Linear Systems Notes (2023/2024)

Griffin Reimerink

Contents

1 Ordinary differential equations

1.1 First-order scalar ODEs

Definition *Scalar first-order ODE*

A **scalar first-order ODE** is of the form $\dot{x}(t) = f(t, x(t))$
with independent variable t, dependent variable x and $f : D \subset \mathbb{R}^2 \to \mathbb{R}$.
A function $x : J \subset \mathbb{R} \to \mathbb{R}$ is a **solution** to the scalar first-order ODE if 3 conditions are satisfied:

1. $(t, x(t)) \in D$ for all $t \in J$
2. x is differentiable
3. $\dot{x}(t) = f(t, x(t))$ for all $t \in J$

This solution is not unique. To fix a specific solution, we can add the **initial value** $x(t_0) = x_0$

Lemma *Solution of a scalar first-order ODE*

Equations of the form $\dot{x} = f(t)$
The solution of $\dot{x} = f(t)$ is the antiderivative of f.
With the initial condition $x(t_0) = x_0$ we have $x(t) = x_0 + \int_{t_0}^{t} f(t)\,\mathrm{dt}$

Equations of the form $\dot{x} = g(x)$ (Autonomous ODE)
For $g(x) \neq 0$, any solution to $\dot{x}(t) = g(x)$ is of the form $H(x) = t + \tau$ with $\frac{dH}{dx}(x) = \frac{1}{g(x)}, \tau \in \mathbb{R}$

Equations of the form $\dot{x} = f(t)g(x)$ (Separable ODE)
Equations of the form $\dot{x} = f(t)g(x)$ can be rewritten as $\frac{1}{g(x)}\,\mathrm{dx} = f(t)\,\mathrm{dt}$
and then solved by integrating both sides: $\int \frac{1}{g(x)}\,\mathrm{dx} = \int f(t)\,\mathrm{dt}$

1.2 Linear scalar ODEs

Definition *Linear scalar ODE*

A **linear scalar ODE** is of the form $\dot{x}(t) = a(t)x(t) + b(t)$ with $a, b : J \to \mathbb{R}$ continuous.
It is **homogeneous** if $b(t) = 0$ for all $t \in J$ and **inhomogeneous** otherwise.
We can define a linear operator $L(x) = \dot{x} - a(t)x$ to rewrite the ODE as $L(x) = b$.

Lemma *Solution of a linear scalar ODE*

Homogeneous case
The solutions to $\dot{x}(t) = a(t)x(t)$ are given by $x(t) = Ce^{F(t)}$, $F(t) = \int a(t)\,\mathrm{dt}$
With the initial condition $x(t_0) = x_0$, we have $x(t) = x_0 e^{F(t)}$, $F(t) = \int_{t_0}^{t} a(\tau)\,\mathrm{d}\tau$

Inhomogeneous case
The solutions to $\dot{x}(t) = a(t)x(t) + b(t)$ are given by $x(t) = Ce^{F(t)} + e^{F(t)} \int e^{-F(t)} b(t)\,\mathrm{dt}$, $F(t) = \int a(t)\,\mathrm{dt}$
With the initial condition $x(t_0) = x_0$, we have $x(t) = x_0 e^{F(t)} + e^{F(t)} \int_{t_0}^{t} e^{-F(t)} b(t)\,\mathrm{dt}$, $F(t) = \int_{t_0}^{t} a(\tau)\,\mathrm{d}\tau$

The solutions to the initial value problems are unique.

Theorem

The difference of two solutions of the inhomogeneous case is a solution to the homogeneous case.

1.3 Systems of differential equations

Definition *System of differential equations*

A **system of differential equations** is of the form $\dot{x}(t) = f(t, x(t))$

where $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ and $f(t, x) = \begin{bmatrix} f_1(t, x_1, x_2, \ldots, x_1) \\ f_2(t, x_1, x_2, \ldots, x_2) \\ \vdots \\ f_n(t, x_1, x_2, \ldots, x_n) \end{bmatrix}$, $f : D \to \mathbb{R}, D \subset \mathbb{R} \times \mathbb{R}^n$

Definition *Higher-order differential equations*

Denote $y^{(k)} = \frac{\mathrm{d}^k y}{\mathrm{d}t^k}(t)$. Consider $y^{(k)}(t) = f\left(t, y(t), \dot{y}(t), \ldots, y^{(k-1)}(t)\right)$.

Introduce the vector $\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} y \\ \dot{y} \\ \vdots \\ y^{(k-1)} \end{bmatrix}$ Then $\dot{\boldsymbol{x}} = \begin{bmatrix} x_2 \\ \vdots \\ x_n \\ f(t, \boldsymbol{x}) \end{bmatrix}$

2 Linear systems

Definition *Linear system (state-space form)*

A **linear system** is a system of the form $\Sigma : \begin{cases} \dot{\boldsymbol{x}}(t) = A\boldsymbol{x}(t) + B\boldsymbol{u}(t) \\ \boldsymbol{y}(t) = C\boldsymbol{x}(t) + D\boldsymbol{u}(t) \end{cases}$
where $\boldsymbol{x}$ is the **state**, $\boldsymbol{u}$ is the **input** and $\boldsymbol{y}$ is the **output**.

