University of Washington Lecture Notes for ME547 Linear Systems

Professor Xu Chen Bryan T. McMinn Endowed Research Professorship Associate Professor Department of Mechanical Engineering University of Washington

Winter 2023

Abstract: ME547 is a first-year graduate course on modern control systems focusing on: statespace description of dynamic systems, linear algebra for controls, solutions of state-space systems, discrete-time models, stability, controllability and observability, state-feedback control, observers, observer state feedback controls, and when time allows, linear quadratic optimal controls. ME547 is a prerequisite to most advanced graduate control courses in the UW ME department.

Copyright: Xu Chen 2021~. Limited copying or use for educational purposes allowed, but please make proper acknowledgement, e.g.,

"Xu Chen, Lecture Notes for University of Washington Linear Systems (ME547), Available at https://faculty.washington.edu/chx/teaching/me547/."

or if you use LATEX:

@Misc{xchen21, key = {controls,systems,lecture notes}, author = {Xu Chen}, title = {Lecture Notes for University of Washington Linear Systems (ME547)}, howpublished = {Available at \url{https://faculty.washington.edu/chx/teaching/me547/}}, month = March, year = 2023 }

Contents

0 Introduction 1 Modeling overview 3 Laplace transform 3 Inverse Laplace transform 4 Z transform 5 State-space overview 6 From the state space to transfer functions 7 State-space realizations (slides) 8 State-space realizations (notes) 9 Solution to state-space systems (slides) 10 Solution to state-space systems (notes) 11 Discretization of state-space models 12 Discretization of transfer-function models 13 Stability 14 Controllability and observability 15 Kalman decomposition 16 State-feedback control 17 Observers and observer-state-feedback control 18 Linear quadratic optimal control A1 Linear Algebra Review

ME547: Linear Systems Introduction

Xu Chen

University of Washington

UW Linear Systems (X. Chen, ME547)

Introduction

1/16

Topic

- 1. Introduction of controls
- 2. Introduction of the course
- 3. Beyond the course

The power of controls

- nanometer precision control
 - information storage
 - semiconductor manufacturing
- optics and laser beam steering
- robotics for manufacturing
- laser-material interaction in additive manufacturing
- everyday life: driving, cooking, showering, to name just a few

UW Linear Systems (X. Chen, ME547)

Introduction

3/16

Topic

- 1. Introduction of controls
- 2. Introduction of the course
- 3. Beyond the course

Course Scope: analysis and control of linear dynamic systems

- System: an interconnection of elements and devices for a desired purpose
- Control System: an interconnection of components forming a system configuration that will provide a desired response
- Feedback: the use of information of the past or the present to influence behaviors of a system

UW Linear Systems (X. Chen, ME547)

Introduction

Why automatic control?

A system can be either manually or automatically controlled. Why automatic control?

- Stability/Safety: difficult/impossible for humans to control the process or would expose humans to risk
- Performance: cannot be done "as well" by humans
- Cost: Humans are more expensive and can get bored
- Robustness: can deliver the requisite performance even if process behaves slightly differently

5/16

Terminologies



Block diagram: visualizes system structure and the flow information in control systems

UW Linear Systems (X. Chen, ME547) Introduction

Open-loop control v.s. closed-loop control



- the output of the plant does not influence the input to the controller
- input and output as *signals*: functions of time, e.g., speed of a car, temperature in a room, voltage applied to a motor, price of a stock, electrical-cardiograph, all as functions of time.

7/16

Open-loop control v.s. closed-loop control



Closed-loop control: regulation example



Regulation control example: automobile cruise control



- What is the control objective?
- What are the process, process output, actuator, sensor, reference, and disturbance?



Control objectives

- Better stability
- Improved response characteristics
- Regulation of output in the presence of disturbances and noises
- Robustness to plant uncertainties
- Tracking time varying desired output

There are some aspects of control objectives that are universal. For example, we would always want our control system to result in closed-loop dynamics that are insensitive to disturbances. This is the disturbance rejection problem. Also, as pointed out previously, we would want the controller to be robust to plant modeling errors.

Means to achieve the control objectives

- Model the controlled plant
- Analyze the characteristics of the plant
- Design control algorithms (controllers)
- Analyze performance and robustness of the control system
- Implement the controller

UW Linear Systems (X. Chen, ME547)

Introduction

13/16

Topic

- 1. Introduction of controls
- 2. Introduction of the course
- 3. Beyond the course

Resources for control education: societies

- AIAA (American Institute of Aeronautics and Astronautics)
 - Publications: AIAA Journal of Guidance, Control and Navigation
- ASME (American Society of Mechanical Engineers)
 - Publications: ASME Journal of Dynamic Systems, Measurement and Control¹
- IEEE (Institute of Electrical and Electronics Engineers)
 - www.ieee.org
 - Control System Society
 - Publications:
 - IEEE Control Systems Magazine¹
 - IEEE Transactions on Control Technology
 - IEEE Transactions on Automatic Control
- IFAC (International Federation of Automatic Control)
 - Publications: Automatica, Control Engineering Practice

¹start looking at these, online or at library UW Linear Systems (X. Chen, ME547) Introduction

15 / 16

IEEE Control Systems Magazine



ME547: Linear Systems Modeling of Dynamic Systems

Xu Chen

University of Washington

UW Linear Systems (X. Chen, ME547)

Modeling

1/13

Topic

1. Introduction

- 2. Methods of Modeling
- 3. Model Properties

Why modeling?

Modeling of physical systems:

- a vital component of modern engineering
- often consists of complex coupled differential equations
- only when we have good understanding of a system can we optimally control it:
 - can simulate and predict actual system response, and
 - design model-based controllers
 - example: nanometer precision control

UW Linear Systems (X. Chen, ME547)

Modeling

3/13

History

- Newton developed Newton's laws in 1686. He is an extremely brilliant scientist and in the meantime very eccentric. He was described as "...so absorbed in his studies that he forgot to eat".
- Faraday discovered induction (Faraday's law), which led to electric motors. He was born into a poor family and had virtually no schooling. He read many books and self-taught himself when he became an apprentice to a bookbinder at the age of 14.

Topic

1. Introduction

2. Methods of Modeling

3. Model Properties

UW Linear Systems (X. Chen, ME547)

Modeling

5/13

Two general approaches of modeling

based on physics:

- using fundamental engineering principles such as Newton's laws, energy conservation, etc
- focus in this course

based on measurement data:

- using input-output response of the system
- a field itself known as system identification

Example: Mass spring damper



Newton's second law gives

$$m\ddot{y}(t) + b\dot{y}(t) + ky(t) = u(t), \ y(0) = y_0, \ \dot{y}(0) = \dot{y}_0$$

- modeled as a second-order ODE with input u(t) and output y(t)
- ▶ if instead, that velocity is the desired output, the model will be different
- Application example: semiconductor wafer scanner

UW Linear Systems (X. Chen, ME547) Modeling

Example: HDDs, SCARA robots

Newton's second law for rotation



- single-stage HDD
- dual-stage HDD, SCARA (Selective Compliance Assembly Robot Arm) robots

7/13

Topic

1. Introduction

2. Methods of Modeling

3. Model Properties

UW Linear Systems (X. Chen, ME547)

Modeling

Model properties: static v.s. dynamic, causal v.s. acausal

Consider a general system \mathcal{M} with input u(t) and output y(t):

$$u \longrightarrow \mathcal{M} \longrightarrow y$$

 $\ensuremath{\mathcal{M}}$ is said to be

- memoryless or static if y(t) depends only on u(t).
- dynamic (has memory) if y at time t depends on input values at other times. e.g.: $y(t) = \mathcal{M}(u(t)) = \gamma u(t)$ is memoryless; $y(t) = \int_0^t u(\tau) d\tau$ is dynamic.
- causal if y(t) depends on $u(\tau)$ for $\tau \leq t$,
- strictly causal if y(t) depends on $u(\tau)$ for $\tau < t$, e.g.: y(t) = u(t-10).
- Exercise: is differentiation causal?

9/13

Linearity and time-invariance

The system ${\mathcal M}$ is called

linear if satisfying the superposition property:

 $\mathcal{M}(\alpha_1 u_1(t) + \alpha_2 u_2(t)) = \alpha_1 \mathcal{M}(u_1(t)) + \alpha_2 \mathcal{M}(u_2(t))$

for any input signals $u_1(t)$ and $u_2(t)$, and any real numbers α_1 and α_2 .

- time-invariant if its properties do not change with respect to time.
- Assuming the same initial conditions, if we shift u(t) by a constant time interval, i.e., consider $\mathcal{M}(u(t + \tau_0))$, then \mathcal{M} is time-invariant if the output $\mathcal{M}(u(t + \tau_0)) = y(t + \tau_0)$.
- e.g., $\dot{y}(t) = Ay(t) + Bu(t)$ is linear and time-invariant; $\dot{y}(t) = 2y(t) - \sin(y(t))u(t)$ is nonlinear, yet time-invariant; $\dot{y}(t) = 2y(t) - t\sin(y(t))u(t)$ is time-varying.



Modeling

11 / 13

Models of continuous-time systems

The systems we will be dealing with are mostly composed of ordinary differential or difference equations.

General continuous-time systems:

$$\frac{d^n y(t)}{dt^n} + a_{n-1} \frac{d^{n-1} y(t)}{dt^{n-1}} + \dots + a_0 y(t) = b_m \frac{d^m u(t)}{dt^m} + b_{m-1} \frac{d^{m-1} u(t)}{dt^{m-1}} + \dots + b_0 u(t)$$

with the initial conditions $y(0) = y_0, \ldots, y^{(n)}(0) = y_0^{(n)}$. For the systems to be causal, it must be that $n \ge m$. (Check, e.g., the case with n = 0 and m = 1.)

Models of discrete-time systems

General discrete-time systems:

- inputs and outputs defined at discrete time instances k = 1, 2, ...
- described by ordinary difference equations in the form of

$$y(k) + a_{n-1}y(k-1) + \cdots + a_0y(k-n) = b_mu(k+m-n) + \cdots + b_0u(k-n)$$

Example: bank statements

- ▶ k year counter; p interest rate; x(k) wealth at the beginning of year k; u(k) money saved at the end of year k; x₀ initial wealth in account
- $x(k+1) = (1+\rho)x(k) + u(k), x(0) = x_0$

UW Linear Systems (X. Chen, ME547)

Modeling

13/13

ME547: Linear Systems Modeling: Review of Laplace Transform

Xu Chen

University of Washington

UW Linear Systems (X. Chen, ME547)

Laplace

1/29

Topic

1. The Big Picture

2. Definitions

3. Examples

4. Properties

Introduction

- The Laplace transform is a powerful tool to solve a wide variety of Ordinary Differential Equations (ODEs).
- Pierre-Simon Laplace (1749-1827):
 - often referred to as the French Newton or Newton of France
 - 13 years more junior than Lagrange
 - developer / pioneer of astronomical stability, mechanics based on calculus, Bayesian interpretation of probability, mathematical physics, just to name a few.
 - studied under Jean le Rond d'Alembert (co-discovered fundamental theorem of algebra, aka d'Alembert/Gauss theorem)

UW Linear Systems (X. Chen, ME547)

Laplace

3/29

The Laplace approach to ODEs



Topic

1. The Big Picture

2. Definitions

- 3. Examples
- 4. Properties

UW Linear Systems (X. Chen, ME547)

Laplace

5/29

Sets of numbers and the relevant domains

- ▶ set: a well-defined collection of distinct objects, e.g., $\{1, 2, 3\}$
- \blacktriangleright \mathbb{R} : the set of real numbers
- \blacktriangleright \mathbb{C} : the set of complex numbers
- \blacktriangleright \in : belong to, e.g., $1 \in \mathbb{R}$
- ▶ \mathbb{R}_+ : the set of positive real numbers
- ▶ \triangleq : defined as, e.g., $y(t) \triangleq 3x(t) + 1$

Continuous-time functions

Formal notation:

$$f: \mathbb{R}_+ \to \mathbb{R}$$

where the input of f is in \mathbb{R}_+ , and the output in \mathbb{R}

- we will mostly use f(t) to denote a continuous-time function
- domain of f is time
- assume that f(t) = 0 for all t < 0

UW Linear Systems (X. Chen, ME547)

Laplace

Laplace transform definition

For a continuous-time function

$$f: \mathbb{R}_+ \to \mathbb{R}$$

define Laplace Transform:

$$F(s) = \mathcal{L}{f(t)} \triangleq \int_0^\infty f(t) e^{-st} dt$$

 $\pmb{s} \in \mathbb{C}$

7 / 29

Existence: Sufficient condition 1



Existence: Sufficient condition 2

▶ f(t) does not grow faster than an exponential as $t \to \infty$:

$$|f(t)| < ke^{at}$$
, for all $t \ge t$

for some constants: $k, \alpha, t \in \mathbb{R}_+$.



Topic

- 1. The Big Picture
- 2. Definitions
- 3. Examples
- 4. Properties

UW Linear Systems (X. Chen, ME547)

Laplace

11/29

Examples: Exponential

$$f(t) = e^{-at}, \ a \in \mathbb{C}$$

 $F(s) = rac{1}{s+a}$

Examples: Exponential

$$egin{aligned} f(t) &= 1(t) = \left\{ egin{aligned} 1, & t \geq 0 \ 0, & t < 0 \end{aligned}
ight. \ F(s) &= rac{1}{s} \end{aligned}$$

UW Linear Systems (X. Chen, ME547)

Laplace

13 / 29

Examples: Sine

$$f(t) = \sin(\omega t)$$

 $F(s) = rac{\omega}{s^2 + \omega^2}$
Use: $\sin(\omega t) = rac{e^{j\omega t} - e^{-j\omega t}}{2j}$, $\mathcal{L}\{e^{j\omega t}\} = rac{1}{s - j\omega}$

Recall: Euler formula

$$e^{ja} = \cos a + j \sin a$$

Leonhard Euler (04/15/1707 - 09/18/1783):

- Swiss mathematician, physicist, astronomer, geographer, logician and engineer
- studied under Johann Bernoulli
- teacher of Lagrange
- wrote 380 articles within 25 years at Berlin
- produced on average one paper per week at age 67, when almost blind!

UW Linear Systems (X. Chen, ME547)

Laplace

15 / 29

Examples: Cosine

$$f(t) = \cos(\omega t)$$
$$F(s) = \frac{s}{s^2 + \omega^2}$$

Examples: Dirac impulse



 Background: a generalized function or distribution, e.g., for derivatives of step functions

Properties:

$$\int_0^\infty \delta(t-T)dt = 1$$

$$\int_0^\infty \delta(t-T)f(t)dt = f(T)$$

UW Linear Systems (X. Chen, ME547)

Laplace

17 / 29

Examples: Dirac impulse

$$\blacktriangleright f(t) = \delta(t)$$

Calculation:

$$\mathcal{L}\{\delta(t)\} = \int_0^\infty e^{-st} \delta(t) dt = e^{-s0} = 1$$

because $\int_0^\infty \delta(t) f(t) dt = f(0)$

Topic

- 1. The Big Picture
- 2. Definitions
- 3. Examples
- 4. Properties

UW Linear Systems (X. Chen, ME547)

Laplace

19/29

Linearity

For any $\alpha, \beta \in \mathbb{C}$ and functions f(t), g(t), let

$$F(s) = \mathcal{L}{f(t)}, \ G(s) = \mathcal{L}{g(t)}$$

then

$$\mathcal{L}\{\alpha f(t) + \beta g(t)\} = \alpha F(s) + \beta G(s)$$

Differentiation

Defining

$$\dot{f}(t) = rac{df(t)}{dt}$$

 $F(s) = \mathcal{L}{f(t)}$

then

$$\mathcal{L}\{\dot{f}(t)\} = sF(s) - f(0)$$

via integration by parts:

$$\mathcal{L}{\dot{f}(t)} = \int_0^\infty e^{-st} \dot{f}(t) dt$$

= $-\int_0^\infty \frac{de^{-st}}{dt} f(t) dt + {e^{-st}f(t)}_{t=0}^{t \to \infty}$ (1)
= $s \int_0^\infty e^{-st} f(t) dt - f(0) = sF(s) - f(0)$

UW Linear Systems (X. Chen, ME547)

Laplace

21/29

Integration

Defining

$$F(s) = \mathcal{L}{f(t)}$$

then

$$\mathcal{L}\left\{\int_0^t f(\tau)d\tau\right\} = \frac{1}{s}F(s)$$

Multiplication by e^{-at}

Defining

$$F(s) = \mathcal{L}\{f(t)\}$$

then

$$\mathcal{L}\left\{e^{-at}f(t)\right\}=F(s+a)$$

Example:

$$\mathcal{L}\{1(t)\} = \frac{1}{s} \quad \mathcal{L}\{e^{-at}\} = \frac{1}{s+a}$$
$$\mathcal{L}\{\sin(\omega t)\} = \frac{\omega}{s^2 + \omega^2} \quad \mathcal{L}\{e^{-at}\sin(\omega t)\} = \frac{\omega}{(s+a)^2 + \omega^2}$$

Multiplication by t

Defining

$$F(s) = \mathcal{L}{f(t)}$$

then

$$\mathcal{L}\left\{tf(t)\right\} = -\frac{dF(s)}{ds}$$

Example:

$$\mathcal{L}{1(t)} = rac{1}{s}$$
 $\mathcal{L}{t} = rac{1}{s^2}$

Time delay τ

Defining

then

$$F(s) = \mathcal{L}{f(t)}$$
$$\mathcal{L}{f(t-\tau)} = e^{-s\tau}F(s)$$

UW Linear Systems (X. Chen, ME547)

Laplace

25 / 29

Convolution

Given f(t), g(t), and

$$(f\star g)(t) = \int_0^t f(t-\tau)g(\tau)d\tau = (g\star f)(t)$$

then

$$\mathcal{L}\left\{(f\star g)(t)\right\}=F(s)G(s)$$

► Proof: exercise

Hence we have

$$\delta(t) \longrightarrow G(s) \longrightarrow g(t) = \mathcal{L}^{-1} \{G(s)\}$$

because

$$1 \longrightarrow \overline{G(s)} \longrightarrow Y(s) = G(s) \times 1$$

Initial Value Theorem

If $f(0_+) = \lim_{t \to 0_+} f(t)$ exists, then

$$f(0_+) = \lim_{s \to \infty} sF(s)$$

UW Linear Systems (X. Chen, ME547)

