{q}} + \dot{\mathbf{q}}^T \frac{\partial U}{\partial \mathbf{q}} \tag{13.442}$$

$$= \dot{\mathbf{q}}^T \left(\boldsymbol{\tau} + \frac{1}{2}\dot{\mathbf{M}}\dot{\mathbf{q}} - \mathbf{C}(\mathbf{q}, \dot{\mathbf{q}})\dot{\mathbf{q}} - \mathbf{g}(\mathbf{q}) \right) + \dot{\mathbf{q}}^T \mathbf{g}(\mathbf{q}) \tag{13.443}$$

$$= \dot{\mathbf{q}}^T \boldsymbol{\tau} + \frac{1}{2}\dot{\mathbf{q}}^T \left(\dot{\mathbf{M}} - 2\mathbf{C}(\mathbf{q}, \dot{\mathbf{q}}) \right) \dot{\mathbf{q}}, \tag{13.444}$$

where $\dot{\mathbf{M}} - 2\mathbf{C}(\mathbf{q}, \dot{\mathbf{q}})$ is a skew symmetric matrix (for details and proof see: "Robotics: Modelling, Planning and Control" by Bruno Siciliano, Lorenzo Sciavicco, Luigi Villani, and Giuseppe Oriolo)

The Rate of Change of Energy of Rigid Body Systems (contd.)

Let $\mathbf{M} \in \mathbb{R}^{n\times n}$ be a skew symmetric matrix, that is, $\mathbf{M}^T = -\mathbf{M}$. For any $\mathbf{x} \in \mathbb{R}^n$, the quadratic form $\mathbf{x}^T\mathbf{M}\mathbf{x} \in \mathbb{R}$. Hence

$$a = \mathbf{x}^T\mathbf{M}\mathbf{x} = \left(\mathbf{x}^T\mathbf{M}\mathbf{x}\right)^T = \mathbf{x}^T\mathbf{M}^T\mathbf{x} = -\mathbf{x}^T\mathbf{M}\mathbf{x} = -a,$$

which implies $a = 0$. Thus, for any $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{x}^T\mathbf{M}\mathbf{x} = 0$.

Since $\dot{\mathbf{M}} - 2\mathbf{C}(\mathbf{q}, \dot{\mathbf{q}})$ is a skew symmetric matrix, we get

$$\frac{1}{2}\dot{\mathbf{q}}^T\left(\dot{\mathbf{M}} - 2\mathbf{C}(\mathbf{q}, \dot{\mathbf{q}})\right)\dot{\mathbf{q}} = 0, \quad \forall \dot{\mathbf{q}} \in \mathbb{R}^n. \tag{13.445}$$

Accordingly, for the rate of change of the energy we have

$$\dot{E} = \dot{\mathbf{q}}^T\boldsymbol{\tau}. \tag{13.446}$$

Stability Analysis of Gravity Compensated PD Controller

To analyze the stability of the equilibrium point $(\mathbf{q} = \mathbf{r}, \dot{\mathbf{q}} = \mathbf{0})$, let

$$V(\mathbf{q}, \dot{\mathbf{q}}) = \frac{1}{2}\dot{\mathbf{q}}^T \mathbf{M} \dot{\mathbf{q}} + \frac{1}{2}(\mathbf{r} - \mathbf{q})^T \mathbf{K}_p (\mathbf{r} - \mathbf{q}) \tag{13.447}$$

Note that $V(\mathbf{q}, \dot{\mathbf{q}}) \geq 0$ and $V(\mathbf{q}, \dot{\mathbf{q}}) = 0 \iff \mathbf{q} = \mathbf{r}$ and $\dot{\mathbf{q}} = \mathbf{0}$. Thus,

$$\begin{aligned}
\dot{V} &= \dot{\mathbf{q}}^T \mathbf{M} \ddot{\mathbf{q}} + \frac{1}{2}\dot{\mathbf{q}}^T \dot{\mathbf{M}} \dot{\mathbf{q}} - \dot{\mathbf{q}}^T \mathbf{K}_p (\mathbf{r} - \mathbf{q}) && (13.448)\\
&= \dot{\mathbf{q}}^T \left(\boldsymbol{\tau} - \mathbf{C}(\mathbf{q}, \dot{\mathbf{q}})\dot{\mathbf{q}} - \mathbf{g}(\mathbf{q}) + \frac{1}{2}\dot{\mathbf{M}}\dot{\mathbf{q}} - \mathbf{K}_p(\mathbf{r} - \mathbf{q}) \right) && (13.449)\\
&= \dot{\mathbf{q}}^T \left(\mathbf{K}_p(\mathbf{r} - \mathbf{q}) - \mathbf{K}_d\dot{\mathbf{q}} + \mathbf{g}(\mathbf{q}) - \mathbf{C}(\mathbf{q}, \dot{\mathbf{q}})\dot{\mathbf{q}} - \mathbf{g} + \frac{1}{2}\dot{\mathbf{M}}\dot{\mathbf{q}} - \mathbf{K}_p(\mathbf{r} - \mathbf{q}) \right) && (13.450)\\
&= -\dot{\mathbf{q}}^T \mathbf{K}_d \dot{\mathbf{q}} + \frac{1}{2}\dot{\mathbf{q}}^T \left(\dot{\mathbf{M}} - 2\mathbf{C}(\mathbf{q}, \dot{\mathbf{q}}) \right) \dot{\mathbf{q}} = -\dot{\mathbf{q}}^T \mathbf{K}_d \dot{\mathbf{q}} && (13.451)
\end{aligned}$$

Since $(\mathbf{q} = \mathbf{r}, \dot{\mathbf{q}} = \mathbf{0})$ is the only invariant set for the system, based on Corollary 12.1.2, we can conclude that $(\mathbf{q} = \mathbf{r}, \dot{\mathbf{q}} = \mathbf{0})$ is globally asymptotically stable.

Contents

Feedback Linearization

Idea

The core idea of the *feedback linearization* is to algebraically transform a nonlinear system into a fully or partially linear system via a *feedback* loop.

Feedback Linearization vs. Linearization (Linearized Model)

- In linearization, we obtain a linear *approximation* of the system at an equilibrium point.
- Feedback linearization transforms the system to a linear system by compensating for nonlinear terms via feedback loop.

Consider the system

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}). \tag{14.452}$$

If we could find a control function $\mathbf{u} = \mathbf{u}(\mathbf{x}, \mathbf{v})$ such that

$$\dot{\mathbf{x}} = \mathbf{f}\left(\mathbf{x}, \mathbf{u}(\mathbf{x}, \mathbf{v})\right) = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{v}, \tag{14.453}$$

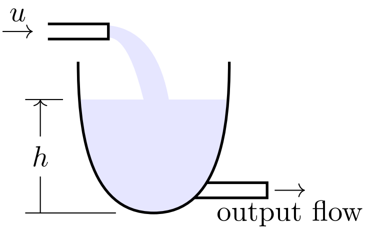
then using linear control theory, we can find $\mathbf{v}$ that satisfies the control objective.

Example - Controlling the Fluid Level in a Tank

Adopted from Slotine JJ, Li W. "Applied Nonlinear Control". 1991

The objective is to control the fluid level in the tank, h, via the input flow rate u, assuming the output flow is known. The system model is

$$A(h)\dot{h} = u - a\sqrt{2gh}, \qquad (14.454)$$



where $A(h)$ is the cross section of the tank and a is the cross section of the outlet pipe. Let

$$u(t) := a\sqrt{2gh} + A(h)v, \qquad (14.455)$$

Substituting (14.455) in (14.454) gives the closed loop system dynamics as

$$A(h)\dot{h} = u - a\sqrt{2gh} = A(h)v \implies \dot{h} = v. \qquad (14.456)$$

Example - Controlling the Fluid Level in a Tank (contd.)

Adopted from Slotine JJ, Li W. “Applied Nonlinear Control”. 1991

Thus, through feedback law (14.455) we could transform the system to a linear system

$$\dot{h} = v, \tag{14.457}$$

Let h_d be the desired fluid level and define $v := k(h_d - h)$, for some $k > 0$. We have

$$\dot{h} = k(h_d - h). \tag{14.458}$$

Let $h(0) = h_0$, the solution to the differential equation is

$$h(t) = \left(1 - e^{-kt}\right) h_d + e^{-kt} h_0 \implies \lim_{t \to \infty} h(t) = h_d. \tag{14.459}$$

Note that the control input to the system is

$$u(t) = a\sqrt{2gh} + A(h)k(h_d - h). \tag{14.460}$$

Example - Controlling the Fluid Level in a Tank (contd.)

Adopted from Slotine JJ, Li W. "Applied Nonlinear Control". 1991

See that the control input $u(t) = a\sqrt{2gh} + A(h)k(h_d - h)$ is composed of two parts:

- $a\sqrt{2gh}$ term adjusts for the output flow,
- $A(h)k(h_d - h)$ is a proportional controller with a non-constant gain.

Tracking Control If the desired level is a known time-varying function $h_d(t)$, we can choose v as

$$v = \dot{h}_d(t) + k(h_d - h). \tag{14.461}$$

Let $e = h_d - h$ and $\dot{e} = \dot{h}_d - \dot{h}$. We have

$$\dot{h} = \dot{h}_d(t) + k(h_d - h) \implies \dot{e} + ke = 0. \tag{14.462}$$

The solution to the error dynamics for the initial condition $e(0) = h_d(0) - h(0)$ is

$$e(t) = e(0)\exp(-kt) \implies \lim_{t\to\infty} e(t) = 0. \tag{14.463}$$

Companion Form (Controllable Canonical Form)

Companion Form (Controllable Canonical Form)

A system is in *companion form* if its dynamics is of the form

$$x^{(n)} = f(\mathbf{x}) + g(\mathbf{x})u, \tag{14.464}$$

where u is the scalar control input and x is scalar output of interest. The vector

$$\mathbf{x} = \left(x,\ \dot{x},\ \ddot{x}, \ldots, x^{(n-1)}\right), \tag{14.465}$$

is the state vector. $f(\mathbf{x})$ and $g(\mathbf{x})$ are nonlinear functions. See that (14.464) is equivalent to

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} x_1 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} = \begin{bmatrix} x_2 \\ \vdots \\ x_n \\ f(\mathbf{x}) + g(\mathbf{x})u \end{bmatrix}. \tag{14.466}$$

Companion Form (Controllable Canonical Form) (contd.)

let us now consider the system

$$x_1^{(n)} = f_1(\mathbf{x}_1, \ldots, \mathbf{x}_p) + \mathbf{g}_1(\mathbf{x}_1, \ldots, \mathbf{x}_p)\mathbf{u}, \tag{14.467a}$$

$$x_2^{(n)} = f_2(\mathbf{x}_1, \ldots, \mathbf{x}_p) + \mathbf{g}_2(\mathbf{x}_1, \ldots, \mathbf{x}_p)\mathbf{u}, \tag{14.467b}$$

$$\vdots$$

$$x_p^{(n)} = f_p(\mathbf{x}_1, \ldots, \mathbf{x}_p) + \mathbf{g}_p(\mathbf{x}_1, \ldots, \mathbf{x}_p)\mathbf{u}, \tag{14.467c}$$

where $\mathbf{u} = (u_1, \ldots,\ u_m)$ is the vector of m inputs. For $i = 1, \ldots, p$, the row vector $\mathbf{g}_i(\mathbf{x}_1, \ldots, \mathbf{x}_p) \in \mathbb{R}^{1\times m}$ is the mapping between $\mathbf{u}$ and $x_i^{(n)}$ where

$$\mathbf{x}_i = \left(x_i,\ \dot{x}_i,\ \ddot{x}_i, \ldots, x_i^{(n-1)}\right). \tag{14.468}$$

Companion Form (Controllable Canonical Form) (contd.)

Accordingly, let $\mathbf{z}_i$ be the vector composed of i-th derivatives of all state variables

$$\mathbf{z}_i = \left(x_1^{(i)}, \ldots, x_p^{(i)}\right), \tag{14.469}$$

and let $\mathbf{z} = (\mathbf{z}_1, \ldots, \mathbf{z}_n)$. Then the system (14.467) is equivalent to

$$\begin{bmatrix} \dot{\mathbf{z}}_1 \\ \vdots \\ \dot{\mathbf{z}}_{n-1} \\ \dot{\mathbf{z}}_n \end{bmatrix} = \begin{bmatrix} \mathbf{z}_2 \\ \vdots \\ \mathbf{z}_n \\ \mathbf{f}(\mathbf{z}) + \mathbf{G}(\mathbf{z})\mathbf{u} \end{bmatrix}, \tag{14.470}$$

where the vector $\mathbf{f}$ and matrix $\mathbf{G}$ are defined as

$$\mathbf{f}(\mathbf{z}) = \begin{bmatrix} f_1(\mathbf{x}_1, \ldots, \mathbf{x}_p) \\ \vdots \\ f_p(\mathbf{x}_1, \ldots, \mathbf{x}_p) \end{bmatrix}, \quad \mathbf{G}(\mathbf{z}) = \begin{bmatrix} \mathbf{g}_1(\mathbf{x}_1, \ldots, \mathbf{x}_p) \\ \vdots \\ \mathbf{g}_p(\mathbf{x}_1, \ldots, \mathbf{x}_p) \end{bmatrix}. \tag{14.471}$$

Feedback Linearization of Systems in Companion Form

Feedback Linearization in Companion Form for SISO Systems

If in (14.464), $g(\mathbf{x}) \neq 0$, then the control function

$$u = \frac{1}{g(\mathbf{x})} \left(v - f(\mathbf{x}) \right), \tag{14.472}$$

transforms the closed-loop system to linear form

$$x^{(n)} = f(\mathbf{x}) + g(\mathbf{x})u = f(\mathbf{x}) + g(\mathbf{x}) \frac{1}{g(\mathbf{x})} \left(v - f(\mathbf{x}) \right) = v, \tag{14.473}$$

which is equivalent to

$$\dot{\mathbf{x}} = \left[\begin{array}{c|c} \mathbf{0}_{n-1\times 1} & \mathbf{I}_{n-1} \\ \hline 0 & \mathbf{0}_{1\times n-1} \end{array} \right] \mathbf{x} + \left[\begin{array}{c} \mathbf{0}_{n-1\times 1} \\ \hline 1 \end{array} \right] v = \mathbf{A}\mathbf{x} + \mathbf{B}v, \tag{14.474}$$

Note that the system $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}v$ is full state controllable.

Feedback Linearization in Companion Form for SISO Systems (contd.)

Tracking Control Let $x_d(t)$ denote a desired trajectory with defined derivatives up to order n. Let the error $e := x_d - x$. For $k_0, \ldots, k_{n-1} > 0$, let v defined as

$$v = x_d^{(n)} + k_{n-1}e^{(n-1)} + \cdots + k_1\dot{e} + k_0 e = x_d^{(n)} + \sum_{i=0}^{n-1} k_i e^{(i)}. \tag{14.475}$$

Accordingly, we get the error dynamics as

$$e^{(n)} + k_{n-1}e^{(n-1)} + \cdots + k_1\dot{e} + k_0 e = 0, \tag{14.476}$$

which exponentially converges to zero in time.

Setpoint Control If $x_d(t) = x_d$ is constant in time, then $x_d^{(i)} = 0$ for all $i \in \mathbb{N}$. Hence, the above control function reduces to

$$v = x_d^{(n)} + k_{n-1}e^{(n-1)} + \cdots + k_1\dot{e} + k_0 e = -\sum_{i=0}^{n-1} k_i x^{(i)} = -\mathbf{K}\mathbf{x}. \tag{14.477}$$

Feedback Linearization in Companion Form for MIMO Systems

For the case of MIMO system (14.470), if $\mathbf{G}(\mathbf{z})$ is non-singular, then via feedback law

$$\mathbf{u} = \mathbf{G}^{-1}(\mathbf{z}) \left(\mathbf{v} - \mathbf{f}(\mathbf{z})\right), \tag{14.478}$$

gives $\mathbf{z}_n = \mathbf{v}$. Let $\mathbf{x}_d(t)$ be a desired trajectory and let $\mathbf{K}_0$ to $\mathbf{K}_{n-1}$ be *diagonal* matrices with positive entries. Choosing $\mathbf{v}$ as

$$\mathbf{v} = \mathbf{x}_d^{(n)}(t) + \sum_{i=0}^{n-1} \mathbf{K}_i \left(\mathbf{x}_d^{(i)} - \mathbf{z}_i\right), \tag{14.479}$$

provides exponential convergence of $\mathbf{e} = \mathbf{x}_d - \mathbf{z}_1$ to zero in time. In fact for $\mathbf{e} = \left(e_1, \ldots, e_p\right)$ we have:

$$e_1^{(n)} + k_{n-1}^1 e_1^{(n-1)} + \cdots + k_1^1 \dot{e}_1 + k_0^1 e_1 = 0, \tag{14.480a}$$

$$\vdots$$

$$e_p^{(n)} + k_{n-1}^p e_p^{(n-1)} + \cdots + k_1^p \dot{e}_p + k_0^p e_p = 0. \tag{14.480b}$$

Feedback Linearization For Robotic Platforms

Feedback Linearization for Lagrangian Systems

Let $\mathbf{q}$ denote the vector of generalized coordinates and $\mathbf{u}$ denote the vector of input force and moments. Consider the dynamics

$$\mathbf{M}(\mathbf{q})\ddot{\mathbf{q}} + \mathbf{\Phi}(\mathbf{q}, \dot{\mathbf{q}}) = \mathbf{T}(\mathbf{q})\mathbf{u}. \tag{14.481}$$

Let $\mathbf{x}_1 = \mathbf{q}$ and $\mathbf{x}_2 = \dot{\mathbf{q}}$. Then the system (14.481) could be presented as

$$\dot{\mathbf{x}}_1 = \mathbf{x}_2, \tag{14.482a}$$

$$\dot{\mathbf{x}}_2 = \mathbf{M}^{-1}\left(\mathbf{T}\mathbf{u} - \mathbf{\Phi}\right). \tag{14.482b}$$

If $\mathbf{T}$ *is invertible*, choosing $\mathbf{u}$ as

$$\mathbf{u} = \mathbf{T}^{-1}\left(\mathbf{M}\mathbf{v} + \mathbf{\Phi}\right), \tag{14.483}$$

gives

$$\dot{\mathbf{x}}_1 = \mathbf{x}_2, \tag{14.484a}$$

$$\dot{\mathbf{x}}_2 = \mathbf{M}^{-1}\left(\mathbf{T}\mathbf{T}^{-1}\left(\mathbf{M}\mathbf{v} + \mathbf{\Phi}\right) - \mathbf{\Phi}\right) = \mathbf{v}. \tag{14.484b}$$

Feedback Linearization for Lagrangian Systems (contd.)

Thus, the closed loop system is

$$\begin{bmatrix} \dot{\mathbf{x}}_1 \\ \dot{\mathbf{x}}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ \mathbf{I} \end{bmatrix} \mathbf{v}, \tag{14.485}$$

which is full state controllable. Hence, we can use linear control design techniques, such as pole-placement, to define $\mathbf{v}$ and accordingly $\mathbf{u} = \mathbf{T}^{-1}(\mathbf{M}\mathbf{v} + \mathbf{\Phi})$.

Tracking Control Let $\mathbf{q}_d(t)$ be a desired trajectory with $\dot{\mathbf{q}}_d(t)$ and $\ddot{\mathbf{q}}_d(t)$ are well defined and let $\mathbf{e}(t) := \mathbf{q}_d(t) - \mathbf{q}(t)$. Choosing

$$\mathbf{v} = \ddot{\mathbf{q}}_d + \mathbf{K}_p(\mathbf{q}_d - \mathbf{x}_1) + \mathbf{K}_d(\dot{\mathbf{q}}_d - \mathbf{x}_2) \tag{14.486}$$

gives

$$\ddot{\mathbf{e}} + \mathbf{K}_d\dot{\mathbf{e}} + \mathbf{K}_p\mathbf{e} = \mathbf{0}. \tag{14.487}$$

For a proper choice of $\mathbf{K}_d$ and $\mathbf{K}_p$ we get exponential convergence of error to zero in time.

Systems That Are Not in Controllable Form

What If the System Is Not in Controllable Form?

When the system is not in companion form, we may be able to transform the system using a change of variable.

As an example, consider a single input system

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, u). \tag{14.488}$$

If there exists a transformation $\mathbf{z} = \mathbf{z}(\mathbf{x})$ such that

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, u) \implies z^{(n)} = f(\mathbf{z}) + g(\mathbf{z})u, \tag{14.489}$$

then we can choose $u = u(\mathbf{z})$ such that

$$\dot{\mathbf{z}} = \mathbf{A}\mathbf{z} + \mathbf{B}v. \tag{14.490}$$

Let us explore this idea in an example.

Example

Adopted from Slotine JJ, Li W. "Applied Nonlinear Control". 1991

Consider the system

$$\dot{x}_1 = -2x_1 + ax_2 + \sin(x_1), \tag{14.491a}$$
$$\dot{x}_2 = -x_2\cos(x_1) + u\cos(2x_1). \tag{14.491b}$$

Note that the nonlinearity in (14.491a) cannot be directly canceled via the input u.
Let us introduce a change of variable

$$z_1 = x_1, \qquad x_1 = z_1, \tag{14.492a}$$
$$z_2 = ax_2 + \sin(x_1), \qquad x_2 = \frac{1}{a}\left(z_2 - \sin(z_1)\right). \tag{14.492b}$$

Note that $\mathbf{x} = \mathbf{0} \iff \mathbf{z} = \mathbf{0}$. The state equations for $\mathbf{z} = (z_1, z_2)$ is

$$\dot{z}_1 = -2z_1 + z_2, \tag{14.493a}$$
$$\dot{z}_2 = \left(\sin(z_1) - 2z_1\right)\cos(z_1) + au\cos(2z_1). \tag{14.493b}$$

Example (contd.)