2.1 Nonlinear systems

Definition *Nonlinear system (state-space form)*

A **nonlinear system** is a system of the form $\Sigma : \begin{cases} \dot{\boldsymbol{x}}(t) = f(\boldsymbol{x}(t), \boldsymbol{u}(t)) \\ \boldsymbol{y}(t) = h(\boldsymbol{x}(t), \boldsymbol{u}(t)) \end{cases}$
where $\boldsymbol{x}$ is the **state**, $\boldsymbol{u}$ is the **input** and $\boldsymbol{y}$ is the **output**.

Definition *Equilibrium*

Let $u(t) = \overline{u}$ be constant. Then, $\overline{x} \in \mathbb{R}^n$ is an **equilibrium** for $\overline{u}$ if $f(\overline{x}, \overline{u}) = 0$

Definition *Linearization*

The **linearization** of the nonlinear system Σ is of the form $\begin{cases} \dot{\widetilde{x}}(t) = A\widetilde{x}(t) + B\widetilde{u}(t) \\ \widetilde{y}(t) = C\widetilde{x}(t) + C\widetilde{u}(t) \end{cases}$
with $\widetilde{x} = x(t) - \overline{x}$, $\widetilde{u} = u(t) - \overline{u}$, $\widetilde{y} = y(t) - \overline{y}$, $A = \frac{\partial f}{\partial x}(\overline{x}, \overline{u})$, $B = \frac{\partial f}{\partial u}(\overline{x}, \overline{u})$, $C = \frac{\partial h}{\partial x}(\overline{x}, \overline{u})$, $D = \frac{\partial h}{\partial u}(\overline{x}, \overline{u})$
where $\overline{y} = h(\overline{x}, \overline{u})$ and A, B, C, D are (not necessarily square) Jacobian matrices.

3 Solutions of linear systems

3.1 The matrix exponential

Definition *Matrix exponential*

For $A \in \mathbb{R}^{n \times n}, t \in \mathbb{R}$, we have $e^{At} = \sum_{k=0}^{\infty} \frac{A^k t^k}{k!}$

Definition *Matrix norm*

Let $A \in \mathbb{F}^{n \times n}$. Then $\|A\| = \sup\left\{ \frac{|Ax|}{|x|} \,\middle|\, x \neq 0 \right\} = \sup\{|Ax| \mid |x| = 1\}$

Lemma *Properties of the matrix exponential*

Let $A, B \in \mathbb{C}^{n\times n}$ and $t, s \in \mathbb{R}$.

1. $\frac{\mathrm{d}}{\mathrm{d}t}e^{At} = Ae^{At} = e^{At}A$
2. e^{At} is invertible, and $(e^{At})^{-1} = e^{-At}$
3. If $AB = BA$, then $e^{At}B = Be^{At}$
4. If $AB = BA$, then $e^{At}e^{Bt} = e^{(A+B)t}$
5. $e^{At}e^{As} = e^{A(t+s)}$
6. $(e^{At})^T = e^{A^Tt}$

Lemma

Let $T \in \mathbb{C}^{n\times n}$ be nonsingular and $A \in \mathbb{C}^{n\times n}$. Then $e^{TAT^{-1}t} = Te^{At}T^{-1}$

Lemma

Let A be diagonalizable. Then $A = T\Lambda T^{-1}$, where T has the eigenvectors of A as columns, and Λ is a diagonal matrix containing the eigenvalues of A.

Theorem

Let $A \in \mathbb{R}^{n\times n}$. Then,

$$e^{A(t-t_0)}x_0 = \sum_{i=1}^{n} c_i v_i e^{\lambda_i(t-t_0)}$$

for some constants $c_1, c_2, \ldots, c_n \in \mathbb{C}$ and (λ_i, v_i) eigenpairs of A, i.e. $Av_i = \lambda_i v_i$

3.2 Jordan canonical form

Definition *Spectrum*

The **spectrum** of $A \in \mathbb{C}^{m\times n}$, denoted $\sigma(A)$, is the set of eigenvalues of A.
We denote the **characteristic polynomial** $\det(\lambda I - A)$ by $\Delta_A(\lambda)$.

Definition *Jordan block*

A **Jordan block** $J_k(\lambda) \in \mathbb{C}^{k\times k}$ is equal to $\lambda I + N$, where N has ones on the upper diagonal. $N^k = 0$.

Lemma

$$e^{J_k(\lambda)t} = e^{\lambda t}\begin{bmatrix} 1 & t & \frac{t^2}{2!} & \cdots & \frac{t^{k-1}}{(k-1)!} \\ 0 & 1 & t & \cdots & \frac{t^{k-2}}{(k-2)!} \\ 0 & 0 & 1 & \cdots & \frac{t^{k-3}}{(k-3)!} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

Theorem *Jordan canonical form*

For any $A \in \mathbb{R}^{n\times n}$, there exists a nonsingular $T \in \mathbb{C}^{n\times n}$ such that $A = TJT^{-1}$ with

$$J = \begin{bmatrix} J_{k_1}(\lambda_1) & & & \\ & J_{k_2}(\lambda_2) & & \\ & & \ddots & \\ & & & J_{k_r}(\lambda_r) \end{bmatrix}$$

with $\lambda_i \in \sigma(A)$. Conversely, if $\lambda \in \sigma(A)$, then $\lambda = \lambda_i$ for some $i \in \{1, 2, \ldots, r\}$