Laplace

27 / 29

Final Value Theorem

If $\lim_{t \to \infty} f(t)$ exists, then

$$\lim_{t\to\infty}f(t)=\lim_{s\to0}sF(s)$$

• Example: find the final value of the system corresponding to:

$$Y_1(s) = rac{3(s+2)}{s(s^2+2s+10)}, \ \ Y_2(s) = rac{3}{s-2}$$

Common Laplace transform pairs

f(t)	<i>F</i> (<i>s</i>)	f(t)	F(s)
$\sin \omega t$	$\frac{\omega}{s^2 + \omega^2}$	e^{-at}	$\frac{1}{s+a}$
$\cos \omega t$	$\frac{s}{s^2 + \omega^2}$	t	$\frac{1}{s^2}$
$t \times (t)$	$-\frac{dX(s)}{ds}$	t ²	$\frac{2}{s^3}$
$\frac{x(t)}{t}$	$\int_{s}^{\infty} X(s) ds$	te ^{-at}	$\frac{1}{(s+a)^2}$
$\delta(t)$	1	$e^{-at}\sin(\omega t)$	$\frac{\omega}{(s+a)^2+\omega^2}$
1 (<i>t</i>)	$\frac{1}{s}$	$e^{-at}\cos\left(\omega t\right)$	$\frac{\frac{s+a}{s+a}}{(s+a)^2+\omega^2}$

UW Linear Systems (X. Chen, ME547)

Laplace

29 / 29

ME547: Linear Systems Modeling: Inverse Laplace Transform

Xu Chen

University of Washington

UW Linear Systems (ME)

Inverse Laplace

1/14

Topic

- 1. Inverse Laplace transform and partial fraction expansion
- 2. Applications of Laplace transform
- 3. From Laplace transform to transfer functions

Common Laplace transform pairs

f(t)	<i>F</i> (<i>s</i>)	f(t)	F(s)
$\sin \omega t$	$\frac{\omega}{s^2 + \omega^2}$	e^{-at}	$\frac{1}{s+a}$
$\cos \omega t$	$\frac{s}{s^2 + \omega^2}$	t	$\frac{1}{s^2}$
tx(t)	$-\frac{dX(s)}{ds}$	t^2	$\frac{2}{s^3}$
$\frac{x(t)}{t}$	$\int_{s}^{\infty} X(s) ds$	te ^{-at}	$\frac{1}{(s+a)^2}$
$\delta(t)$	1	$e^{-at}\sin(\omega t)$	$\frac{\omega}{(s+a)^2+\omega^2}$
1 (<i>t</i>)	$\frac{1}{s}$	$e^{-at}\cos\left(\omega t\right)$	$\frac{\frac{(s+a)^2+\omega}{s+a}}{(s+a)^2+\omega^2}$

```
UW Linear Systems (ME)
```

Inverse Laplace

3/14

Overview of inverse Laplace transform: modularity and decomposition

Goal: to break a large Laplace transform into small blocks, so that we can use elemental examples of Laplace transfer functions:

$$G(s) = \frac{B(s)}{A(s)} = \frac{B_1(s)}{A_1(s)} + \frac{B_2(s)}{A_2(s)} + \dots$$

We will use examples to demonstrate strategies for common fractional expansions.

Real and distinct roots in A(s)

Example 1

$$G(s) = \frac{B(s)}{A(s)} = \frac{32}{s(s+4)(s+8)} = \frac{K_1}{s} + \frac{K_2}{s+4} + \frac{K_3}{s+8}$$

UW Linear Systems (ME)

Inverse Laplace

5/14

Real and repeated roots in A(s)

Example 2

$$G(s) = \frac{2}{(s+1)(s+2)^2} = \frac{K_1}{s+1} + \frac{K_2}{s+2} + \frac{K_3}{(s+2)^2}$$

- $K_3 = \lim_{s \to -2} (s+2)^2 G(s) = -2$ • $K_1 = \lim_{s \to -1} (s+1) G(s) = 2$
- For K_2 , we hit both sides with $(s+2)^2$ then differentiate once w.r.t. s, to get

$$K_2 = \lim_{s \to -2} \frac{d}{ds} (s+2)^2 G(s) = -2$$
Topic

1. Inverse Laplace transform and partial fraction expansion

2. Applications of Laplace transform

3. From Laplace transform to transfer functions

UW Linear Systems (ME)

Inverse Laplace

Solution of a first-order ODE

Example 1: Let $a > 0, b > 0, y(0) = y_0 \in \mathbb{R}$, obtain the solution to the ODE:

$$\dot{y}(t) = -ay(t) + b1(t)$$

where $1(t) = \left\{ egin{array}{cc} 1, & t \geq 0 \\ 0, & t < 0 \end{array}
ight.$

- Laplace transform: $\mathcal{L}{\dot{y}(t)} = sY(s) y(0)$
- \blacktriangleright \Rightarrow Solution in Laplace domain:

$$Y(s) = \frac{1}{s+a}y(0) + \frac{b}{s(s+a)} = \frac{1}{s+a}y(0) + \frac{b}{a}\left(\frac{1}{s} - \frac{1}{s+a}\right)$$

▶ Apply inverse Laplace transform: $y(t) = \mathcal{L}^{-1}{Y(s)} = ...$

Solution:

$$y(t) = e^{-at}y(0) + \frac{b}{a}(1(t) - e^{-at})$$

Solution of a first-order ODE

Example 1: $a > 0, b > 0, y(0) = y_0 \in \mathbb{R}$:

$$\dot{y}(t) = -ay(t) + b1(t) \Rightarrow Y(s) = \frac{1}{s+a}y(0) + \frac{b}{s(s+a)}$$

 $y(t) = e^{-at}y(0) + \frac{b}{a}(1(t) - e^{-at})$

Observations:

- From the ODE, $y(\infty) = \frac{b}{a}$
- Using final value theorem,

$$\lim_{t\to\infty} y(t) = \lim_{s\to 0} sY(s) = \frac{b}{a}$$

UW Linear Systems (ME)

Inverse Laplace

9/14

Solution of a first-order ODE

Example 2: Let a > 0, b > 0, $y(0) = y_0 \in \mathbb{R}$, obtain the solution to the ODE:

$$\dot{y}(t) = -ay(t) + b\delta(t)$$

- Laplace transform: $\mathcal{L}{\dot{y}(t)} = sY(s) y_0$
- ► ⇒ Solution in Laplace domain: $Y(s) = \frac{1}{s+a}(y_0 + b)$
- Apply inverse Laplace transform: $y(t) = \mathcal{L}^{-1}{Y(s)} = e^{-at}(y_0 + b)$
- Q: what's the initial value from initial value theorem? what does the impulse do to the initial condition?

Topic

- 1. Inverse Laplace transform and partial fraction expansion
- 2. Applications of Laplace transform
- 3. From Laplace transform to transfer functions



Connecting two domains

N-th order differential equation:

$$\frac{d^{n}y}{dt^{n}} + a_{n-1}\frac{d^{n-1}y}{dt^{n-1}} + \dots + a_{1}\dot{y} + a_{0}y = b_{m}\frac{d^{m}u}{dt^{m}} + b_{m-1}\frac{d^{m-1}u}{dt^{m-1}} + \dots + b_{1}\dot{u} + b_{o}u$$

where y(0) = 0, $\frac{dy}{dt}|_{t=0} = 0$, ..., $\frac{d^{n-1}y}{dt^{n-1}}|_{t=0} = 0$ Applying Laplace transform yields

$$(s^n + a_{n-1}s^{n-1} + \dots + a_0)Y(s) = (b_ms^m + b_{m-1}s^{m-1} + \dots + b_0)U(s)$$

 $\Rightarrow Y(s) = rac{b_ms^m + b_{m-1}s^{m-1} + \dots + b_0}{s^n + a_{n-1}s^{n-1} + \dots + a_0}U(s)$

Transfer functions

$$G(s) = \frac{Y(s)}{U(s)} = \frac{b_m s^m + b_{m-1} s^{n-1} + \dots + b_0}{s^n + a_{n-1} s^{n-1} + \dots + a_0}$$

- A(s) = 0: Characteristic equation (C.E.)
- Roots of C.E.: poles of G(s)
- Roots of B(s) = 0: zeros of G(s)
- $m \le n$: realizability condition; G(s) is called
 - proper if $n \ge m$
 - strictly proper if n > m

• Examples:
$$G_1(s) = K$$
, $G_2(s) = rac{k}{s+a}$

UW Linear Systems (ME)

Inverse Laplace

The DC gain

$$G(s) = rac{Y(s)}{U(s)} = rac{b_m s^m + b_{m-1} s^{n-1} + \dots + b_0}{s^n + a_{n-1} s^{n-1} + \dots + a_0}$$

- DC gain: the ratio of the output of a system to its input (presumed constant, e.g., Laplace transform = 1/s) after all transients have decayed.
- can use the Final Value Theorem to find the DC gain of a system:

DC gain of
$$G(s) = \lim_{s \to 0} sY(s) = \lim_{s \to 0} sG(s)\frac{1}{s} = \lim_{s \to 0} G(s)$$

Example: find the DC gain of G₁(s) = K and G₂(s) = k/(s+a). Try (i) solve the ODE and (ii) the Final Value Theorem

ME547: Linear Systems Modeling: Z Transform

Xu Chen

University of Washington

UW Linear Systems (X. Chen, ME547)

Z transform

1/23

Topic

1. Introduction

- 2. Common Z transform pairs
- 3. Properties and application to dynamic systems
- 4. Application to system analysis
- 5. Application
- 6. More on the final value theorem

Overview

- The Z transformation is a powerful tool to solve a wide variety of Ordinary difference Equations (OdEs)
- Analogous to Laplace transform for continuous-time signals

UW Linear Systems (X. Chen, ME547)

Z transform

Definition

Let x(k) be a real discrete-time sequence that is zero if k < 0. The (one-sided) Z transform of x(k) is

$$X(z) \triangleq \mathcal{Z}\{x(k)\} = \sum_{k=0}^{\infty} x(k)z^{-k}$$

= $x(0) + x(1)z^{-1} + x(2)z^{-2} + \dots$

where $z \in \mathbb{C}$.

- ► a linear operator: $\mathcal{Z} \{ \alpha f(k) + \beta g(k) \} = \alpha \mathcal{Z} \{ f(k) \} + \beta \mathcal{Z} \{ g(k) \}$
- ► the series 1 + x + x² + ... converges to ¹/_{1-x} for |x| < 1 [region of convergence (ROC)]</p>
- (also, recap that $\sum_{k=0}^{N} \beta^k = \frac{1-\beta^{N+1}}{1-\beta}$ if $\beta \neq 1$)

Example: Derive the Z transform of the geometric sequence $\{a^k\}_{k=0}^{\infty}$

$$\mathcal{Z}\{a^k\} = \sum_{k=0}^{\infty} a^k z^{-k} = \boxed{\frac{1}{1 - az^{-1}}} = \frac{z}{z - a}$$
for $|az^{-1}| < 1 \Leftrightarrow |z| > |a|$.

UW Linear Systems (X. Chen, ME547)

Z transform

Example: Step sequence (discrete-time unit step function)

$$1(k) = \begin{cases} 1, & \forall k = 1, 2, \dots \\ 0, & \forall k = \dots, -1, 0 \end{cases}$$
$$\mathcal{Z}\{1(k)\} = \mathcal{Z}\{a^k\}\Big|_{k=1} = \boxed{\frac{1}{1-z^{-1}}} = \frac{z}{z-1}$$

for |z| > 1

Example: Discrete-time impulse

$$\delta(k) = egin{cases} 1, & k = 0 \ 0, & ext{otherwise} \ \mathcal{Z}\{\delta(k)\} = 1 \end{cases}$$

UW Linear Systems (X. Chen, ME547)

Z transform

7 / 23

Exercise: Derive the Z transform of $cos(\omega_0 k)$

Topic

- 1. Introduction
- 2. Common Z transform pairs
- 3. Properties and application to dynamic systems
- 4. Application to system analysis
- 5. Application
- 6. More on the final value theorem

UW Linear Systems (X. Chen, ME547)	Z transform		9/23
f(k)	F(z)	ROC	
$\delta(k)$	1	All z	
$a^{k}1\left(k ight)$	$\frac{1}{1 - az^{-1}}$	z > a	
$-a^k 1(-k-1)$	$\frac{1}{1 - az^{-1}}$	z < a	
$ka^{k}1(k)$	$\frac{az^{-1}}{(1-az^{-1})^2}$	z > a	
$-ka^k1(-k-1)$	$rac{az^{-1}}{(1-az^{-1})^2}$	z < a	
$\cos(\omega_0 k)$	$\frac{1-z^{-1}\cos(\omega_0)}{1-2z^{-1}\cos(\omega_0)+z^{-2}}$	z > 1	
$\sin(\omega_0 k)$	$\frac{z^{-1}\sin(\omega_0)}{1-2z^{-1}\cos(\omega_0)+z^{-2}}$	z > 1	
$a^k \cos(\omega_0 k)$	$rac{1-az^{-1}\cos(\omega_0)}{1-2az^{-1}\cos(\omega_0)+a^2z^{-2}}$	z > a	
$a^k \sin(\omega_0 k)$	$\frac{az^{-1}\sin(\omega_0)}{1-2az^{-1}\cos(\omega_0)+a^2z^{-2}}$	z > a	

Topic

- 1. Introduction
- 2. Common Z transform pairs
- 3. Properties and application to dynamic systems
- 4. Application to system analysis
- 5. Application
- 6. More on the final value theorem

UW Linear Systems (X. Chen, ME547)

Z transform

10 / 23

Properties of Z transform: time shift

- Similar to Laplace transform, Z transform processes nice properties that provide conveniences in system analysis.
- Let $\mathcal{Z}{x(k)} = X(z)$ and $x(k) = 0 \forall k < 0$.
- one-step delay:

$$\begin{aligned} \mathcal{Z}\{x(k-1)\} &= \sum_{k=0}^{\infty} x(k-1)z^{-k} = \sum_{k=1}^{\infty} x(k-1)z^{-k} + x(-1) \\ &= \sum_{k=1}^{\infty} x(k-1)z^{-(k-1)}z^{-1} + x(-1) \\ &= z^{-1}X(z) + x(-1) = \boxed{z^{-1}X(z)} \end{aligned}$$

▶ analogously, <u>Z{x(k+1)}</u> = ∑_{k=0}[∞] x(k+1)z^{-k} = zX(z) - zx(0)
 ▶ Thus, if x(k+1) = Ax(k) + Bu(k) and the initial conditions are zero,

$$zX(z) = AX(z) + BU(z) \Rightarrow X(z) = (zI - A)^{-1}BU(z)$$

provided that (zI - A) is invertible.

Solving difference equations

Solve the difference equation

$$y(k) + 3y(k-1) + 2y(k-2) = u(k-2)$$

where y(-2) = y(-1) = 0 and u(k) = 1(k).

$$\mathcal{Z}\{y(k-2)\} = \mathcal{Z}\{y(k-1-1)\} = z^{-1}\mathcal{Z}\{y(k-1)\} + y(-2)$$
$$= z^{-1}(z^{-1}Y(z) + y(-1)) + y(-2)$$

Taking Z transforms on both sides and applying the initial conditions \Rightarrow

$$(1 + 3z^{-1} + 2z^{-2})Y(z) = z^{-2}U(z)$$

$$\Rightarrow Y(z) = \frac{1}{z^2 + 3z + 2}U(z)$$

UW Linear Systems (X. Chen, ME547)

Z transform

12/23

Solving difference equations

Solve the difference equation

$$y(k) + 3y(k-1) + 2y(k-2) = u(k-2)$$

where y(-2) = y(-1) = 0 and u(k) = 1(k). Cont'd

$$Y(z) = \frac{1}{z^2 + 3z + 2} U(z) = \frac{1}{(z+2)(z+1)} U(z)$$

 $u(k) = 1(k) \Rightarrow U(z) = 1/(1-z^{-1}) \Rightarrow$

$$Y(z) = \frac{z}{(z-1)(z+2)(z+1)} = \frac{1}{6}\frac{z}{z-1} + \frac{1}{3}\frac{z}{z+2} - \frac{1}{2}\frac{z}{z+1}$$

(careful with the partial fraction expansion) Inverse Z transform then gives

$$\gamma(k) = rac{1}{6}1(k) + rac{1}{3}(-2)^k - rac{1}{2}(-1)^k, \ k \geq 0$$

Topic

- 1. Introduction
- 2. Common Z transform pairs
- 3. Properties and application to dynamic systems
- 4. Application to system analysis
- 5. Application
- 6. More on the final value theorem

UW Linear Systems (X. Chen, ME547)

Z transform

14 / 23

From difference equation to transfer functions

We now generalize the concept in the last example. Assume that y(k) = 0 $\forall k < 0$. Applying Z transform to the ordinary difference equation

$$y(k) + a_{n-1}y(k-1) + \cdots + a_0y(k-n) = b_mu(k+m-n) + \cdots + b_0u(k-n)$$

yields

$$(z^n + a_{n-1}z^{n-1} + \cdots + a_0) Y(z) = (b_m z^m + b_{m-1}z^{m-1} + \cdots + b_0) U(z)$$

Hence

$$Y(z) = \underbrace{\frac{b_m z^m + b_{m-1} z^{m-1} \cdots + b_1 z + b_0}{z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0}}_{C_n(z)} U(z)$$

 $G_{yu}(z)$: discrete-time transfer function

Analogous to the continuous-time case, you can write down the concepts of characteristic equation, poles, and zeros of the discrete-time transfer function $G_{yu}(z)$.