Adopted from Slotine JJ, Li W. “Applied Nonlinear Control”. 1991

Now if we chose

$$u = \frac{1}{a\cos(2z_1)}\left[v - \sin(z_1)\cos(z_1) + 2z_1\cos(z_1)\right], \tag{14.494}$$

we obtain a linearized model for $\mathbf{z}$ as

$$\begin{bmatrix}\dot{z}_1\\ \dot{z}_2\end{bmatrix} = \begin{bmatrix}-2 & 1\\ 0 & 0\end{bmatrix}\begin{bmatrix}z_1\\ z_2\end{bmatrix} + \begin{bmatrix}0\\ 1\end{bmatrix}v, \tag{14.495}$$

which is full state controllable. Hence, based on the design requirements, we can find a gain matrix such that

$$v = -\begin{bmatrix}k_1 & k_2\end{bmatrix}\mathbf{z}. \tag{14.496}$$

satisfies the control objectives.

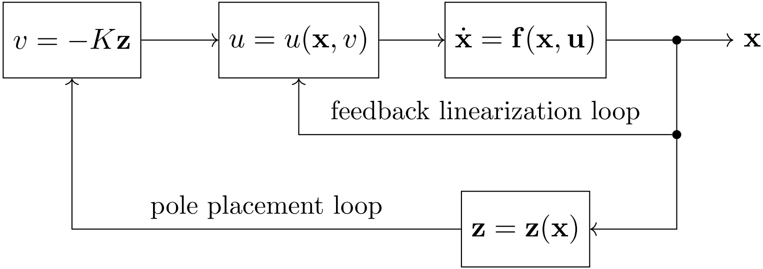
Example (contd.)

Adopted from Slotine JJ, Li W. “Applied Nonlinear Control”. 1991

The final control law is then

$$u = \frac{1}{a\cos(2z_1)}\left[v - \sin(z_1)\cos(z_1) + 2z_1\cos(z_1)\right] \tag{14.497}$$

$$= \frac{1}{a\cos(2x_1)}\left[-k_1x_1 - k_2\left(ax_2 + \sin(x_1)\right) - \sin(x_1)\cos(x_1) + 2x_1\cos(x_1)\right]. \tag{14.498}$$



Input-Output Linearization Motivating Examples

Output Tracking Problem

Consider the system

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}), \tag{14.499}$$

$$\mathbf{y} = \mathbf{h}(\mathbf{x}). \tag{14.500}$$

In most tracking problems, we are interested in controlling the output variable $\mathbf{y}$ to track a sufficiently smooth reference input $\mathbf{r}(t)$, *while keeping the state vector* $\mathbf{x}$ *bounded.*

There are two major concerns

- Linearizing the state equation does not necessarily linearize the output equation;
- The output $\mathbf{y}$ is only indirectly related to the input $\mathbf{u}$.

Let us explore these concerns in the following examples

Example 1

Adopted from Khalil, H.K. "Nonlinear Systems". third edition. 2002

Consider the system

$$\dot{x}_1 = a\sin(x_2), \tag{14.501a}$$
$$\dot{x}_2 = -x_1^2 + u, \tag{14.501b}$$
$$y = x_2. \tag{14.501c}$$

Introducing the following change of variable, we get

$$z_1 = x_1, \qquad \dot{z}_1 = a\sin(x_2) = z_2, \tag{14.502a}$$
$$z_2 = a\sin(x_2), \qquad \dot{z}_2 = a\cos(x_2)\dot{x}_2 = a\cos\left(x_2\right)\left[-x_1^2 + u\right]. \tag{14.502b}$$

Accordingly, we can linearize the system via the control

$$u = \frac{v}{a\cos(x_2)} + x_1^2. \tag{14.503}$$

Example 1 (contd.)

Adopted from Khalil, H.K. “Nonlinear Systems”. third edition. 2002

Using the control (14.503), the linearized system model is

$$\dot{x}_1 = a\sin(x_2), \qquad \dot{z}_1 = z_2, \tag{14.504a}$$

$$\dot{x}_2 = -x_1^2 + u, \qquad \dot{z}_2 = v, \tag{14.504b}$$

$$y = x_2, \qquad y = \sin^{-1}\left(z_2/a\right). \tag{14.504c}$$

Although the state equation in $\mathbf{z}$ is linear, the output y is a nonlinear function of $\mathbf{z}$. Hence, solving a tracking control problem for y is still remains as a nonlinear problem. Alternatively, if we pick u as

$$u = x_1^2 + v, \tag{14.505}$$

we obtain the form and the tracking controller for the reference input $r(t)$ as

$$\begin{cases} \dot{x}_2 & = v, \\ y & = x_2, \end{cases} \implies \dot{y} = v, \quad v = \dot{r} + k\left(r - y\right) \implies \lim_{t\to\infty} |r(t) - y(t)| = 0. \tag{14.506}$$

Example 1 (contd.)

Adopted from Khalil, H.K. “Nonlinear Systems”. third edition. 2002

Input-Output Linearizable Systems

Sometimes it may be more beneficial to linearize the input-output map at the expense of leaving part of the state equation nonlinear. If this is possible, the system is said to be *input-output* linearizable.

⚠ The linearized input-output map may not account for all the dynamics of the system. For the case of this example we have

$$\dot{x}_1 = a\sin(x_2), \tag{14.507a}$$

$$\dot{x}_2 = v, \tag{14.507b}$$

$$y = x_2. \tag{14.507c}$$

Note that x_1 is not connected to y; that is, the linearizing feedback control has made x_1 *unobservable* from y.

Example 1 (contd.)

Adopted from Khalil, H.K. "Nonlinear Systems". third edition. 2002

Assume the objective of the tracking control is to keep $y = x_2$ at a constant reference $r \neq 0$. Hence for $x_1(t)$ we have

$$\dot{x}_1 = a\sin(x_2) = a\sin(r) \implies x_1(t) = x_1(0) + t\,a\sin(r) \tag{14.508}$$

Note that, for $\sin(r) \neq 0$, the state $x_1(t)$ will grow unbounded, which leads to an internal stability issue.

When we design input-output linearization control, we should make sure that the unobserved internal states are well behaved; that is, the internal states are stable or bounded.

Example 2

Adopted from Slotine JJ, Li W. "Applied Nonlinear Control". 1991

Consider the system

$$\dot{x}_1 = \sin(x_2) + (x_2 + 1)x_3, \tag{14.509a}$$

$$\dot{x}_2 = x_1^5 + x_3, \tag{14.509b}$$

$$\dot{x}_3 = x_1^2 + u, \tag{14.509c}$$

$$y = x_1. \tag{14.509d}$$

To obtain a direct relation between u and y, let us differentiate y with respect to t to get

$$\dot{y} = \dot{x}_1 = \sin(x_2) + (x_2 + 1)x_3 \tag{14.510}$$

The output y is still not a direct function of u. Hence, we differentiate y once more to get

$$\ddot{y} = \cos(x_2)\dot{x}_2 + x_3\dot{x}_2 + (x_2 + 1)\dot{x}_3 = f(\mathbf{x}) + (x_2 + 1)u. \tag{14.511}$$

$$f(\mathbf{x}) = \left[\cos(x_2) + x_3\right]\left(x_1^5 + x_3\right) + (x_2 + 1)x_1^2. \tag{14.512}$$

Example 2 (contd.)

Adopted from Slotine JJ, Li W. "Applied Nonlinear Control". 1991

Now we can choose

$$u = \frac{1}{x_2 + 1}\left(v - f(\mathbf{x})\right), \tag{14.513}$$

to obtain the linearized input-output system as

$$\ddot{y} = f(\mathbf{x}) + (x_2 + 1)u \implies \ddot{y} = v. \tag{14.514}$$

Let $r(t)$ be a reference trajectory, we can design a simple tracking controller for the double-integrator system (14.514) as

$$v = \ddot{r} + k_1\dot{e} + k_2 e, \tag{14.515}$$

where the error $e := r - y$ and $k_1, k_2 > 0$. The above control law provides the perfect tracking where $|e(t)| \to 0$ in time.

Example 2 (contd.)

Adopted from Slotine JJ, Li W. "Applied Nonlinear Control". 1991

- The control law is defined everywhere, except when $x_2 = -1$.
- Full state measurement is necessary in implementing the control (computations of y and $\dot{y}$ and u require $\mathbf{x}$.
- We need to consider the behavior of the internal dynamics (states that are *unobservable* by the input)

Relative Degree

If ρ differentiation of the output is needed to yield an explicit relationship between the output and the input, the system is said to have *relative degree* ρ.

For any controllable system of order n, it will need at most n differentiations of any output for the control input to appear. Hence $\rho \leq n$. Intuitively, if it needs more than n differentiations, then the system must be of order higher than n. If the control input never appears, the system is not controllable.

Formal Introduction

Feedback Linearization

Consider a class of nonlinear systems of the form

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{G}(\mathbf{x})\mathbf{u}, \tag{14.516}$$

$$\mathbf{y} = \mathbf{h}(\mathbf{x}). \tag{14.517}$$

Feedback linearization aims to address if there exists a control

$$\mathbf{u} = \boldsymbol{\alpha}(\mathbf{x}) + \boldsymbol{\beta}(\mathbf{x})\mathbf{v} \tag{14.518}$$

along with a change of variable

$$\mathbf{z} = \boldsymbol{\phi}(\mathbf{x}), \tag{14.519}$$

such that $\dot{\mathbf{z}} = \mathbf{A}\mathbf{z} + \mathbf{B}\mathbf{v}$ is linear and stabilizable.

Requirements on the Change of Variable $\mathbf{z} = \boldsymbol{\phi}(\mathbf{x})$

Consider the change of variable $\mathbf{z} = \boldsymbol{\phi}(\mathbf{x})$. We need:

- $\boldsymbol{\phi} : D \to D_z$ needs to be invertible: there must exists an inverse map $\boldsymbol{\phi}^{-1} : D_z \to D$ such that $\mathbf{x} = \boldsymbol{\phi}^{-1}(\mathbf{z})$ for all $\mathbf{z} \in D_z = \boldsymbol{\phi}(D)$.

- Since the derivatives of $\mathbf{z}$ and $\mathbf{x}$ should be continuous, then $\boldsymbol{\phi}$ and $\boldsymbol{\phi}^{-1}$ need to be continuously differentiable.

Diffeomorphism

Given two differentiable manifolds M and N, a differentiable map $f : M \to N$ is a *diffeomorphism* if it is a bijection and its inverse $f^{-1} : N \to M$ is differentiable as well. If these functions are r times continuously differentiable, f is called a C^r-diffeomorphism.

Input-State Linearizable Systems

Definition 14.1 (Feedback Linearizable System)

A nonlinear system

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{G}(\mathbf{x})\mathbf{u}, \tag{14.520}$$

where $\mathbf{f} : D \to \mathbb{R}^n$ and $\mathbf{G} : D \to \mathbb{R}^{n\times m}$ are sufficiently smooth, is said to be feedback linearizable (or input-state linearizable) if there exists a diffeomorphism $\mathbf{T}$ such that $\mathbf{0} \in \mathbf{T}(D)$ and the change of variable $\mathbf{z} = \mathbf{T}(\mathbf{x})$ transforms (14.520) into the form

$$\dot{\mathbf{z}} = \mathbf{A}\mathbf{z} + \mathbf{B}\boldsymbol{\gamma}(\mathbf{x})\left[\mathbf{u} - \boldsymbol{\alpha}(\mathbf{x})\right], \tag{14.521}$$

with $(\mathbf{A}, \mathbf{B})$ controllable and $\boldsymbol{\gamma}(\mathbf{x})$ nonsingular for all $\mathbf{x} \in \mathbf{D}$.

If the system is in the form of (14.521), then we can linearize it via the control

$$\mathbf{u} = \boldsymbol{\alpha}(\mathbf{x}) + \boldsymbol{\beta}(\mathbf{x})\mathbf{v} = \boldsymbol{\alpha}(\mathbf{x}) + \boldsymbol{\gamma}^{-1}(\mathbf{x})\mathbf{v} \tag{14.522}$$

Function Derivatives

Given a smooth function $f : \mathbb{R}^n \to \mathbb{R}$, the *gradient* of f with respect to the input $\mathbf{x} \in \mathbb{R}^n$ is

$$\nabla_{\mathbf{x}} f = \frac{\partial f}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} & \cdots & \frac{\partial f}{\partial x_n} \end{bmatrix}. \tag{14.523}$$

Let $\mathbf{g} : \mathbb{R}^n \to \mathbb{R}^m$ be a vector field. Hence

$$\mathbf{g}(\mathbf{x}) = \begin{bmatrix} g_1(\mathbf{x}) \\ \vdots \\ g_m(\mathbf{x}) \end{bmatrix}. \tag{14.524}$$

The *Jacobian* of $\mathbf{g}$ with respect to the input $\mathbf{x} \in \mathbb{R}^n$ is

$$\nabla_{\mathbf{x}} \mathbf{g} = \frac{\partial \mathbf{g}}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial g_1}{\partial \mathbf{x}} \\ \vdots \\ \frac{\partial g_m}{\partial \mathbf{x}} \end{bmatrix} = \begin{bmatrix} \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial g_m}{\partial x_1} & \cdots & \frac{\partial g_m}{\partial x_n} \end{bmatrix} \tag{14.525}$$

Lie Derivative

⚠ Here, we discuss the Lie Derivative solely in the context of feedback linearization. Please refer to references in *differential geometry* for a rigorous treatment of the subject.

Definition 14.2

Let $h : \mathbb{R}^n \to \mathbb{R}$ be a smooth function, and $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^n$ be a vector field.
The *Lie derivative* of h with respect to $\mathbf{f}$ is a mapping from $\mathbb{R}^n \to \mathbb{R}$ defined as

$$L_{\mathbf{f}} h(\mathbf{x}) = \nabla_{\mathbf{x}} h(\mathbf{x})\ \mathbf{f}(\mathbf{x}) \tag{14.526}$$

We can think of $L_{\mathbf{f}} h(\mathbf{x})$ as directional derivative of h along the vector field directions defined by $\mathbf{f}$.

Lie Derivative (contd.)

To find the k-th Lie derivative of h over the vector field $\mathbf{f}$ we have

$$L_{\mathbf{f}}^{0}h(\mathbf{x}) = h(\mathbf{x}), \tag{14.527}$$

$$L_{\mathbf{f}}^{1}h(\mathbf{x}) = L_{\mathbf{f}}h(\mathbf{x}) = \nabla_{\mathbf{x}}h(\mathbf{x})\ \mathbf{f}(\mathbf{x}), \tag{14.528}$$

$$L_{\mathbf{f}}^{2}h(\mathbf{x}) = L_{\mathbf{f}}L_{\mathbf{f}}h(\mathbf{x}) = \nabla_{\mathbf{x}}\left[L_{\mathbf{f}}h(\mathbf{x})\right]\mathbf{f}(\mathbf{x}), \tag{14.529}$$

$$\vdots$$

$$L_{\mathbf{f}}^{k}h(\mathbf{x}) = L_{\mathbf{f}}L_{\mathbf{f}}^{k-1}h(\mathbf{x}) = \nabla_{\mathbf{x}}\left[L_{\mathbf{f}}^{i-1}h(\mathbf{x})\right]\mathbf{f}(\mathbf{x}). \tag{14.530}$$

if $\mathbf{g}(\mathbf{x})$ is also a vector field, then

$$L_{\mathbf{g}}L_{\mathbf{f}}h(\mathbf{x}) = \nabla_{\mathbf{x}}\left[L_{\mathbf{f}}h(\mathbf{x})\right]\mathbf{g}(\mathbf{x}). \tag{14.531}$$

Time Derivatives of the Output in SISO System

Consider the single-input single-output system

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})u, \qquad y = h(\mathbf{x}), \tag{14.532a}$$

The derivative of y with respect to time is

$$\dot{y} = \frac{\partial h}{\partial \mathbf{x}} \dot{\mathbf{x}} = \nabla h \left[\mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})u\right] = \nabla h \ \mathbf{f}(\mathbf{x}) + \nabla h \ \mathbf{g}(\mathbf{x})u = L_{\mathbf{f}} h(\mathbf{x}) + L_{\mathbf{g}} h(\mathbf{x}) \ u. \tag{14.533}$$

If $L_{\mathbf{g}} h(\mathbf{x}) = 0$, then $\dot{y} = L_{\mathbf{f}} h(\mathbf{x})$, which is independent of u. If we continue to calculate the second derivative of y, that is $\ddot{y} = y^{(2)}$, we have

$$y^{(2)} = \nabla_{\mathbf{x}} \left(L_{\mathbf{f}} h(\mathbf{x})\right) \left[\mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})u\right] = L_{\mathbf{f}}^2 h(\mathbf{x}) + L_{\mathbf{g}} L_{\mathbf{f}} h(\mathbf{x}) \ u. \tag{14.534}$$

Similarly if $L_{\mathbf{g}} L_{\mathbf{f}} h(\mathbf{x}) = 0$, we can continue and find ρ-th derivative as

$$y^{(\rho)} = L_{\mathbf{f}}^{\rho} h(\mathbf{x}) + L_{\mathbf{g}} L_{\mathbf{f}}^{\rho-1} h(\mathbf{x}) \ u. \tag{14.535}$$

Input-Output Linearization (Single Input)

Definition 14.3 (Input-Output Linearizable SISO Systems)

Consider SISO system with sufficiently smooth maps $\mathbf{f}, \mathbf{g} : D \to \mathbb{R}^n$ and $h : D \to \mathbb{R}$

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})u, \qquad y = h(\mathbf{x}). \tag{14.536a}$$

If u first explicitly appears in the ρ-th derivative of the output as

$$y^{(\rho)} = L_{\mathbf{f}}^{\rho} h(\mathbf{x}) + L_{\mathbf{g}} L_{\mathbf{f}}^{\rho-1} h(\mathbf{x}) u, \tag{14.537}$$

then the system is input-output linearizable and the control function

$$u = \frac{1}{L_{\mathbf{g}} L_{\mathbf{f}}^{\rho-1} h(\mathbf{x})} \left[v - L_{\mathbf{f}}^{\rho} h(\mathbf{x}) \right], \tag{14.538}$$

reduces the input-output map to a chain of ρ integrators $y^{(\rho)} = v$.

Relative Degree

Definition 14.4 (Relative Degree)

The nonlinear system, with n dimensional state,

$$\begin{aligned}\dot{\mathbf{x}} &= \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})u, \\ y &= h(\mathbf{x}),\end{aligned}$$

is said to have *relative degree* ρ in a region $D_0 \subset D$, $1 \leq \rho \leq n$, if for all $\mathbf{x} \in D_0$

$$\begin{aligned}L_{\mathbf{g}} L_{\mathbf{f}}^{i-1} h(\mathbf{x}) &= \mathbf{0}, \quad \forall i \in \{1, 2, \ldots, \rho - 1\} \\ L_{\mathbf{g}} L_{\mathbf{f}}^{\rho-1} h(\mathbf{x}) &\neq \mathbf{0}.\end{aligned}$$

Zero Dynamics

Motivation From Linear Systems

[The following discussion is adopted from Khalil, H.K. "Nonlinear Systems". third edition. 2002]

Consider a single input, single output linear system

$$H(s) = \frac{b_m s^m + b_{m-1}s^{m-1} + \cdots + b_1 s + b_0}{s^n + a_{n-1}s^{n-1} + \cdots + a_1 s + a_0}, \tag{14.540}$$

where $m < n$ and $b_m \neq 0$. A possible state model for the system is

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}u, \tag{14.541a}$$

$$y = \mathbf{C}\mathbf{x}, \tag{14.541b}$$

where $\mathbf{C} = \begin{bmatrix} b_0 & b_1 & \cdots & b_m & 0 & \cdots & 0 \end{bmatrix}$ and

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-1} \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}. \tag{14.542}$$

Motivation From Linear Systems (contd.)

See that (14.541) is special case of (14.536), where $\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x}$, $\mathbf{g}(\mathbf{x}) = \mathbf{B}$ and $h(\mathbf{x}) = \mathbf{C}\mathbf{x}$. In order to find the relative degree of the system we compute the derivative of y as

$$\dot{\mathbf{y}} = \mathbf{C}\dot{\mathbf{x}} = \mathbf{C}\mathbf{A}\mathbf{x} + \mathbf{C}\mathbf{B}u. \tag{14.543}$$

If $m = n - 1$, then $\mathbf{CB} = b_{n-1} \neq 0$ and the system has relative degree one. Otherwise, $\mathbf{CB} = 0$.

$$\mathbf{y}^{(2)} = \mathbf{C}\mathbf{A}\dot{\mathbf{x}} = \mathbf{C}\mathbf{A}^2\mathbf{x} + \mathbf{C}\mathbf{A}\mathbf{B}u. \tag{14.544}$$

For $k < n - m - 1$, $\mathbf{C}\mathbf{A}^k$ is obtained by shifting $\mathbf{C}$ k-times to the right. Hence

$$\mathbf{C}\mathbf{A}^{k-1}\mathbf{B} = 0, \text{ for } k = 1, 2, ..., n - m - 1, \quad \text{and} \quad \mathbf{C}\mathbf{A}^{n-m-1}\mathbf{B} = b_m \neq 0 \tag{14.545}$$

Thus, u first appears in $y^{(n-m)}$ indicating the relative degree of the system is $n - m$:

$$y^{(n-m)} = \mathbf{C}\mathbf{A}^{n-m}\mathbf{x} + \mathbf{C}\mathbf{A}^{n-m-1}\mathbf{B}u. \tag{14.546}$$

Motivation From Linear Systems (contd.)