Definition *Multiplicities*

For $\lambda \in \sigma(A)$,

1. its **algebraic multiplicity** a_λ is its multiplicity as a root of Δ_A
2. its **geometric multiplicity** g_λ is the dimension of its **eigenspace** E_λ where $E_\lambda = \{v \in \mathbb{C}^n \mid (A - \lambda I)v = 0\}$

Properties of multiplicities

$$\sum_{\lambda \in \sigma(A)} a_\lambda = n \qquad 1 \leq g_\lambda \leq a_\lambda \quad \forall \lambda \in \sigma(A)$$

Theorem

A is diagonalizable if and only if $a_\lambda = g_\lambda$ for all $\lambda \in \sigma(A)$

Definition *Generalized eigenspace*

The **generalized eigenspace** of $\lambda \in \sigma(A)$ is

$$K_\lambda = \{v \in \mathbb{C}^n \mid (A - \lambda I)^p v = 0 \text{ for some integer } p > 0\}$$

Theorem *Properties of the generalized eigenspace*

Let $A \in \mathbb{C}^{n \times n}$ and $\lambda \in \sigma(A)$.

1. $\dim K_\lambda = a_\lambda$
2. $K_\lambda \cap K_{\lambda'} = \{0\}$ for any distinct $\lambda, \lambda' \in \sigma(A)$
3. K_λ is A-**invariant**, i.e. $v \in K_\lambda \implies Av \in K_\lambda$

Definition *Cycle*

Let $\lambda \in \sigma(A), v \in K_\lambda$. Let k be the smallest integer such that $(A - \lambda I)^k v = 0$. Then, $\{(A - \lambda I)^{k-1} v, (A - \lambda I)^{k-2} v, \ldots, v\}$ is called a **cycle** of generalized eigenvectors.

Theorem

Let $\lambda \in \sigma(A)$. Then K_λ has a basis consisting of cycles.

Definition *Dot diagram*

Let $\lambda \in \sigma(A)$ and

- $r_{\lambda,1} = n - \text{rank}(A - \lambda I) = g_\lambda$
- $r_{\lambda,1} = \text{rank}(A - \lambda I)^{j-1} - \text{rank}(A - \lambda I)^j$

To construct the dot diagram, we put $r_{\lambda,i}$ dots (left-aligned) in the ith row for all i.

Theorem *Properties of the dot diagram*

Let $\lambda \in \sigma(A)$ and consider the dot diagram.

1. # columns = # independent cycles in $K_\lambda = g_\lambda$
2. # dots in each column = length of the corresponding cycle
3. # dots = a_λ

Algorithm *Computation of the Jordan canonical form*

If only J needs to be computed, skip steps 2c and 4.

1. Compute λ_i and a_{λ_i}
2. For each λ_i:
   (a) compute $\text{rank}(A - \lambda I)^j$ for integers $j > 0$, stop when the difference between ranks is 0
   (b) construct the dot diagram by definition
   (c) compute each cycle by definition (such that the v's are linearly independent)
3. Use the dot diagram for each λ to form J
   (a) The number of Jordan blocks corresponding to λ_i is the number of columns of its dot diagram.
   (b) The size of each Jordan block is the number of dots in its corresponding column.
4. Use the corresponding cycles to form T (the columns of T are the vectors contained in the cycles)

3.3 Solutions of linear systems

Algorithm *Computation of e^{At}*

1. Compute the eigenvalues of A. If they are distinct, also compute the eigenvectors.
2. If A is diagonalizable:
   (a) Compute the diagonalization $A = T\Lambda T^{-1}$
   T has the eigenvectors as columns and Λ has the eigenvalues on the diagonal.
   (b) Compute e^{At} using $e^{At} = e^{T\Lambda T^{-1}t} = Te^{\Lambda t}T^{-1}$
   Λ is a diagonal matrix and therefore $e^{\Lambda t}$ can be computed entry-wise.
3. If A is not diagonalizable:
   (a) Compute the Jordan canonical form $A = TJT^{-1}$
   (b) For each Jordan block, compute $e^{J_{k_i}(\lambda_i)t}$
   (c) Compute e^{At} using $e^{At} = e^{TJT^{-1}t} = Te^{Jt}T^{-1} = T \begin{bmatrix} e^{J_{k_1}(\lambda_1)t} & & 0 \\ & \ddots & \\ 0 & & e^{J_{k_r}(\lambda_r)t} \end{bmatrix} T^{-1}$

Theorem *Unique solution of a linear system*

Homogeneous case
Consider $\dot{x}(t) = Ax(t), x(t_0) = x_0$ for $A \in \mathbb{R}^{n\times n}$.
The unique solution is:

$$x(t; t_0, x_0) = e^{A(t-t_0)}x_0$$

Inhomogeneous case
Consider $\dot{x}(t) = Ax(t) + Bu(t)$, $x(t_0) = x_0$ for $u : J \to \mathbb{R}^m, A \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{n\times m}$.
The unique solution is:

$$x(t; t_0, x_0, u) = e^{A(t-t_0)}x_0 + \int_{t_0}^{t} e^{A(t-\tau)}Bu(\tau)\,\mathrm{d}\tau$$

4 Stability, controllability and observability

4.1 Stability

Definition *Stability*

The system $\dot{x}(t) = Ax(t)$ is called
1. **stable** if for all $x_0 \in \mathbb{R}$, there exists $M > 0$ such that $|x(t : x_0)| \leq M$ for all $t \geq 0$
2. **asymptotically stable** if for all $x_0 \in \mathbb{R}$, $\lim_{n\to\infty} tx(t; x_0) = 0$