Transfer functions in two domains

$$y(k) + a_{n-1}y(k-1) + \dots + a_0y(k-n) = b_mu(k+m-n) + \dots + b_0u(k-n)$$
$$\iff G_{yu}(z) = \frac{B(z)}{A(z)} = \frac{b_mz^m + b_{m-1}z^{m-1} + \dots + b_1z + b_0}{z^n + a_{n-1}z^{n-1} + \dots + a_1z + a_0}$$

v.s.

$$\frac{d^{n}y(t)}{dt^{n}} + a_{n-1}\frac{d^{n-1}y(t)}{dt^{n-1}} + \dots + a_{0}y(t) = b_{m}\frac{d^{m}u(t)}{dt^{m}} + b_{m-1}\frac{d^{m-1}u(t)}{dt^{m-1}} + \dots + b_{0}u(t)$$
$$\iff G_{yu}(s) = \frac{B(s)}{A(s)} = \frac{b_{m}s^{m} + \dots + b_{1}s + b_{0}}{s^{n} + a_{n-1}s^{n-1} + \dots + a_{1}s + a_{0}}$$

Properties	$G_{yu}(s)$	$G_{yu}(z)$
poles and zeros	roots of <i>A</i> (<i>s</i>) and <i>B</i> (<i>s</i>)	roots of $A(z)$ and $B(z)$
causality condition	$n \ge m$	$n \ge m$
DC gain / steady-state response to unit step	<i>G_{yu}</i> (0)	$G_{yu}(1)$

UW Linear Systems (X. Chen, ME547)

Z transform

Additional useful properties of Z transform

• time shifting (assuming x(k) = 0 if k < 0):

$$\mathcal{Z}\left\{x(k-n_d)\right\}=z^{-n_d}X(z)$$

- ► Z-domain scaling: $\mathcal{Z}\left\{a^{k}x(k)\right\} = X\left(a^{-1}z\right)$
- differentiation: $\mathcal{Z} \{kx(k)\} = -z \frac{dX(z)}{dz}$
- time reversal: $\mathcal{Z} \{x(-k)\} = X(z^{-1})$
- convolution: let $f(k) * g(k) \triangleq \sum_{j=0}^{k} f(k-j) g(j)$, then

$$\mathcal{Z}\left\{f(k) * g(k)\right\} = F(z) G(z)$$

- initial value theorem: $f(0) = \lim_{z \to \infty} F(z)$
- ▶ final value theorem: $\lim_{k\to\infty} f(k) = \lim_{z\to 1} (z-1) F(z)$ if f(k) exists and is finite

Topic

- 1. Introduction
- 2. Common Z transform pairs
- 3. Properties and application to dynamic systems
- 4. Application to system analysis
- 5. Application
- 6. More on the final value theorem

UW Linear Systems (X. Chen, ME547) Z transform

Mortgage payment

- image you borrow \$100,000 (e.g., for a mortgage)
- annual percent rate: APR = 4.0%
- plan to pay off in 30 years with fixed monthly payments
- interest computed monthly
- What is your monthly payment?

Mortgage payment

- ▶ borrow \$100,000 \Rightarrow initial debt y(0) = 100,000
- $APR = 4.0\% \Rightarrow MPR = \frac{4.0\%}{12} = 0.0033$
- ▶ pay off in 30 years ($N = 30 \times 12 = 360$ months) $\Rightarrow y(N) = 0$
- monthly mortgage dynamics

$$y(k+1) = \underbrace{(1 + MPR)}_{a} y(k) - \underbrace{b}_{\text{monthly payment}} 1(k)$$

$$\implies Y(z) = \frac{z}{z-a} y(0) + \frac{1}{z-a} \frac{b}{1-z^{-1}}$$

$$= \frac{1}{1-az^{-1}} y(0) + \frac{b}{1-a} \left(\frac{1}{1-az^{-1}} - \frac{1}{1-z^{-1}} \right)$$

$$\Rightarrow y(k) = a^{k} y(0) + \frac{b}{1-a} \left(a^{k} - 1 \right)$$

• need $y(N) = 0 \Rightarrow a^N y(0) = -\frac{b}{1-a} (a^N - 1) \Rightarrow b = \frac{a^N y(0)(a-1)}{a^N - 1} = $477.42.$

Z transform

UW Linear Systems (X. Chen, ME547)

Topic

- 1. Introduction
- 2. Common Z transform pairs
- 3. Properties and application to dynamic systems
- 4. Application to system analysis
- 5. Application
- 6. More on the final value theorem

 $\lim_{k\to\infty} f(k) = \lim_{z\to 1} (z-1) F(z)$ if f(k) exists and is finite.

Proof:
If
$$\lim_{k \to \infty} f(k) = f_{\infty}$$
 exists then $\sum_{k=0}^{\infty} [f(k+1) - f(k)] = f_{\infty} - f(0)$.
on one hand: $\lim_{z \to 1} \sum_{k=0}^{\infty} z^{-k} [f(k+1) - f(k)] = \sum_{k=0}^{\infty} [f(k+1) - f(k)]$
on the other hand: $\sum_{k=0}^{\infty} z^{-k} [f(k+1) - f(k)] = zF(z) - zf(0) - F(z)$
Thus $f_{\infty} - f(0) = \lim_{z \to 1} [zF(z) - zf(0) - F(z)] \Rightarrow \lim_{k \to \infty} f(k) = \lim_{z \to 1} (z-1)F(z)$

UW Linear Systems (X. Chen, ME547)

Z transform

 $\lim_{k\to\infty} f(k) = \lim_{z\to 1} (z-1) F(z)$ if f(k) exists and is finite.

Intuition:

• f(k) is the impulse response of

$$\delta(k) \longrightarrow F(z) \longrightarrow f(k)$$

namely, the step response of

$$\frac{1}{1-z^{-1}} \longrightarrow \left(1-z^{-1}\right) F(z) \longrightarrow y(k)$$

▶ if the system is stable, then the final value of f(k) is the steady-state value of the step response of (1 - z⁻¹) F(z), i.e.

$$\lim_{k \to \infty} f(k) = \lim_{z \to 1} (1 - z^{-1}) F(z) = \lim_{z \to 1} (z - 1) F(z)$$

ME547: Linear Systems Modeling: State-Space Models

Xu Chen

University of Washington

UW Linear Systems (X. Chen, ME547)

State Space Intro

1/12

Topic

1. Motivation

2. State-space description of a system

Why state space?

- Static/memoryless system: present output depends only on its present input: y(k) = f(u(k))
- Dynamic system: present output depends on past and its present input,
 - e.g., y(k) = f(u(k), u(k-1), ..., u(k-n), ...)
 - described by differential or difference equations, or have time delays
- How much information from the past is needed?

UW Linear Systems (X. Chen, ME547)

State Space Intro

The concept of states of a dynamic system

The state x(t) at time t is the information you need at time t that together with future values of the input, will let you compute future values of the output y.

loosely speaking:

- the "aggregated effect of past inputs"
- the necessary "memory" that the dynamic system keeps at each time instance

Example

- a point mass subject to a force input
- ▶ to predict the future motion of the particle, we need to know
 - current position and velocity of the particle
 - the *future* force
- \blacktriangleright \Rightarrow states: position and velocity

UW Linear Systems (X. Chen, ME547)

State Space Intro

5/12

The order of a dynamic system

- the number, n of state variables that is necessary and sufficient to uniquely describe the system.
- For a given dynamic system,
 - the choice of state variables is not unique.
 - However, its order n is fixed; i.e. you need not more than n but not less than n state variables.

Topic

1. Motivation

2. State-space description of a system

UW Linear Systems (X. Chen, ME547)

State Space Intro

7/12

States of a discrete-time system

Consider a discrete-time dynamic system:



The state of the system at any instance k_o is the minimum set of variables,

$$x_1(k_o), x_2(k_o), \cdots, x_n(k_o)$$

that fully describe the system and its response for $k \ge k_o$ to any given set of inputs.

Loosely speaking, $x_1(k_o), x_2(k_o), \dots, x_n(k_o)$ defines the system's memory.

Discrete-time state-space description



UW Linear Systems (X. Chen, ME547)

State Space Intro

Continuous-time state-space description



$$\frac{dx(t)}{dt} = f(x(t), u(t), t)$$

$$y(t) = h(x(t), u(t), t)$$

$$u(t): \text{ input}$$

$$y(t): \text{ output}$$

x(t): state

UW Linear Systems (X. Chen, ME547)

y(t) = Cx(t) + Du(t)

Example: mass-spring-damper





UW Linear Systems (X. Chen, ME547)

State Space Intro

11/12

Example: mass-spring-damper



ME547: Linear Systems Modeling: Relationship Between State-Space Models and Transfer Functions

Xu Chen

University of Washington

UW Linear Systems (X. Chen, ME547)

State Space to Transfer Functions

1/15

Topic

- 1. Transform techniques in the state space
- 2. Linear algebra recap
- 3. Example

Continuous-time LTI state-space description

$$\begin{array}{c|c} u(t) & & y(t) \\ \hline & & \\ x_1, x_2, \dots, x_n \end{array} \rightarrow$$

$$\frac{d}{dt}x(t) = Ax(t) + Bu(t)$$

$$y(t) = Cx(t) + Du(t)$$
(1)

u(t): input
 y(t): output
 x(t): state

 UW Linear Systems (X. Chen, ME547)
 State Space to Transfer Functions
 3/15

Recap: LTI input/output description

$$\begin{array}{c|c} u(t) & & y(t) \\ \hline & & \\ \hline & & \\ x_1, x_2, \dots, x_n \end{array} \rightarrow$$

Let $u(t) \in \mathbb{R}$ and $y(t) \in \mathbb{R}$, then

$$y(t) = (g \star u)(t)$$

= $\int_0^t g(t - \tau) u(\tau) d\tau$ (2)

where g(t) is the system's impulse response. Laplace domain:

$$Y(s) = G(s)U(s)$$

$$Y(s) = \mathcal{L}{y(t)}, U(s) = \mathcal{L}{u(t)}, G(s) = \mathcal{L}{g(t)}$$

From state space to transfer function

Given $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times 1}$, $C \in \mathbb{R}^{1 \times n}$, $D \in \mathbb{R}$,

$$\frac{d}{dt}x(t) = Ax(t) + Bu(t)$$

$$y(t) = Cx(t) + Du(t)$$
(3)

 $\stackrel{\mathcal{L}}{\Rightarrow}$

$$sX(s) - x(0) = AX(s) + BU(s)$$

$$Y(s) = CX(s) + DU(s)$$
(4)

When x(0) = 0, we have

$$\frac{Y(s)}{U(s)} = C(sI - A)^{-1}B + D \triangleq : G(s)$$
(5)

-the transfer function between u and y.

UW Linear Systems (X. Chen, ME547) State Space to Transfer Functions 5/15

Analogously for discrete-time systems

For $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times 1}$, $C \in \mathbb{R}^{1 \times n}$, $D \in \mathbb{R}$, taking the Z transform yields

$$x(k+1) = Ax(k) + Bu(k)$$

$$y(k) = Cx(k) + Du(k)$$
(6)

 $\stackrel{\mathcal{Z}}{\Rightarrow}$

$$zX(z) - zx(0) = AX(z) + BU(z)$$

$$Y(z) = CX(z) + DU(z)$$
(7)

 \Rightarrow Thus, when x(0) = 0, we have

$$\frac{Y(z)}{U(z)} = C(zI - A)^{-1}B + D \triangleq : G(z)$$

-the transfer function between u and y.

From state space to transfer function: Observations

$$\frac{d}{dt}x(t) = A_{n \times n}x(t) + B_{n \times 1}u(t)$$

$$y(t) = C_{1 \times n}x(t) + Du(t)$$
(8)

Dimensions:

$$G(s) = \underbrace{C}_{1 \times n} \underbrace{(sI - A)^{-1}}_{n \times n} \underbrace{B}_{n \times 1} + D$$

• Uniqueness: G(s) is unique given the state-space model

Topic

- 1. Transform techniques in the state space
- 2. Linear algebra recap
- 3. Example

Matrix inverse

$$M^{-1} = \frac{1}{\det(M)} \operatorname{Adj}(M)$$

where $\operatorname{Adj}(M) = \{\operatorname{Cofactor matrix of } M\}^{T}$
e.g.: $M = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 1 & 0 & 6 \end{bmatrix}$, $\{\operatorname{Cofactor matrix of } M\} = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix}$
where $c_{11} = \begin{vmatrix} 4 & 5 \\ 0 & 6 \end{vmatrix} = 24$, $c_{12} = -\begin{vmatrix} 0 & 5 \\ 1 & 6 \end{vmatrix} = 5$, $c_{13} = \begin{vmatrix} 0 & 4 \\ 1 & 0 \end{vmatrix} = -4$,
 $c_{21} = -\begin{vmatrix} 2 & 3 \\ 0 & 6 \end{vmatrix} = -12$, $c_{22} = \begin{vmatrix} 1 & 3 \\ 1 & 6 \end{vmatrix} = 3$, $c_{23} = -\begin{vmatrix} 1 & 2 \\ 1 & 0 \end{vmatrix} = 2$,
 $c_{31} = \begin{vmatrix} 2 & 3 \\ 4 & 5 \end{vmatrix} = -2$, $c_{32} = -\begin{vmatrix} 1 & 3 \\ 0 & 5 \end{vmatrix} = -5$, $c_{33} = \begin{vmatrix} 1 & 2 \\ 0 & 4 \end{vmatrix} = 4$

UW Linear Systems (X. Chen, ME547) State Space to Transfer Functions

9/15

Topic

1. Transform techniques in the state space

2. Linear algebra recap

3. Example

Mass-spring-damper





Mass-spring-damper

 \Rightarrow

$$\frac{d}{dt} \begin{bmatrix} p(t) \\ v(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{b}{m} \end{bmatrix} \begin{bmatrix} p(t) \\ v(t) \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{1}{m} \end{bmatrix} u(t) \quad (10)$$

$$y(t) = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} p(t) \\ v(t) \end{bmatrix}$$

$$\frac{G(s) = C(sI - A)^{-1}B + D}{G(s) = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} s & 0 \\ 0 & s \end{bmatrix} - \begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{b}{m} \end{bmatrix} \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ \frac{1}{m} \end{bmatrix}$$

Mass-spring-damper

$$\begin{bmatrix} \begin{bmatrix} s & 0 \\ 0 & s \end{bmatrix} - \begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{b}{m} \end{bmatrix} \end{bmatrix}^{-1} = \begin{bmatrix} s & -1 \\ \frac{k}{m} & s + \frac{b}{m} \end{bmatrix}^{-1}$$

$$= \frac{1}{s^2 + \frac{b}{m}s + \frac{k}{m}} \begin{bmatrix} s + \frac{b}{m} & 1 \\ -\frac{k}{m} & s \end{bmatrix}$$
(11)

UW Linear Systems (X. Chen, ME547) State Space to Transfer Functions

13 / 15

Mass-spring-damper

Putting the inverse in yields

$$G(s) = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} s & -1 \\ \frac{k}{m} & s + \frac{b}{m} \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ \frac{1}{m} \end{bmatrix}$$
$$= \frac{\begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} s + \frac{b}{m} & 1 \\ -\frac{k}{m} & s \end{bmatrix} \begin{bmatrix} 0 \\ \frac{1}{m} \end{bmatrix}}{s^2 + \frac{b}{m}s + \frac{k}{m}}$$
(12)

namely

$$G(s) = rac{rac{1}{m}}{s^2 + rac{b}{m}s + rac{k}{m}}$$

Exercise

Given the following state-space system parameters: $A = \begin{bmatrix} 0 & -6 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$, $B = \begin{bmatrix} -6 & 0 & -3 \\ -2 & 1 & 0 \\ 0 & 2 & 3 \end{bmatrix}$, $C = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}$, obtain the transfer function G(s).

UW Linear Systems (X. Chen, ME547)

State Space to Transfer Functions

ME 547: Linear Systems State-Space Canonical Forms

Xu Chen

University of Washington

UW Linear Systems (X. Chen, ME547)

State-space canonical forms

1/31

Goal

The realization problem:

existence and uniqueness: the same system can have infinite amount of state-space representations: e.g.

$$\begin{cases} \dot{x} = Ax + Bu \\ y = Cx \end{cases} \qquad \begin{cases} \dot{x} = Ax + \frac{1}{2}Bu \\ y = 2Cx \end{cases}$$

$$\left(\text{packed representation: } \Sigma_1 = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array} \right], \Sigma_2 = \left[\begin{array}{c|c} A & \frac{B}{2} \\ \hline 2C & D \end{array} \right] \right)$$

- canonical realizations exist
- relationship between different realizations?
- example problem for this set of notes:

$$G(s) = \frac{b_2 s^2 + b_1 s + b_0}{s^3 + a_2 s^2 + a_1 s + a_0}.$$
 (1)

Outline

- 1. CT controllable canonical form
- 2. CT observable canonical form
- 3. CT diagonal and Jordan canonical forms
- 4. Modified canonical form
- 5. DT state-space canonical forms
- 6. Similar realizations

UW Linear Systems (X. Chen, ME547)

State-space canonical forms

3/31

Controllable canonical form (ccf)

Choose
$$x_1$$
 such that

$$u \longrightarrow \left| \frac{1}{s^{3} + a_{2}s^{2} + a_{1}s + a_{0}} \right|^{x_{1}} \longrightarrow \left| b_{2}s^{2} + b_{1}s + b_{0} \right| \longrightarrow y$$
(2)
$$X_{1}(s) = \frac{U(s)}{s^{3} + a_{2}s^{2} + a_{1}s + a_{0}} \Rightarrow \ddot{x}_{1} + a_{2}\ddot{x}_{1} + a_{1}\dot{x}_{1} + a_{0}x_{1} = u$$

Let
$$x_2 = \dot{x}_1, \ x_3 = \dot{x}_2$$
. Then $\dot{x}_3 = -a_2x_3 - a_1x_2 - a_0x_1 + u$.

$$Y(s) = (b_2 s^2 + b_1 s + b_0) X_1(s) \Rightarrow y = b_2 \underbrace{\ddot{x}_1}_{x_3} + b_1 \underbrace{\dot{x}_1}_{x_2} + b_0 x_1$$

Putting in matrix form yields

$$\frac{d}{dt} \begin{bmatrix} x_{1}(t) \\ x_{2}(t) \\ x_{3}(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -a_{0} & -a_{1} & -a_{2} \end{bmatrix} \begin{bmatrix} x_{1}(t) \\ x_{2}(t) \\ x_{3}(t) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} u(t)$$
(3)
$$y(t) = \begin{bmatrix} b_{0} & b_{1} & b_{2} \end{bmatrix} \begin{bmatrix} x_{1}(t) \\ x_{2}(t) \\ x_{3}(t) \end{bmatrix}$$

Block diagram realization of controllable canonical forms



General ccf

For a single-input single-output transfer function

$$G(s) = \frac{b_{n-1}s^{n-1} + \cdots + b_1s + b_0}{s^n + a_{n-1}s^{n-1} + \cdots + a_1s + a_0} + d,$$

we can verify that

$$\Sigma_{c} = \begin{bmatrix} A_{c} & B_{c} \\ \hline C_{c} & D_{c} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & \ddots & \ddots & \vdots & 0 \\ \vdots & \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & 0 & \dots & 0 & 1 & 0 \\ \hline -a_{0} & -a_{1} & \dots & -a_{n-2} & -a_{n-1} & 1 \\ \hline b_{0} & b_{1} & \dots & b_{n-2} & b_{n-1} & d \end{bmatrix}$$
(5)

realizes G(s). This realization is called the *controllable canonical form*.