To further study the control of input-output linearizable systems and the issues of internal stability, let us continue with our linear system example where

$$H(s) = \frac{N(s)}{D(s)}, \quad \deg(D) = n \text{ and } \deg(N) = m < n, \tag{14.547}$$

with the relative degree $\rho = n - m$. Using Euclidean division, we can write $D(s)$ as

$$D(s) = Q(s)N(s) + R(s), \quad \deg(Q) = n - m = \rho, \text{ and } \deg(R) < m, \tag{14.548}$$

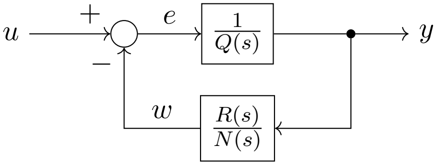
and the leading coefficient of $Q(s)$ is $1/b_m$. Accordingly, we can write $H(s)$ as

$$H(s) = \frac{N(s)}{Q(s)N(s) + R(s)} = \frac{\frac{1}{Q(s)}}{1 + \frac{1}{Q(s)}\frac{R(s)}{N(s)}} \tag{14.549}$$

Motivation From Linear Systems (contd.)

Note that (14.549) represents a negative feedback connection with $1/Q(s)$ in the forward path and $R(s)/N(s)$ in the feedback path

$$H(s) = \frac{\dfrac{1}{Q(s)}}{1 + \dfrac{1}{Q(s)}\dfrac{R(s)}{N(s)}}$$



Since the transfer function $1/Q(s)$ has no zeros, it can be realized by the ρth-order state

$$\boldsymbol{\xi} = \left(y,\ \dot{y}, \ldots, y^{(\rho-1)}\right) \tag{14.550}$$

Motivation From Linear Systems (contd.)

which leads to the state equation

$$\dot{\boldsymbol{\xi}} = \left(\mathbf{A}_c + \mathbf{B}_c \boldsymbol{\lambda}^T\right) \boldsymbol{\xi} + \mathbf{B}_c b_m e, \tag{14.551a}$$

$$y = \mathbf{C}_c \boldsymbol{\xi}, \tag{14.551b}$$

where $\boldsymbol{\lambda} \in \mathbb{R}^\rho$ and

$$\mathbf{A}_c = \left[\begin{array}{c|c} \mathbf{0} & \mathbf{I} \\ \hline 0 & \mathbf{0} \end{array}\right], \quad \mathbf{B}_c = \left[\begin{array}{c} \mathbf{0} \\ \hline 1 \end{array}\right], \quad \mathbf{C}_c = \left[\begin{array}{c|c} 1 & \mathbf{0} \end{array}\right] \tag{14.552}$$

Let (A_0, B_0, C_0) be a minimal realization of the transfer function $R(s)/N(s)$; Accordingly

$$\dot{\boldsymbol{\eta}} = \mathbf{A}_0 \boldsymbol{\eta} + \mathbf{B}_0 y \tag{14.553a}$$

$$w = \mathbf{C}_0 \boldsymbol{\eta}. \tag{14.553b}$$

Motivation From Linear Systems (contd.)

From the feedback connection, we can realize $H(s)$ by the state model

$$\dot{\boldsymbol{\eta}} = \mathbf{A}_0\boldsymbol{\eta} + \mathbf{B}_0\mathbf{C}_c\boldsymbol{\xi} \tag{14.554a}$$

$$\dot{\boldsymbol{\xi}} = \mathbf{A}_c\boldsymbol{\xi} + \mathbf{B}_c\left(\boldsymbol{\lambda}^T\boldsymbol{\xi} - b_m\mathbf{C}_0\boldsymbol{\eta} + b_m u\right), \tag{14.554b}$$

$$y = \mathbf{C}_c\boldsymbol{\xi}. \tag{14.554c}$$

See that eigenvalues of A_0 are the roots of the polynomial $N(s)$, which are the zeros of the transfer function $H(s)$. Using the structure of $\mathbf{A}_c$, $\mathbf{B}_c$ and $\mathbf{C}_c$, we can verify that

$$y^{(p)} = \boldsymbol{\lambda}^T\boldsymbol{\xi} - b_m\mathbf{C}_0\boldsymbol{\eta} + b_m u. \tag{14.555}$$

Motivation From Linear Systems (contd.)

Choosing the feedback control u as

$$u = \frac{1}{b_m}\left(v - \boldsymbol{\lambda}^T\boldsymbol{\xi} + b_m\mathbf{C}_0\boldsymbol{\eta}\right), \tag{14.556}$$

results in the closed-loop system

$$\dot{\boldsymbol{\eta}} = \mathbf{A}_0\boldsymbol{\eta} + \mathbf{B}_0\mathbf{C}_c\boldsymbol{\xi} \tag{14.557a}$$

$$\dot{\boldsymbol{\xi}} = \mathbf{A}_c\boldsymbol{\xi} + \mathbf{B}_c v, \tag{14.557b}$$

$$y = \mathbf{C}_c\boldsymbol{\xi}, \tag{14.557c}$$

where the input output map is a chain of ρ integrators and the state subvector $\boldsymbol{\eta}$ is *unobservable* from the output y.

Motivation From Linear Systems (contd.)

Suppose we want to stabilize the output at a constant reference r. This requires stabilizing $\boldsymbol{\xi}$ at $\boldsymbol{\xi}_d = (r, 0, \ldots, 0)$. Accordingly, let $\boldsymbol{\zeta} := \boldsymbol{\xi} - \boldsymbol{\xi}_d$. Noting that $\mathbf{A}_c \boldsymbol{\xi}_d = \mathbf{0}$, we have

$$\dot{\boldsymbol{\zeta}} = \dot{\boldsymbol{\xi}} - \dot{\boldsymbol{\xi}}_d = \mathbf{A}_c \boldsymbol{\xi} - \mathbf{B}_c v = \mathbf{A}_c \boldsymbol{\zeta} + \mathbf{A}_c \boldsymbol{\xi}_d + \mathbf{B}_c v \implies \dot{\boldsymbol{\zeta}} = \mathbf{A}_c \boldsymbol{\zeta} + \mathbf{B}_c v \tag{14.558}$$

Choosing $v = -\mathbf{K}\boldsymbol{\zeta} = -\mathbf{K}(\boldsymbol{\xi} - \boldsymbol{\xi}_d)$ such that $\mathbf{A}_c - \mathbf{B}_c\mathbf{K}$ is Hurwitz, completes the design of the control law as

$$u = \frac{1}{b_m}\left[-\mathbf{K}(\boldsymbol{\xi} - \boldsymbol{\xi}_d) - \boldsymbol{\lambda}^T \boldsymbol{\xi} + b_m \mathbf{C}_0 \boldsymbol{\eta}\right], \tag{14.559}$$

with the corresponding closed-loop system

$$\dot{\boldsymbol{\eta}} = \mathbf{A}_0 \boldsymbol{\eta} + \mathbf{B}_0 \mathbf{C}_c (\boldsymbol{\zeta} + \boldsymbol{\xi}_d) \tag{14.560a}$$

$$\dot{\boldsymbol{\zeta}} = (\mathbf{A}_c - \mathbf{B}_c\mathbf{K})\, \boldsymbol{\zeta}. \tag{14.560b}$$

Since $\mathbf{A}_c - \mathbf{B}_c\mathbf{K}$, $\boldsymbol{\zeta} \to \mathbf{0}$ in time for any $\boldsymbol{\zeta}(0)$. What about $\boldsymbol{\eta}$?

Motivation From Linear Systems (contd.)

A continuous signal $x(t)$ is *bounded* if $\exists L > 0$, such that $|x(t)| \leq L$ for all t.

Definition 14.5 (bounded-input, bounded-output (BIBO) stability)

A system is BIBO stable if the output remains bounded for any bounded input.

Necessary and sufficient condition of BIBO stability in LTI systems is that the impulse response, $h(t)$, be absolutely integrable, that is, it's L^1 norm exists:

$$\int_{-\infty}^{\infty} |h(t)|\, dt = \|h\|_1 \in \mathbb{R}$$

For SISO systems with transfer function $H(s)$, a sufficient condition for BIBO stability is that the region of convergence (ROC) of $H(s)$ includes the imaginary axis.

- For causal systems, the ROC is the open region in $\mathbb{C}$ to the right of a vertical line whose abscissa is the real part of the "largest pole" (the pole with the greatest real part)
- Hence, all poles of the system must be in the strict left half of the s-plane.

Motivation From Linear Systems (contd.)

Recall that (A_0, B_0, C_0) is assumed to be a minimal realization of the transfer function $R(s)/N(s)$. Hence, to ensure that $\boldsymbol{\eta}$ will be bounded for all bounded $y(t)$ and all bounded initial states $\boldsymbol{\eta}(0)$, we need $\mathbf{A}_0$ to be Hurwitz:

$$\mathbf{A}_0 \text{ is Hurwitz} \implies \text{roots of } N(s) \text{ have negative real part} \implies \text{zeros of } H(s) \text{ have negative real part.}$$

A transfer function having all zeros in the open-left half plane is called *minimum phase*.

From a pole placement viewpoint, our state feedback control function assigns the closed-loop eigenvalues into two groups:

- ρ eigenvalues are assigned in the open-left half plane as the eigenvalues of $\mathbf{A}_c - \mathbf{B}_c\mathbf{K}$,
- $n - \rho$ eigenvalues are assigned at the open-loop zeros.

Zero Dynamics

We now want to develop a nonlinear version of (14.554) for the nonlinear system (14.536) with relative degree ρ. For $\boldsymbol{\xi}$ we could follow the same process, since the input-output map will still be a chain of ρ integrators. We need a nonlinear version of $\dot{\boldsymbol{\eta}} = \mathbf{A}_0\boldsymbol{\eta} + \mathbf{B}_0\mathbf{C}_c\boldsymbol{\xi}$ where u is absent. A change of variables that would give us the desired outcome is

$$\mathbf{z} = \mathbf{T}(\mathbf{x}) = \begin{bmatrix} \phi_1(\mathbf{x}) \\ \vdots \\ \phi_{n-\rho}(\mathbf{x}) \\ \hline h(\mathbf{x}) \\ \vdots \\ L_{\mathbf{f}}^{\rho-1}h(\mathbf{x}) \end{bmatrix} := \begin{bmatrix} \boldsymbol{\Phi}(\mathbf{x}) \\ \hline \boldsymbol{\psi}(\mathbf{x}) \end{bmatrix} := \begin{bmatrix} \boldsymbol{\eta} \\ \hline \boldsymbol{\xi} \end{bmatrix}, \tag{14.561}$$

where ϕ_1 to $\phi_{n-\rho}$ are chosen such that $T(\mathbf{x})$ is a diffeomorphism on a domain $D_0 \subset D$ and

$$L_{\mathbf{g}}\phi_i(\mathbf{x}) = 0, \quad \text{for } 1 \leq i \leq n-\rho, \quad \forall \mathbf{x} \in D_0. \tag{14.562}$$

Zero Dynamics (contd.)

Theorem 14.1

Consider the system with relative degree $\rho \leq n$ *in* D *defined as*

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})u, \qquad y = h(\mathbf{x}). \tag{14.563}$$

If $\rho = n$*, then for every* $\mathbf{x}_0 \in D$*, there exists a neighborhood* N *of* $\mathbf{x}_0$ *such that the map*

$$T(\mathbf{x}) = \begin{bmatrix} h(\mathbf{x}) \\ \vdots \\ L_{\mathbf{f}}^{n-1} h(\mathbf{x}) \end{bmatrix} \tag{14.564}$$

restricted to N*, is a diffeomorphism. If* $\rho < n$*, then there exists a neighborhood* N *of* $\mathbf{x}_0$ *and smooth functions* $\phi_1(\mathbf{x}), \ldots, \phi_{n-\rho}(\mathbf{x})$ *such that (14.562) is satisfied for all* $\mathbf{x} \in N$ *and the map* $T(\mathbf{x})$ *of (14.561), restricted to* N*, is a diffeomorphism on* N*.*

Zero Dynamics (contd.)

For proof see Khalil, H.K. "Nonlinear Systems". third edition. 2002

The condition (14.562) ensures that u cancels out in

$$\dot{\boldsymbol{\eta}} = \frac{\partial \boldsymbol{\Phi}}{\partial \mathbf{x}} \left[\mathbf{f}(\mathbf{x}) + \mathbf{g}(x)u\right] = \frac{\partial \boldsymbol{\Phi}}{\partial \mathbf{x}} \mathbf{f}(\mathbf{x}). \tag{14.565}$$

We can verify that the change of variable (14.561) transforms (14.536) to

$$\dot{\boldsymbol{\eta}} = \mathbf{f}_0(\boldsymbol{\eta}, \boldsymbol{\xi}) \tag{14.566a}$$

$$\dot{\boldsymbol{\xi}} = \mathbf{A}_c \boldsymbol{\xi} + \mathbf{B}_c \gamma(\mathbf{x}) \left[u - \alpha(\mathbf{x})\right] \tag{14.566b}$$

$$y = \mathbf{C}_c \boldsymbol{\xi} \tag{14.566c}$$

where $\boldsymbol{\xi} \in \mathbb{R}^{\rho}$, $\boldsymbol{\eta} \in \mathbb{R}^{n-\rho}$, $(\mathbf{A}_c, \mathbf{B}_c, \mathbf{C}_c)$ is a canonical form of a chain of ρ integrators,

$$\mathbf{f}_0(\boldsymbol{\eta}, \boldsymbol{\xi}) = \left.\frac{\partial \boldsymbol{\Phi}}{\partial \mathbf{x}} \mathbf{f}(\mathbf{x})\right|_{\mathbf{x}=\mathbf{T}^{-1}(\mathbf{z})}, \quad \gamma(\mathbf{x}) = L_{\mathbf{g}} L_{\mathbf{f}}^{\rho-1} h(\mathbf{x}) \quad \text{and} \quad \alpha(\mathbf{x}) = -\frac{L_{\mathbf{f}}^{\rho} h(\mathbf{x})}{L_{\mathbf{g}} L_{\mathbf{f}}^{\rho-1} h(\mathbf{x})}. \tag{14.567}$$

Zero Dynamics (contd.)

The system (14.566) is said to be in the *normal form*, which decomposes the system into

- an external part $\boldsymbol{\xi}$ where it could be linearized by the state feedback control $u = \alpha(\mathbf{x}) + \beta(\mathbf{x})v$ with $\beta(\mathbf{x}) = \gamma^{-1}(\mathbf{x})$;
- an internal part $\boldsymbol{\eta}$ that is *unobservable* by the control.

Zero Dynamics

The internal dynamics are described by $\dot{\boldsymbol{\eta}} = \mathbf{f}_0(\boldsymbol{\eta}, \boldsymbol{\xi})$. Setting $\boldsymbol{\xi} = \mathbf{0}$ in this equation gives

$$\dot{\boldsymbol{\eta}} = \mathbf{f}_0(\boldsymbol{\eta}, \mathbf{0}) \tag{14.568}$$

which is called the *zero dynamics*

The naming of zero dynamics matches nicely with the fact that, for linear systems, (14.568) is given by $\dot{\boldsymbol{\eta}} = \mathbf{A}_0\boldsymbol{\eta}$, where the eigenvalues of $\mathbf{A}_0$ are the zeros of the transfer function $H(s)$.

Zero Dynamics (contd.)

Similar to linear systems:

Minimum Phase System

The system is said to be *minimum phase* if

$$\dot{\boldsymbol{\eta}} = \mathbf{f}_0(\boldsymbol{\eta}, \mathbf{0})$$

has an asymptotically stable equilibrium point in the domain of interest.

In particular, if $\mathbf{T}(\mathbf{x})$ is chosen such that the origin $(\boldsymbol{\eta} = \mathbf{0}, \boldsymbol{\xi} = \mathbf{0})$ is an equilibrium point of (14.566), then the system is said to be minimum phase if the origin of the zero dynamics (14.568) is asymptotically stable.

Zero Dynamics (contd.)

Zero dynamics can also be characterized in the original coordinates. If y is identically zero, the solution of the state equation must be confined to the set

$$Z^* = \{\mathbf{x} \in D_0 \,|\, h(\mathbf{x}) = L_f h(\mathbf{x}) = \cdots = L_{\mathbf{f}}^{\rho-1} h(\mathbf{x}) = 0\}$$

Since $y = 0$ implies $h(\mathbf{x})$ and all the derivatives of y are zero and the input must be

$$u = u^*(\mathbf{x}) := \alpha(\mathbf{x})\Big|_{\mathbf{x} \in Z^*}$$

The restricted motion of the system is described by

$$\dot{\mathbf{x}} = \mathbf{f}^*(\mathbf{x}) := \Big[\mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})\alpha(\mathbf{x})\Big]\Big|_{\mathbf{x} \in Z^*}$$

In the special case $\rho = n$, the normal form reduces to

$$\dot{\mathbf{z}} = \mathbf{A}_c \mathbf{z} + \mathbf{B}_c \gamma(z)[u - \alpha(\mathbf{x})]$$
$$y = \mathbf{C}_c \mathbf{z}$$

Zero Dynamics (contd.)

where $\mathbf{z} = \boldsymbol{\xi} = \left(h(\mathbf{x}), \ldots, L_{\mathbf{f}}^{n-1} h(\mathbf{x})\right)$ and $\boldsymbol{\eta}$ does not exist. In this case, the system has no zero dynamics and, by default, is said to be minimum phase.

Contents

Further Readings

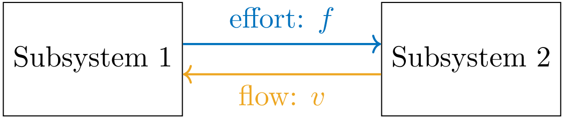
Most of the discussions in this section are adopted from the following references.

- Hogan, N. (1985). Impedance control: an approach to manipulation. Part I: Theory. ASME Journal of Dynamic Systems Measurement and Control, 107(1), 1-7.
- Hogan, N. (1985). Impedance control: an approach to manipulation. Part II: Implementation. ASME Journal of Dynamic Systems Measurement and Control, 107(1), 8-16.
- Hogan, N. (1985). Impedance control: an approach to manipulation. Part III: Application. ASME Journal of Dynamic Systems Measurement and Control, 107(1), 17-24.

Causality

instantaneous power flow between two physical systems is always definable as the product of two conjugate variables:

- an *effort*, such as a force or a voltage, and
- a *flow*, such as a velocity or a current

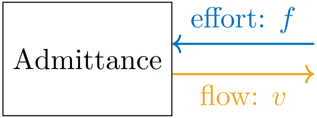


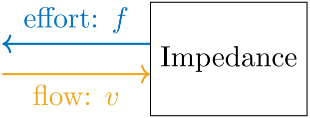
A subsystem cannot determine both variables:

- Along any degree of freedom, a manipulator may exert a force on its environment or impose a motion on it, but not both.

Impedance and Admittance

Along any degree of freedom, physical systems come in two types:





Admittance subsystem accepts effort input and yields flow output

Impedance subsystem accepts flow input and yields effort output

Impedance Control

The Idea

The underlying idea behind Impedance control could be summarized as

- A constrained inertial object can always be pushed on, but it may not always move! Hence, the world is an admittance from a manipulator point of view.
- To ensure physical compatibility with the environmental admittance, the manipulator should admit the behavior of an impedance.
- The behavior of the manipulator should be adaptable: the mechanical interaction and the coupling between the manipulator and the environment may change.
- If we impose motion control in the form of an impedance, the dynamic interaction between manipulator and environment is control1able by adjusting the defined impedance, hence the term *impedance control*.

Derivation

Let $\mathbf{x} \in \mathbb{R}^c$ denote the pose of the end-effector in the workspace, which may include linear and angular terms. Let $\mathbf{L} : \mathbb{R}^n \to \mathbb{R}^c$ denote the forward kinematics. Accordingly

$$\mathbf{J}(\mathbf{q}) = \frac{\partial \mathbf{x}}{\partial \mathbf{q}} = \frac{\partial}{\partial \mathbf{q}} \mathbf{L}(\mathbf{q}), \quad \dot{\mathbf{x}} = \mathbf{J}(\mathbf{q})\dot{\mathbf{q}}. \tag{15.569}$$

Let $\mathbf{f}_{int}$ denote the *interface* force between the environment and the end-effector, based on principal of virtual work we have

$$\boldsymbol{\tau} = \mathbf{J}^T(\mathbf{q})\, \mathbf{f}_{int}. \tag{15.570}$$

Let $\mathbf{x}_0$ define the *desired equilibrium position* for the end-effector in the absence of environmental forces. Some possible values for $\mathbf{f}$ are

- Stiffness: $\mathbf{f}_{int} = \mathbf{k}(\mathbf{x}_0 - \mathbf{x}) = \mathbf{k}\left(\mathbf{x}_0 - \mathbf{L}(\mathbf{q})\right)$,
- Damping: $\mathbf{f}_{int} = \mathbf{b}(\dot{\mathbf{x}}_0 - \dot{\mathbf{x}}) = \mathbf{b}\left(\dot{\mathbf{x}}_0 - \mathbf{J}(\mathbf{q})\,\dot{\mathbf{q}}\right)$.