Observations

Let $(e^{At})_{ij}$ denote element i, j of e^{At}
1. $|(e^{At})_{ij}| \leq m \quad \forall i, j \implies \dot{x}(t) = Ax(t)$ is stable
2. $\lim_{n\to\infty} t(e^{At})_{ij} = 0 \quad \forall i, j \implies \dot{x}(t) = Ax(t)$ is asymptotically stable
3. By the Jordan canonical form, each $(e^{At})_{ij}$ is a sum of terms of the from $t^k e^{\lambda t}$ where λ is an eigenvalue of A and k is a nonnegative integer.

Lemma

Consider $t \mapsto t^k e^{\lambda t}$ with $k \geq 0$
1. If $\Re(\lambda) < 0$, then $\lim_{n\to\infty} tt^k e^{\lambda t} = 0$ and $|t^k e^{\lambda t}| \leq M$ for all $t \geq 0$
2. For any $\alpha \in \mathbb{R}$ such that $\Re(\lambda) < \alpha$, $|t^k e^{\lambda t}| \leq Me^{\alpha t}$ for all $t \geq 0$

Definition

$$\mathbb{C}_- = \{z \in \mathbb{C} \mid \Re(z) < 0\} \qquad \overline{\mathbb{C}}_- = \{z \in \mathbb{C} \mid \Re(z) \leq 0\}$$

Theorem

The system $\dot{x}(t) = Ax(t)$ is

1. stable if and only if $\sigma(A) \subset \overline{\mathbb{C}}_-$ and every λ with $\Re(\lambda) = 0$ is **semisimple**, i.e. $\alpha(\lambda) = \gamma(\lambda)$
2. asymptotically stable if and only if $\sigma(A) \subset \mathbb{C}_-$

Definition *Stability of polynomials*

A polynomial $p(s) = a_n s^n + a_{n-1}s^{n-1} + \ldots + a_1 s + a_0$ with $a_i \in \mathbb{R}, a_n \neq 0$ is called **stable**
if all roots have negative real parts.

Theorem

The system $\dot{x}(t) = Ax(t)$ is asymptotically stable $\iff$ the characteristic polynomial of A is stable

Theorem *Routh-Hurwitz condition*

The polynomial $p(s)$ is stable if and only if:

1. $a_{n-1} \neq 0$ and has the same sign as a_n.
2. the polynomial $q(s) = a_{n-1}p(s) - a_n(a_{n-1}s^n + a_{n-3}s^{n-2} + a_{n-5}s^{n-4} + \ldots)$ is stable

Lemma

Let $p(s)$ with $a_i \in \mathbb{R}$ be stable. Then all a_i are nonzero and have the same sign.

Theorem *Kharitonov's theorem*

Let $a_i^-, a_i^+ \in \mathbb{R}$, $i \in \{0, 1, \ldots, n\}$ satisfy $a_i^- \leq a_i^+$.
Define $\mathcal{P}(s) = \left\{a_n s^n + a_{n-1}s^{n-1} + \ldots + a_1 s + a_0 \mid a_i^- \leq a_1 \leq a_i^+ \text{ for all } i \in \{0, 1, \ldots, n\}\right\}$
$\mathcal{P}$ is stable (i.e. all polynomials in $\mathcal{P}$ are stable) if and only if the following 4 polynomials are all stable:

$$\begin{aligned}
p^{++}(s) &= a_0^+ + a_1^+ s + a_2^- s^2 + a_3^- s^3 + a_4^+ s^4 + a_5^+ s^5 + a_6^- s^6 + \ldots \\
p^{+-}(s) &= a_0^+ + a_1^- s + a_2^- s^2 + a_3^+ s^3 + a_4^+ s^4 + a_5^- s^5 + a_6^- s^6 + \ldots \\
p^{-+}(s) &= a_0^- + a_1^+ s + a_2^+ s^2 + a_3^- s^3 + a_4^- s^4 + a_5^+ -^5 + a_6^+ s^6 + \ldots \\
p^{--}(s) &= a_0^- + a_1^- s + a_2^+ s^2 + a_3^+ s^3 + a_4^- s^4 + a_5^- s^5 + a_6^+ s^6 + \ldots
\end{aligned}$$

4.2 Controllability

Consider the linear system $\Sigma : \dot{\boldsymbol{x}}(t) = A\boldsymbol{x}(t) + B\boldsymbol{u}(t)$

Definition *Reachability*

$x_f \in \mathbb{R}^n$ is **reachable** at time $T > 0$
if there exists an input function $u : [0, T] \to \mathbb{R}^m$ such that $x(T; 0, u) = x_f$

Definition *Reachable subspace*

$$W_T = \{x_f \in \mathbb{R}^n \mid x_f \text{ is reachable at } T\}$$

Definition *Reachablility of a system*

Σ is reachable at time $T > 0$ if $W_T = \mathbb{R}^n$

Theorem

Let $v \in \mathbb{R}$ and $T > 0$. The following statements are equivalent:

1. $v^T x = 0$ for all $x \in W_T$
2. $v^T e^{At} B = 0$ for all $0 \leq t \leq T$
3. $v^T A^k B = 0$ for all $k \in \{0, 1, 2, \ldots\}$
4. $v^T \begin{bmatrix} B & AB & A^2B & \cdots & A^{n-1}B \end{bmatrix} = 0$

Theorem

W_T is independent of T for $T > 0$ and $W_T = \operatorname{im} \begin{bmatrix} B & AB & A^2B & \cdots & A^{n-1}B \end{bmatrix}$

Definition *Controllability*

Σ is **controllable** at $T > 0$ if for any $x_0, x_f \in \mathbb{R}^n$, there exists $u : [0, T] \to \mathbb{R}^m \quad \text{s.t.} \quad x(T; x_0, u) = x_f$

Theorem

Σ is controllable at $T > 0$ if and only if Σ is reachable at $T > 0$

Theorem

The following are equivalent:

1. $\exists T > 0 \quad \text{s.t.} \quad \Sigma$ is controllable at T
2. Σ is controllable at T for all $T > 0$
3. $W = \mathbb{R}^n$
4. $\operatorname{rank} \begin{bmatrix} B & AB & A^2B & \cdots & A^{n-1}B \end{bmatrix} = n$

4.3 Observability

Consider the linear system $\Sigma : \dot{\boldsymbol{x}}(t) = A\boldsymbol{x}(t), \boldsymbol{y} = C\boldsymbol{x}(t)$

Definition *Indistinguishable initial conditions*

$x_0, x_0' \in \mathbb{R}^n$ are **indistinguishable** on $[0, T]$ if $y(t; x_0) = y(t; x_0^i)$ for all $t \in [0, T]$

Definition *Unobservable subspace*

$$N_T = \{x \in \mathbb{R}^n \mid x \text{ and } 0 \text{ are indistinguishable on } [0, T]\}$$

Theorem

Let $T > 0$. The following are equivalent:

1. $x \in N_T$
2. $Ce^{At}x = 0$ for all $t \in [0, T]$
3. $CA^k x = 0$ for all $k \in \{0, 1, 2, \ldots\}$
4. $\begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} x = 0$

Theorem

N_T is independent of T for $T > 0$ and $N_T = \ker \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix}$

Definition *Observability*

Σ is **observable** on $[0, T]$ is $x_0, x_0' \in \mathbb{R}^n$ are indistinguishable on $[0, T]$ only if $x_0 = x_0'$

Theorem

The following are equivalent:

1. $\exists T > 0 \quad \text{s.t.} \quad \Sigma$ is observable on $[0, T]$
2. Σ is observable on $[0, T]$ for all $T > 0$
3. $N = \{0\}$
4. $\operatorname{rank} \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} = n$

4.4 Similarity

Definition *Similarity*

$\Sigma(A, B, C, D)$ and $\Sigma(\overline{A}, \overline{B}, \overline{C}, \overline{D})$ are **similar** if there exists a nonsingular T such that

$$\overline{A} = TAT^{-1} \qquad \overline{B} = TB \qquad \overline{C} = CT^{-1} \qquad \overline{D} = D$$

Theorem

Let $\Sigma(A, B)$ and $\Sigma(\overline{A}, \overline{B})$ be similar. Let $\Sigma(A, C)$ and $\Sigma(\overline{A}, \overline{C})$ be similar. Then,

$$\Sigma(A, B) \text{ is controllable} \iff \Sigma(\overline{A}, \overline{B}) \text{ is controllable}$$

$$\Sigma(A, C) \text{ is observable} \iff \Sigma(\overline{A}, \overline{C}) \text{ is observable}$$

Definition *A-invariance*

Let $A \in \mathbb{R}^{n \times n}$ and let $v \subset \mathbb{R}^n$ be a subspace.
Then, V if A-**invariant** if $x \in V \implies Ax \in V$ (notation: $AV \subset V$)

4.5 Canonical forms

Theorem

W is the smallest A-invariant subspace containing the image of B.
(*smallest* means that W is a subset of any other subspace satisfying the condition.)

Theorem *Canonical form for uncontrollable systems*

Consider $\Sigma(A, B)$ and let $k = \dim W < n$. Then there exists a nonsingular $T \in \mathbb{R}^{n \times n}$ such that

$$TAT^{-1} = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} \qquad TB = \begin{bmatrix} B_1 \\ 0 \end{bmatrix}$$

with $A_{11} \in \mathbb{R}^{k \times k}, B_1 \in \mathbb{R}^{k \times m}$ and $\Sigma(A_{11}, B_1)$ controllable.

Theorem *Canonical form for unobservable systems*

Consider $\Sigma(A, C)$ and let $k = \dim N < n$. Then there exists a nonsingular $T \in \mathbb{R}^{n \times n}$ such that

$$TAT^{-1} = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} \qquad CT^{-1} = \begin{bmatrix} B_1 \\ 0 \end{bmatrix}$$

with $A_{11} \in \mathbb{R}^{k \times k}, C_1 \in \mathbb{R}^{p \times k}$ and $\Sigma(A_{11}, C_1)$ observable.