- 1. CT controllable canonical form
- 2. CT observable canonical form
- 3. CT diagonal and Jordan canonical forms
- 4. Modified canonical form
- 5. DT state-space canonical forms
- 6. Similar realizations

UW Linear Systems (X. Chen, ME547)

State-space canonical forms

7/31

Observable canonical form (ocf)

$$Y(s) = \frac{b_2 s^2 + b_1 s + b_0}{s^3 + a_2 s^2 + a_1 s + a_0} U(s)$$

$$\Rightarrow Y(s) = -\frac{a_2}{s} Y(s) - \frac{a_1}{s^2} Y(s) - \frac{a_0}{s^3} Y(s) + \frac{b_2}{s} U(s) + \frac{b_1}{s^2} U(s) + \frac{b_0}{s^3} U(s).$$

In a block diagram, the above looks like



Observable canonical form



Here, the states are connected by

$$\begin{array}{ll} Y(s) = X_1(s) & y(t) = x_1(t) \\ sX_1(s) = -a_2X_1(s) + X_2(s) + b_2U(s) & \dot{x}_1(t) = -a_2x_1(t) + x_2(t) + b_2u(t) \\ sX_2(s) = -a_1X_1(s) + X_3(s) + b_1U(s) & \Rightarrow & \dot{x}_2(t) = -a_1x_1(t) + x_3(t) + b_1u(t) \\ sX_3(s) = -a_0X_1(s) + b_0U(s) & \dot{x}_3(t) = -a_0x_1(t) + b_0u(t) \end{array}$$

UW Linear Systems (X. Chen, ME547)

State-space canonical forms

Observable canonical form

$$\begin{cases} \dot{x}_{1}(t) = -a_{2}x_{1}(t) + x_{2}(t) + b_{2}u(t) \\ \dot{x}_{2}(t) = -a_{1}x_{1}(t) + x_{3}(t) + b_{1}u(t) \\ \dot{x}_{3}(t) = -a_{0}x_{1}(t) + b_{0}u(t) \\ y(t) = x_{1}(t) \end{cases}$$
$$\Rightarrow \dot{x}(t) = \underbrace{\begin{bmatrix} -a_{2} & 1 & 0 \\ -a_{1} & 0 & 1 \\ -a_{0} & 0 & 0 \end{bmatrix}}_{A_{o}} x(t) + \underbrace{\begin{bmatrix} b_{2} \\ b_{1} \\ b_{0} \end{bmatrix}}_{B_{o}} u(t) \\ y(t) = \underbrace{\begin{bmatrix} 1 & 0 & 0 \end{bmatrix}}_{C_{o}} x(t)$$

The above is called the *observable canonical form* realization of G(s).

Exercise

Verify that
$$C_o(sI - A_o)^{-1}B_o = G(s)$$

General ocf

In the general case, the observable canonical form of the transfer function

$$G(s) = \frac{b_{n-1}s^{n-1} + \dots + b_1s + b_0}{s^n + a_{n-1}s^{n-1} + \dots + a_1s + a_0} + d$$

is

$$\Sigma_{o} = \begin{bmatrix} A_{o} & B_{o} \\ \hline C_{o} & D_{o} \end{bmatrix} = \begin{bmatrix} -a_{n-1} & 1 & 0 & \dots & 0 & b_{n-1} \\ -a_{n-2} & 0 & \ddots & \ddots & \vdots & b_{n-2} \\ \vdots & \vdots & \ddots & \ddots & 0 & \vdots \\ -a_{1} & 0 & \cdots & 0 & 1 & b_{1} \\ \hline -a_{0} & 0 & \cdots & 0 & 0 & b_{0} \\ \hline 1 & 0 & \cdots & 0 & 0 & d \end{bmatrix}.$$
(6)

UW Linear Systems (X. Chen, ME547)

State-space canonical forms

11/31

1. CT controllable canonical form

- 2. CT observable canonical form
- 3. CT diagonal and Jordan canonical forms
- 4. Modified canonical form
- 5. DT state-space canonical forms
- 6. Similar realizations
Diagonal form

When

$$G(s) = \frac{B(s)}{A(s)} = \frac{b_2 s^2 + b_1 s + b_0}{s^3 + a_2 s^2 + a_1 s + a_0}$$

and the poles $p_1
eq p_2
eq p_3$, partial fractional expansion yields

$$G(s) = \frac{k_1}{s-p_1} + \frac{k_2}{s-p_2} + \frac{k_3}{s-p_3}, \ k_i = \lim_{p \to p_i} (s-p_i) \frac{B(s)}{A(s)},$$

namely



UW Linear Systems (X. Chen, ME547)

State-space canonical forms

13/31

Diagonal form



The state-space realization of the above is

$$A = \begin{bmatrix} p_1 & 0 & 0 \\ 0 & p_2 & 0 \\ 0 & 0 & p_3 \end{bmatrix}, B = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, C = \begin{bmatrix} k_1 & k_2 & k_3 \end{bmatrix}, D = 0.$$

Jordan form

If poles repeat, say,

$$G(s)=rac{b_2s^2+b_1s+b_0}{s^3+a_2s^2+a_1s+a_0}=rac{b_2s^2+b_1s+b_0}{(s-p_1)(s-p_m)^2}, \ p_1
eq p_m\in\mathbb{R},$$

then partial fraction expansion gives

$$G(s) = \frac{k_1}{s - p_1} + \frac{k_2}{(s - p_m)^2} + \frac{k_3}{s - p_m} \quad w/ \begin{cases} k_1 &= \lim_{s \to p_1} G(s)(s - p_1) \\ k_2 &= \lim_{s \to p_m} G(s)(s - p_m)^2 \\ k_3 &= \lim_{s \to p_m} \frac{d}{ds} \left\{ G(s)(s - p_m)^2 \right\} \end{cases}$$

UW Linear Systems (X. Chen, ME547)

State-space canonical forms

15 / 31

Jordan form

$$G(s) = rac{k_1}{s - p_1} + rac{k_2}{(s - p_m)^2} + rac{k_3}{s - p_m}$$

has the block diagram realization:



Jordan form



The state-space realization of the above, called the Jordan canonical form, is

	p_1	0	0 -		1			
A =	0 0	р _т 0	1 p _m _	, <i>B</i> =	0 1	$, C = \begin{bmatrix} k_1 \end{bmatrix}$	<i>k</i> ₂	k_3], $D = 0.$

UW Linear Systems (X. Chen, ME547)

State-space canonical forms

17/31

- 1. CT controllable canonical form
- 2. CT observable canonical form
- 3. CT diagonal and Jordan canonical forms
- 4. Modified canonical form
- 5. DT state-space canonical forms
- 6. Similar realizations

Modified canonical form

If the system has complex poles, say,

$$G(s) = \frac{b_2 s^2 + b_1 s + b_0}{s^3 + a_2 s^2 + a_1 s + a_0} = \frac{k_1}{s - p_1} + \frac{\alpha s + \beta}{(s - \sigma)^2 + \omega^2}$$

then we have



where $k_2 = (\beta + \alpha \sigma)/\omega$ and $k_3 = \alpha$. UW Linear Systems (X. Chen, ME547) State-space canonical forms

19/31

Modified canonical form



 \Rightarrow modified Jordan form:

$$A = \begin{bmatrix} p_1 & 0 & 0 \\ 0 & \sigma & \omega \\ 0 & -\omega & \sigma \end{bmatrix}, B = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, C = \begin{bmatrix} k_1 & k_2 & k_3 \end{bmatrix}, D = 0.$$

- 1. CT controllable canonical form
- 2. CT observable canonical form
- 3. CT diagonal and Jordan canonical forms
- 4. Modified canonical form
- 5. DT state-space canonical forms
- 6. Similar realizations

UW Linear Systems (X. Chen, ME547) State-s

State-space canonical forms

21/31

DT state-space canonical forms

The procedures for finding state space realizations in discrete time is similar to the continuous time cases. The only difference is that we use

$$\mathcal{Z}\left\{x(k-n)\right\}=z^{-n}X(z),$$

instead of

$$\mathcal{L}\left\{\frac{d^n}{dt^n}x(t)\right\}=s^nX(s),$$

assuming zero state initial conditions.

Fundamental relationships:

$$x(k) \longrightarrow z^{-1} \longrightarrow x(k-1)$$

$$X(z) \longrightarrow z^{-1} \longrightarrow z^{-1}X(z)$$

$$x(k+n) \longrightarrow z^{-1} \longrightarrow x(k+n-1)$$

DT ccf

$$G(z) = \frac{b_2 z^2 + b_1 z + b_0}{z^3 + a_2 z^2 + a_1 z + a_0}$$

Same transfer-function structure \Rightarrow same A, B, C, D matrices of the canonical forms as those in continuous-time cases

Controllable canonical form:

$$\begin{bmatrix} x_{1}(k+1) \\ x_{2}(k+1) \\ x_{3}(k+1) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -a_{0} & -a_{1} & -a_{2} \end{bmatrix} \begin{bmatrix} x_{1}(k) \\ x_{2}(k) \\ x_{3}(k) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} u(k)$$
$$y(k) = \begin{bmatrix} b_{0} & b_{1} & b_{2} \end{bmatrix} \begin{bmatrix} x_{1}(k) \\ x_{2}(k) \\ x_{3}(k) \end{bmatrix}$$

UW Linear Systems (X. Chen, ME547) State-space canonical forms

23/31

DT ocf

$$G(z) = \frac{b_2 z^2 + b_1 z + b_0}{z^3 + a_2 z^2 + a_1 z + a_0}$$

Observable canonical form:

$$\begin{bmatrix} x_{1} (k+1) \\ x_{2} (k+1) \\ x_{3} (k+1) \end{bmatrix} = \begin{bmatrix} -a_{2} & 1 & 0 \\ -a_{1} & 0 & 1 \\ -a_{0} & 0 & 0 \end{bmatrix} \begin{bmatrix} x_{1} (k) \\ x_{2} (k) \\ x_{3} (k) \end{bmatrix} + \begin{bmatrix} b_{2} \\ b_{1} \\ b_{0} \end{bmatrix} u (k)$$
$$y (k) = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_{1} (k) \\ x_{2} (k) \\ x_{3} (k) \end{bmatrix}$$

DT diagonal form

$$G(z) = \frac{b_2 z^2 + b_1 z + b_0}{z^3 + a_2 z^2 + a_1 z + a_0}$$

Diagonal form (distinct poles):

$$G(z) = \frac{k_1}{z - p_1} + \frac{k_2}{z - p_2} + \frac{k_3}{z - p_3}$$
$$\begin{bmatrix} x_1 (k+1) \\ x_2 (k+1) \\ x_3 (k+1) \end{bmatrix} = \begin{bmatrix} p_1 & 0 & 0 \\ 0 & p_2 & 0 \\ 0 & 0 & p_3 \end{bmatrix} \begin{bmatrix} x_1 (k) \\ x_2 (k) \\ x_3 (k) \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} u (k)$$
$$y (k) = \begin{bmatrix} k_1 & k_2 & k_3 \end{bmatrix} \begin{bmatrix} x_1 (k) \\ x_2 (k) \\ x_3 (k) \end{bmatrix}$$

UW Linear Systems (X. Chen, ME547)

State-space canonical forms

25 / 31

DT Jordan form 1

$$G(z) = \frac{b_2 z^2 + b_1 z + b_0}{z^3 + a_2 z^2 + a_1 z + a_0}$$

► Jordan form (2 repeated poles):

$$G(z) = \frac{k_1}{z - p_1} + \frac{k_2}{(z - p_m)^2} + \frac{k_3}{z - p_m}$$

$$\begin{bmatrix} x_{1} (k+1) \\ x_{2} (k+1) \\ x_{3} (k+1) \end{bmatrix} = \begin{bmatrix} p_{1} & 0 & 0 \\ 0 & p_{m} & 1 \\ 0 & 0 & p_{m} \end{bmatrix} \begin{bmatrix} x_{1} (k) \\ x_{2} (k) \\ x_{3} (k) \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} u (k)$$
$$y (k) = \begin{bmatrix} k_{1} & k_{2} & k_{3} \end{bmatrix} \begin{bmatrix} x_{1} (k) \\ x_{2} (k) \\ x_{3} (k) \end{bmatrix}$$

DT Jordan form 2

$$G(z) = \frac{b_2 z^2 + b_1 z + b_0}{z^3 + a_2 z^2 + a_1 z + a_0}$$

► Jordan form (2 complex poles):

$$G(s) = \frac{k_1}{z - p_1} + \frac{\alpha z + \beta}{(z - \sigma)^2 + \omega^2}$$

$$\begin{bmatrix} x_{1} (k+1) \\ x_{2} (k+1) \\ x_{3} (k+1) \end{bmatrix} = \begin{bmatrix} p_{1} & 0 & 0 \\ 0 & \sigma & \omega \\ 0 & -\omega & \sigma \end{bmatrix} \begin{bmatrix} x_{1} (k) \\ x_{2} (k) \\ x_{3} (k) \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} u (k)$$
$$y (k) = \begin{bmatrix} k_{1} & k_{2} & k_{3} \end{bmatrix} \begin{bmatrix} x_{1} (k) \\ x_{2} (k) \\ x_{3} (k) \end{bmatrix}$$

where $k_2 = (\beta + \alpha \sigma)/\omega$, $k_3 = \alpha$.

UW Linear Systems (X. Chen, ME547) State-space canonical forms

27 / 31

Exercise

obtain the controllable canonical form:

•
$$G(s) = \frac{z^{-1}-z^{-3}}{1+2z^{-1}+z^{-2}}$$

- 1. CT controllable canonical form
- 2. CT observable canonical form
- 3. CT diagonal and Jordan canonical forms
- 4. Modified canonical form
- 5. DT state-space canonical forms

6. Similar realizations

UW Linear Systems (X. Chen, ME547) State-space canonical forms

Relation between different realizations

Similar realizations

Given one realization Σ of a transfer function G(s) and a nonsingular $T \in \mathbb{R}^{n \times n}$, we can define *new* states:

$$Tx^* = x$$

Then

$$\dot{x}(t) = Ax(t) + Bu(t) \Rightarrow rac{d}{dt} (Tx^*(t)) = ATx^*(t) + Bu(t),$$

 $\Rightarrow \Sigma^*: \begin{cases} \dot{x}^*(t) = T^{-1}ATx^*(t) + T^{-1}Bu(t) \\ y(t) = CTx^*(t) + Du(t) \end{cases}$

namely

$$\Sigma^* = \begin{bmatrix} T^{-1}AT \mid T^{-1}B \\ \hline CT \mid D \end{bmatrix}$$

also realizes G(s) and is said to be *similar* to Σ .

29/31

Relation between different realizations Similar realizations

Exercise (Another observable canonical form.)

Verify that the following realize the same system

$$\Sigma = egin{bmatrix} -a_2 & 1 & 0 & b_2 \ -a_1 & 0 & 1 & b_1 \ -a_0 & 0 & 0 & b_0 \ \hline 1 & 0 & 0 & d \ \end{bmatrix}, \ \Sigma^* = egin{bmatrix} 0 & 0 & -a_0 & b_0 \ 1 & 0 & -a_1 & b_1 \ 0 & 1 & -a_2 & b_2 \ \hline 0 & 0 & 1 & d \ \end{bmatrix}$$

UW Linear Systems (X. Chen, ME547)

State-space canonical forms

31/31

1 From Transfer Function to State Space: State-Space Canonical Forms

It is straightforward to derive the *unique* transfer function corresponding to a state-space model. The inverse problem, i.e., building internal descriptions from transfer functions, is less trivial and is the subject of *realization* theory.

A single transfer function has infinite amount of state-space representations. Consider, for example, the two models

which share the same transfer function $C(sI - A)^{-1}B$.

We start with the most common realizations: controller canonical form, observable canonical form, and Jordan form, using the following unit problem:

$$G(s) = \frac{b_2 s^2 + b_1 s + b_0}{s^3 + a_2 s^2 + a_1 s + a_0}.$$
(1)

1.1 Controllable Canonical Form.

Consider first:

$$Y(s) = \frac{1}{s^3 + a_2 s^2 + a_1 s + a_0} U(s).$$
⁽²⁾

Similar to choosing position and velocity in the spring-mass-damper example, we can choose

$$x_1 = y, \ x_2 = \dot{x}_1 = \dot{y}, \ x_3 = \dot{x}_2 = \ddot{y},$$
(3)

which gives

$$\frac{d}{dt} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -a_0 & -a_1 & -a_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} u$$

$$y = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$
(4)

For the general case in (1), i.e., $\ddot{y} + a_2\ddot{y} + a_1\dot{y} + a_0y = b_2\ddot{u} + b_1\dot{u} + b_0u$, there are terms with respect to the derivative of the input. Choosing simply (3) does not generate a proper state equation. However, we can decompose (1) as

$$u \longrightarrow \frac{1}{s^3 + a_2 s^2 + a_1 s + a_0} \longrightarrow b_2 s^2 + b_1 s + b_0 \longrightarrow y$$

$$\tag{5}$$

The first part of the connection

$$u \longrightarrow \frac{1}{s^3 + a_2 s^2 + a_1 s + a_0} \longrightarrow \tilde{y} \tag{6}$$

looks exactly like what we had in (2). Denote the output here as \tilde{y} . Then we have

$$\frac{d}{dt} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -a_0 & -a_1 & -a_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} u,$$

where

$$x_1 = \tilde{y}, \ x_2 = \dot{x}_1, \ x_3 = \dot{x}_2.$$
 (7)

Introducing the states in (7) also addresses the problem of the rising differentiations in u. Notice now, that the second part of (5) is nothing but

$$x_1 \longrightarrow b_2 s^2 + b_1 s + b_0 \longrightarrow y$$

 So

$$y = b_2 \ddot{x}_1 + b_1 \dot{x}_1 + b_0 x_1 = b_2 x_3 + b_1 x_2 + b_0 x_1 = \begin{bmatrix} b_0 & b_1 & b_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

The above procedure constructs the *controllable canonical form* of the third-order transfer function (1):

$$\frac{d}{dt} \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -a_0 & -a_1 & -a_2 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} u(t)$$

$$y(t) = \begin{bmatrix} b_0 & b_1 & b_2 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix}$$
(8)

In a block diagram, the state-space system looks like



Example 1. Obtain the controllable canonical forms of the following systems

• $G(s) = \frac{s^2 + 1}{s^3 + 2s + 10}$

- Comparing the transfer function with the general form yields $A = \begin{bmatrix} 1 \\ 1 \\ -10 \\ -2 \end{bmatrix}$, $B = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$, $C = \begin{bmatrix} 1 \\ -10 \\ -2 \end{bmatrix}$

- $\begin{bmatrix} 1 & 0 & 1 \end{bmatrix}$
- $G(s) = \frac{b_0 s^2 + b_1 s + b_2}{s^3 + a_0 s^2 + a_1 s + a_2}$

- Notice the difference in the coefficients. We have
$$A = \begin{bmatrix} 1 \\ -a_2 & -a_1 & a_0 \end{bmatrix}$$
, $B = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$, $C = \begin{bmatrix} b_2 & b_1 & b_0 \end{bmatrix}$

General Case.