The mappings $\mathbf{k}, \mathbf{b} : \mathbb{R}^c \to \mathbb{R}^c$ do not need to be linear. The control (15.570) does not account for the inertial, frictional and gravitational dynamics of the manipulator.

Derivation (contd.)

We can choose control such that it "masks" the true nonlinear dynamics of the manipulator and imposes a simpler desired dynamics.

Assume a *desired* end-effector behavior to be imposed on the manipulator is given by

$$\mathbf{W}\ddot{\mathbf{x}} - \mathbf{B}(\dot{\mathbf{x}}_0 - \dot{\mathbf{x}}) - \mathbf{K}(\mathbf{x}_0 - \mathbf{x}) = \mathbf{f}_{int}, \tag{15.571}$$

where the positive definite matrix $\mathbf{W}$ denotes a *desired inertia* of the end-effector as observed by the environment. The mappings $\mathbf{B}$ and $\mathbf{K}$ denote *desired* damping and stiffness of end-effector, respectively. The point $\mathbf{x}_0(t)$ serves as the reference for the desired damping and stiffness.

Consider the manipulator model

$$\mathbf{M}(\mathbf{q})\ddot{\mathbf{q}} + \mathbf{h}(\mathbf{q}, \dot{\mathbf{q}}) = \boldsymbol{\tau} + \mathbf{J}^T\mathbf{f}_{int}, \tag{15.572}$$

where $\mathbf{M}$ is the positive definite mass matrix and $\mathbf{J}^T\mathbf{f}_{int}$ is the joint torques due to $\mathbf{f}_{int}$.

Derivation (contd.)

Based on feedback linearization, choosing the control

$$\boldsymbol{\tau} = \mathbf{M}\mathbf{v} + \mathbf{h}(\mathbf{q}, \dot{\mathbf{q}}) - \mathbf{J}^T \mathbf{f}_{int}, \tag{15.573}$$

results in a linearized generalized acceleration map $\ddot{\mathbf{q}} = \mathbf{v}$. Noting that the desired end-effector behavior is defined in the workspace, we have

$$\ddot{\mathbf{x}} = \mathbf{J}(\mathbf{q})\ddot{\mathbf{q}} + \dot{\mathbf{J}}(\mathbf{q})\dot{\mathbf{q}} \quad \Longrightarrow \quad \ddot{\mathbf{q}} = \mathbf{J}^{-1}(\mathbf{q}) \left[\ddot{\mathbf{x}} - \dot{\mathbf{J}}(\mathbf{q})\dot{\mathbf{q}}\right]. \tag{15.574}$$

Solving (15.571) for $\ddot{\mathbf{x}}$ and substituting in (15.574) gives

$$\boldsymbol{\tau} = \mathbf{M}\mathbf{J}^{-1} \left[\ddot{\mathbf{x}} - \dot{\mathbf{J}}\dot{\mathbf{q}}\right] + \mathbf{h} - \mathbf{J}^T \mathbf{f}_{int} \tag{15.575}$$

$$= \mathbf{M}\mathbf{J}^{-1}\mathbf{W}^{-1} \left[\mathbf{f}_{int} + \mathbf{B}(\dot{\mathbf{x}}_0 - \mathbf{J}\dot{\mathbf{q}}) + \mathbf{K}\left(\mathbf{x}_0 - \mathbf{L}(\mathbf{q})\right)\right] - \mathbf{M}\mathbf{J}^{-1}\dot{\mathbf{J}}\dot{\mathbf{q}} + \mathbf{h} - \mathbf{J}^T \mathbf{f}_{int}. \tag{15.576}$$

To simplify, we omitted the function dependencies in the equation above.

Contents

Motivation

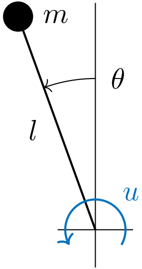
Consider the equation of motion of a simple point-mass pendulum

$$\ddot{\theta} = \frac{g}{l}\sin(\theta) + \frac{u}{ml^2}. \tag{16.577}$$

Let $x_1 := \theta$, $x_2 = \dot{\theta}$ and $m = l = g = 1$, then

$$\dot{x}_1 = x_2 \tag{16.578a}$$

$$\dot{x}_2 = \sin(x_1) + u \tag{16.578b}$$



Our objective is to find the control u, such that, from any initial state $\mathbf{x}(0)$ the pendulum swings to the equilibrium point $x_1 = x_2 = 0$ (the upward standing configuration).

Motivation (contd.)

First, let us assume that we can apply any piecewise continuous input u to the system. Let the error vector $\mathbf{e}$ defined as $\mathbf{e} = \mathbf{x}_d - \mathbf{x} = \mathbf{0} - \mathbf{x}$. For the error dynamics we have

$$\dot{e}_1 = e_2 \tag{16.579a}$$

$$\dot{e}_2 = \sin(e_1) - u \tag{16.579b}$$

Note that $\mathbf{x} = \mathbf{0} \implies \mathbf{e} = \mathbf{0}$. Linearizing $\dot{\mathbf{e}}$ about the equilibrium point $\mathbf{e} = \mathbf{0}$ and $u = 0$ gives

$$\begin{bmatrix} \dot{e}_1 \\ \dot{e}_2 \end{bmatrix} \approx \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \end{bmatrix} + \begin{bmatrix} 0 \\ -1 \end{bmatrix} u, \tag{16.580}$$

which is full state controllable. Hence, let $u = k_1 e_1 + k_2 e_2$ then the closed-loop system for the linear approximation as

$$\begin{bmatrix} \dot{e}_1 \\ \dot{e}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 - k_1 & -k_2 \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}.$$

Motivation (contd.)

Accordingly, for any $k_1 > 1$ and $k_2 > 0$ we get global asymptotic stability of the origin (for the linear approximation of the system).

Now, let us study the region of attraction of the equilibrium for the nonlinear closed loop system when the input is constrained to a feasible set U.

To limit the input to U we use the following projection. For any $U \subset \mathbb{R}^m$, let

$$\operatorname*{proj}_U \mathbf{u}(t) := \arg\min_{\mathbf{v}\in U} \|\mathbf{u}(t) - \mathbf{v}\|. \tag{16.581}$$

For our pendulum example, let $U = [-w, w]$ for $w > 0$, then

$$\operatorname*{proj}_U u = \begin{cases} u, & |u| < w, \\ \operatorname{sgn}(u)\cdot w, & |u| \geq w. \end{cases} \tag{16.582}$$

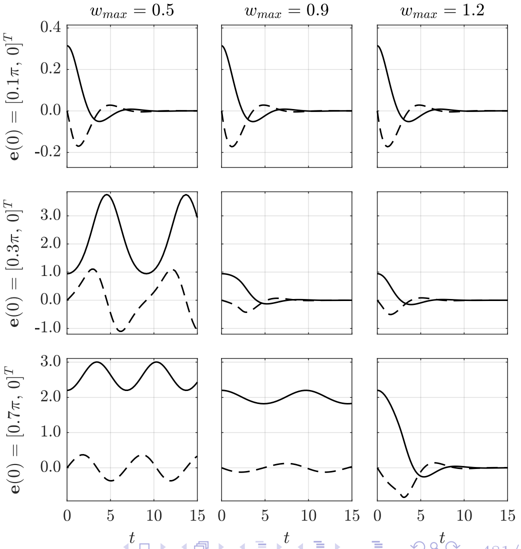
Accordingly, we denote the region of attraction of the equilibrium point with $R_A(w)$.

Motivation (contd.)

Based on the defined feasible control input, the nonlinear closed-loop system dynamics is

$$\dot{x}_1 = x_2$$
$$\dot{x}_2 = \sin(x_1) + \underset{[-w,w]}{\text{proj}}\,(-k_1 x_1 - k_2 x_2)$$

From the simulation results, we can see the convergence of the trajectories to the origin is both a function of initial condition $\mathbf{x}_0$ and input saturation w.



Motivation (contd.)

To analytically define $R_A(w)$, we divide the state space into three subsets of

$$\begin{aligned} \Gamma &:= \left\{|u| \leq w\right\} & &\text{where} & u &= k_1e_1 + k_2e_2, \\ \Sigma_L &:= \{u < -w\} & &\text{where} & u &= -w, \\ \Sigma_R &:= \{u > w\} & &\text{where} & u &= w. \end{aligned}$$

For trajectories in Γ, the total energy of the system is

$$E_\Gamma(\mathbf{e}) = \frac{1}{2}e_2^2 + \cos(e_1) - 1 + \frac{1}{2}k_1e_1^2, \implies \dot{E}_\Gamma(\mathbf{e}) = -k_2e_2^2 \leq 0. \tag{16.584}$$

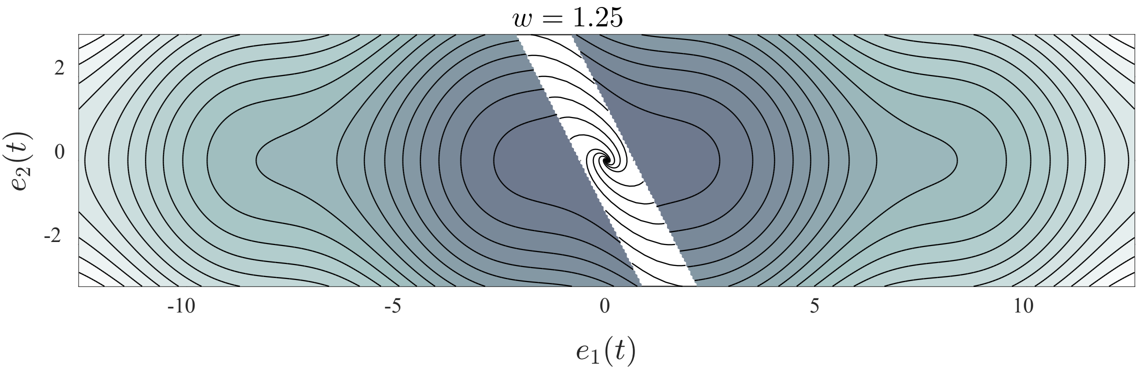
Since the input is constant in Σ_L and Σ_R, we can formulate the Hamiltonian functions for these regions as

$$H_{\Sigma_L}(\mathbf{e}) := \tfrac{1}{2}e_2^2 + \cos(e_1) - w_{max}e_1, \tag{16.585}$$

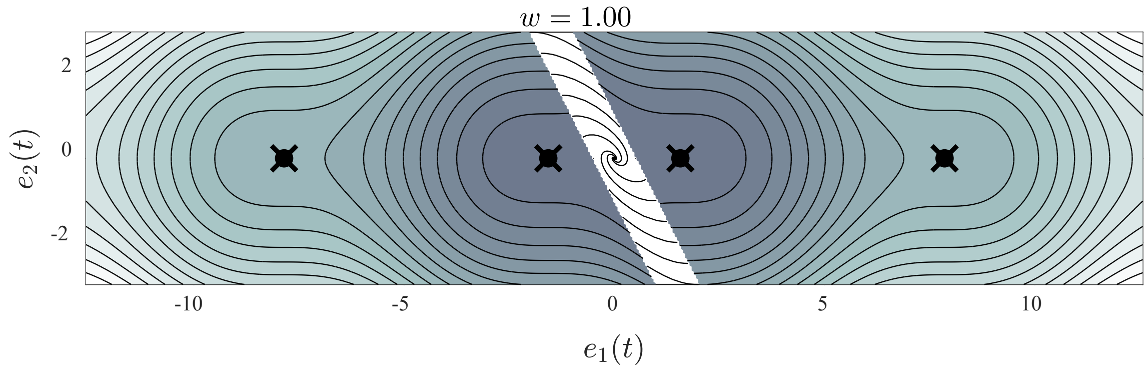
$$H_{\Sigma_R}(\mathbf{e}) := \tfrac{1}{2}e_2^2 + \cos(e_1) + w_{max}e_1, \tag{16.586}$$

where based on the work-energy principle we have $\dot{H}_{\Sigma_L}\left(\mathbf{e}(t)\right) = \dot{H}_{\Sigma_R}\left(\mathbf{e}(t)\right) = 0.$

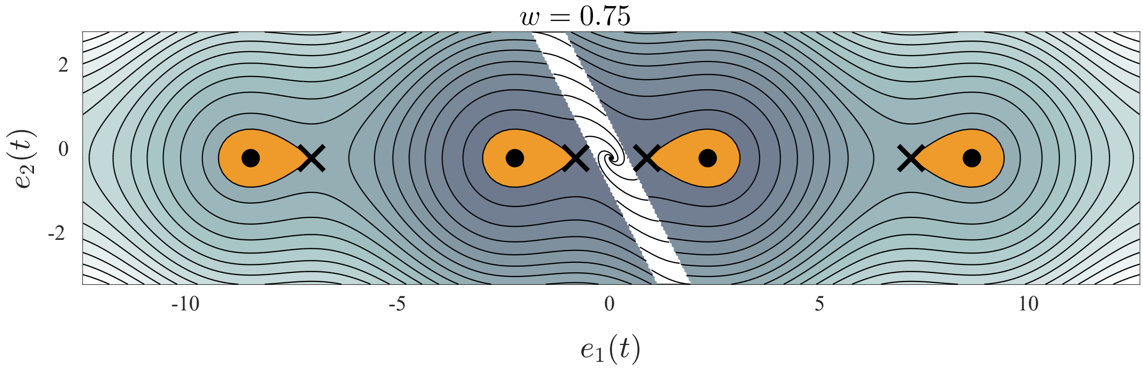
Motivation (contd.)



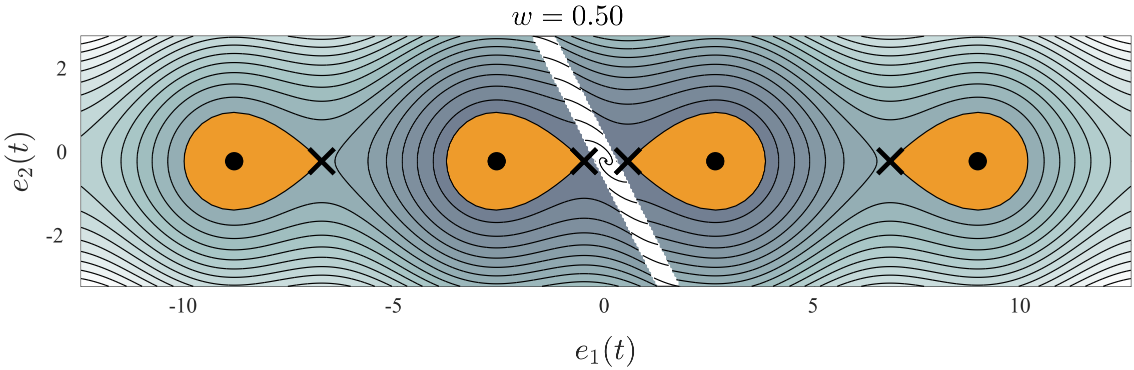
Motivation (contd.)



Motivation (contd.)



Motivation (contd.)



Energy Control

Energy Control Idea

Most of the discussions in this section are adopted from the work of K. J. Aström, in particular "Hybrid control of inverted pendulums".

Idea

Let $\mathbf{x}_d$ be a desired state for the system

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}). \tag{16.587}$$

In energy control we reformulate the problem as finding $\mathbf{u}$ such that $E(\mathbf{x}) \to E(\mathbf{x}_d)$ in time. That is, instead of directly controlling the state vector we design the control so that the energy of the system reaches to the energy level of the desired state.

Derivation For Lagrangian Systems

As discussed in (13.446), for a Lagrangian system

$$\mathbf{M}(\mathbf{q})\ddot{\mathbf{q}} + \mathbf{b}(\mathbf{q}, \dot{\mathbf{q}}) + \mathbf{g}(\mathbf{q}) = \mathbf{T}(\mathbf{q})\mathbf{u}, \tag{16.588}$$

where

$$\mathbf{b}(\mathbf{q}, \dot{\mathbf{q}}) = \mathbf{C}(\mathbf{q}, \dot{\mathbf{q}})\dot{\mathbf{q}} \qquad \text{and} \qquad \mathbf{g}(\mathbf{q}) = \frac{\partial U(\mathbf{q})}{\partial \mathbf{q}}, \tag{16.589}$$

with $U(\mathbf{q})$ being the total potential energy, the total energy of the system is

$$E = \frac{1}{2}\dot{\mathbf{q}}^T \mathbf{M}\dot{\mathbf{q}} + U(\mathbf{q}), \tag{16.590}$$

that has the time rate of change of

$$\dot{E} = \dot{\mathbf{q}}^T \mathbf{T}(\mathbf{q})\mathbf{u}. \tag{16.591}$$

Derivation For Lagrangian Systems (contd.)

Our goal is to find $\mathbf{u}$ such that $E(t)$ approaches to a desired energy level E_d in time. Consider a state representation for (16.588) with the state vector $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2) = (\mathbf{q}, \dot{\mathbf{q}})$ as

$$\dot{\mathbf{x}}_1 = \mathbf{x}_2, \tag{16.592a}$$

$$\dot{\mathbf{x}}_2 = \mathbf{M}^{-1}\mathbf{T}\mathbf{u} - \mathbf{M}^{-1}\left(\mathbf{b}(\mathbf{x}_1, \mathbf{x}_2) + \mathbf{g}(\mathbf{x}_1)\right) \tag{16.592b}$$

To derive the energy to a desired energy $E_d = E(\mathbf{x}_d)$, we pick Control-Lyapunov function

$$V(\mathbf{x}) = \frac{1}{2}\left(E(\mathbf{x}) - E_d\right)^2 . \tag{16.593}$$

Taking the derivative of V with respect to time yields

$$\dot{V}(\mathbf{x}) = \left(E(\mathbf{x}) - E_d\right)\dot{E}(\mathbf{x}) = \left(E(\mathbf{x}) - E_d\right)\mathbf{x}_2^T\mathbf{T}\mathbf{u}. \tag{16.594}$$

Hence, we need to find $\mathbf{u}$ such that, for all $\mathbf{x} \in D$ with $\mathbf{x}_d \in D$, $\dot{V}(\mathbf{x})$ be negative definite.

Hybrid Control Approach

Given the state vector $\mathbf{x} \in \mathbb{R}^n$, the energy of the system is a mapping from $\mathbb{R}^n \to \mathbb{R}$. Thus, E could not be a one-to-one mapping. Accordingly, controlling the energy of the system confines the state to a subset of

$$\left\{\mathbf{x} \in \mathbb{R}^n \mid E(\mathbf{x}) = E_d\right\}, \tag{16.595}$$

which is not necessarily a singleton.

If our objective is to stabilize the system at a desired state $\mathbf{x}_d$, we can utilize a hybrid approach, where we "catch" the trajectory near the desired state $\mathbf{x}_d$.

Let $\mathbf{u}_E$ denote the energy controller and $\mathbf{u}_S$ to be a local stabilizing controller for $\mathbf{x}_d$ with the region of attraction R_A; a possible hybrid controller for the system is

$$\mathbf{u}(\mathbf{x}) = \begin{cases} \mathbf{u}_S(\mathbf{x}) & \text{if} \quad \mathbf{x} \in \Omega \subset R_A, \\ \mathbf{u}_E(\mathbf{x}) & \text{otherwise.} \end{cases} \tag{16.596}$$

Example

Let $x_1 = \theta$ and $x_2 = \dot{\theta}$, the energy of the system is

$$E = \tfrac{1}{2}\dot{\theta}^2 + \cos(\theta) = \tfrac{1}{2}x_2^2 + \cos(x_1) \implies \dot{E} = u\, x_2$$

Since $E\,(\mathbf{x} = \mathbf{0}) = 1$ we set the Lyapunov control function as

$$V = \frac{1}{2}(E-1)^2 \implies \dot{V} = (E-1)\, u\, x_2.$$

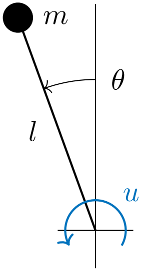
Let $k > 0$, we can define the control function $u_E(\mathbf{x})$ as

$$u_E(\mathbf{x}) := -k\left(E(\mathbf{x}) - 1\right)\mathrm{sgn}(x_2) \implies \dot{V} = -k\left(E(\mathbf{x}) - 1\right)^2 |x_2| \leq 0.$$

Let $u_S = -k_1 x_1 - k_2 x_2$, for $k_1, k_2 > 0$, we have

$$u(\mathbf{x}) := \underset{[-w,w]}{\mathrm{proj}} \begin{cases} u_S(\mathbf{x}) = -k_1 x_1 - k_2 x_2, & \text{if} \quad 1 - \cos(x_1) + x_2^2 \leq \epsilon, \\ u_E(\mathbf{x}) = k(E-1)\,\mathrm{sgn}(x_2), & \text{otherwise.} \end{cases}$$

$$x_1 = \theta,$$
$$x_2 = \dot{\theta}.$$



Contents

Introduction to Calculus of Variations

What is a Functional

[**Linear Algebra**] A functional is a linear mapping from a vector space V into its field of scalars. That is, a functional is an element of the dual space V^*.