Theorem *Controllability canonical form*

Let $\Sigma(A, B)$ with $u(t) \in \mathbb{R}, x(t) \in \mathbb{R}^{n \times n}$ be controllable.
Then, there exists a nonsingular $T \in \mathbb{R}^{n \times n}$ such that

$$TAT^{-1} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-2} & -a_{n-1} \end{bmatrix} \qquad TB = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

where $a_i \in \mathbb{R}$ and $\Delta_A(s) = s^n + a_{a-1}s^{n-1} + \ldots + a_1 s + a_0$

Theorem *Observability canonical form*

Let $\Sigma(A,C)$ with $y(t)\in\mathbb{R}, x(t)\in\mathbb{R}^{n\times n}$ be controllable.
Then, there exists a nonsingular $T\in\mathbb{R}^{n\times n}$ such that

$$TAT^{-1}=\begin{bmatrix}0&0&0&\cdots&0&-a_0\\1&0&0&\cdots&0&-a_1\\0&1&0&\ddots&0&-a_2\\\vdots&\vdots&\ddots&\ddots&\vdots&\vdots\\0&0&\cdots&1&0&-a_{n-2}\\0&0&\cdots&0&1&-a_{n-1}\end{bmatrix}\qquad CT^{-1}=\begin{bmatrix}0&0&\cdots&0&0&1\end{bmatrix}$$

where $a_i\in\mathbb{R}$ and $\Delta_A(s)=s^n+a_{n-1}s^{n-1}+\ldots+a_1s+a_0$

Algorithm *Computation of the controllability canonical form*

1. Verify controllability
2. Compute the characteristic polynomial of A
3. Compute T using $T^{-1}=\begin{bmatrix}q_1&q_2&\cdots&q_n\end{bmatrix}$
   where $q_n=B,\quad q_{n-1}=AB+a_{n-1}B,\quad q_{n-2}=A^2B+a_{n-1}AB+a_{n-2}B,\quad\ldots$
   and $\Delta_A(s)=s^n+a_{n-1}s^{n-1}+\ldots+a_1s+a_0$

4.6 Controllable and observable eigenvalues

Definition *Controllable and observable eigenvalues*

$\lambda\in\sigma(A)$ is (A,B) controllable if rank $\begin{bmatrix}A-\lambda I & B\end{bmatrix}=n$

$\lambda\in\sigma(A)$ is (A,C) observable if rank $\begin{bmatrix}A-\lambda I\\ C\end{bmatrix}=n$

Theorem *Hautus test*

$\Sigma(A,B)$ is controllable if and only if rank $\begin{bmatrix}A-\lambda I & B\end{bmatrix}=n$ for all $\lambda\in\sigma(A)$.

$\Sigma(A,B)$ is observable if and only if rank $\begin{bmatrix}A-\lambda I\\ C\end{bmatrix}=n$ for all $\lambda\in\sigma(A)$.

5 Stabilization

5.1 Stabilization by static state feedback

Definition *Closed-loop system*

Let $u(t)=Fx(t)$ with $F\in\mathbb{R}^{m\times n}$. Then $\Sigma(A,B):\dot{x}(t)=(A+BF)x(t)$ is a **closed-loop system**.

Theorem *Pole placement theorem*

The following are equivalent:
1. $\Sigma(A,B)$ is controllable
2. For any monic polynomial p of degree n, there exists $F\in\mathbb{R}^{m\times n}$ such that $\Delta_{A+BF}(s)=p(s)$

Algorithm *Computation of the static state feedback controller given $\sigma(A+BF)$*

1. Verify controllability
2. Compute the characteristic polynomial of A
3. Compute T (controllability canonical form)
4. Find the characteristic polynomial p that corresponds to $\sigma(A+BF)$
5. Compute $\overline{F}=FT^{-1}=\begin{bmatrix}f_0&f_1&\cdots&f_{n-2}&f_{n-1}\end{bmatrix}$ where $f_i=a_i-p_i$,
   p_i are the coefficients of p and a_i are the coefficients of Δ_A
6. Transform $\overline{F}$ to the original coordinates (solve $\overline{F}=FT^{-1}$ for F)

Definition *Stabilizability*

Σ is **stabilizable** if there exists $F \in \mathbb{R}^{m\times n}$ such that $\sigma(A+BF) \subset \mathbb{C}_-$

Theorem

$\Sigma(A, B)$ is stabilizable if and only if every $\lambda \in \sigma(A)$ satisfying $\lambda \notin \mathbb{C}_-$ is (A, B)-controllable.

Theorem *Hautus test for stabilizability*

$\Sigma(A, B)$ is stabilizable if and only if rank $\begin{bmatrix} A-\lambda I & B \end{bmatrix} = n$ for all $\lambda \in \sigma(A)$ such that $\Re(\lambda) \geq 0$

5.2 State observers

Definition *State observer*

Let Ω be a linear system of the form $\begin{cases} \dot{w}(t) = Pw(t) + Qu(t) + Ry(t) \\ \xi(t) = Sw(t) \end{cases}$

with $w(t) \in \mathbb{R}^{n_w}$, $\xi \in \mathbb{R}^n$ and $e(t) = \xi(t) - x(t)$.
Ω is a **state observer** if for any pair of initial conditions x_0, w_0 such that $e(0) = Sw_0 - x_0 = 0$ and for any input function $u(t)$, $\quad e(t) = 0$ for all $t \geq 0$.

Definition *Stable state observer*

A state observer Ω is **stable** if for any pair of initial conditions and any input function $u(t)$, $\lim_{n\to\infty} te(t) = 0$

Theorem

Ω is a state observer if and only if $SQ = B$ and $SP = AS - SRCS$.