For a single-input single-output transfer function

$$G(s) = \frac{b_{n-1}s^{n-1} + \dots + b_1s + b_0}{s^n + a_{n-1}s^{n-1} + \dots + a_1s + a_0} + d$$

we can verify that

$$\Sigma_{c} = \begin{bmatrix} A_{c} & B_{c} \\ \hline C_{c} & D_{c} \end{bmatrix} = \begin{bmatrix} 0 & 1 & \cdots & 0 & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 1 & 0 \\ \hline -a_{0} & -a_{1} & \cdots & -a_{n-2} & -a_{n-1} & 1 \\ \hline b_{0} & b_{1} & \cdots & b_{n-2} & b_{n-1} & d \end{bmatrix}$$
(9)

realizes G(s). This realization is called the *controllable canonical form*.

1.2 Observable Canonical Form.

Consider again

$$Y(s) = G(s)U(s) = \frac{b_2s^2 + b_1s + b_0}{s^3 + a_2s^2 + a_1s + a_0}U(s).$$

Expanding and dividing by s^3 yield

$$\left(1 + a_2\frac{1}{s} + a_1\frac{1}{s^2} + a_0\frac{1}{s^3}\right)Y(s) = \left(b_2\frac{1}{s} + b_1\frac{1}{s^2} + b_0\frac{1}{s^3}\right)U(s)$$

and therefore

$$Y(s) = -a_2 \frac{1}{s} Y(s) - a_1 \frac{1}{s^2} Y(s) - a_0 \frac{1}{s^3} Y(s) + b_2 \frac{1}{s} U(s) + b_1 \frac{1}{s^2} U(s) + b_0 \frac{1}{s^3} U(s).$$

In a block diagram, the above looks like



or more specifically,



Here, the states are connected by

$$\begin{split} Y(s) &= X_1(s) & y(t) = x_1(t) \\ sX_1(s) &= -a_2 X_1(s) + X_2(s) + b_2 U(s) & \dot{x}_1(t) = -a_2 x_1(t) + x_2(t) + b_2 u(t) \\ sX_2(s) &= -a_1 X_1(s) + X_3(s) + b_1 U(s) & \Rightarrow & \dot{x}_2(t) = -a_1 x_1(t) + x_3(t) + b_1 u(t) \\ sX_3(s) &= -a_0 X_1(s) + b_0 U(s) & \dot{x}_3(t) = -a_0 x_1(t) + b_0 u(t) \end{split}$$

or in matrix form:

$$\dot{x}(t) = \underbrace{\begin{bmatrix} -a_2 & 1 & 0 \\ -a_1 & 0 & 1 \\ -a_0 & 0 & 0 \end{bmatrix}}_{A_o} x(t) + \underbrace{\begin{bmatrix} b_2 \\ b_1 \\ b_0 \end{bmatrix}}_{B_o} u(t)$$
(10)
$$y(t) = \underbrace{\begin{bmatrix} 1 & 0 & 0 \end{bmatrix}}_{C_o} x(t)$$

The above is called the *observable canonical form* realization of G(s).

Exercise 1. Verify that $C_o(sI - A_o)^{-1}B_o = G(s)$.

General Case.

In the general case, the *observable canonical form* of the transfer function

$$G(s) = \frac{b_{n-1}s^{n-1} + \dots + b_1s + b_0}{s^n + a_{n-1}s^{n-1} + \dots + a_1s + a_0} + d$$

is

$$\Sigma_{o} = \begin{bmatrix} A_{o} & B_{o} \\ \hline C_{o} & D_{o} \end{bmatrix} = \begin{bmatrix} -a_{n-1} & 1 & \cdots & 0 & 0 & b_{n-1} \\ -a_{n-2} & 0 & \cdots & 0 & 0 & b_{n-2} 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\ -a_{1} & 0 & \cdots & 0 & 1 & b_{1} \\ \hline -a_{0} & & \cdots & 0 & b_{0} \\ \hline 1 & & \cdots & & d \end{bmatrix}.$$
(11)

Exercise 2. Obtain the controllable and observable canonical forms of

$$G(s) = \frac{k_1}{s - p_1}$$

1.3 Diagonal and Jordan canonical forms.

1.3.1 Diagonal form.

When

$$G(s) = \frac{B(s)}{A(s)} = \frac{b_2 s^2 + b_1 s + b_0}{s^3 + a_2 s^2 + a_1 s + a_0}$$

and the poles of the transfer function $p_1 \neq p_2 \neq p_3$, we can write, using partial fractional expansion,

$$G(s) = \frac{k_1}{s - p_1} + \frac{k_2}{s - p_2} + \frac{k_3}{s - p_3}, \ k_i = \lim_{p \to p_i} (s - p_i) \frac{B(s)}{A(s)},$$

namely



The state-space realization of the above is

$$A = \begin{bmatrix} p_1 & 0 & 0 \\ 0 & p_2 & 0 \\ 0 & 0 & p_3 \end{bmatrix}, B = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, C = \begin{bmatrix} k_1 & k_2 & k_3 \end{bmatrix}, D = 0.$$

1.3.2 Jordan form.

If poles repeat, say,

$$G(s) = \frac{b_2 s^2 + b_1 s + b_0}{s^3 + a_2 s^2 + a_1 s + a_0} = \frac{b_2 s^2 + b_1 s + b_0}{(s - p_1)(s - p_m)^2}, \ p_1 \neq p_m \in \mathbb{R},$$

then partial fraction expansion gives

$$G(s) = \frac{k_1}{s - p_1} + \frac{k_2}{(s - p_m)^2} + \frac{k_3}{s - p_m},$$

where

$$k_1 = \lim_{s \to p_1} G(s)(s - p_1)$$

$$k_2 = \lim_{s \to p_m} G(s)(s - p_m)^2$$

$$k_3 = \lim_{s \to p_m} \frac{d}{ds} \left\{ G(s)(s - p_m)^2 \right\}$$

In state space, we have



The state-space realization of the above, called the Jordan canonical form,¹ is

$$A = \begin{bmatrix} p_1 & 0 & 0 \\ 0 & p_m & 1 \\ 0 & 0 & p_m \end{bmatrix}, B = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, C = \begin{bmatrix} k_1 & k_2 & k_3 \end{bmatrix}, D = 0.$$

1.4 Modified canonical form.

If the system has complex poles, say,

$$G(s) = \frac{b_2 s^2 + b_1 s + b_0}{s^3 + a_2 s^2 + a_1 s + a_0} = \frac{b_2 s^2 + b_1 s + b_0}{(s - p_1) \left[(s - \sigma)^2 + \omega^2\right]},$$

then partial fraction expansion gives

$$G\left(s\right) = \frac{k_{1}}{s - p_{1}} + \frac{\alpha s + \beta}{\left(s - \sigma\right)^{2} + \omega^{2}},$$

which has the graphical representation as below:



Here $k_2 = (\beta + \alpha \sigma)/\omega$ and $k_3 = \alpha$.

You should be able to check that the block diagram matches with the transfer function realization. The above can be realized by the modified Jordan form in state space:

$$A = \begin{bmatrix} p_1 & 0 & 0\\ 0 & \sigma & \omega\\ 0 & -\omega & \sigma \end{bmatrix}, B = \begin{bmatrix} 1\\ 0\\ 1 \end{bmatrix}, C = \begin{bmatrix} k_1 & k_2 & k_3 \end{bmatrix}, D = 0$$

1.5 Discrete-Time Transfer Functions and Their State-Space Canonical Forms

The procedures for finding state space realizations in discrete time is similar to the continuous time cases. The only difference is that we use

$$\mathcal{Z}\left\{x(k-n)\right\} = z^{-n}X(z)$$

instead of

$$\mathcal{L}\left\{\frac{d^n}{dt^n}x(t)\right\} = s^n X(s),$$

assuming zero state initial conditions.

We have the fundamental relationships:

$$x\left(k\right) \longrightarrow \boxed{z^{-1}} \longrightarrow x\left(k-1\right)$$

¹The A matrix is called a Jordan matrix.

$$\begin{array}{c} X\left(z\right) & & \overbrace{z^{-1}} \longrightarrow z^{-1}X\left(z\right) \\ x\left(k+n\right) & & \overbrace{z^{-1}} \longrightarrow x\left(k+n-1\right) \end{array}$$

The discrete-time state-space description of a general transfer function G(z) is

$$x (k + 1) = Ax (k) + Bu (k)$$
$$y (k) = Cx (k) + Du (k)$$

and satisfies $G(z) = C(zI - A)^{-1}B + D$.

Take again a third-order system as the example:

$$G\left(z\right) = \frac{b_{2}z^{2} + b_{1}z + b_{0}}{z^{3} + a_{2}z^{2} + a_{1}z + a_{0}} = \frac{b_{2}z^{-1} + b_{1}z^{-2} + b_{0}z^{-3}}{1 + a_{2}z^{-1} + a_{1}z^{-2} + a_{0}z^{-3}}.$$

The A, B, C, D matrices of the canonical forms are exactly the same as those in continuous-time cases.

Controllable canonical form:

$$\begin{bmatrix} x_1 (k+1) \\ x_2 (k+1) \\ x_3 (k+1) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -a_0 & -a_1 & -a_2 \end{bmatrix} \begin{bmatrix} x_1 (k) \\ x_2 (k) \\ x_3 (k) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} u (k)$$
$$y (k) = \begin{bmatrix} b_0 & b_1 & b_2 \end{bmatrix} \begin{bmatrix} x_1 (k) \\ x_2 (k) \\ x_3 (k) \end{bmatrix}$$

Observable canonical form:

$$\begin{bmatrix} x_1 (k+1) \\ x_2 (k+1) \\ x_3 (k+1) \end{bmatrix} = \begin{bmatrix} -a_2 & 1 & 0 \\ -a_1 & 0 & 1 \\ -a_0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 (k) \\ x_2 (k) \\ x_3 (k) \end{bmatrix} + \begin{bmatrix} b_2 \\ b_1 \\ b_0 \end{bmatrix} u (k)$$
$$y (k) = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 (k) \\ x_2 (k) \\ x_3 (k) \end{bmatrix}$$

Diagonal form (distinct poles):

$$G(z) = \frac{k_1}{z - p_1} + \frac{k_2}{z - p_2} + \frac{k_3}{z - p_3}$$

$$\begin{bmatrix} x_1 (k+1) \\ x_2 (k+1) \\ x_3 (k+1) \end{bmatrix} = \begin{bmatrix} p_1 & 0 & 0 \\ 0 & p_2 & 0 \\ 0 & 0 & p_3 \end{bmatrix} \begin{bmatrix} x_1 (k) \\ x_2 (k) \\ x_3 (k) \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} u (k)$$

$$y (k) = \begin{bmatrix} k_1 & k_2 & k_3 \end{bmatrix} \begin{bmatrix} x_1 (k) \\ x_2 (k) \\ x_3 (k) \end{bmatrix}$$

Jordan form (2 repeated poles):

$$G(z) = \frac{k_1}{z - p_1} + \frac{k_2}{(z - p_m)^2} + \frac{k_3}{z - p_m}$$

$$\begin{bmatrix} x_1 (k+1) \\ x_2 (k+1) \\ x_3 (k+1) \end{bmatrix} = \begin{bmatrix} p_1 & 0 & 0 \\ 0 & p_m & 1 \\ 0 & 0 & p_m \end{bmatrix} \begin{bmatrix} x_1 (k) \\ x_2 (k) \\ x_3 (k) \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} u (k)$$

$$y (k) = \begin{bmatrix} k_1 & k_2 & k_3 \end{bmatrix} \begin{bmatrix} x_1 (k) \\ x_2 (k) \\ x_3 (k) \end{bmatrix}$$

Jordan form (2 complex poles):

$$G(s) = \frac{k_1}{z - p_1} + \frac{\alpha z + \beta}{(z - \sigma)^2 + \omega^2}$$

$$x_1(k+1) \\ x_2(k+1) \\ x_3(k+1) \end{bmatrix} = \begin{bmatrix} p_1 & 0 & 0 \\ 0 & \sigma & \omega \\ 0 & -\omega & \sigma \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} u(k)$$

$$y(k) = \begin{bmatrix} k_1 & k_2 & k_3 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \end{bmatrix}$$

where $k_2 = (\beta + \alpha \sigma)/\omega$, $k_3 = \alpha$.

Exercise: obtain the controllable canonical form for the following systems

- $G(s) = \frac{z^{-1} z^{-3}}{1 + 2z^{-1} + z^{-2}}$
- $G(s) = \frac{b_0 z^2 + b_1 z + b_2}{z^3 + a_0 z^2 + a_1 z + a_2}$

1.6 Similar Realizations

Besides the canonical forms, other system realizations exist. Let us begin with the realization Σ of some transfer function G(s). Let $T \in \mathbb{C}^{n \times n}$ be nonsingular. We can define *new* states by:

$$Tx^* = x$$

We can rewrite the differential equations defining Σ in terms of these new states by plugging in $x = Tx^*$:

$$\frac{d}{dt}\left(Tx^{*}(t)\right) = ATx^{*}(t) + Bu(t),$$

to obtain

$$\Sigma^*: \begin{cases} \dot{x}^*(t) &= T^{-1}ATx^*(t) + T^{-1}Bu(t) \\ y(t) &= CTx^*(t) + Du(t) \end{cases}$$

This new realization

$$\Sigma^* = \begin{bmatrix} T^{-1}AT & T^{-1}B \\ \hline CT & D \end{bmatrix},\tag{12}$$

also realizes G(s) and is said to be *similar* to Σ .

Similar realizations are fundamentally the same. Indeed, we arrived at Σ_{new} from Σ via nothing more than a change of variables.

Exercise 3 (Another observable canonical form.). Verify that

$$\Sigma = \begin{bmatrix} -a_2 & 1 & 0 & b_2 \\ -a_1 & 0 & 1 & b_1 \\ -a_0 & 0 & 0 & b_0 \\ \hline 1 & 0 & 0 & d \end{bmatrix}$$

is similar to

$$\Sigma^* = \begin{bmatrix} 0 & 0 & -a_0 & b_0 \\ 1 & 0 & -a_1 & b_1 \\ 0 & 1 & -a_2 & b_2 \\ \hline 0 & 0 & 1 & d \end{bmatrix}$$

ME 547: Linear Systems Solution of Time-Invariant State-Space Equations

Xu Chen

University of Washington

UW Linear Systems (X. Chen, ME547)

SS Solution

1/56

Outline

1. Solution Formula

Continuous-Time Case The Solution to $\dot{x} = ax + bu$ The Solution to n^{th} -order LTI Systems The State Transition Matrix e^{At} Computing e^{At} when A is diagonal or in Jordan form Discrete-Time LTI Case The State Transition Matrix A^k Computing A^k when A is diagonal or in Jordan form

- 2. Explicit Computation of the State Transition Matrix e^{At}
- 3. Explicit Computation of the State Transition Matrix A^k
- 4. Transition Matrix via Inverse Transformation

Introduction

The Solution to $\dot{x} = ax + bu$

► To solve the vector equation $\dot{x} = Ax + Bu$, we start with the scalar case when $x, a, b, u \in \mathbb{R}$.

fundamental property of exponential functions

$$\frac{d}{dt}e^{at} = ae^{at}, \quad \frac{d}{dt}e^{-at} = -ae^{-at}$$

 $\dot{x}(t) = ax(t) + bu(t), \ a \neq 0 \stackrel{\because e^{-at} \neq 0}{\Longrightarrow} e^{-at} \dot{x}(t) - e^{-at} ax(t) = e^{-at} bu(t), \text{ namely,}$

$$\frac{d}{dt} \left\{ e^{-at} x(t) \right\} = e^{-at} bu(t) \Leftrightarrow d \left\{ e^{-at} x(t) \right\} = e^{-at} bu(t) dt$$
$$\Longrightarrow \left[e^{-at} x(t) = e^{-at_0} x(t_0) + \int_{t_0}^t e^{-a\tau} bu(\tau) d\tau \right]$$

UW Linear Systems (X. Chen, ME547)

SS Solution

3 / 56

The solution to $\dot{x} = ax + bu$

$$e^{-at}x(t) = e^{-at_0}x(t_0) + \int_{t_0}^t e^{-a\tau}bu(\tau) d\tau$$

when $t_0 = 0$, we have

$$x(t) = \underbrace{e^{at}x(0)}_{\text{free response}} + \underbrace{\int_{0}^{t} e^{a(t-\tau)} bu(\tau) d\tau}_{\text{forced response}}$$
(1)

The solution to $\dot{x} = ax + bu$

Solution Concepts of $e^{at}x(0)$



 $e^{-1} \approx 37\%$, $e^{-2} \approx 14\%$, $e^{-3} \approx 5\%$, $e^{-4} \approx 2\%$ Time Constant $\triangleq \frac{1}{|a|}$ when a < 0: After three time constants, $e^{at}x(0)$ reduces to 5% of its initial value ~ the free response has died down.

UW Linear Systems (X. Chen, ME547)

SS Solution

5 / 56

* Fundamental Theorem of Differential Equations

addresses the question of whether a dynamical system has a unique solution or not.

Theorem

Consider $\dot{x} = f(x, t)$, $x(t_0) = x_0$, with:

- f (x, t) piecewise continuous in t (continuous except at finite points of discontinuity)

then there exists a unique function of time $\phi(\cdot) : \mathbb{R}_+ \to \mathbb{R}^n$ which is continuous almost everywhere and satisfies

$$\blacktriangleright \phi(t_0) = x_0$$

▶ $\dot{\phi}(t) = f(\phi(t), t), \forall t \in \mathbb{R}_+ \setminus D$, where D is the set of discontinuity points for f as a function of t.