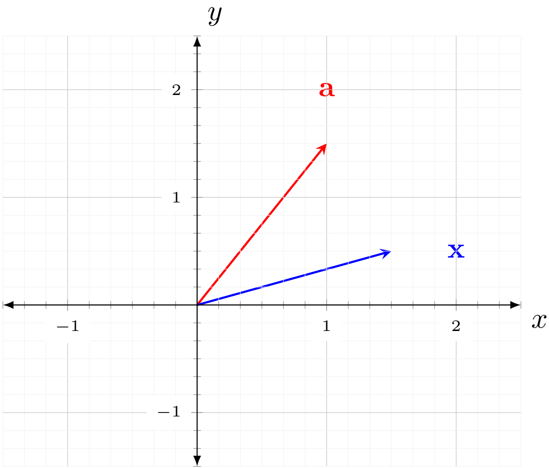
[**Functional Analysis**] A functional is a mapping from a space X into the field of real or complex numbers

$$f : X \to \mathbb{R}, \quad \text{or} \quad f : X \to \mathbb{C}$$

[**Computer Science**] functional is synonymous with a higher-order function, which is a function that takes one or more functions as arguments or returns them.

Examples

Let $\mathbf{x} \in \mathbb{R}^n$, then for a fixed $\mathbf{a} \in \mathbb{R}^n$, the mapping $\mathbf{x} \mapsto \mathbf{a}^T\mathbf{x} \in \mathbb{R}$ is a functional.



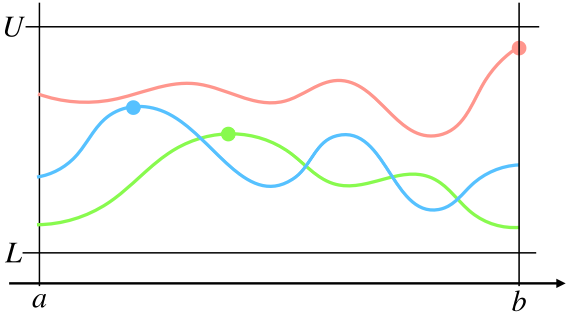
$$\mathbf{x} = \begin{bmatrix} 1.5 \\ 0.5 \end{bmatrix}$$

$$\mathbf{a} = \begin{bmatrix} 1 \\ 1.5 \end{bmatrix}$$

$$\mathbf{a}^T\mathbf{x} = \begin{bmatrix} 1 & 1.5 \end{bmatrix} \begin{bmatrix} 1.5 \\ 0.5 \end{bmatrix} = 2.25$$

Examples (contd.)

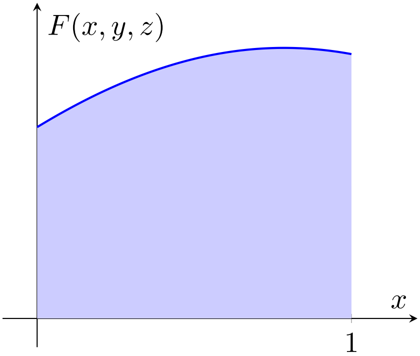
Let $V := \{f : [a, b] \to \mathbb{R} \,|\, L \leq f(x) \leq U\}$, then $f \mapsto \sup(f)$ is a functional



Examples (contd.)

Let $C^k(I)$ denote the space of functions with k continuous derivatives on a bounded interval I, then the following map is a functional.

$$u \mapsto \int_I F\left(u,\, u',\, u'',\, \dots,\, u^{(k)}\right) \mathrm{d}x,$$



$$\begin{aligned}
F(x,\, y,\, z) &= |y(x)| + |z(x)|, \\
y &= \sin(x), \\
y' &= \cos(x), \\
F(x,\, y,\, y') &= |\sin(x)| + |\cos(x)|, \\
\int_0^1 F(x,\, y,\, y')\, dx &= \int_0^1 \left(|\sin(x)| + |\cos(x)|\right)\, dx \\
&= \sin(1) - \cos(1) + 1 \approx 1.3012.
\end{aligned}$$

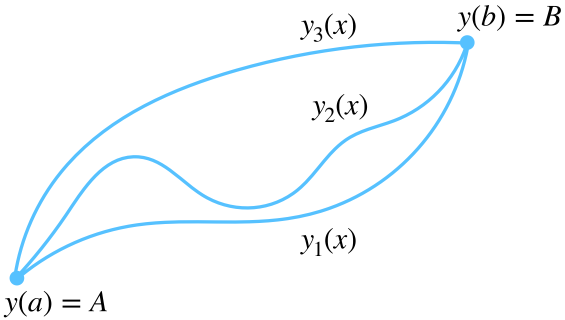
Variational Problem Examples

Example: Shortest Path

Find the shortest plane curve joining points A and B. That is: find $y = y(x)$ for which

$$\int_a^b \sqrt{1 + (y')^2}\, \mathrm{d}\, x,$$

achieves its minimum.



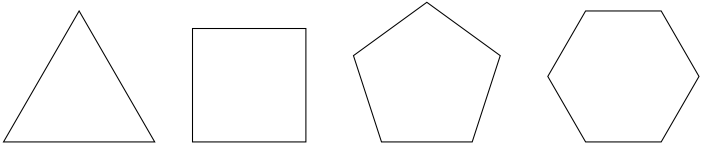
Example: Brachistochrone Problem [1696 John Bernoulli]

Let A and B be two fixed points, then the time it takes a particle to slide under the influence of gravity along some path joining A and B depends on the choice of the path.

Solved by: John Bernoulli, James Bernoulli, Newton, L'Hospital

Example: Isoperimetric [solved by Euler (1707 to 1783)]

Among all the closed curves of a given length l, find the curve enclosing the greatest area.



Necessary Condition for Extrema

$f : \mathbb{R}^n \to \mathbb{R}$ is said to have a relative (local) extremum at $\mathbf{x}^*$ if

$$\Delta f = f(\mathbf{x}) - f(\mathbf{x}^*)$$

has the same sign for all $\mathbf{x}$ in some neighborhood of $\mathbf{x}^*$.

$$\Delta f \geq 0 \implies x^* \text{ is a minimum}$$
$$\Delta f \leq 0 \implies x^* \text{ is a maximum}$$

Analogously, a functional $J(y)$ has a relative extremum at y^* if $J(y) - J(y^*)$ does not change sign in some neighborhood of y^*.

Variation (Differential) of a Functional

Let $J(y)$ be functional defined on some metric space** and, for a fixed y, let

$$\Delta J(h) := J(y+h) - J(y),$$

be its *increment* corresponding to $h = h(x)$. Suppose

$$\Delta J(h) = \phi(h) + \epsilon \|h\|,$$

where $\phi(h)$ is a linear functional and $\epsilon \to 0$ as $\|h\| \to 0$. Then, $J(y)$ is *differentiable*, and the principal linear part $\phi(h)$, denoted by $\delta J(h)$, is the *variation* of $J(y)$.

Theorem 17.1

The differential of a differentiable functional is unique.

For proof see: "Calculus of Variations" by I. M. Gelfand, S. V. Fomin

**Here we assume a normed vector space to simplify derivations.

Example

Let $y \in C^0(0,\,1)$. Find the variation of the functional

$$J(y) = \int_0^1 y^2(x) + 2\,y(x)\,dx.$$

Solution

$$\begin{aligned}
\Delta J(h) = J(y+h) - J(y) &= \int_0^1 (y+h)^2 + 2\,\left(y+h\right)\,dx - \int_0^1 y^2 + 2\,y\,dx \\
&= \int_0^1 y^2 + h^2 + 2\,y\,h + 2\,y + 2\,h - y^2 - 2y\,dx \\
&= \int_0^1 \left(2\,y + 2\right)\,h + h^2\,dx \\
&= \int_0^1 \left(2\,y + 2\right)\,h\,dx + \int_0^1 h^2\,dx.
\end{aligned}$$

Example (contd.)

We need to check if $\Delta J(h) = \phi(h) + \epsilon\|h\|$ where ϕ is a linear functional and $\|h\| \to 0$ as $\epsilon \to 0$. In fact,

$$\int_0^1 h^2\,dx = \frac{\|h\|}{\|h\|}\int_0^1 h^2\,dx = \|h\|\int_0^1 \frac{h^2}{\|h\|}\,dx = \|h\|\int_0^1 \frac{|h|\,|h|}{\|h\|}\,dx$$

Since $h \in C^0(0,1)$, let $\|h\| := \max\left|h(x)\right|$. Hence $|h| \leq \|h\|$, that implies $|h|/\|h\| \leq 1$ and

$$\int_0^1 \frac{|h|\,|h|}{\|h\|}\,dx \leq \int_0^1 |h|\,dx \leq \int_0^1 \|h\|\,dx = \|h\|.$$

Hence $\|h\| \to 0$ implies $\epsilon = \int_0^1 h^2/\|h\|\,dx \to 0$. Thus, we can conclude:

$$\delta J(y,\, h) = \int_0^1 (2\,y + 2)\;h\,dx.$$

The Fundamental Theorem of the Calculus of Variations

Theorem 17.2

A necessary condition for the differentiable functional $J(y)$ to have an extremum at y^* is

$$\delta J(h) = 0, \quad \textit{at} \quad y = y^*$$

Proof. Suppose $J(y)$ has a minimum at $y = y^*$. According to the definition

$$\Delta J(h) = J(y+h) - J(y) = \delta J(h) + \epsilon \|h\|, \quad \|h\| \to 0 \implies \epsilon \to 0.$$

Thus for sufficiently small $\|h\|$, $\text{sign}\left[\delta J(h)\right] = \text{sign}\left[\Delta J(h)\right]$. Suppose $\delta J(h_0) \neq 0$ for some admissible h_0, then for any $\alpha > 0$

$$\delta J(-\alpha\, h_0) = -\delta J(\alpha\, h_0).$$

Hence, $\Delta J(h)$ can be made to have either sign for arbitrarily small $\|h\|$. That is a contradiction to y^* being an extremum. □

The Simplest Variational Problem - Euler's Equation

Let $f = f(x, y, z)$ be a function with continuous first and second partial derivatives with respect to all its arguments. Then for all $y(x)$ which are continuously differentiable for $a \leq x \leq b$ and satisfy boundary conditions

$$y(a) = A, \quad y(b) = B,$$

find the function for which the functional

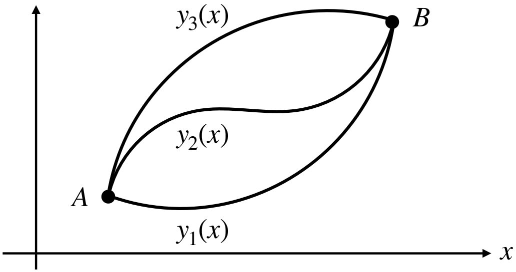
$$J(y) = \int_a^b f\left(x, y, y'\right) \mathrm{d}x$$

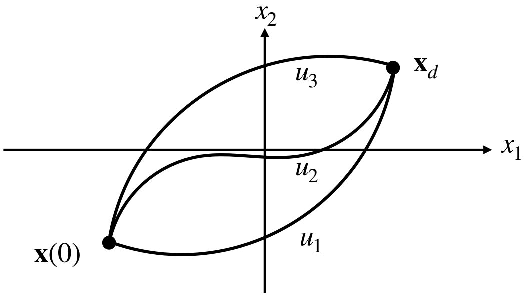
has a weak extremum.

The Simplest Variational Problem: Connection to Control

$$J(y) = \int_a^b f\left(x,\, y,\, y'\right) \mathrm{d}x,$$

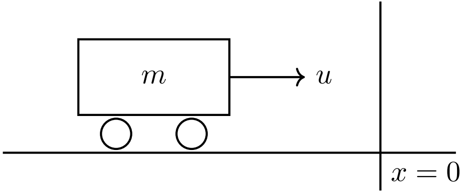
$$y(a) = A, \quad y(b) = B.$$

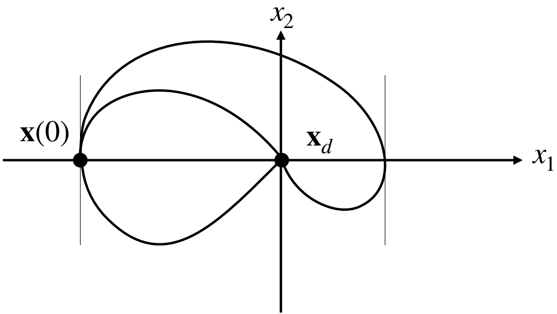




The Simplest Variational Problem: Connection to Control (contd.)

$$\dot{x}_1 = x_2,$$
$$\dot{x}_2 = u/m.$$





Clearly, not all paths in the state-space is a feasible trajectory!

Tools Needed To Find the Answer to the Simplest Problem

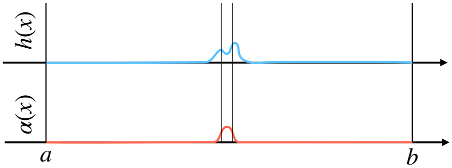
Lemma 17.3

If $\alpha(x)$ is continuous in $[a,\, b]$, and if

$$\int_a^b \alpha(x)\, h(x)\, dx = 0$$

for every function $h \in C^0(a, b)$ such that $h(a) = h(b) = 0$, then $\alpha(x) = 0$ for all $x \in [a,\, b]$.

Sketch of the proof:



Tools Needed To Find the Answer to the Simplest Problem (contd.)

Lemma 17.4

If $\alpha(x)$ *is continuous in* $[a,\, b]$*, and if*

$$\int_a^b \alpha(x)\, h'(x)\, dx = 0$$

for every function $h \in C^1(a,b)$ *such that* $h(a) = h(b) = 0$*, then* $\alpha(x) = c$ *for all* $x \in [a,\, b]$*.*

Proof. Let c be the constant defined in the Lemma. Accordingly

$$\int_a^b \big(\alpha(x) - c\big)\; dx = 0.$$

Let

$$h(x) = \int_a^x \big(\alpha(\xi) - c\big)\; d\xi.$$

Tools Needed To Find the Answer to the Simplest Problem (contd.)

We can see that $h(x) \in C^1(a,\, b)$ and $h(a) = h(b) = 0$. Then from one hand we get

$$\int_a^b \big(\alpha(x) - c\big)\; h'(x)\, dx = \int_a^b \alpha(x)\, h'(x)\, dx - c\; \big(h(b) - h(a)\big) = 0.$$

while on the other hand we have

$$0 = \int_a^b \big(\alpha(x) - c\big)\, h'(x)\, dx = \int_a^b \big(\alpha(x) - c\big)^2\; dx \implies \big(\alpha(x) - c\big)^2 = 0 \implies \alpha(x) = c.$$

$\square$

Tools Needed To Find the Answer to the Simplest Problem (contd.)

Lemma 17.5

If $\alpha(x)$ *and* $\beta(x)$ *are continuous in* $[a,\, b]$*, and if*

$$\int_a^b \left[\alpha(x)\, h(x) + \beta(x)\, h'(x)\right]\, dx = 0$$

for every function $h \in C^1(a, b)$ *such that* $h(a) = h(b) = 0$*, then* $\beta(x)$ *is differentiable and* $\beta'(x) = \alpha(x)$ *for all* $x \in [a,\, b]$*.*

Proof. Let

$$A(x) = \int_a^x \alpha(\xi)\, d\xi.$$

Using integration by parts we get

$$\int_a^b \alpha(x)\, h(x)\, dx = -\int_a^b A(x)\, h'(x)\, dx.$$

Tools Needed To Find the Answer to the Simplest Problem (contd.)

Accordingly

$$\int_a^b \left[\alpha(x)\, h(x) + \beta(x)\, h'(x)\right]\, dx = \int_a^b \left[-A(x) + \beta(x)\right]\, h'(x)\, dx = 0$$

According to Lemma 17.4

$$\beta(x) - A(x) = c \implies \beta'(x) = \alpha(x),$$

for all $x \in [a,\, b]$. Note that the differentiability of $\beta(x)$ was not assumed in advance. $\square$

The Simplest Variational Problem - Solution

Let $h(x)$ be continuously differentiable for $a \leq x \leq b$ and $h(a) = h(b) = 0$. Then

$$\begin{aligned}\Delta J = J(y+h) - J(y) &= \int_a^b f\left(x,\, y+h,\, y'+h'\right) \mathrm{d}x - \int_a^b f\left(x,\, y,\, y'\right)\, dx \\ &= \int_a^b \Big(f\left(x,\, y+h,\, y'+h'\right) - f\left(x,\, y,\, y'\right) \Big)\, dx.\end{aligned}$$

Using Taylor's theorem we get

$$\Delta J = \int_a^b \Big(f_y\left(x,\, y,\, y'\right) h + f_{y'}\left(x,\, y,\, y'\right) h' \Big)\, dx + \cdots\,,$$

where $f_y = \frac{\partial f}{\partial y}$ and $f_{y'} = \frac{\partial f}{\partial y'}$.

The Simplest Variational Problem - Solution (contd.)

The first integral of ΔJ represents the principal linear part. Hence

$$\delta J = \int_a^b \left[f_y \left(x,\, y,\, y' \right) h + f_{y'} \left(x,\, y,\, y' \right) h' \right] dx.$$

Based on the Theorem 17.2, a necessary condition for $J(y)$ to have an exremum is

$$\delta J = \int_a^b \left[f_y \left(x,\, y,\, y' \right) h + f_{y'} \left(x,\, y,\, y' \right) h' \right] dx = 0,$$

for all admissible h. According to Lemma 17.5, we have

$$\delta J = 0 \implies f_y - \frac{d}{dx} f_{y'} = 0.$$

Euler's Equation

Theorem 17.6

Let $J(y)$ be a functional of the form

$$J(y) = \int_a^b f(x,\, y,\, y')\, dx,$$

defined on the set of functions $y(x)$ which have continuous first derivatives in $[a,\, b]$ and satisfy the boundary conditions $y(a) = A$, $y(b) = B$. Then a necessary condition for $J(y)$ to have an extremum for a given function $y(x)$ is that $y(x)$ satisfies Euler's equation

$$f_y - \frac{d}{dx} f_{y'} = 0.$$

Example

Find an extremal for the functional

$$J(x) = \int_0^{\pi/2} f\left(t,\, x(t),\, \dot{x}(t)\right)\, dx = \int_0^{\pi/2} \dot{x}^2(t) - x^2(t)\, dt$$

over $x \in C^2\left(0, \frac{\pi}{2}\right)$ such that $x(0) = 0$ and $x\left(\frac{\pi}{2}\right) = 1$.

Solution

$$f_x = \frac{\partial f}{\partial x} = -2\, x, \quad f_{\dot{x}} = \frac{\partial f}{\partial x'} = 2\, \dot{x}, \quad \frac{\mathrm{d}}{\mathrm{d}x}\, f_{\dot{x}} = 2\, \ddot{x}.$$

Based on Euler's equation the extremal is the solution to the Boundary Value Problem (BVP)

$$\begin{cases} f_x - \dfrac{\mathrm{d}}{\mathrm{d}x}\, f_{\dot{x}} = -2\, x - 2\ddot{x} = 0 \\ x(0) = 0, \\ x\left(\frac{\pi}{2}\right) = 0. \end{cases}$$

$$\begin{aligned} \ddot{x} + x = 0 &\implies x(t) = c_1 \cos(t) + c_2 \sin(t), \\ x(0) = 0 &\implies c_1 = 0, \\ x\left(\pi/2\right) = 1 &\implies c_2 = 1, \end{aligned}$$

Thus, the extremal is $x(t) = \sin(t)$.

Connection to Classical Analysis (functions of n variables)
From Infinite to Finite Dimensions

Method of Finite Differences [Euler]

Consider the functional

$$J(y) = \int_a^b f\left(x,\, y,\, y'\right) \mathrm{d}x, \quad y(a) = A, \quad y(b) = B.$$

Using the points

$$a = x_0,\, x_1,\, \ldots,\, x_n,\, x_{n+1} = b,$$

we divide the interval $[a,\, b]$ to $n+1$ parts. Then we replace $y = y(x)$ by the polygonal line with vertices

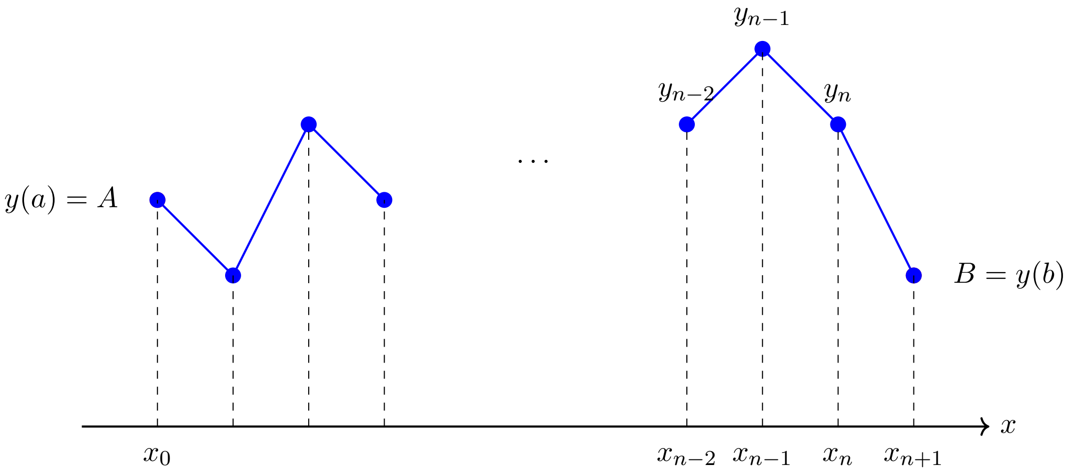
$$\left(x_0,\, A\right),\, \left(x_1,\, y\left(x_1\right)\right),\, \ldots\, \left(x_n,\, y\left(x_n\right)\right),\, \left(x_{n+1},\, B\right),$$

and we approximate $J(y)$ with the sum

$$J\left(y_1,\, \ldots,\, y_n\right) = \sum_{i=1}^{n+1} f\left(x_i,\, y_i,\, \frac{y_i - y_{i-1}}{h}\right) h_i,$$

where $y_i = y\left(x_i\right)$, $h_i = x_i - x_{i-1}$.