Theorem *General form of a state observer*

The general form of a state observer for Σ is $\dot{\xi}(t) = (A - GC)\xi(t) + Bu(t) + Gy(t)$.
The estimation error satisfies $\dot{e}(t) = (A - GC)e(t)$.
The state observer is stable if and only if $\sigma(A - GC) \in \mathbb{C}_-$.

Definition *Detectability*

Σ is **detectable** if there exists $G \in \mathbb{R}^{n\times p}$ such that $\sigma(A - GC) \subset \mathbb{C}_-$

Lemma

$$(A, C) \text{ is detectable} \iff (A^T, C^T) \text{ is stabilizable}$$

Theorem *Hautus test for detectability*

$\Sigma(A, B)$ is detectable if and only if rank $\begin{bmatrix} A-\lambda I \\ C \end{bmatrix} = n$ for all $\lambda \in \sigma(A)$ such that $\Re(\lambda) \geq 0$

Corollary

Consider Σ. The following are equivalent:

1. There exists a stable state observer for Σ
2. Σ is detectable
3. Every $\lambda \in \sigma(A)$ such that $\lambda \notin \mathbb{C}_-$ is (A, C) observable

5.3 Stabilization by dynamic output feedback

Dynamic output feedback controller

Controller: $\begin{cases} \dot{w}(t) = Kw(t) + Ly(t) \\ u(t) = Mw(t) + Ny(t) \end{cases}$ Closed-loop dynamics: $\begin{bmatrix} \dot{x}(t) \\ \dot{w}(t) \end{bmatrix} = \begin{bmatrix} A + BNC & BM \\ LC & K \end{bmatrix} \begin{bmatrix} x(t) \\ w(t) \end{bmatrix}$

Definition *Stabilization problem*

Given Σ, find K, L, M, N such that $A_{cl} = \begin{bmatrix} A + BNC & BM \\ LC & K \end{bmatrix}$ satisfies $\sigma(A_{cl}) \subset \mathbb{C}_-$

Lemma *Solution to the stabilization problem*

Let

- Ω be a stable state observer for Σ
- F solves the stabilization problem by static state feedback
- $\Gamma : \begin{cases} \dot{\xi}(t) = (A - GC + BF)\xi(t) + Gy(t) \\ u(t) = F\xi(t) \end{cases}$

Then Γ solves the stabilization problem by dynamic output feedback.

Theorem

There exists a dynamic output feedback controller Γ stabilizing $\Sigma(A, B, C)$ if and only if the matrix pair (A, B) is stabilizable and the matrix pair (A, C) is detectable.

Algorithm *Designing a dynamic output feedback controller for a model*

1. Write the model in state-space form
2. If nonlinear:
   (a) Find an equilibrium point
   (b) Linearize the system around the equilibrium
3. Check controllability/stabilizability
4. Check observability/detectability
5. Find F such that $\sigma(A + BF) \subset \mathbb{C}_-$
6. Find G such that $\sigma(A - GC) \subset \mathbb{C}_-$
7. Construct Γ

6 Input-output properties

6.1 Impulse response matrix

Unique output solution of a linear system

$$y(t; t_0, x_0, u) = Ce^{A(t-t_0)} + \int_{t_0}^{t} Ce^{A(t-\tau)} Bu(\tau)\, \mathrm{d}\tau + Du(t)$$

Theorem

Let $\Sigma(A, B, C, D)$ and $\Sigma(\overline{A}, \overline{B}, \overline{C}, \overline{D})$ be similar.
Then for any input function $u : [0, \infty) \to \mathbb{R}^m$, $y(t; 0, u) = \overline{y}(t; 0, u)$

Lemma

Define the function $u_\varepsilon : \mathbb{R} \to \mathbb{R}$ as $u_\varepsilon(t) = \begin{cases} \frac{1}{2\varepsilon} & -\varepsilon \le t \le \varepsilon \\ 0 & \text{otherwise} \end{cases}$

Define $y_\varepsilon(t) = \int_{-\varepsilon}^{t} Ce^{A(t-\tau)} Bu_\varepsilon(\tau)\, \mathrm{d}\tau$. Then $\lim_{\varepsilon \to 0^+} y_\varepsilon(t) = \begin{cases} Ce^{At}B & \text{if } t > 0 \\ 0 & \text{if } t < 0 \end{cases}$

Definition *Dirac delta function*

A "function" $\delta : \mathbb{R} \to \mathbb{R}$ with defining properties:

1. $\delta(t) = 0$ for all $t \neq 0$
2. For any continuous $\phi : \mathbb{R} \to \mathbb{R}$, $\displaystyle\int_{-\infty}^{\infty} \phi(t-\tau)\delta(t)\,\mathrm{d}t = \phi(t)$

Definition *Impulse response matrix*

The **impulse response matrix** for Σ is $H : \mathbb{R} \to \mathbb{R}^{p\times m}$ defined as

$$H(t) = \begin{cases} Ce^{At}B + D\delta(t) & t \geq 0 \\ 0 & t < 0 \end{cases}$$

Theorem

Consider Σ with impulse response matrix H. Then

$$y(t; 0, u) = \int_0^t H(t-\tau)u(\tau)\,\mathrm{d}\tau$$

6.2 Transfer functions

Definition *Exponentially bounded function*

A function $f : [0, \infty) \to \mathbb{R}$ is **exponentially bounded** if there exist some $M, \alpha \in \mathbb{R}$ such that $|f(t)| \leq Me^{\alpha t}$ for all t.