The solution to n^{th} -order LTI systems

Analogous to scalar case, the general state-space equations

$$\Sigma: \begin{cases} \dot{x}(t) = Ax(t) + Bu(t) \\ y(t) = Cx(t) + Du(t) \end{cases} \quad x(t_0) = x_0 \in \mathbb{R}^n, \ A \in \mathbb{R}^{n \times n}$$

have the solution

$$x(t) = \underbrace{e^{A(t-t_0)}x_0}_{\text{free response}} + \underbrace{\int_{t_0}^t e^{A(t-\tau)}Bu(\tau)d\tau}_{\text{forced response}}$$
(2)

$$y(t) = Ce^{A(t-t_0)}x_0 + C\int_{t_0}^t e^{A(t-\tau)}Bu(\tau)d\tau + Du(t)$$

In both the free and the forced responses, computing e^{At} is key.
 e^{A(t-t_0)}: called the transition matrix

UW Linear Systems (X. Chen, ME547)

SS Solution

7 / 56

The state transition matrix e^{At}

For the scalar case with $a \in \mathbb{R}$, Taylor expansion gives

$$e^{at} = 1 + at + \frac{1}{2}(at)^2 + \dots + \frac{1}{n!}(at)^n + \dots$$
 (3)

The transition scalar $\Phi(t, t_0) = e^{a(t-t_0)}$ satisfies

$$\Phi(t, t) = 1$$
 (transition to itself)
 $\Phi(t_3, t_2)\Phi(t_2, t_1) = \Phi(t_3, t_1)$ (consecutive transition)
 $\Phi(t_2, t_1) = \Phi^{-1}(t_1, t_2)$ (reverse transition)

The state transition matrix e^{At}

For the matrix case with $A \in \mathbb{R}^{n \times n}$

$$e^{At} = I_n + At + \frac{1}{2}A^2t^2 + \dots + \frac{1}{n!}A^nt^n + \dots$$
(4)

As I_n and Aⁱ are matrices of dimension n × n, e^{At} must ∈ ℝ^{n×n}.
 The transition matrix Φ(t, t₀) = e^{A(t-t₀)} satisfies

$$e^{A0} = I_n$$
 $\Phi(t, t) = I_n$
 $e^{At_1}e^{At_2} = e^{A(t_1+t_2)}$ $\Phi(t_3, t_2)\Phi(t_2, t_1) = \Phi(t_3, t_1)$
 $e^{-At} = [e^{At}]^{-1}$. $\Phi(t_2, t_1) = \Phi^{-1}(t_1, t_2)$.

Note, however, that e^{At}e^{Bt} = e^{(A+B)t} if and only if AB = BA. (Check by using Taylor expansion.)

UW Linear Systems (X. Chen, ME547)

SS Solution

Computing a structured e^{At} via Taylor expansion convenient when A is a diagonal or Jordan matrix

The case with a diagonal matrix
$$A = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}$$
:
 $A^2 = \begin{bmatrix} \lambda_1^2 & 0 & 0 \\ 0 & \lambda_2^2 & 0 \\ 0 & 0 & \lambda_3^2 \end{bmatrix}, \dots, A^n = \begin{bmatrix} \lambda_1^n & 0 & 0 \\ 0 & \lambda_2^n & 0 \\ 0 & 0 & \lambda_3^n \end{bmatrix}$

all matrices on the right side of

$$e^{At} = I + At + \frac{1}{2}A^2t^2 + \dots + \frac{1}{n!}A^nt^n + \dots$$

are easy to compute

Computing a structured e^{At} via Taylor expansion convenient when A is a diagonal or Jordan matrix The case with a diagonal matrix $A = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}$: $e^{At} = I + At + \frac{1}{2}A^2t^2 + \dots + \frac{1}{n!}A^nt^n + \dots$ $\begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \lambda_1t & 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{2}\lambda_1^2t^2 & 0 & 0 \end{bmatrix}$

$$\begin{split} e^{At} &= I + At + \frac{1}{2}A^{2}t^{2} + \dots + \frac{1}{n!}A^{n}t^{n} + \dots \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} \lambda_{1}t & 0 & 0 \\ 0 & \lambda_{2}t & 0 \\ 0 & 0 & \lambda_{3}t \end{bmatrix} + \begin{bmatrix} \frac{1}{2}\lambda_{1}^{2}t^{2} & 0 & 0 \\ 0 & \frac{1}{2}\lambda_{2}^{2}t^{2} & 0 \\ 0 & 0 & \frac{1}{2}\lambda_{3}^{2}t^{2} \end{bmatrix} + \dots \\ &= \begin{bmatrix} 1 + \lambda_{1}t + \frac{1}{2}\lambda_{1}^{2}t^{2} + \dots & 0 & 0 \\ 0 & 1 + \lambda_{2}t + \frac{1}{2}\lambda_{2}^{2}t^{2} + \dots & 0 \\ 0 & 1 + \lambda_{3}t + \frac{1}{2}\lambda_{3}^{2}t^{2} + \dots \end{bmatrix} \\ &= \begin{bmatrix} e^{\lambda_{1}t} & 0 & 0 \\ 0 & e^{\lambda_{2}t} & 0 \\ 0 & 0 & e^{\lambda_{3}t} \end{bmatrix}. \end{split}$$

UW Linear Systems (X. Chen, ME547)

SS Solution

11/56

Computing a structured e^{At} via Taylor expansion

The case with a Jordan matrix
$$A = \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix}$$
:
• Decompose $A = \underbrace{\begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix}}_{\lambda I_3} + \underbrace{\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}}_{N}$. Then
 $e^{At} = e^{(\lambda I_3 t + Nt)}$.
• also, $(\lambda I_3 t) (Nt) = \lambda Nt^2 = (Nt) (\lambda I_3 t)$ and hence
 $e^{(\lambda I_3 t + Nt)} = e^{\lambda I t} e^{Nt}$

thus

$$\underline{e^{At} = e^{(\lambda I_3 t + Nt)}} = e^{\lambda It} e^{Nt} \stackrel{\text{``e}^{\lambda It} = e^{\lambda t}I}{=} e^{\lambda t} e^{Nt}$$

Computing a structured e^{At} via Taylor expansion

$$A = \underbrace{\begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix}}_{\lambda I_{3}} + \underbrace{\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}}_{N} e^{At} = e^{\lambda t} e^{Nt}$$

$$N \text{ is nilpotent}^{1} : N^{3} = N^{4} = \dots = 0I_{3}, \text{ yielding}$$

$$e^{Nt} = I_{3} + Nt + \frac{1}{2}N^{2}t^{2} + \frac{1}{3!}N^{3}t^{3} + \dots = \begin{bmatrix} 1 & t & \frac{t^{2}}{2} \\ 0 & 1 & t \\ 0 & 0 & 1 \end{bmatrix}.$$

$$Thus$$

$$e^{At} = \begin{bmatrix} e^{\lambda t} & te^{\lambda t} & \frac{t^{2}}{2}e^{\lambda t} \\ 0 & e^{\lambda t} & te^{\lambda t} \end{bmatrix}. \quad (5)$$

1 "nil" \sim zero; "potent"	\sim taking powers.	
UW Linear Systems (X. Chen, ME547)	SS Solution	13 / 56

Computing a structured e^{At} via Taylor expansion

Example (mass moving on a straight line with zero friction and no external force)

$$\frac{d}{dt} \left[\begin{array}{c} x_1 \\ x_2 \end{array} \right] = \underbrace{\left[\begin{array}{c} 0 & 1 \\ 0 & 0 \end{array} \right]}_{A} \left[\begin{array}{c} x_1 \\ x_2 \end{array} \right]$$

$$x(t) = e^{At}x(0) \text{ where}$$

$$e^{At} = I + \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} t + \frac{1}{2!} \underbrace{\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}}_{=\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}} t^2 + \dots = \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}.$$

Computing low-order e^{At} via column solutions

We discuss an intuition of the matrix entries in e^{At} . Consider example:

$$\dot{x} = Ax = \begin{bmatrix} 0 & 1 \\ 0 & -1 \end{bmatrix} x, \quad x(0) = x_0$$

$$x(t) = e^{At}x(0) = \begin{bmatrix} 1 \text{st column} \\ a_1(t) \end{bmatrix} \begin{bmatrix} 2 \text{nd column} \\ a_2(t) \end{bmatrix} \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix}$$
(6)
$$= a_1(t)x_1(0) + a_2(t)x_2(0)$$
(7)

Observation

$$egin{aligned} x(0) &= \left[egin{aligned} 1 \ 0 \end{array}
ight] \Rightarrow x(t) = a_1(t), \ x(0) &= \left[egin{aligned} 0 \ 1 \end{array}
ight] \Rightarrow x(t) = a_2(t). \end{aligned}$$

UW Linear Systems (X. Chen, ME547)

SS Solution

15 / 56

Computing low-order e^{At} via column solutions

$$\dot{x} = Ax = \begin{bmatrix} 0 & 1 \\ 0 & -1 \end{bmatrix} x, \quad x(0) = x_0$$

Hence, we can obtain e^{At} from the following

1. write out
$$\dot{x}_1(t) = x_2(t)$$

 $\dot{x}_2(t) = -x_2(t) \Rightarrow x_1(t) = e^{0t}x_1(0) + \int_0^t e^{0(t-\tau)}x_2(\tau)d\tau$
 $x_2(t) = -x_2(t) \Rightarrow x_2(t) = e^{-t}x_2(0)$
2. let $x(0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, then $\begin{array}{c} x_1(t) \equiv 1 \\ x_2(t) \equiv 0 \end{array}$, namely $x(t) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$
3. let $x(0) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, then $x_2(t) = e^{-t}$ and $x_1(t) = 1 - e^{-t}$, or more compactly, $x(t) = \begin{bmatrix} 1 - e^{-t} \\ e^{-t} \end{bmatrix}$
4. using (6), write out directly $\begin{array}{c} e^{At} = \begin{bmatrix} 1 & 1 - e^{-t} \\ 0 & e^{-t} \end{bmatrix}$

Computing low-order e^{At} via column solutions

Exercise

Compute e^{At} where

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

UW Linear Systems (X. Chen, ME547)

SS Solution

17 / 56

Discrete-time LTI case

For the discrete-time system

$$x(k+1) = Ax(k) + Bu(k), x(0) = x_0,$$

iteration of the state-space equation gives

$$x(k) = A^{k-k_0}x(k_o) + [A^{k-k_0-1}B, A^{k-k_0-2}B, \cdots, B] \begin{bmatrix} u(k_0) \\ u(k_0+1) \\ \vdots \\ u(k-1) \end{bmatrix}$$

$$\Leftrightarrow x(k) = \underbrace{A^{k-k_0}x(k_o)}_{\text{free response}} + \underbrace{\sum_{j=k_0}^{k-1} A^{k-1-j}Bu(j)}_{\text{forced response}}$$

Discrete-time LTI case

$$x(k) = \underbrace{A^{k-k_0}x(k_o)}_{\text{free response}} + \underbrace{\sum_{j=k_0}^{k-1} A^{k-1-j}Bu(j)}_{\text{forced response}}$$

the transition matrix, defined by $\Phi(k,j) = A^{k-j}$, satisfies

$$egin{aligned} \Phi(k,k) &= 1 \ \Phi(k_3,k_2) \Phi(k_2,k_1) &= \Phi(k_3,k_1) \ \Phi(k_2,k_1) &= \Phi^{-1}(k_1,k_2) \end{aligned}$$
 if and only if A is nonsingular



The state transition matrix A^k

Similar to the continuous-time case, when A is a diagonal or Jordan matrix, the Taylor expansion formula readily generates A^k .

Diagonal matrix A =
$$\begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}$$

$$: A^k = \begin{bmatrix} \lambda_1^k & 0 & 0 \\ 0 & \lambda_2^k & 0 \\ 0 & 0 & \lambda_3^k \end{bmatrix}$$

Computing a structured A^k via Taylor expansion

Jordan canonical form
$$A = \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix} = \underbrace{\begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix}}_{\lambda I_3} + \underbrace{\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}}_{N}$$



Computing a structured A^k via Taylor expansion

Exercise
Recall that
$$\begin{pmatrix} k \\ 3 \end{pmatrix} = \frac{1}{3!}k(k-1)(k-2)$$
. Show

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

$$\Rightarrow A^{k} = \begin{bmatrix} \lambda^{k} & k\lambda^{k-1} & \frac{1}{2!}k(k-1)\lambda^{k-2} & \frac{1}{3!}k(k-1)(k-2)\lambda^{k-3} \\ 0 & \lambda^{k} & k\lambda^{k-1} & \frac{1}{2!}k(k-1)\lambda^{k-2} \\ 0 & 0 & \lambda^{k} & k\lambda^{k-1} \\ 0 & 0 & 0 & \lambda^{k} \end{bmatrix}$$

1. Solution Formula

Continuous-Time Case The Solution to $\dot{x} = ax + bu$ The Solution to n^{th} -order LTI Systems The State Transition Matrix e^{At} Computing e^{At} when A is diagonal or in Jordan form Discrete-Time LTI Case The State Transition Matrix A^k Computing A^k when A is diagonal or in Jordan form

2. Explicit Computation of the State Transition Matrix e^{At}

- 3. Explicit Computation of the State Transition Matrix A^k
- 4. Transition Matrix via Inverse Transformation

UW Linear Systems (X. Chen, ME547)

SS Solution

23 / 56

Explicit computation of a general e^{At}

- Why another method: general matrices may not be diagonal or Jordan
- Approach: transform a general matrix to a diagonal or Jordan form, via similarity transformation

Computing e^{At} via similarity transformation

Principle Concept.

1. Given

$$\dot{x}(t) = Ax(t) + Bu(t), \ x(t_0) = x_0 \in \mathbb{R}^n, \ A \in \mathbb{R}^{n \times n}$$

find a nonsingular $T \in \mathbb{R}^{n \times n}$ such that a coordinate transformation defined by $x(t) = Tx^*(t)$ yields

$$\frac{d}{dt}(Tx^*(t)) = ATx^*(t) + Bu(t)$$
$$\frac{d}{dt}x^*(t) = \underbrace{T^{-1}AT}_{\triangleq \Lambda: \text{ diagonal or Jordan}} x^*(t) + \underbrace{T^{-1}B}_{B^*}u(t)$$
$$x^*(0) = T^{-1}x_0$$

UW Linear Systems (X. Chen, ME547)

SS Solution

25 / 56

Computing e^{At} via similarity transformation 1. When u(t) = 0

$$\dot{x}(t) = Ax(t) \stackrel{x=Tx^*}{\Longrightarrow} \frac{d}{dt} x^*(t) = \underbrace{T^{-1}AT}_{\triangleq \Lambda: \text{ diagonal or Jordan}} x^*(t)$$

2. Now
$$x^*(t)$$
 can be solved easily: e.g., if $\Lambda = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$, then
 $x^*(t) = e^{\Lambda t} x^*(0) = \begin{bmatrix} e^{\lambda_1 t} & 0 \\ 0 & e^{\lambda_2 t} \end{bmatrix} \begin{bmatrix} x_1^*(0) \\ x_2^*(0) \end{bmatrix} = \begin{bmatrix} e^{\lambda_1 t} x_1^*(0) \\ e^{\lambda_2 t} x_2^*(0) \end{bmatrix}$.
3. $x(t) = Tx^*(t)$ then yields

$$x(t) = Te^{\Lambda t}x^*(0) = Te^{\Lambda t}T^{-1}x_0$$

4. On the other hand, $x(t) = e^{At}x_0$. Hence

$$e^{At} = T e^{\Lambda t} T^{-1}$$

Similarity transformation

- Existence of Solutions: T comes from the theory of eigenvalues and eigenvectors in linear algebra.
- If two matrices A, $B \in \mathbb{C}^{n \times n}$ are similar:
 - $A = TBT^{-1}, \ T \in \mathbb{C}^{n imes n}$, then
 - ▶ their A^n and B^n are also similar: e.g.,

$$A^2 = TBT^{-1}TBT^{-1} = TB^2T^{-1}$$

their exponential matrices are also similar

$$e^{At} = T e^{Bt} T^{-1}$$

as

$$Te^{Bt} T^{-1} = T(I_n + Bt + \frac{1}{2}B^2t^2 + \dots)T^{-1}$$

= $TI_n T^{-1} + TBtT^{-1} + \frac{1}{2}TB^2t^2T^{-1} + \dots$
= $I + At + \frac{1}{2}A^2t^2 + \dots = e^{At}$

UW Linear Systems (X. Chen, ME547)

SS Solution

27 / 56

Similarity transformation

▶ For $A \in \mathbb{R}^{n \times n}$, an eigenvalue $\lambda \in C$ of A is the solution to the characteristic equation

$$\det\left(A-\lambda I\right)=0\tag{8}$$

The corresponding eigenvectors are the nonzero solutions to

$$At = \lambda t \Leftrightarrow (A - \lambda I) t = 0 \tag{9}$$

Similarity transformation

The case with distinct eigenvalues (diagonalization). Recall: When $A \in \mathbb{R}^{n \times n}$ has *n* distinct eigenvalues such that

$$Ax_1 = \lambda_1 x_1$$

:
$$Ax_n = \lambda_n x_n$$

$$\Gamma \lambda_1$$

or equivalently $A\underbrace{[x_1, x_2, \dots, x_n]}_{\triangleq T} = [x_1, x_2, \dots, x_n] \underbrace{\begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{bmatrix}}_{\Lambda}$

 $[x_1, x_2, \ldots, x_n]$ is square and invertible. Hence

$$A = T\Lambda T^{-1}, \ \Lambda = T^{-1}AT$$

UW Linear Systems (X. Chen, ME547)

SS Solution

29 / 56

Example (Mechanical system with strong damping)

$$\frac{d}{dt} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix}}_{A} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Find eigenvalues: $det(A - \lambda I) = det \begin{bmatrix} -\lambda & 1 \\ -2 & -\lambda - 3 \end{bmatrix} = (\lambda + 2) (\lambda + 1) \Rightarrow \lambda_1 = -2, \lambda_2 = -1$

Find associate eigenvectors:

$$\lambda_{1} = -2: (A - \lambda_{1}I) t_{1} = 0 \Rightarrow t_{1} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

$$\lambda_{1} = -1: (A - \lambda_{2}I) t_{2} = 0 \Rightarrow t_{2} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$\text{Define } T \text{ and } \Lambda: T = \begin{bmatrix} t_{1} & t_{2} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -2 & -1 \end{bmatrix},$$

$$\Lambda = \begin{bmatrix} \lambda_{1} & 0 \\ 0 & \lambda_{2} \end{bmatrix} = \begin{bmatrix} -2 & 0 \\ 0 & -1 \end{bmatrix}$$

Example (Mechanical system with strong damping)

$$\frac{d}{dt} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix}}_{A} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

•
$$T = \begin{bmatrix} 1 & 1 \\ -2 & -1 \end{bmatrix}$$
, $\Lambda = \begin{bmatrix} -2 & 0 \\ 0 & -1 \end{bmatrix}$
• Compute $T^{-1} = \begin{bmatrix} 1 & 1 \\ -2 & -1 \end{bmatrix}^{-1} = \begin{bmatrix} -1 & -1 \\ 2 & 1 \end{bmatrix}$
• Compute $e^{At} = Te^{\Lambda t}T^{-1} = T\begin{bmatrix} e^{-2t} & 0 \\ 0 & e^{-1t} \end{bmatrix} T^{-1} = \begin{bmatrix} -e^{-2t} + 2e^{-t} & -e^{-2t} + e^{-t} \\ 2e^{-2t} - 2e^{-t} & 2e^{-2t} - e^{-t} \end{bmatrix}$

UW Linear Systems (X. Chen, ME547)

SS Solution

31 / 56

Similarity transform: diagonalization

Physical interpretations

diagonalized system:
$$x^{*}(t) = \begin{bmatrix} e^{\lambda_{1}t} & 0\\ 0 & e^{\lambda_{2}t} \end{bmatrix} \begin{bmatrix} x_{1}^{*}(0)\\ x_{2}^{*}(0) \end{bmatrix} = \begin{bmatrix} e^{\lambda_{1}t}x_{1}^{*}(0)\\ e^{\lambda_{2}t}x_{2}^{*}(0) \end{bmatrix}$$

x(t) = Tx*(t) = e^{λ₁t}x₁*(0)t₁ + e^{λ₂t}x₂*(0)t₂ then decomposes the state trajectory into two modes parallel to the two eigenvectors.