Conceptual image



Introduction to Optimal Control

Optimal Control Problem

Let $\dot{\mathbf{y}} = f(t, \mathbf{y}, \mathbf{u})$ be a given control system. Informally, an optimal control problem is to find $\mathbf{u}^* \in U$ (set of admissible control) that minimizes the performance measure $J\left(\mathbf{y_u},\, \mathbf{u}\right)$, that is

$$\mathbf{u}^* = \arg\min_{\mathbf{u}\in U} J(\mathbf{y_u}, \mathbf{u}),$$

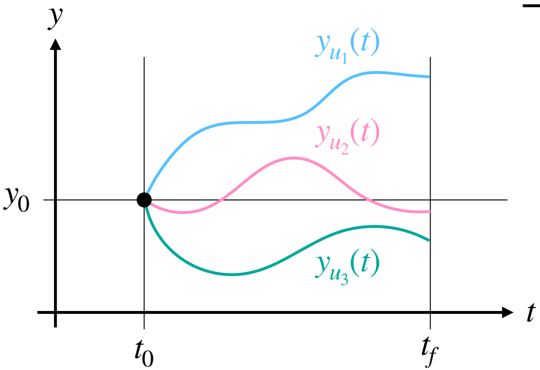
such that for for all $t \in T = \left[t_0,\, t_f\right]$ we have

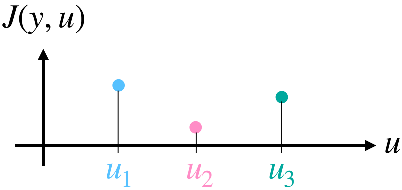
$$\begin{aligned} G(t, \mathbf{y}, \mathbf{u}) &\leq \mathbf{0}, \\ E(t, \mathbf{y}, \mathbf{u}) &= \mathbf{0}, \end{aligned}$$

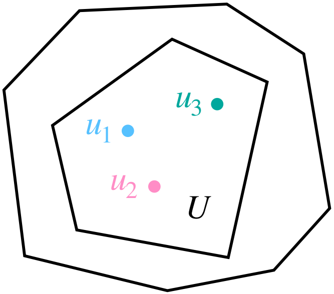
where, for a fixed $\mathbf{u}$, the function $\mathbf{y_u}$ is the solution to the initial-value problem

$$\begin{cases} \dot{\mathbf{y}} = f(t, \mathbf{y}, \mathbf{u}), \quad t \in \left(t_0,\, t_f\right), \\ \mathbf{y}(t_0) = \mathbf{y}_0. \end{cases}$$

Conceptual Image

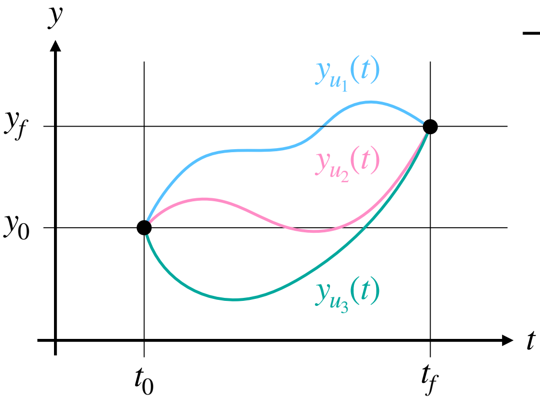


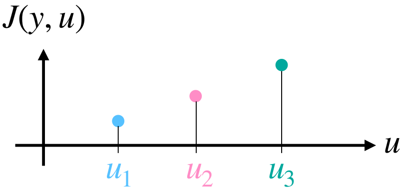


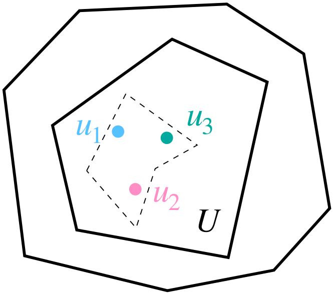


Conceptual Image

$$E = y(t_f) - y_f = 0$$



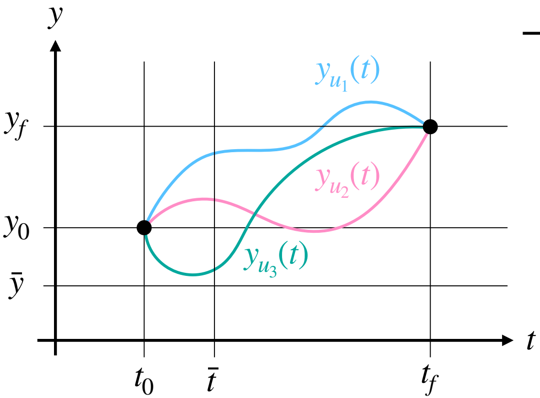


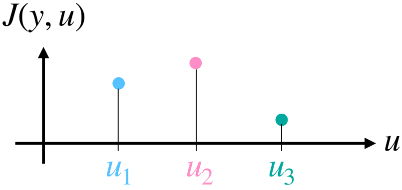


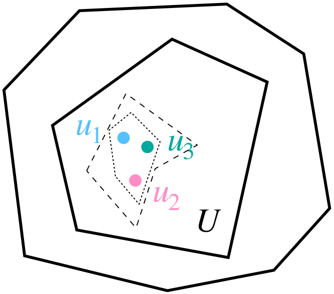
Conceptual Image

$$E = y(t_f) - y_f = 0$$

$$G = \bar{y} - y(\bar{t}) \leq 0$$







Pontryagin's Minimum Principle

Theorem 17.7 (Pontryagin's Minimum Principle)

Given $\dot{\mathbf{y}} = f(t, \mathbf{y}, \mathbf{u})$, *a time interval* $T = [t_0, t_f]$ *and a set of feasible control functions* U. *Let*

$$J(\mathbf{u}) := h\big(t_f, \mathbf{y}(t_f)\big) + \int_{t_0}^{t_f} g(t, \mathbf{y}, \mathbf{u})\ \mathrm{d}t,$$

and let $H(t, \mathbf{y}, \boldsymbol{\lambda}, \mathbf{u})$ *represent the Hamiltonian function defined as*

$$H(t, \mathbf{y}, \boldsymbol{\lambda}, \mathbf{u}) := g(t, \mathbf{y}, \mathbf{u}) + \boldsymbol{\lambda}^T f(t, \mathbf{y}, \mathbf{u}),$$

Then, $\forall t \in T$ *and* $\forall \mathbf{u} \in U$, *the control* $\mathbf{u}^*$ *that minimizes* J *must satisfy the necessary conditions:*

Pontryagin's Minimum Principle (contd.)

Theorem 17.7 (Pontryagin's Minimum Principle)

1. $\dot{\mathbf{y}}^*(t) = H_{\boldsymbol{\lambda}}(t, \mathbf{y}^*, \boldsymbol{\lambda}^*, \mathbf{u}^*)$
2. $\dot{\boldsymbol{\lambda}}^* = -H_{\mathbf{y}}(t, \mathbf{y}^*, \boldsymbol{\lambda}^*, \mathbf{u}^*)$
3. $H(t, \mathbf{y}^*, \boldsymbol{\lambda}^*, \mathbf{u}^*) \leq H(t, \mathbf{y}^*, \boldsymbol{\lambda}^*, \mathbf{u})$
4. $\left[h_{\mathbf{y}}\left(t_f, \mathbf{y}^*(t_f)\right) - \boldsymbol{\lambda}^*(t_f)\right]^T \delta\mathbf{y}_f + \left[h_t\left(t_f, \mathbf{y}^*(t_f)\right) + H\left(t_f, \cdots\right)\right] \delta t_f = 0;$

where $\delta\mathbf{y}_f$ and δt_f are respectively the variations of the final state and time. Moreover, if there is no constraint on the range of functions in U, that is for all $\mathbf{u} \in U$, $\mathbf{u}(\cdot) \in \mathbb{R}^m$, then condition (3) simplifies to

3. $H_{\mathbf{u}}(t, \mathbf{y}^*, \boldsymbol{\lambda}^*, \mathbf{u}^*) = \mathbf{0}.$

In all the above statements $H_x = \frac{\partial H}{\partial x}$.

Derivation of Linear Quadratic Regulator (LQR)

The derivation here is adopted from *Donald E Kirk. "Optimal control theory: an introduction"* and are primarily due to the work of R. E. Kalman.
We want to find a linear control $\mathbf{v} := \mathbf{K}\,\mathbf{e}$ for

$$\dot{\mathbf{e}} = \mathbf{A}\,\mathbf{e} + \mathbf{B}\,\mathbf{v},$$

that minimizes

$$J := \frac{1}{2}\mathbf{e}^T(t_f)\,\mathbf{C}\,\mathbf{e}(t_f) + \frac{1}{2}\int_{t_0}^{t_f} \mathbf{e}^T\mathbf{Q}\mathbf{e} + \mathbf{v}^T\mathbf{R}\mathbf{v}\;\mathrm{d}t,$$

Assuming:

- $t_f < \infty$ is fixed;
- $\mathbf{C}$, $\mathbf{Q} \in \mathbb{R}^{n\times n}$ are *symmetric positive semi-definite*;
- $\mathbf{R} \in \mathbb{R}^{m\times m}$ is a *symmetric positive definite matrix*;
- $\mathbf{e}(\cdot) \in \mathbb{R}^n$ and $\mathbf{v}(\cdot) \in \mathbb{R}^m$ are not bounded and $\mathbf{e}(t_f)$ is not constrained

Derivation of Linear Quadratic Regulator (LQR) (contd.)

The Hamiltonian for the problem is

$$H(t, \mathbf{e}, \boldsymbol{\lambda}, \mathbf{v}) = \frac{1}{2}\mathbf{e}^T \mathbf{Q}\, \mathbf{e} + \frac{1}{2}\mathbf{v}^T \mathbf{R} \mathbf{v} + \boldsymbol{\lambda}^T \left(\mathbf{A}\, \mathbf{e} + \mathbf{B}\, \mathbf{v}\right),$$

and the necessary conditions of Pontryagin's minimum principle are

$$\dot{\mathbf{e}}^* = H_{\boldsymbol{\lambda}}\left(t,\, \mathbf{e}^*,\, \boldsymbol{\lambda}^*,\, \mathbf{v}^*\right) = \mathbf{A}\, \mathbf{e}^* + \mathbf{B}\, \mathbf{v}^*, \tag{17.597a}$$

$$\dot{\boldsymbol{\lambda}}^* = -H_{\mathbf{e}}\left(t,\, \mathbf{e}^*,\, \boldsymbol{\lambda}^*,\, \mathbf{v}^*\right) = -\mathbf{Q}\, \mathbf{e}^* - \mathbf{A}^T \boldsymbol{\lambda}^*, \tag{17.597b}$$

$$\mathbf{0} = H_{\mathbf{v}}\left(t,\, \mathbf{e}^*,\, \boldsymbol{\lambda}^*,\, \mathbf{v}^*\right) = \mathbf{R}\, \mathbf{v}^* + \mathbf{B}^T \boldsymbol{\lambda}^*. \tag{17.597c}$$

Solving (17.597c) for $\mathbf{v}^*$ and substituting in condition (17.597a) gives

$$\dot{\mathbf{e}}^* = \mathbf{A}\, \mathbf{e}^* - \mathbf{B}\, \mathbf{R}^{-1} \mathbf{B}^T \boldsymbol{\lambda}^*. \tag{17.598}$$

Derivation of Linear Quadratic Regulator (LQR) (contd.)

Equations (17.597b) and (17.598) form a set of linear homogeneous differential equations

$$\begin{bmatrix} \dot{\mathbf{e}}^* \\ \dot{\boldsymbol{\lambda}}^* \end{bmatrix} = \begin{bmatrix} \mathbf{A} & -\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^T \\ -\mathbf{Q} & -\mathbf{A}^T \end{bmatrix} \begin{bmatrix} \mathbf{e}^* \\ \boldsymbol{\lambda}^* \end{bmatrix}, \tag{17.599}$$

with the solution

$$\begin{bmatrix} \mathbf{e}^*(t_f) \\ \boldsymbol{\lambda}^*(t_f) \end{bmatrix} = \begin{bmatrix} \boldsymbol{\phi}_{11}(t, t_f) & \boldsymbol{\phi}_{12}(t, t_f) \\ \boldsymbol{\phi}_{21}(t, t_f) & \boldsymbol{\phi}_{22}(t, t_f) \end{bmatrix} \begin{bmatrix} \mathbf{e}^*(t) \\ \boldsymbol{\lambda}^*(t) \end{bmatrix}. \tag{17.600}$$

From condition (4) of Pontryagin's Minimum Principle: setting $\delta t_f = 0$ (since the final time is fixed) we get

$$\boldsymbol{\lambda}^*(t_f) = \mathbf{C}\,\mathbf{e}^*(t_f). \tag{17.601}$$

Derivation of Linear Quadratic Regulator (LQR) (contd.)

Substituting (17.601) in (17.600) gives

$$\mathbf{e}^*(t_f) = \boldsymbol{\phi}_{11}(t, t_f)\,\mathbf{e}^*(t) + \boldsymbol{\phi}_{12}(t, t_f)\,\boldsymbol{\lambda}^*(t), \tag{17.602}$$

$$\mathbf{C}\,\mathbf{e}^*(t_f) = \boldsymbol{\phi}_{21}(t, t_f)\,\mathbf{e}^*(t) + \boldsymbol{\phi}_{22}(t, t_f)\,\boldsymbol{\lambda}^*(t). \tag{17.603}$$

Substituting the upper equation into the lower gives

$$\mathbf{C}\,\boldsymbol{\phi}_{11}\,\mathbf{e}^*(t) + \mathbf{C}\,\boldsymbol{\phi}_{12}\,\boldsymbol{\lambda}^*(t) = \boldsymbol{\phi}_{21}\,\mathbf{e}^*(t) + \boldsymbol{\phi}_{22}\,\boldsymbol{\lambda}^*(t). \tag{17.604}$$

Solving the above equation for $\boldsymbol{\lambda}^*$ gives

$$\boldsymbol{\lambda}^* = (\boldsymbol{\phi}_{22} - \mathbf{C}\,\boldsymbol{\phi}_{12})^{-1}\,(\mathbf{C}\,\boldsymbol{\phi}_{11} - \boldsymbol{\phi}_{21})\,\mathbf{e}^*(t).$$

Kalman has shown that the required inverse exists for all $t \in [t_0, t_f]$. The above equation can be written as

$$\boldsymbol{\lambda}^*(t) = \mathbf{K}(t,\, t_f)\,\mathbf{e}(t)$$

Derivation of Linear Quadratic Regulator (LQR) (contd.)

From (17.597c) we have $\mathbf{R}\,\mathbf{v}^* + \mathbf{B}^T\,\boldsymbol{\lambda}^*$. Substituting $\boldsymbol{\lambda}^*(t) = \mathbf{K}(t,\,t_f)\,\mathbf{e}(t)$ gives

$$\begin{aligned}\mathbf{v}^* &= -\mathbf{R}^{-1}\,\mathbf{B}^T\,\mathbf{K}\,\mathbf{e}\\ &= \mathbf{F}(t,\,t_f)\,\mathbf{e},\end{aligned}$$

Note:

- The optimal control law is a linear, albeit time varying, combination of the system states;
- Even if the plant is fixed, the feedback gain matrix $\mathbf{F}$ is time-varying
- Generally, in order to implement $\mathbf{v}^*$, we need to resort to numerical procedures to evaluate $\boldsymbol{\phi}_{ij}(t, t_f)$.

Derivation of Linear Quadratic Regulator (LQR) (contd.)

Alternatively, it can be shown that the matrix $\mathbf{K}$ satisfies the matrix differential equation

$$\dot{\mathbf{K}}(t) = -\mathbf{K}(t)\,\mathbf{A}(t) - \mathbf{A}^T(t)\,\mathbf{K}(t) - \mathbf{Q}(t) + \mathbf{K}(t)\,\mathbf{B}(t)\,\mathbf{R}^{-1}(t)\,\mathbf{B}^T(t)\,\mathbf{K}(t),$$
$$\mathbf{K}(t_f) = \mathbf{C}$$

This matrix differential equation is of the *Riccati* type; and we call this equation the *Riccati equation*.

Note:

- Since $\mathbf{K}$ is an $n \times n$ matrix, this equation is a system of n^2 first order differential equations.
- it can be shown that $\mathbf{K}$ is symmetric; hence, not n^2, but $n(n+1)/2$ first order differential equations must be solved

Derivation of Time Invariant LQR

If

- $\mathbf{A}$, $\mathbf{B}$, $\mathbf{R}$, $\mathbf{Q}$ are time invariant
- $\mathbf{C} = \mathbf{0}$
- $t_f = \infty$

then

$$\lim_{t \to \infty} \mathbf{K}(t) = \mathbf{K},$$

is a constant matrix. Consequently, the linear optimal control strategy is a constant linear map applied to $\mathbf{e}(t)$. In this case, $\mathbf{K}$ could be obtained by solving

$$\mathbf{K}\,\mathbf{A} + \mathbf{A}^T\,\mathbf{K} + \mathbf{Q} - \mathbf{K}\,\mathbf{B}\,\mathbf{R}^{-1}\,\mathbf{B}^T\,\mathbf{K} = \mathbf{0},$$

which is obtained by setting $\dot{\mathbf{K}}(t) = \mathbf{0}$ in *Riccati equation*.

Numerical Approaches in Optimal Control

Indirect Methods

setting up the necessary conditions for optimality (such as Pontryagin's Maximum Principle) and solving the resulting boundary-value problem

Direct Methods

convert the optimal control problem into a finite-dimensional parameter optimization problem by discretizing the control and state trajectories.

Finite Differences in State and Control

$$\begin{cases} \dot{\mathbf{x}} = f\left(t,\, \mathbf{x},\, \mathbf{u}\right), \\ \mathbf{x}(a) = \mathbf{x}_0, \end{cases}$$

where $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{u} \in \mathbb{R}^m$. Consider the Given $N \in \mathbb{N}$, consider discretization of time

$$a = t_0 < t_1 < t_2 < \cdots < t_N < t_{N+1} = b.$$

For $k \in \{0, \ldots, N+1\}$

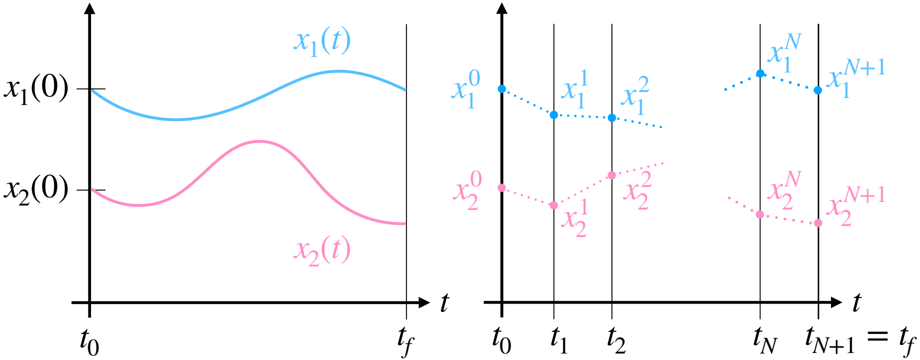
$$\mathbf{x}^k = \left(x_1^k,\, x_2^k,\, \ldots,\, x_n^k\right),$$
$$\mathbf{u}^k = \left(u_1^k,\, u_2^k,\, \ldots,\, u_m^k\right).$$

Discretization of Space

Assume fixed $\mathbf{x}^0 = \mathbf{x}(a)$, a sample discretization of state is

$$\hat{\mathbf{x}} = \left(\mathbf{x}^1, \ldots, \mathbf{x}^N, \mathbf{x}^{N+1}\right) \in \mathbb{R}^{n(N+1)}$$

The following sketch shows a sample for $n = 2$, $a = 0$ and $b = t_f$:

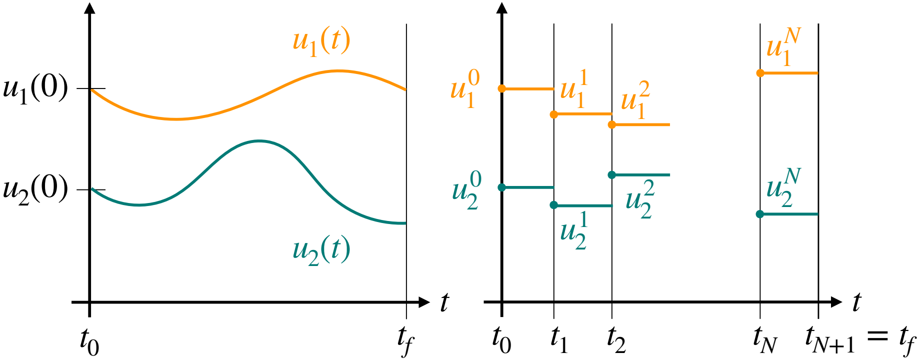


Discretization of Control

A sample discretization of control is

$$\hat{\mathbf{u}} = \left(\mathbf{u}^0, \mathbf{u}^1, \ldots, \mathbf{u}^N,\right) \in \mathbb{R}^{m(N+1)}$$

The following sketch shows a sample for $m = 2$, $a = 0$, $b = t_f$ and zero-order hold:



Differential Equation Constraint

Given $\hat{\mathbf{u}} \in \mathbb{R}^{m(N+1)}$ and assuming construction of control $\mathbf{u}(t)$ via zero-order hold, we need the point $\hat{\mathbf{x}} \in \mathbb{R}^{n(N+1)}$ to coincide with the solution of

$$\begin{cases} \dot{\mathbf{x}} = f\left(t, \mathbf{x}, \mathbf{u}\right), \\ \mathbf{x}(a) = \mathbf{x}_0, \end{cases}$$

From Taylor's theorem we have

$$\begin{aligned} \mathbf{x}(t + \delta t) &= \mathbf{x}(t) + \frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t}\,\delta t + \cdots, \\ &= \mathbf{x}(t) + f\left(t, \mathbf{x}(t), \mathbf{u}(t)\right)\,\delta t + \cdots \end{aligned}$$

Adopting the above identity to $\hat{\mathbf{x}}$ we get

$$\mathbf{x}^k + f\left(t, \mathbf{x}^k, \mathbf{u}^k\right) - \mathbf{x}^{k+1} = \mathbf{0}, \quad \forall k \geq 0.$$

Approximation of The Cost Functional

For the time interval $[a,\, b]$, let the cost functional for the optimal control problem be

$$J(\mathbf{x},\, \mathbf{u}) := h\left(b,\, \mathbf{x}(b)\right) + \int_a^b g\left(t, \mathbf{y}, \mathbf{u}\right)\, dt,$$

Using the discretization of state and control, we can approximate the cost functional as the sum

$$J\left(\hat{\mathbf{x}},\, \hat{\mathbf{u}}\right) = h\left(b,\, \mathbf{x}^{N+1}\right) + \sum_{k=0}^{N} g\left(t_k, \mathbf{x}^k, \mathbf{u}^k\right)\left(t_{k+1} - t_k\right).$$

Simple Finite-Dimensional Optimal Control Problem

find $(\hat{\mathbf{x}}^*,\, \hat{\mathbf{u}}^*) \in \mathbb{R}^{(n+m)(N+1)}$ that minimizes the cost functional

$$J\left(\hat{\mathbf{x}},\, \hat{\mathbf{u}}\right) = h\left(b,\, \mathbf{x}^{N+1}\right) + \sum_{k=0}^{N} g\left(t_k, \mathbf{x}^k, \mathbf{u}^k\right)\, \left(t_{k+1} - t_k\right).$$

subjected to

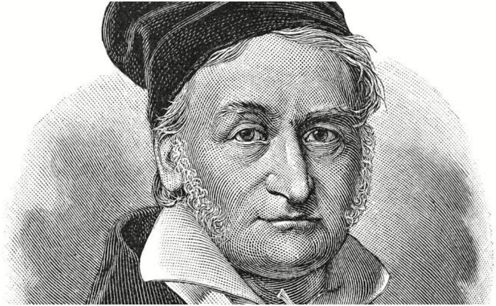
$$\begin{aligned} \mathbf{x}^k + f\left(t,\, \mathbf{x}^k,\, \mathbf{u}^k\right) - \mathbf{x}^{k+1} &= \mathbf{0}, \\ E\left(t,\, \mathbf{x}^k,\, \mathbf{u}^k\right) &= 0, \\ G\left(t,\, \mathbf{x}^k,\, \mathbf{u}^k\right) &\leq 0. \end{aligned}$$

Contents

Method of Least Squares

The most probable value of the unknown quantities will be that in which the sum of the squares of the differences between the actually observed and the computed values multiplied by numbers that measure the degree of precision is a minimum.

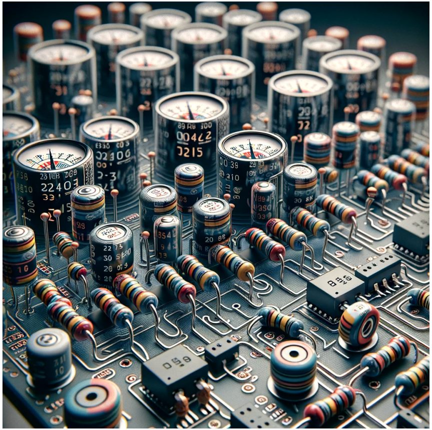
— Karl Friedrich Gauss (30 April 1777 – 23 February 1855)



Method of Least Squares

Assume that each element of the measurement $\mathbf{y} = (y_1, \ldots, y_k)$ is a linear combination of the *unknown parameter vector* $\mathbf{x} = (x_1, \ldots, x_n)$, with the addition of some measurement noise:

$$
\begin{aligned}
y_1 &= H_{11}\, x_1 + H_{12}\, x_2 + \cdots + H_{1n}\, x_n + v_1 \\
&\vdots \\
y_k &= H_{k1}\, x_1 + H_{k2}\, x_2 + \cdots + H_{kn}\, x_n + v_k
\end{aligned}
$$



In matrix form

$$\mathbf{y} = \mathbf{H}\,\mathbf{x} + \mathbf{v}$$

Method of Least Squares - Derivation

The question is: How can we find the "best" estimate $\hat{\mathbf{x}}$ of $\mathbf{x}$?

Let $\boldsymbol{\epsilon}_{\mathbf{y}} = \left(\epsilon_{y_1}, \ldots, \epsilon_{y_k}\right)$ be the difference between the noisy measurements and the vector $\mathbf{H}\,\hat{\mathbf{x}}$

$$\boldsymbol{\epsilon}_{\mathbf{y}} = \mathbf{y} - \mathbf{H}\,\hat{\mathbf{x}}.$$

$\boldsymbol{\epsilon}_{\mathbf{y}}$ is called the *measurement residual*. Based on Gauss, $\hat{\mathbf{x}}$ is the vector that minimizes the cost functional

$$J = {\epsilon_{y_1}}^2 + \cdots + {\epsilon_{y_k}}^2 = \boldsymbol{\epsilon}_{\mathbf{y}}^T\,\boldsymbol{\epsilon}_{\mathbf{y}}$$

See that if there is no measurement noise or measurement error (that is $\mathbf{v} = \mathbf{0}$), then

$$\boldsymbol{\epsilon}_{\mathbf{y}} = \mathbf{y} - \mathbf{H}\,\hat{\mathbf{x}} = \mathbf{H}\,\mathbf{x} + \mathbf{v} - \mathbf{H}\,\hat{\mathbf{x}} = \mathbf{H}\left(\mathbf{x} - \hat{\mathbf{x}}\right),$$

and if $\mathrm{null}(\mathbf{H}) = \{\mathbf{0}\}$, then $\boldsymbol{\epsilon}_{\mathbf{y}} = \mathbf{0}$ implies $\mathbf{x} = \hat{\mathbf{x}}$.

Method of Least Squares - Derivation (contd.)

$$J = \boldsymbol{\epsilon}_{\mathbf{y}}^T \, \boldsymbol{\epsilon}_{\mathbf{y}} = (\mathbf{y} - \mathbf{H}\,\hat{\mathbf{x}})^T \, (\mathbf{y} - \mathbf{H}\,\hat{\mathbf{x}}) = \mathbf{y}^T \, \mathbf{y} - \mathbf{y}^T \, \mathbf{H}\,\hat{\mathbf{x}} - \hat{\mathbf{x}}^T \, \mathbf{H}^T \, \mathbf{y} + \hat{\mathbf{x}}^T \, \mathbf{H}^T \, \mathbf{H}\,\hat{\mathbf{x}}$$

To find extremum of J, we need to set

$$\frac{\partial J}{\partial \hat{\mathbf{x}}} = \mathbf{0} \implies \frac{\partial J}{\partial \hat{\mathbf{x}}} = -\mathbf{y}^T \, \mathbf{H} - \mathbf{y}^T \, \mathbf{H} + 2\,\mathbf{x}^T \, \mathbf{H}^T \, \mathbf{H} = \mathbf{0}.$$

Solving the above equation for $\hat{\mathbf{x}}$ gives

$$\hat{\mathbf{x}} = \left(\mathbf{H}^T \, \mathbf{H}\right)^{-1} \mathbf{H}^T \, \mathbf{y} = \mathbf{H}^L \, \mathbf{y}$$

The matrix

$$\mathbf{H}^L = \left(\mathbf{H}^T \, \mathbf{H}\right)^{-1} \mathbf{H}^T$$

is known as *left pseudo inverse* of $\mathbf{H}$ and exists if $k \geq n$ and $\mathbf{H}$ is full rank.

Example - Direct Measurements

Let $x \in \mathbb{R}$ equal to the measured noisy output y (our measurement is directly measuring the unknown parameter with no scaling)

$$y_1 = x + v_1$$
$$\vdots$$
$$y_k = x + v_k$$

or equivalently

$$\begin{bmatrix} y_1 \\ \vdots \\ y_k \end{bmatrix} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} x + \begin{bmatrix} v_1 \\ \vdots \\ v_k \end{bmatrix}$$

Example - Direct Measurements (contd.)

$$
\begin{aligned}
\hat{x} &= \left(\mathbf{H}^T\,\mathbf{H}\right)^{-1}\,\mathbf{H}^T\,\mathbf{y} \\
&= \left(\begin{bmatrix} 1 & \cdots & 1 \end{bmatrix}\begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}\right)^{-1}\begin{bmatrix} 1 & \cdots & 1 \end{bmatrix}\begin{bmatrix} y_1 \\ \vdots \\ y_k \end{bmatrix} \\
&= \frac{1}{k}\left(y_1 + \cdots + y_k\right)
\end{aligned}
$$

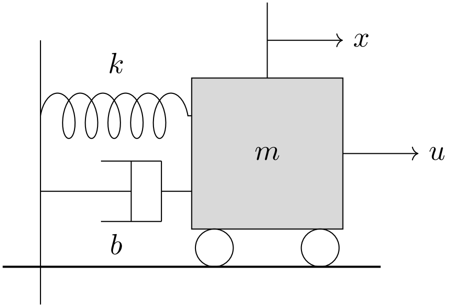
Example - Parameter Identification

For a mass-spring-damper system, we have measurements of position, velocity, acceleration and the input force at different instances $i \in \{1, \dots, N\}$.
At each instance i we have

$$m\,\ddot{x}_i + b\,\dot{x}_i + k\,x_i + v_i = u_i.$$

Collecting all N measurements we have

$$\begin{bmatrix} u_1 \\ \vdots \\ u_N \end{bmatrix} = \begin{bmatrix} \ddot{x}_1 & \dot{x}_1 & x_1 \\ \vdots & \vdots & \vdots \\ \ddot{x}_N & \dot{x}_N & x_N \end{bmatrix} \begin{bmatrix} \hat{m} \\ \hat{b} \\ \hat{k} \end{bmatrix} + \begin{bmatrix} v_1 \\ \vdots \\ v_N \end{bmatrix}$$



Weighted Least Square

What if we have more confidence in some measurements than others?

- We can get some information from less reliable measurements!
- We should never throw away measurements, no matter how unreliable they may be.

In mathematical terms, let $\mathbf{x} = (x_1, \ldots, x_n)$ be a constant but unknown vector and $\mathbf{y} = (y_1, \ldots, y_k)$, a k-element noisy measurement vector, is a linear combination of x_i. The variance of the measurement noise may be different for each element of $\mathbf{y}$.

$$\begin{bmatrix} y_1 \\ \vdots \\ y_k \end{bmatrix} = \begin{bmatrix} H_{11} & \cdots & H_{1n} \\ \vdots & \ddots & \vdots \\ H_{k1} & \cdots & H_{kn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} v_1 \\ \vdots \\ v_k \end{bmatrix}$$

$$E\left(v_i^2\right) = \sigma_i^2 \quad i \in \{1, \ldots, k\}$$

Weighted Least Square

Assume that the noise for each measurement is zero-mean and independent. The measurement covariance matrix is

$$\mathbf{R} = E\left(\mathbf{v}\,\mathbf{v}^T\right) = \begin{bmatrix} \sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_k^2 \end{bmatrix}$$

Accordingly, we modify the cost functional to be the *weighted* sum of squares as

$$J = \epsilon_{y_1}^2/\sigma_1^2 + \cdots + \epsilon_{y_k}^2/\sigma_k^2$$

For example, if y_1 is a relatively noisy measurement, then we do not care as much about minimizing the difference between y_1 and the first element of $\mathbf{H}\,\hat{\mathbf{x}}$ because we do not have much confidence in y_1.

Weighted Least Square

The cost functional can be written as

$$\begin{aligned} J &= \boldsymbol{\epsilon}_{\mathbf{y}}^{T}\,\mathbf{R}^{-1}\,\boldsymbol{\epsilon}_{\mathbf{y}} \\ &= \left(\mathbf{y} - \mathbf{H}\hat{\mathbf{x}}\right)^{T}\,\mathbf{R}^{-1}\,\left(\mathbf{y} - \mathbf{H}\hat{\mathbf{x}}\right) \\ &= \mathbf{y}^{T}\,\mathbf{R}^{-1}\,\mathbf{y} - \hat{\mathbf{x}}^{T}\,\mathbf{H}^{T}\,\mathbf{R}^{-1}\,\mathbf{y} - \mathbf{y}^{T}\,\mathbf{R}^{-1}\,\mathbf{H}\,\hat{\mathbf{x}} + \hat{\mathbf{x}}^{T}\,\mathbf{H}^{T}\,\mathbf{R}^{-1}\,\mathbf{H}\,\hat{\mathbf{x}} \end{aligned}$$

with partial derivative of J with respect to $\hat{\mathbf{x}}$ equal to

$$\frac{\partial J}{\partial \hat{\mathbf{x}}} = -2\mathbf{y}^{T}\,\mathbf{R}^{-1}\,\mathbf{H} + 2\hat{\mathbf{x}}^{T}\,\mathbf{H}^{T}\,\mathbf{R}^{-1}\,\mathbf{H}.$$

Setting $\frac{\partial J}{\partial \hat{\mathbf{x}}}$ equal to zero gives

$$\hat{\mathbf{x}} = \left(\mathbf{H}^{T}\,\mathbf{R}^{-1}\,\mathbf{H}\right)^{-1}\,\mathbf{H}^{T}\,\mathbf{R}^{-1}\,\mathbf{y}$$

This method requires $\mathbf{R}$ be non-singular. Meaning each of the measurements y_i must be corrupted by at least some noise.

Example - Direct Measurement with a Covariance Matrix

$$y_i = x + v_i$$
$$E\left(v_i^2\right) = \sigma_i^2$$

$$\begin{bmatrix} y_1 \\ \vdots \\ y_k \end{bmatrix} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} x + \begin{bmatrix} v_1 \\ \vdots \\ v_k \end{bmatrix}$$
$$\mathbf{R} = \mathrm{diag}\left(\sigma_1^2, \ldots, \sigma_k^2\right)$$

Example - Direct Measurement with a Covariance Matrix (contd.)

$$
\begin{aligned}
\hat{x} &= \left(\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{R}^{-1}\mathbf{y} \\
&= \left(\begin{bmatrix}1 & \cdots & 1\end{bmatrix}\begin{bmatrix}\sigma_1^2 & & 0\\ & \ddots & \\ 0 & & \sigma_k^2\end{bmatrix}^{-1}\begin{bmatrix}1\\ \vdots\\ 1\end{bmatrix}\right)^{-1}\begin{bmatrix}1 & \cdots & 1\end{bmatrix}\begin{bmatrix}\sigma_1^2 & & 0\\ & \ddots & \\ 0 & & \sigma_k^2\end{bmatrix}^{-1}\begin{bmatrix}y_1\\ \vdots\\ y_k\end{bmatrix} \\
&= \left(\sum_{i=1}^{k}\frac{1}{\sigma_i^2}\right)^{-1}\left(\frac{y_1}{\sigma_1^2}+\cdots+\frac{y_k}{\sigma_k^2}\right)
\end{aligned}
$$

We see that the optimal estimate $\hat{\mathbf{x}}$ is a weighted sum of the measurements, where each measurement is weighted by the inverse of its uncertainty.

Recursive Least Squares Estimation

- If we obtain measurements sequentially, we need to update our estimate of $\hat{\mathbf{x}}$;
- with each new measurement, we need to augment the H matrix and completely recompute the estimate $\hat{\mathbf{x}}$;
- $\mathbf{H} \in \mathbb{R}^{k \times n}$. Hence, as k increases, the computational load increases and it could become prohibitive.

Question:

- suppose we have $\hat{\mathbf{x}}$ after $k-1$ measurements,
- and we obtain a new measurement y_k;
- How can we update $\hat{\mathbf{x}}$ without completely reworking $\left(\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H}\right)^{-1} \mathbf{H}^T \mathbf{R}^{-1} \mathbf{y}$?

Recursive Least Squares Estimation

A linear recursive estimator can be written in the form

$$\begin{aligned}\mathbf{y}_k &= \mathbf{H}_k\,\mathbf{x} + \mathbf{v}_k \\ \hat{\mathbf{x}}_k &= \hat{\mathbf{x}}_{k-1} + \mathbf{K}_k\left(\mathbf{y}_k - \mathbf{H}_k\,\hat{\mathbf{x}}_{k-1}\right)\end{aligned}$$

- the quantity $(\mathbf{y}_k - \mathbf{H}_k\,\hat{\mathbf{x}}_{k-1})$ is called *the correction term*
- $\mathbf{K}_k$ is *estimator gain matrix* and needs to be determined.

If the correction term or $\mathbf{K}_k$ is zero, then the estimate does not change from $(k-1)$ to k.

Recursive Least Squares Estimation - Mean of the Estimation Error

Before finding the optimal gain matrix $\mathbf{K}_k$, let us find the mean of the estimation error:

$$
\begin{aligned}
E\left(\boldsymbol{\epsilon}_{\mathbf{x},\,k}\right) &= E\left(\mathbf{x} - \hat{\mathbf{x}}_k\right) \\
&= E\left[\mathbf{x} - \hat{\mathbf{x}}_{k-1} - \mathbf{K}_k\left(\mathbf{y}_k - \mathbf{H}_k\,\hat{\mathbf{x}}_{k-1}\right)\right] \\
&= E\left[\boldsymbol{\epsilon}_{\mathbf{x},k-1} - \mathbf{K}_k\left(\mathbf{H}_k\,\mathbf{x} + \mathbf{v}_k - \mathbf{H}_k\,\hat{\mathbf{x}}_{k-1}\right)\right] \\
&= E\left[\boldsymbol{\epsilon}_{\mathbf{x},k-1} - \mathbf{K}_k\,\mathbf{H}_k\left(\mathbf{x} - \hat{\mathbf{x}}_{k-1}\right) - \mathbf{K}_k\,\mathbf{v}_k\right] \\
&= \left(\mathbf{I} - \mathbf{K}_k\,\mathbf{H}_k\right)E\left(\boldsymbol{\epsilon}_{\mathbf{x},k-1}\right) - \mathbf{K}_k\,E\left(\mathbf{v}_k\right)
\end{aligned}
$$

So, if $E\left(\mathbf{v}_k\right) = \mathbf{0}$ and $E\left(\boldsymbol{\epsilon}_{\mathbf{x},k-1}\right) = \mathbf{0}$, then $E\left(\boldsymbol{\epsilon}_{\mathbf{x},k}\right) = \mathbf{0}$, regardless of the $\mathbf{K}_k$.

if $E(\mathbf{v}_k) = \mathbf{0}$ for all k and $\hat{\mathbf{x}}_0 = E(\mathbf{x})$ then $E(\hat{\mathbf{x}}_k) = E(\mathbf{x}_k)$ for all k. This implies the linear recursive least square estimator is an *unbiased estimator* $\implies$ *on average*, the estimate $\hat{\mathbf{x}}$ will be equal to the true value $\mathbf{x}$.

Recursive Least Squares Estimation - Optimal $\mathbf{K}_k$

Since the estimator is unbiased regardless $\mathbf{K}_k$, we must choose some other optimality criterion. We can aim to minimize the sum of the variances of the estimation errors at k:

$$\begin{aligned} J_k &= E\left[(x_1 - \hat{x}_1)^2\right] + \cdots + E\left[(x_n - \hat{x}_n)^2\right] \\ &= E\left(\epsilon^2_{x_1,\,k} + \cdots + \epsilon^2_{x_n,\,k}\right) \\ &= E\left(\boldsymbol{\epsilon}^T_{\mathbf{x},\,k}\,\boldsymbol{\epsilon}_{\mathbf{x},\,k}\right) \\ &= E\left[\mathrm{Tr}\left(\boldsymbol{\epsilon}_{\mathbf{x},\,k}\,\boldsymbol{\epsilon}^T_{\mathbf{x},\,k}\right)\right] \\ &= \mathrm{Tr}\left(\mathbf{P}_k\right) \end{aligned}$$

where $\mathbf{P}_k = E\left(\boldsymbol{\epsilon}_{\mathbf{x},\,k}\,\boldsymbol{\epsilon}^T_{\mathbf{x},\,k}\right)$ is the estimation-error covariance.

Recursive Least Squares Estimation - Optimal $\mathbf{K}_k$ (contd.)