Definition *Laplace transform*

For f exponentially bounded, its **Laplace transform** is

$$\mathcal{L}(f)(s) = \int_0^{\infty} f(t)e^{-st}\,\mathrm{d}t \qquad \text{for } s \in \mathbb{C} \text{ with } \Re(s) > \alpha$$

Theorem *Properties of the Laplace transform*

1. $\mathcal{L}(f+g) = \mathcal{L}(f) + \mathcal{L}(g)$
2. $\mathcal{L}(\alpha f) = \alpha\mathcal{L}(f)$
3. if f differentiable and f' exp. bounded, $\mathcal{L}(f') = s\mathcal{L}(f) - f(0)$
4. if $u, h : [0, \infty) \to \mathbb{R}$ exp. bounded,
   then $y(t) = \displaystyle\int_0^t j(t-\tau)u(t)\,\mathrm{d}\tau$ is exp. bounded and $\mathcal{L}(y) = \mathcal{L}(h)\mathcal{L}(u)$

Definition *Transfer function matrix*

The **transfer function** of Σ is $T(s) = C(sI - A)^{-1}B + D$

Relation between Laplace transforms of input and output

$\hat{y}(s) = T(s)\hat{u}(s)$ where $\hat{u}(s) = \mathcal{L}(u)(s)$ and $\hat{y} = \mathcal{L}(y(\,\cdot\,; 0, u))(s)$

Theorem

$$T(s) = \mathcal{L}(H)(s)$$

for all $s \in \mathbb{C}$ such that $\Re(s) > \Lambda(A)$, where $\Lambda(A) = \max\{\Re(\lambda) \mid \lambda \in \sigma(A)\}$

Theorem

If $\Sigma(A, B, C, D)$ and $\Sigma(\overline{A}, \overline{B}, \overline{C}, \overline{D})$ are similar, then they have the same transfer function.

6.3 Transfer functions for SISO systems

A **SISO system** has a single input and a single output, i.e. $m = 1$, $p = 1$

Theorem

Consider the system $\Sigma(A, B, C, D)$ with

$$A = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-2} & -a_{n-1} \end{bmatrix} \qquad B = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 1 \end{bmatrix} \qquad C = \begin{bmatrix} c_0 & c_1 & c_2 & \dots & c_{n-2} & c_{n-1} \end{bmatrix}$$

Then, $T(s) = \dfrac{c_{n-1}s^{n-1} + c_{n-2}s^{n-2} + \ldots + c_1 s + c_0}{s^n + a_{n-1}s^{n-1} + a_{n-2}s^{n-2} + \ldots + a_1 s + a_0}$

Theorem

Consider the SISO system Σ and let (A, B) be controllable. Then, the polynomials

$$p(s) = C \operatorname{adj}(sI - A)B \qquad q(s) = \Delta_A(s)$$

are coprime (i.e. they have no common roots) if and only if (A, C) is observable.

6.4 External stability

Transfer functions

Scalar transfer function (SISO): $T(s) = \dfrac{p(s)}{q(s)} = \dfrac{p'(s)}{q'(s)}$ $\quad p', q'$ coprime

Matrix transfer function: $T(s) = \dfrac{1}{q(s)} P(s)$

Definition *Pole*

(**Scalar case**) $\lambda \in \mathbb{C}$ is a **pole** of T if is a root of q'
(**Matrix case**) $\lambda \in \mathbb{C}$ is a **pole** of T if it is a pole of at least 1 of its elements

Theorem

1. If $\lambda \in \mathbb{C}$ is a pole of T, then $\lambda \in \sigma(A)$
2. If $\lambda \in \mathbb{C} \in \sigma(A)$, (A, B) is controllable and (A, C) is observable, then λ is a pole of T

Definition *Internal and external stability*

A system $\Sigma(A, B, C, D)$ is

- **internally stable** if $\dot{x}(t) = Ax(t)$ is asymptotically stable
- **externally stable** if there exists $\gamma > 0$ such that, for any bounded input function $u : [0, \infty) \to \mathbb{R}^m$,

$$\sup_{t \in \mathbb{R}_+} |y(t; 0, u)| \leq \gamma \sup_{t \in \mathbb{R}_+} |u(t)| \quad \text{or} \quad |u(t)| \leq 1 \quad \forall t \in [0, \infty) \implies |y(t; 0, u)| \leq \gamma \quad \forall t \in [0, \infty)$$

Lemma

$$\Sigma(A, B, C, D) \text{ is externally stable} \iff \Sigma(A, B, C, 0) \text{ is externally stable}$$

Theorem

The following are equivalent:

1. Σ is externally stable
2. $\displaystyle\int_0^\infty \|Ce^{At}B\|\,\mathrm{d}t < \infty$ (recall: $\|M\| = \sup\{\frac{|Mx|}{|x|} \mid x \neq 0\}$)
3. $\lim\limits_{t\to\infty} Ce^{At}B = 0$
4. all poles of T are in $\mathbb{C}_-$

Theorem

1. If Σ is internally stable, then it is externally stable.
2. If Σ is externally stable, the matrix pair (A, B) is controllable, and the matrix pair (A, C) is observable, then Σ is internally stable.