Similarity transform: diagonalization

Physical interpretations

- If x(0) is aligned with one eigenvector, say, t₁, then x²(0) = 0 and x(t) = e^{λ₁t}x¹*(0)t₁ + e^{λ₂t}x²*(0)t₂ dictates that x(t) will stay in the direction of t₁.
- i.e., if the state initiates along the direction of one eigenvector, then the free response will stay in that direction without "making turns".
- If λ₁ < 0, then x(t) will move towards the origin of the state space; if λ₁ = 0, x(t) will stay at the initial point; and if positive, x(t) will move away from the origin along t₁.</p>
- Furthermore, the magnitude of \(\lambda_1\) determines the speed of response.

UW Linear Systems (X. Chen, ME547)

SS Solution

33 / 56

Similarity transform: diagonalization

Physical interpretations: example



Similarity transformation

The case with complex eigenvalues

Consider the undamped spring-mass system

$$\frac{d}{dt} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}}_{A} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \ \det(A - \lambda I) = \lambda^2 + 1 = 0 \Rightarrow \lambda_{1,2,} = \pm j.$$

The eigenvectors are

$$\lambda_1 = j: (A - jI)t_1 = 0 \Rightarrow t_1 = \begin{bmatrix} 1 \\ j \end{bmatrix}$$
$$\lambda_2 = -j: (A + jI)t_2 = 0 \Rightarrow t_2 = \begin{bmatrix} 1 \\ -j \end{bmatrix} \text{ (complex conjugate of } t_1\text{)}.$$

Hence

$$T = \begin{bmatrix} 1 & 1 \\ j & -j \end{bmatrix}, \ T^{-1} = \frac{1}{2} \begin{bmatrix} 1 & -j \\ 1 & j \end{bmatrix}$$

UW Linear Systems (X. Chen, ME547)

SS Solution

35 / 56

Similarity transformation

The case with complex eigenvalues

$$\frac{d}{dt} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}}_{A} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$\lambda_{1,2,} = \pm j$$

$$T = \begin{bmatrix} 1 & 1 \\ j & -j \end{bmatrix}, \quad T^{-1} = \frac{1}{2} \begin{bmatrix} 1 & -j \\ 1 & j \end{bmatrix}$$

► we have

$$e^{At} = T e^{\Lambda t} T^{-1} = T \left[egin{array}{cc} e^{jt} & 0 \ 0 & e^{-jt} \end{array}
ight] T^{-1} = \left[egin{array}{cc} \cos t & \sin t \ -\sin t & \cos t \end{array}
ight].$$
Similarity transformation

The case with complex eigenvalues

As an exercise, for a general $A \in \mathbb{R}^{2 \times 2}$ with complex eigenvalues $\sigma \pm j\omega$, you can show that by using $T = [t_R, t_I]$ where t_R and t_I are the real and the imaginary parts of t_1 , an eigenvector associated with $\lambda_1 = \sigma + j\omega$, $x = Tx^*$ transforms $\dot{x} = Ax$ to

$$\dot{x}^*(t) = \left[egin{array}{cc} \sigma & \omega \ -\omega & \sigma \end{array}
ight] x^*(t)$$

and

$$e^{\begin{bmatrix} \sigma & \omega \\ -\omega & \sigma \end{bmatrix}^{t}} = \begin{bmatrix} e^{\sigma t} \cos \omega t & e^{\sigma t} \sin \omega t \\ -e^{\sigma t} \sin \omega t & e^{\sigma t} \cos \omega t \end{bmatrix}$$

UW Linear Systems (X. Chen, ME547)

SS Solution

37 / 56

Similarity transformation

The case with repeated eigenvalues via generalized eigenvectors

Consider $A = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$: two repeated eigenvalues $\lambda (A) = 1$, and $(A - \lambda I) t_1 = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix} t_1 = 0 \Rightarrow t_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

No other linearly independent eigenvectors exist. What next?

A is already very similar to the Jordan form. Try instead

$$A\begin{bmatrix} t_1 & t_2 \end{bmatrix} = \begin{bmatrix} t_1 & t_2 \end{bmatrix} \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix},$$

which requires $At_2 = t_1 + \lambda t_2$, i.e.,

$$(A - \lambda I) t_2 = t_1 \Leftrightarrow \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix} t_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \Rightarrow t_2 = \begin{bmatrix} 0 \\ 0.5 \end{bmatrix}$$

 t_2 is linearly independent from $t_1 \Rightarrow t_1$ and t_2 span \mathbb{R}^2 . (t_2 is called a generalized eigenvector.)

Similarity transformation

The case with repeated eigenvalues via generalized eigenvectors

For general 3 × 3 matrices with det $(\lambda I - A) = (\lambda - \lambda_m)^3$, i.e., $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_m$, we look for T such that

$$A = TJT^{-1}$$

where J has three canonical forms:

$$i), \begin{bmatrix} \lambda_{m} & 0 & 0 \\ 0 & \lambda_{m} & 0 \\ 0 & 0 & \lambda_{m} \end{bmatrix}, iii), \begin{bmatrix} \lambda_{m} & 1 & 0 \\ 0 & \lambda_{m} & 1 \\ 0 & 0 & \lambda_{m} \end{bmatrix}$$
$$ii), \begin{bmatrix} \lambda_{m} & 1 & 0 \\ 0 & \lambda_{m} & 0 \\ 0 & 0 & \lambda_{m} \end{bmatrix} \text{ or } \begin{bmatrix} \lambda_{m} & 0 & 0 \\ 0 & \lambda_{m} & 1 \\ 0 & 0 & \lambda_{m} \end{bmatrix}$$

UW Linear Systems (X. Chen, ME547)

SS Solution

39 / 56

Similarity transformation

The case with repeated eigenvalues via generalized eigenvectors

i),
$$A = TJT^{-1}$$
, $J = \begin{bmatrix} \lambda_m & 0 & 0 \\ 0 & \lambda_m & 0 \\ 0 & 0 & \lambda_m \end{bmatrix}$

this happens

- ▶ when A has three linearly independent eigenvectors, i.e., $(A \lambda_m I)t = 0$ yields t_1 , t_2 , and t_3 that span \mathbb{R}^3 .
- mathematically: when nullity $(A \lambda_m I) = 3$, namely, rank $(A \lambda_m I) = 3$ nullity $(A \lambda_m I) = 0$.

Similarity transformation

The case with repeated eigenvalues via generalized eigenvectors

ii),
$$A = TJT^{-1}$$
, $J = \begin{bmatrix} \lambda_m & 1 & 0 \\ 0 & \lambda_m & 0 \\ 0 & 0 & \lambda_m \end{bmatrix}$ or $\begin{bmatrix} \lambda_m & 0 & 0 \\ 0 & \lambda_m & 1 \\ 0 & 0 & \lambda_m \end{bmatrix}$

▶ this happens when $(A - \lambda_m I)t = 0$ yields two linearly independent solutions, i.e., when nullity $(A - \lambda_m I) = 2$.

• we then have, e.g., $\Gamma \lambda$,

$$A[t_1, t_2, t_3] = [t_1, t_2, t_3] \begin{bmatrix} \lambda_m & 1 & 0 \\ 0 & \lambda_m & 0 \\ 0 & 0 & \lambda_m \end{bmatrix}$$

1 0]

$$\Leftrightarrow [\lambda_m t_1, t_1 + \lambda_m t_2, \lambda_m t_3] = [At_1, At_2, At_3]$$
(10)

- \blacktriangleright t_1 and t_3 are the directly computed eigenvectors.
- For t_2 , the second column of (10) gives

$$(A-\lambda_m I) t_2 = t_1$$

UW Linear Systems (X. Chen, ME547)

SS Solution

41 / 56

Similarity transformation

The case with repeated eigenvalues via generalized eigenvectors

iii),
$$A = TJT^{-1}$$
, $J = \begin{bmatrix} \lambda_m & 1 & 0 \\ 0 & \lambda_m & 1 \\ 0 & 0 & \lambda_m \end{bmatrix}$

► this is for the case when $(A - \lambda_m I)t = 0$ yields only one linearly independent solution, i.e., when nullity $(A - \lambda_m I) = 1$.

We then have

$$A[t_1, t_2, t_3] = [t_1, t_2, t_3] \begin{bmatrix} \lambda_m & 1 & 0 \\ 0 & \lambda_m & 1 \\ 0 & 0 & \lambda_m \end{bmatrix}$$

$$\Leftrightarrow [\lambda_m t_1, t_1 + \lambda_m t_2, t_2 + \lambda_m t_3] = [At_1, At_2, At_3]$$
yielding

$$(A - \lambda_m I) t_1 = 0$$

$$(A - \lambda_m I) t_2 = t_1, (t_2 : \text{ generalized eigenvector})$$

$$(A - \lambda_m I) t_3 = t_2, (t_3 : \text{ generalized eigenvector})$$

Example

$$A = \begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix}, \text{ det } (A - \lambda I) = \lambda^2 \Rightarrow \lambda_1 = \lambda_2 = 0, J = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

Two repeated eigenvalues with rank $(A - 0I) = 1 \Rightarrow$ only one linearly independent eigenvector: $(A - 0I) t_1 = 0 \Rightarrow t_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$
Generalized eigenvector: $(A - 0I) t_2 = t_1 \Rightarrow t_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$
Coordinate transform matrix:

$$T = [t_1, t_2] = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, T^{-1} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$$

$$e^{At} = Te^{Jt}T^{-1} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} e^{0t} & te^{0t} \\ 0 & e^{0t} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 - t & t \\ -t & 1 + t \end{bmatrix}$$

UW Linear Systems (X. Chen, ME547) SS Solution

43 / 56

Example

$$A = \left| egin{array}{cc|c} -1 & 1 \ -1 & 1 \end{array}
ight|, \ \det \left(A - \lambda I
ight) = \lambda^2 \Rightarrow \lambda_1 = \lambda_2 = 0.$$

Observation

►
$$\lambda_1 = 0$$
, $t_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ implies that if $x_1(0) = x_2(0)$ then the response is characterized by $e^{0t} = 1$

• i.e.,
$$x_1(t) = x_1(0) = x_2(0) = x_2(t)$$
. This makes sense because $\dot{x}_1 = -x_1 + x_2$ from the state equation.

Example (Multiple eigenvectors)

Obtain the eigenvectors of

$$A = \begin{bmatrix} -2 & 2 & -3 \\ 2 & 1 & -6 \\ -1 & -2 & 0 \end{bmatrix} (\lambda_1 = 5, \ \lambda_2 = \lambda_3 = -3).$$

UW Linear Systems (X. Chen, ME547)

SS Solution

45 / 56

Generalized eigenvectors

Physical interpretation.

When
$$\dot{x} = Ax$$
, $A = TJT^{-1}$ with $J = \begin{bmatrix} \lambda_m & 1 & 0 \\ 0 & \lambda_m & 0 \\ 0 & 0 & \lambda_m \end{bmatrix}$, we have
 $x(t) = e^{At}x(0) = T \begin{bmatrix} e^{\lambda_m t} & te^{\lambda_m t} & 0 \\ 0 & e^{\lambda_m t} & 0 \\ 0 & 0 & e^{\lambda_m t} \end{bmatrix} T^{-1}x(0)$
 $= T \begin{bmatrix} e^{\lambda_m t} & te^{\lambda_m t} & 0 \\ 0 & e^{\lambda_m t} & 0 \\ 0 & 0 & e^{\lambda_m t} \end{bmatrix} T^{-1}T_x^{I}(0)$

▶ If the initial condition is in the direction of t_1 , i.e., $x^*(0) = [x_1^*(0), 0, 0]^T$ and $x_1^*(0) \neq 0$, the above equation yields $x(t) = x_1^*(0)t_1e^{\lambda_m t}$.

Generalized eigenvectors

Physical interpretation Cont'd.

When
$$\dot{x} = Ax$$
, $A = TJT^{-1}$ with $J = \begin{bmatrix} \lambda_m & 1 & 0\\ 0 & \lambda_m & 0\\ 0 & 0 & \lambda_m \end{bmatrix}$, we have
 $x(t) = e^{At}x(0) = T \begin{bmatrix} e^{\lambda_m t} & te^{\lambda_m t} & 0\\ 0 & e^{\lambda_m t} & 0\\ 0 & 0 & e^{\lambda_m t} \end{bmatrix} T^{-1}x(0)$
 $= T \begin{bmatrix} e^{\lambda_m t} & te^{\lambda_m t} & 0\\ 0 & e^{\lambda_m t} & 0\\ 0 & 0 & e^{\lambda_m t} \end{bmatrix} T^{-1}T_x^{I}(0)$

If x(0) starts in the direction of t₂, i.e., x^{*}(0) = [0, x₂^{*}(0), 0]^T, then x(t) = x₂^{*}(0)(t₁te^{λ_mt} + t₂e^{λ_mt}). In this case, the response does not remain in the direction of t₂ but is confined in the subspace spanned by t₁ and t₂.

UW Linear Systems (X. Chen, ME547)

SS Solution

47 / 56

Example

Obtain eigenvalues of J and e^{Jt} by inspection:

Explicit computation of A^k

Everything in getting the similarity transform applies to the DT case:



UW Linear Systems (X. Chen, ME547)

SS Solution

49 / 56

Example								
Write	down .	J ^k for	J =	$\left[\begin{array}{c} -1\\ 0\\ 0\end{array}\right]$	$egin{array}{ccc} 0 & 0 \ -1 & 1 \ 0 & -1 \end{array}$) L and 1		
	☐ −10	1	0	0	0			
	0	-10	0	0	0			
J =	0	0	-2	0	0			
	0	0	0	-100	1			
	0	0	0	-1	-100			

1. Solution Formula

Continuous-Time Case The Solution to $\dot{x} = ax + bu$ The Solution to n^{th} -order LTI Systems The State Transition Matrix e^{At} Computing e^{At} when A is diagonal or in Jordan form Discrete-Time LTI Case The State Transition Matrix A^k Computing A^k when A is diagonal or in Jordan form

- 2. Explicit Computation of the State Transition Matrix e^{At}
- 3. Explicit Computation of the State Transition Matrix A^k
- 4. Transition Matrix via Inverse Transformation



Transition matrix via inverse transformation



On the other hand, from Laplace transform:

$$\dot{x}(t) = Ax(t) + Bu(t) \Rightarrow X(s) = \underbrace{(sI - A)^{-1}x(0)}_{l} + \underbrace{(sI - A)^{-1}BU(s)}_{l}$$

free response

forced response

Comparing x(t) and X(s) gives

$$e^{At} = \mathcal{L}^{-1}\left\{ (sI - A)^{-1} \right\}$$
 (11)

SS Solution



Example

$$A = \begin{bmatrix} \sigma & \omega \\ -\omega & \sigma \end{bmatrix}$$
$$e^{At} = \mathcal{L}^{-1} \begin{bmatrix} s - \sigma & -\omega \\ \omega & s - \sigma \end{bmatrix}^{-1}$$
$$= \mathcal{L}^{-1} \left\{ \frac{1}{(s - \sigma)^2 + \omega^2} \begin{bmatrix} s - \sigma & \omega \\ -\omega & s - \sigma \end{bmatrix} \right\}$$
$$= e^{\sigma t} \begin{bmatrix} \cos(\omega t) & \sin(\omega t) \\ -\sin(\omega t) & \cos(\omega t) \end{bmatrix}$$

UW Linear Systems (X. Chen, ME547)

SS Solution

53 / 56

Transition matrix via inverse transformation (DT case)

	Discrete-time system
state eq.	$x(k+1) = Ax(k) + Bu(k), x(0) = x_0$
	$(k{-}1)$
solution	$x(k) = A^{k}x(0) + \sum A^{(k-1-j)}Bu(j)$
	free response $j=0$
	forced response
transition matrix	transition matrix A ^k

On the other hand, from Z transform:

$$X(z) = (zI - A)^{-1} zx(0) + (zI - A)^{-1} BU(s)$$

Hence

$$A^{k} = \mathcal{Z}^{-1} \left\{ (zI - A)^{-1} z \right\}$$
(12)

SS Solution

Example

Example

$$\begin{aligned} A &= \begin{bmatrix} \sigma & \omega \\ -\omega & \sigma \end{bmatrix} \\ A^{k} &= \mathcal{Z}^{-1} \left\{ z \begin{bmatrix} z - \sigma & -\omega \\ \omega & z - \sigma \end{bmatrix}^{-1} \right\} \\ &= \mathcal{Z}^{-1} \left\{ \frac{z}{(z - \sigma)^{2} + \omega^{2}} \begin{bmatrix} z - \sigma & \omega \\ -\omega & z - \sigma \end{bmatrix} \right\} \\ &= \mathcal{Z}^{-1} \left\{ \frac{z}{z^{2} - 2r\cos\theta z + r^{2}} \begin{bmatrix} z - r\cos\theta & r\sin\theta \\ -r\sin\theta & z - r\cos\theta \end{bmatrix} \right\} \\ &, r &= \sqrt{\sigma^{2} + \omega^{2}}, \ \theta &= \tan^{-1}\frac{\omega}{\sigma} \\ &= r^{k} \begin{bmatrix} \cos k\theta & \sin k\theta \\ -\sin k\theta & \cos k\theta \end{bmatrix} \end{aligned}$$

UW Linear Systems (X. Chen, ME547)