To obtain a recursive formula for $\mathbf{P}_k$ we have

$$\begin{aligned}
\mathbf{P}_k &= E\left(\boldsymbol{\epsilon}_{\mathbf{x},\,k}\,\boldsymbol{\epsilon}_{\mathbf{x},\,k}^T\right)\\
&= E\left(\left[(\mathbf{I}-\mathbf{K}_k\,\mathbf{H}_k)\ \boldsymbol{\epsilon}_{\mathbf{x},k-1}-\mathbf{K}_k\,\mathbf{v}_k\right]\left[\cdots\right]^T\right)\\
&= (\mathbf{I}-\mathbf{K}_k\,\mathbf{H}_k)\ E\left(\boldsymbol{\epsilon}_{\mathbf{x},k-1}\,\boldsymbol{\epsilon}_{\mathbf{x},k-1}^T\right)\ (\mathbf{I}-\mathbf{K}_k\,\mathbf{H}_k)^T+\mathbf{K}_k\,E\left(\mathbf{v}_k\,\mathbf{v}_k^T\right)\ \mathbf{K}_k^T\\
&\quad -\mathbf{K}_k\,E\left(\mathbf{v}_k\,\boldsymbol{\epsilon}_{\mathbf{x},k-1}^T\right)\ (\mathbf{I}-\mathbf{K}_k\,\mathbf{H}_k)^T-(\mathbf{I}-\mathbf{K}_k\,\mathbf{H}_k)\ E\left(\boldsymbol{\epsilon}_{\mathbf{x},k-1}\,\mathbf{v}_k^T\right)\ \mathbf{K}_k^T
\end{aligned}$$

The estimation error at time $(k-1)$ is independent of the measurement noise at time k ($\boldsymbol{\epsilon}_{\mathbf{x},k-1}$ is independent of $\mathbf{v}_k$). Therefore,

$$E\left(\mathbf{v}_k\,\boldsymbol{\epsilon}_{\mathbf{x},k-1}^T\right)=E\left(\mathbf{v}_k\right)\,E\left(\boldsymbol{\epsilon}_{\mathbf{x},k-1}\right)=\mathbf{0}$$

Recursive Least Squares Estimation - Optimal $\mathbf{K}_k$ (contd.)

Accordingly, we have

$$\mathbf{P}_k = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k)\ \mathbf{P}_{k-1}\ (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k)^T + \mathbf{K}_k \mathbf{R}_k \mathbf{K}_k^T$$

where $\mathbf{R}_k$ is the covariance of $\mathbf{v}_k$.

- As the measurement noise increases ($\mathbf{R}_k$ increases) the uncertainty in our estimate also increases ($\mathbf{P}_k$ increases), which is consistent with intuition.
- $\mathbf{P}_k$ should be positive definite (it is a covariance matrix). The form presented above guarantees that P_k will be positive definite if $\mathbf{P}_{k-1}$ and $\mathbf{R}_k$ are positive definite.

- Recall that, independent of $\mathbf{K}_k$, if $E\left(\mathbf{v}_k\right) = \mathbf{0}$ and $E\left(\boldsymbol{\epsilon}_{\mathbf{x},k-1}\right) = \mathbf{0}$, then $E\left(\boldsymbol{\epsilon}_{\mathbf{x},k}\right) = \mathbf{0}$ (the mean of the estimation error is zero).
- So if we choose $\mathbf{K}_k$ to make the $J = \mathrm{Tr}\left(\mathbf{P}_k\right)$ small then the estimation error will not only be zero-mean, but also be consistently close to zero.

Recursive Least Squares Estimation - Optimal $\mathbf{K}_k$ (contd.)

Recall from linear algebra that for symmetric $\mathbf{B}$ we have

$$\frac{\partial}{\partial \mathbf{A}} \operatorname{Tr}\left(\mathbf{A}\,\mathbf{B}\,\mathbf{A}^T\right) = 2\,\mathbf{A}\,\mathbf{B}.$$

Since

$$\mathbf{P}_k = \left(\mathbf{I} - \mathbf{K}_k\,\mathbf{H}_k\right)\,\mathbf{P}_{k-1}\,\left(\mathbf{I} - \mathbf{K}_k\,\mathbf{H}_k\right)^T + \mathbf{K}_k\,\mathbf{R}_k\,\mathbf{K}_k^T,$$

then

$$\frac{\partial J_k}{\partial \mathbf{K}_k} = 2\,\left(\mathbf{I} - \mathbf{K}_k\,\mathbf{H}_k\right)\,\mathbf{P}_{k-1}\,\left(-\mathbf{H}_k^T\right) + 2\,\mathbf{K}_k\,\mathbf{R}_k.$$

Setting the above derivative to zero and solving for $\mathbf{K}_k$ gives

$$\mathbf{K}_k = \mathbf{P}_{k-1}\,\mathbf{H}_k^T\left(\mathbf{H}_k\,\mathbf{P}_{k-1}\,\mathbf{H}_k^T + \mathbf{R}_k\right)^{-1}$$

Recursive Least Squares Estimation - Summary

1. Initialize the estimator
   a. Set the initial values of $\hat{\mathbf{x}}_0$ and $\mathbf{P}_0$

$$\hat{\mathbf{x}}_0 = E(\mathbf{x})$$
$$\mathbf{P}_0 = E\left[(\mathbf{x} - \hat{\mathbf{x}}_0)(\mathbf{x} - \hat{\mathbf{x}}_0)^T\right]$$

   b. If initial $\mathbf{x}$ is completely unknown, then $\mathbf{P}_0 = \infty\,\mathbf{I}$. If $\mathbf{x}$ is known certainly then $\mathbf{P}_0 = \mathbf{0}$.
2. For $k = 1, 2 \ldots$ perform:
   a. Obtain the measurement $\mathbf{y}_k$. Assuming that

$$\mathbf{y}_k = \mathbf{H}_k\,\mathbf{x} + \mathbf{v}_k$$

   where $\mathbf{v}_k$ is zero-mean random vector with covariance $\mathbf{R}_k$ and $E\left(\mathbf{v}_i\,\mathbf{v}_k^T\right) = \mathbf{0}$ for $i \neq k$.
   b. Update $\hat{\mathbf{x}}$, $\mathbf{P}_k$, and $\mathbf{K}_k$ based on

$$\mathbf{K}_k = \mathbf{P}_{k-1}\,\mathbf{H}_k^T\left(\mathbf{H}_k\,\mathbf{P}_{k-1}\,\mathbf{H}_k^T + \mathbf{R}_k\right)^{-1},$$
$$\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_{k-1} + \mathbf{K}_k\left(\mathbf{y}_k - \mathbf{H}_k\,\hat{\mathbf{x}}_{k-1}\right),$$
$$\mathbf{P}_k = (\mathbf{I} - \mathbf{K}_k\,\mathbf{H}_k)\,\mathbf{P}_{k-1}\,(\mathbf{I} - \mathbf{K}_k\,\mathbf{H}_k)^T + \mathbf{K}_k\,\mathbf{R}_k\,\mathbf{K}_k^T.$$

General Recursive Least Squares Estimation - With Alternative Forms

Measurements

$$\mathbf{y}_k = \mathbf{H}_k\,\mathbf{x} + \mathbf{v}_k \qquad\qquad \mathbf{x} = \text{constant}$$

$$E(\mathbf{v}_k) = \mathbf{0} \qquad\qquad E(\mathbf{v}_k\,\mathbf{v}_i^T) = \mathbf{R}_k\,\delta_{k-i}$$

The initial estimate

$$\hat{\mathbf{x}}_0 = E(\mathbf{x}) \qquad\qquad \mathbf{P}_0 = E\left[(\mathbf{x}-\hat{\mathbf{x}}_0)(\mathbf{x}-\hat{\mathbf{x}}_0)^T\right]$$

For $k = 1,\, 2,\, \ldots$

$$\begin{aligned}
\mathbf{K}_k &= \mathbf{P}_{k-1}\,\mathbf{H}_k^T\left(\mathbf{H}_k\,\mathbf{P}_{k-1}\,\mathbf{H}_k^T + \mathbf{R}_k\right)^{-1} \\
&= \mathbf{P}_k\,\mathbf{H}_k^T\,\mathbf{R}_k^{-1}, \\
\hat{\mathbf{x}}_k &= \hat{\mathbf{x}}_{k-1} + \mathbf{K}_k\left(\mathbf{y}_k - \mathbf{H}_k\,\hat{\mathbf{x}}_{k-1}\right), \\
\mathbf{P}_k &= \left(\mathbf{I} - \mathbf{K}_k\,\mathbf{H}_k\right)\,\mathbf{P}_{k-1}\,\left(\mathbf{I} - \mathbf{K}_k\,\mathbf{H}_k\right)^T + \mathbf{K}_k\,\mathbf{R}_k\,\mathbf{K}_k^T \\
&= \left(\mathbf{P}_{k-1}^{-1}\mathbf{H}_k^T\,\mathbf{R}_k^{-1}\,\mathbf{H}_k\right)^{-1} \\
&= \left(\mathbf{I} - \mathbf{K}_k\,\mathbf{H}_k\right)\mathbf{P}_{k-1}
\end{aligned}$$

Propagation of State and it's Covariance

Discrete-Time Systems

Suppose we have the following linear discrete-time system

$$\mathbf{x}_k = \mathbf{F}_{k-1}\,\mathbf{x}_{k-1} + \mathbf{G}_{k-1}\,\mathbf{u}_{k-1} + \mathbf{w}_{k-1}$$

where

- $\mathbf{u}_k$ is a known input
- $\mathbf{w}_k$ is Gaussian zero-mean white noise with covariance $\mathbf{Q}_k$.

The Question is:

How does the mean of the state $\mathbf{x}_k$ change with time?
How does the covariance of $\mathbf{x}_k$ change with time?

Discrete-Time Systems (contd.)

If we take the expected value of both sides we obtain

$$\begin{aligned}\bar{\mathbf{x}}_k &= E(\mathbf{x}_k)\\ &= \mathbf{F}_{k-1}\,\bar{\mathbf{x}}_{k-1} + \mathbf{G}_{k-1}\,\mathbf{u}_{k-1}\end{aligned}$$

To obtain the change of covariance of $\mathbf{x}_k$ in time, we have

$$\begin{aligned}(\mathbf{x}_k - \bar{\mathbf{x}}_k)(\mathbf{x}_k - \bar{\mathbf{x}}_k)^T &= (\mathbf{F}_{k-1}\,\mathbf{x}_{k-1} + \mathbf{G}_{k-1}\,\mathbf{u}_{k-1} + \mathbf{w}_{k-1} - \bar{\mathbf{x}}_k)(\cdots)^T\\ &= \Big(\mathbf{F}_{k-1}\,(\mathbf{x}_{k-1} - \bar{\mathbf{x}}_{k-1}) + \mathbf{w}_{k-1}\Big)(\cdots)^T\\ &= \mathbf{F}_{k-1}\,(\mathbf{x}_{k-1} - \bar{\mathbf{x}}_{k-1})\,(\mathbf{x}_{k-1} - \bar{\mathbf{x}}_k)^T\,\mathbf{F}_{k-1}^T + \mathbf{w}_{k-1}^T\,\mathbf{w}_{k-1}\\ &+ \mathbf{w}_{k-1}\,(\mathbf{x}_{k-1} - \bar{\mathbf{x}}_{k-1})^T\,\mathbf{F}_{k-1}^T + \mathbf{F}_{k-1}\,(\mathbf{x}_{k-1} - \bar{\mathbf{x}}_{k-1})\,\mathbf{w}_{k-1}^T\end{aligned}$$

Since $(\mathbf{x}_{k-1} - \bar{\mathbf{x}}_{k-1})$ is uncorrelated with $\mathbf{w}_{k-1}$, we obtain

$$\mathbf{P}_k = E\left[(\mathbf{x}_k - \bar{\mathbf{x}}_k)(\mathbf{x}_k - \bar{\mathbf{x}}_k)^T\right] = \mathbf{F}_{k-1}\,\mathbf{P}_{k-1}\,\mathbf{F}_{k-1}^T + \mathbf{Q}_{k-1}$$

Kalman Filter

Forms of Estimation

Goal is to estimate $\mathbf{x}_k$ based on our knowledge of the system dynamics and the noisy measurements $\mathbf{y}_k$

- *a priori* estimate: estimate of $\mathbf{x}_k$ **before** we process the measurement at time k

$$\hat{\mathbf{x}}_k^- = E\left(\mathbf{x}_k \,|\, \mathbf{y}_1,\, \mathbf{y}_2,\, \dots,\mathbf{y}_{k-1}\right)$$

- *a posteriori* estimate: estimate of $\mathbf{x}_k$ **after** we process the measurement at time k

$$\hat{\mathbf{x}}_k^+ = E\left(\mathbf{x}_k \,|\, \mathbf{y}_1,\, \mathbf{y}_2,\, \dots,\mathbf{y}_{k-1},\, \mathbf{y}_k\right)$$

- *a predicted* estimate: for some $0 < M < k \in \mathbb{N}$

$$\hat{\mathbf{x}}_{k|k-M} = E\left(\mathbf{x}_k \,|\, \mathbf{y}_1,\, \mathbf{y}_2,\, \dots,\, \mathbf{y}_{k-M}\right)$$

- *a smoothed* estimate: for some $0 < N \in \mathbb{N}$

$$\hat{\mathbf{x}}_{k|k+N} = E\left(\mathbf{x}_k \,|\, \mathbf{y}_1,\, \mathbf{y}_2,\, \dots,\, \mathbf{y}_k,\, \dots,\, \mathbf{y}_{k+N}\right)$$

Derivation of the Discrete-Time Kalman Filter

Suppose we have the following linear discrete-time system

$$
\begin{aligned}
\mathbf{x}_k &= \mathbf{F}_{k-1}\,\mathbf{x}_{k-1} + \mathbf{G}_{k-1}\,\mathbf{u}_{k-1} + \mathbf{w}_{k-1} \\
\mathbf{y}_k &= \mathbf{H}_k\,\mathbf{x}_k + \mathbf{v}_k
\end{aligned}
$$

where $\mathbf{w}_k$ and $\mathbf{v}_k$ are uncorrelated zero-mean white noise

$$
\begin{aligned}
\mathbf{w}_k &\sim (0, \mathbf{Q}_k) \\
\mathbf{v}_k &\sim (0, \mathbf{R}_k) \\
E\left[\mathbf{w}_k\,\mathbf{w}_j^T\right] &= \mathbf{Q}_k\,\delta_{k-j} \\
E\left[\mathbf{v}_k\,\mathbf{v}_j^T\right] &= \mathbf{R}_k\,\delta_{k-j} \\
E\left[\mathbf{v}_k\,\mathbf{w}_j^T\right] &= \mathbf{0}
\end{aligned}
$$

δ_{k-j} is the Kronecker delta function: $\delta_{k-j} = 1$ if $k = j$ and $= 0$ otherwise.

Derivation of the Discrete-Time Kalman Filter (contd.)

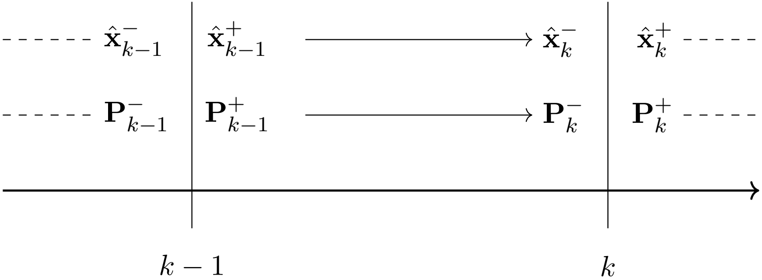
Let $\hat{\mathbf{x}}_0^+$ denote our initial estimate of $\mathbf{x}_0$. Since the first measurement is taken at $k = 1$, we do not have any measurements available to estimate $\mathbf{x}_0$. Accordingly

$$\hat{\mathbf{x}}_0^+ = E\left(\mathbf{x}_0\right)$$

Let $\mathbf{P}_k$ denote the covariance of the estimation error:

$$\mathbf{P}_k^- = E\left[\left(\mathbf{x}_k - \hat{\mathbf{x}}_k^-\right)\left(\mathbf{x}_k - \hat{\mathbf{x}}_k^-\right)^T\right]$$
$$\mathbf{P}_k^+ = E\left[\left(\mathbf{x}_k - \hat{\mathbf{x}}_k^+\right)\left(\mathbf{x}_k - \hat{\mathbf{x}}_k^-\right)^T\right]$$

Derivation of the Discrete-Time Kalman Filter (contd.)



Derivation of the Discrete-Time Kalman Filter (contd.)

We begin the estimation process with $\hat{\mathbf{x}}_0^+$ (our best estimate of the initial state $\mathbf{x}_0$.

Question: Given $\hat{\mathbf{x}}_0^+$ how can we find $\hat{\mathbf{x}}_1^-$?

Recall the propagation of the mean of $\mathbf{x}$:

$$\bar{\mathbf{x}}_k = E(\mathbf{x}_k) = \mathbf{F}_{k-1}\,\bar{\mathbf{x}}_{k-1} + \mathbf{G}_{k-1}\,\mathbf{u}_{k-1}$$

Accordingly we have

$$\hat{\mathbf{x}}_1^- = \mathbf{F}_0\,\hat{x}_0^+ + \mathbf{G}_0\,\mathbf{u}_0$$

The reasoning can be extended to obtain the following more general equation:

$$\hat{\mathbf{x}}_k^- = \mathbf{F}_{k-1}\,\hat{x}_{k-1}^+ + \mathbf{G}_{k-1}\,\mathbf{u}_{k-1}$$

Derivation of the Discrete-Time Kalman Filter (contd.)

Next we need to find an update equation for $\mathbf{P}$ (the covariance of the state estimation error). We begin with $\mathbf{P}_0^+$ (the covariance of our initial estimate of $\mathbf{x}_0$.

If we know the initial state perfectly, then $\mathbf{P}_0^+ = \mathbf{0}$.

If we have absolutely no idea of the value of $\mathbf{x}_0$, then $\mathbf{P}_0^+ = \infty \, \mathbf{I}$

Recall how the covariance of the state of a linear discrete-time system propagates with time:

$$\mathbf{P}_k = \mathbf{F}_{k-1} \, \mathbf{P}_{k-1} \, \mathbf{F}_{k-1}^T + \mathbf{Q}_{k-1}$$

Accordingly we have

$$\mathbf{P}_1^- = \mathbf{F}_0 \, \mathbf{P}_0^+ \, \mathbf{F}_0^T + \mathbf{Q}_0$$

The reasoning can be extended to obtain the following more general equation:

$$\mathbf{P}_k^- = \mathbf{F}_{k-1} \, \mathbf{P}_{k-1}^+ \, \mathbf{F}_{k-1}^T + \mathbf{Q}_{k-1}$$

Derivation of the Discrete-Time Kalman Filter (contd.)

Question: Given $\hat{\mathbf{x}}_k^-$ how can we obtain $\hat{\mathbf{x}}_k^+$

The only difference between $\hat{\mathbf{x}}_k^-$ and $\hat{\mathbf{x}}_k^+$ is that $\hat{\mathbf{x}}_k^+$ takes the measurement $\mathbf{y}_k$ into account. Recall from the recursive least squares development that the availability of the measurement $\mathbf{y}_k$ changes the estimate of a constant $\mathbf{x}$ as:

$$\begin{aligned}
\mathbf{K}_k &= \mathbf{P}_{k-1}\,\mathbf{H}_k^T\left(\mathbf{H}_k\,\mathbf{P}_{k-1}\,\mathbf{H}_k^T + \mathbf{R}_k\right)^{-1} \\
&= \mathbf{P}_k\,\mathbf{H}_k^T\,\mathbf{R}_k^{-1}, \\
\hat{\mathbf{x}}_k &= \hat{\mathbf{x}}_{k-1} + \mathbf{K}_k\left(\mathbf{y}_k - \mathbf{H}_k\,\hat{\mathbf{x}}_{k-1}\right), \\
\mathbf{P}_k &= \left(\mathbf{I} - \mathbf{K}_k\,\mathbf{H}_k\right)\mathbf{P}_{k-1}\left(\mathbf{I} - \mathbf{K}_k\,\mathbf{H}_k\right)^T + \mathbf{K}_k\,\mathbf{R}_k\,\mathbf{K}_k^T \\
&= \left(\mathbf{P}_{k-1}^{-1}\mathbf{H}_k^T\,\mathbf{R}_k^{-1}\,\mathbf{H}_k\right)^{-1} \\
&= \left(\mathbf{I} - \mathbf{K}_k\,\mathbf{H}_k\right)\mathbf{P}_{k-1}
\end{aligned}$$

Derivation of the Discrete-Time Kalman Filter (contd.)

$$
\begin{aligned}
\mathbf{K}_k &= \mathbf{P}_k^- \, \mathbf{H}_k^T \left(\mathbf{H}_k \, \mathbf{P}_k^- \, \mathbf{H}_k^T + \mathbf{R}_k \right)^{-1} \\
&= \mathbf{P}_k^+ \, \mathbf{H}_k^T \, \mathbf{R}_k^{-1}, \\
\hat{\mathbf{x}}_k^+ &= \hat{\mathbf{x}}_k^- + \mathbf{K}_k \left(\mathbf{y}_k - \mathbf{H}_k \, \hat{\mathbf{x}}_k^- \right), \\
\mathbf{P}_k^+ &= \left(\mathbf{I} - \mathbf{K}_k \, \mathbf{H}_k \right) \mathbf{P}_k^- \left(\mathbf{I} - \mathbf{K}_k \, \mathbf{H}_k \right)^T + \mathbf{K}_k \, \mathbf{R}_k \, \mathbf{K}_k^T \\
&= \left(\left(\mathbf{P}_k^- \right)^{-1} \mathbf{H}_k^T \, \mathbf{R}_k^{-1} \, \mathbf{H}_k \right)^{-1} \\
&= \left(\mathbf{I} - \mathbf{K}_k \, \mathbf{H}_k \right) \mathbf{P}_k^-
\end{aligned}
$$