SS Solution

55 / 56

Example
Consider
$$A = \begin{bmatrix} 0.7 & 0.3 \\ 0.1 & 0.5 \end{bmatrix}$$
. We have

$$\begin{aligned}
(zI - A)^{-1} z \\
&= \begin{bmatrix} \frac{z(z-0.5)}{(z-0.8)(z-0.4)} & \frac{0.3z}{(z-0.8)(z-0.4)} \\ \frac{0.1z}{(z-0.8)(z-0.4)} & \frac{z(z-0.7)}{(z-0.8)(z-0.4)} \end{bmatrix} \\
&= \begin{bmatrix} \frac{0.75z}{z-0.8} + \frac{0.25z}{z-0.4} & \frac{0.75z}{z-0.8} - \frac{0.75z}{z-0.4} \\ \frac{0.25z}{z-0.8} - \frac{0.25z}{z-0.4} & \frac{0.25z}{z-0.8} + \frac{0.75z}{z-0.4} \end{bmatrix} \\
\Rightarrow A^{k} = \begin{bmatrix} 0.75(0.8)^{k} + 0.25(0.4)^{k} & 0.75(0.8)^{k} - 0.75(0.4)^{k} \\ 0.25(0.8)^{k} - 0.25(0.4)^{k} & 0.25(0.8)^{k} + 0.75(0.4)^{k} \end{bmatrix}$$

1 Solution of Time-Invariant State-Space Equations

1.1 Continuous-Time State-Space Solutions

1.1.1 The Solution to $\dot{x} = ax + bu$

To solve the vector equation $\dot{x} = Ax + Bu$, we start with the scalar case when $x, a, b, u \in \mathbb{R}$. The solution can be easily derived using one fundamental property of exponential functions, that

$$\frac{d}{dt}e^{at} = ae^{at},$$

and

$$\frac{d}{dt}e^{-at} = -ae^{-at}.$$

Consider the ODE

$$\dot{x}(t) = ax(t) + bu(t), \ a \neq 0.$$

Since $e^{-at} \neq 0$, the above is equivalent to

$$e^{-at}\dot{x}\left(t\right) - e^{-at}ax\left(t\right) = e^{-at}bu\left(t\right),$$

namely,

$$\frac{d}{dt} \left\{ e^{-at} x\left(t\right) \right\} = e^{-at} b u\left(t\right),$$
$$\Rightarrow d \left\{ e^{-at} x\left(t\right) \right\} = e^{-at} b u\left(t\right) dt.$$

Integrating both sides from t_0 to t_1 gives

$$e^{-at_{1}}x(t_{1}) = e^{-at_{0}}x(t_{0}) + \int_{t_{0}}^{t_{1}} e^{-at}bu(t) dt.$$

It does not matter whether we use t or τ in the integration $\int_{t_0}^{t_1} e^{-at} bu(t) dt$. Hence we can change notations and get

$$e^{-at}x(t) = e^{-at_0}x(t_0) + \int_{t_0}^t e^{-a\tau}bu(\tau)\,d\tau,$$

$$\Leftrightarrow x(t) = e^{a(t-t_0)}x(t_0) + \int_{t_0}^t e^{a(t-\tau)}bu(\tau)\,d\tau.$$

Taking $t_0 = 0$ gives

$$x(t) = \underbrace{e^{at}x(0)}_{\text{free response}} + \underbrace{\int_{0}^{t} e^{a(t-\tau)} bu(\tau) d\tau}_{\text{forced response}}$$
(1)

where the free response is the part of the solution due only to initial conditions when no input is applied, and the forced response is the part due to the input alone.

Solution Concepts.

Time Constant. When a < 0, e^{at} is a decaying function. For the free response $e^{at}x(0)$, the exponential function satisfies $e^{-1} \approx 37\%$, $e^{-2} \approx 14\%$, $e^{-3} \approx 5\%$, and $e^{-4} \approx 2\%$. The time constant is defined as

$$T = \frac{1}{|a|}.$$

After three time constants, the free response reduces to 5% of its initial value. Roughly, we say the free response has died down.

Graphically, the exponential function looks like:



Unit Step Response. When a < 0 and u(t) = 1(t) (the step function), the solution is



$$x(t) = \frac{b}{|a|}(1 - e^{at}).$$

1.1.2 * Fundamental Theorem of Differential Equations

The following theorem addresses the question of whether a dynamical system has a unique solution or not.

Theorem 1. Consider $\dot{x} = f(x, t), x(t_0) = x_0$, with:

- f(x,t) piecewise continuous in t
- f(x,t) Lipschitz continuous in x

then there exists a *unique* function of time $\phi(\cdot) : \mathbb{R}_+ \to \mathbb{R}^n$ which is continuous almost everywhere and satisfies

- $\phi(t_0) = x_0$
- $\dot{\phi}(t) = f(\phi(t), t), \forall t \in \mathbb{R}_+ \setminus D$, where D is the set of discontinuity points for f as a function of t.

Remark 1. Piecewise continuous functions are continuous except at finite points of discontinuity.

- example 1: f(t) = |t|
- example 2:

$$f(x,t) = \begin{cases} A_1 x, & t \le t_1 \\ A_2 x, & t > t_1 \end{cases}$$

Lipschitz continuous functions are those that satisfy the cone constraint:

$$||f(x,t) - f(y,t)|| \le k(t) ||x - y|$$

where k(t) is piecewise continuous.

- example: f(x) = Ax + B
- a graphical representation of a Lipschitz function is that it must stay within a cone in the space of (x, f(x))
- a function is Lipschitz continuous if it is continuously differentiable with its derivative bounded everywhere. This is a sufficient condition. Functions can be Lipschitz continuous but not differentiable: e.g., the saturation function and f(x) = |x|.
- A continuous function is not necessarily Lipschitz continuous at all: e.g., a function whose derivative at x = 0 is infinity.

1.1.3 The Solution to nth-order LTI System

Consider the general state-space equation

$$\Sigma: \begin{cases} \dot{x}(t) = Ax(t) + Bu(t) \\ y(t) = Cx(t) + Du(t) \end{cases} \qquad x(t_0) = x_0 \in \mathbb{R}^n, \ A \in \mathbb{R}^{n \times n}$$

Only the first equation here is a differential equation. Once we solve this equation for x(t), we can find y(t) very easily using the second equation. Also, f(x,t) = Ax + Bu satisfies the conditions in Fundamental Theorem for Differential Equations. A unique solution thus exists. The solution of the state-space equations is given in closed form by

$$x(t) = \underbrace{e^{A(t-t_0)}x_0}_{\text{free response}} + \underbrace{\int_{t_0}^t e^{A(t-\tau)}Bu(\tau)d\tau}_{\text{forced response}}$$
(2)

Derivation of the general state-space solution. Since $e^{-At} \neq 0$, $\dot{x}(t) = Ax(t) + Bu(t)$ is equivalent to

$$e^{-At}\dot{x}(t) - e^{-At}Ax(t) = e^{-At}Bu(t)$$

namely

$$\frac{d}{dt} \left(e^{-At} x\left(t \right) \right) = e^{-At} Bu\left(t \right)$$
$$\Leftrightarrow d \left(e^{-At} x\left(t \right) \right) = e^{-At} Bu\left(t \right) dt$$

Integrating both sides from t_0 to t_1 gives

$$e^{-At_{1}}x(t_{1}) = e^{-At_{0}}x(t_{0}) + \int_{t_{0}}^{t_{1}} e^{-At}Bu(t) dt$$

Changing notations from t to τ in the integral yields

$$e^{-At}x(t) = e^{-At_0}x(t_0) + \int_{t_0}^t e^{-A\tau}Bu(\tau)\,d\tau$$

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

In both the free and the forced responses, computing the matrix e^{At} is key. $e^{A(t-t_0)}$ is called the transition matrix, and can be computed using a few handy results in linear algebra.

1.1.4 The State Transition Matrix e^{At}

For the scalar case with $a \in \mathbb{R}$, Tylor expansion gives

$$e^{at} = 1 + at + \frac{1}{2}(at)^2 + \dots + \frac{1}{n!}(at)^n + \dots$$
 (3)

The transition scalar $\Phi(t, t_0) = e^{a(t-t_0)}$ satisfies

$$\begin{split} \Phi(t,t) &= 1 & (\text{transition to itself}) \\ \Phi(t_3,t_2) \Phi(t_2,t_1) &= \Phi(t_3,t_1) & (\text{consecutive transition}) \\ \Phi(t_2,t_1) &= \Phi^{-1}(t_1,t_2) & (\text{reverse transition}) \end{split}$$

For the matrix case with $A \in \mathbb{R}^{n \times n}$

$$e^{At} = I + At + \frac{1}{2}A^2t^2 + \dots + \frac{1}{n!}A^nt^n + \dots$$
(4)

As I and A^i are matrices of dimension $n \times n$, we confirm that $e^{At} \in \mathbb{R}^{n \times n}$. The state transition matrix $\Phi(t, t_0) = e^{A(t-t_0)}$ satisfies

$$e^{A0} = I_n$$

 $e^{At_1}e^{At_2} = e^{A(t_1+t_2)}$
 $e^{-At} = [e^{At}]^{-1}.$

Similar to the scalar case, it can be shown that

$$\Phi(t,t) = I$$

$$\Phi(t_3,t_2)\Phi(t_2,t_1) = \Phi(t_3,t_1)$$

$$\Phi(t_2,t_1) = \Phi^{-1}(t_1,t_2).$$

Note, however, that $e^{At}e^{Bt} = e^{(A+B)t}$ if and only if AB = BA. (Check by using Tylor expansion.)

When A is a diagonal or Jordan matrix, the Tylor expansion formula readily generates e^{At} :

Diagonal matrix
$$A = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}$$
. In this case $A^n = \begin{bmatrix} \lambda_1^n & 0 & 0 \\ 0 & \lambda_2^n & 0 \\ 0 & 0 & \lambda_3^n \end{bmatrix}$ is also diagonal and hence

$$e^{At} = I + At + \frac{1}{2}A^{2}t^{2} + \dots + \frac{1}{n!}A^{n}t^{n} + \dots$$
(5)

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} \lambda_1 t & 0 & 0 \\ 0 & \lambda_2 t & 0 \\ 0 & 0 & \lambda_3 t \end{bmatrix} + \begin{bmatrix} \frac{1}{2}\lambda_1^2 t^2 & 0 & 0 \\ 0 & \frac{1}{2}\lambda_2^2 t^2 & 0 \\ 0 & 0 & \frac{1}{2}\lambda_3^2 t^2 \end{bmatrix} + \dots$$
(6)

$$= \begin{bmatrix} 1+\lambda_1t+\frac{1}{2}\lambda_1^2t^2+\dots & 0 & 0\\ 0 & 1+\lambda_2t+\frac{1}{2}\lambda_2^2t^2+\dots & 0\\ 0 & 0 & 1+\lambda_3t+\frac{1}{2}\lambda_3^2t^2+\dots \end{bmatrix}$$
(7)

$$= \begin{bmatrix} e^{\lambda_1 t} & 0 & 0\\ 0 & e^{\lambda_2 t} & 0\\ 0 & 0 & e^{\lambda_3 t} \end{bmatrix}.$$
 (8)

Jordan canonical form
$$A = \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix}$$
. Decompose
$$A = \underbrace{\begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix}}_{\lambda I_3} + \underbrace{\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & \lambda \\ 0 & 0 & \lambda \end{bmatrix}}_{N}$$

Then

$$e^{At} = e^{(\lambda I_3 t + Nt)}.$$

As $(\lambda It)(Nt) = \lambda Nt^2 = (Nt)(\lambda It)$, we have $e^{At} = e^{\lambda It}e^{Nt} = e^{\lambda t}e^{Nt}$. Also, N has the special property of $N^3 = N^4 = \cdots = 0I_3$, yielding $e^{Nt} = I + Nt + \frac{1}{-}N^2t^2 = \begin{bmatrix} 1 & t & \frac{t^2}{2} \\ 0 & 1 & t \end{bmatrix}$

Thus

$$e^{At} = \underbrace{\begin{bmatrix} e^{\lambda t} & te^{\lambda t} & \frac{t^2}{2}e^{\lambda t} \\ 0 & e^{\lambda t} & te^{\lambda t} \end{bmatrix}}_{0 & 0 & e^{\lambda t}}.$$
(9)

Remark 2 (Nilpotent matrices). The matrix

$$N = \left[\begin{array}{rrr} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{array} \right]$$

is a *nilpotent* matrix that equals to zero when raised to a positive integral power. ("nil" \sim zero; "potent" \sim taking powers.) When taking powers of N, the off-diagonal 1 elements march to the top right corner and finally vanish.

Example. Consider a mass moving on a straight line with zero friction and no external force. Let x_1 and x_2 be be the position and the velocity of the mass, respectively. The state-space description of the system is

$$\frac{d}{dt} \left[\begin{array}{c} x_1 \\ x_2 \end{array} \right] = \underbrace{\left[\begin{array}{c} 0 & 1 \\ 0 & 0 \end{array} \right]}_A \left[\begin{array}{c} x_1 \\ x_2 \end{array} \right].$$

Then $x(t) = e^{At}x(0)$ and

$$e^{At} = I + \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} t + \frac{1}{2!} \underbrace{\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}}_{=\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}} t^2 + \dots \underbrace{= \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}}_{=\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}}.$$

Columns of the state-transition matrix. We discuss an intuition of the matrix entries in the e^{At} matrix. Consider the system equation

$$\dot{x} = Ax = \begin{bmatrix} 0 & 1 \\ 0 & -1 \end{bmatrix} x, \quad x(0) = x_0,$$

with the solution

$$x(t) = e^{At}x(0) = \begin{bmatrix} | & | & | \\ a_1(t) & | & a_2(t) \\ | & | & | \end{bmatrix} \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} = a_1(t)x_1(0) + a_2(t)x_2(0),$$
(10)

where $a_1(t)$ and $a_2(t)$ are columns of e^{At} . They satisfy

$$x(0) = \begin{bmatrix} 1\\0 \end{bmatrix} \Rightarrow x(t) = a_1(t),$$
$$x(0) = \begin{bmatrix} 0\\1 \end{bmatrix} \Rightarrow x(t) = a_2(t).$$

Hence, we can obtain e^{At} from the following, without using explicitly the Tylor expansion,

1. write out
$$\frac{\dot{x}_1(t) = x_2(t)}{\dot{x}_2(t) = -x_2(t)} \Rightarrow \frac{x_1(t) = e^{0t}x_1(0) + \int_0^t e^{0(t-\tau)}x_2(\tau)d\tau}{x_2(t) = e^{-t}x_2(0)} \neq \frac{x_1(t) = e^{0t}x_1(0) + \int_0^t e^{-\tau}x_2(0)d\tau}{x_2(t) = e^{-t}x_2(0)}$$

2. let
$$x(0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$
, then $\begin{array}{c} x_1(t) \equiv 1 \\ x_2(t) \equiv 0 \end{array}$, namely $x(t) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$
3. let $x(0) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, then $x_2(t) = e^{-t}$ and $x_1(t) = 1 - e^{-t}$, or more compactly, $x(t) = \begin{bmatrix} 1 - e^{-t} \\ e^{-t} \end{bmatrix}$

4. using the property of (10), write out directly

$$e^{At} = \left[\begin{array}{cc} 1 & 1 - e^{-t} \\ 0 & e^{-t} \end{array} \right]$$

Exercise. Use the above method to compute e^{At} where

$$A = \left[\begin{array}{ccc} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{array} \right].$$

1.2 Discrete-Time LTI State-Space Solutions

For the discrete-time system

$$x(k+1) = Ax(k) + Bu(k), \ x(0) = x_0,$$

iteration of the state-space equation gives

$$x(k+1) = Ax(k) + Bu(k)$$

$$\begin{bmatrix} u(k_0) \end{bmatrix}$$
(11)

$$\Rightarrow x(k) = A^{k-k_0} x(k_o) + \begin{bmatrix} A^{k-k_0-1} B & A^{k-k_0-2} B & \cdots & B \end{bmatrix} \begin{bmatrix} u(k_0) \\ u(k_0+1) \\ \vdots \\ u(k-1) \end{bmatrix}$$
(12)

$$\Leftrightarrow x(k) = \underbrace{A^{k-k_0}x(k_o)}_{\text{free response}} + \underbrace{\sum_{j=k_0}^{k-1} A^{k-1-j}Bu(j)}_{\text{forced response}}$$
(13)

where the transition matrix is defined by $\Phi(k, j) = A^{k-j}$ and satisfies

1.2.1 The State Transition Matrix A^k

Similar to the continuous-time case, when A is a diagonal or Jordan matrix, the Tylor expansion formula readily generates A^k . We have

• Diagonal matrix
$$A = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}$$
: $A^k = \begin{bmatrix} \lambda_1^k & 0 & 0 \\ 0 & \lambda_2^k & 0 \\ 0 & 0 & \lambda_3^k \end{bmatrix}$
• Jordan canonical form $A = \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix} = \underbrace{\begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix}}_{\lambda I_3} + \underbrace{\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}}_{N}$: With the nilpotent N and the

commutative property $(\lambda I_3) N = N (\lambda I_3)$, we have

$$\begin{aligned} A^{k} &= (\lambda I_{3} + N)^{k} = (\lambda I_{3})^{k} + k (\lambda I_{3})^{k-1} N + \underbrace{\begin{pmatrix} k \\ 2 \end{pmatrix}}_{2 \text{ combination}} (\lambda I_{3})^{k-2} N^{2} + \underbrace{\begin{pmatrix} k \\ 3 \end{pmatrix}}_{N^{3} = N^{4} = \dots = 0I_{3}} \\ \overset{N^{3} = N^{4} = \dots = 0I_{3}}{\overset{N^{3} = N^{4} = \dots = 0I_{3}} \end{aligned}$$
$$= \begin{bmatrix} \lambda^{k} & 0 & 0 \\ 0 & \lambda^{k} \\ 0 & 0 & \lambda^{k} \end{bmatrix} + k\lambda^{k-1} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} + \frac{k(k-1)}{2}\lambda^{k-2} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$
$$= \begin{bmatrix} \lambda^{k} & k\lambda^{k-1} & \frac{1}{2!}k (k-1)\lambda^{k-2} \\ 0 & \lambda^{k} & k\lambda^{k-1} \\ 0 & 0 & \lambda^{k} \end{bmatrix}$$

Exercise. Show that

$$A = \begin{bmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{bmatrix} \Rightarrow A^{k} = \begin{bmatrix} \lambda^{k} & k\lambda^{k-1} & \frac{1}{2!}k(k-1)\lambda^{k-2} & \frac{1}{3!}k(k-1)(k-2)\lambda^{k-3} \\ 0 & \lambda^{k} & k\lambda^{k-1} & \frac{1}{2!}k(k-1)\lambda^{k-2} \\ 0 & 0 & \lambda^{k} & k\lambda^{k-1} \\ 0 & 0 & 0 & \lambda^{k} \end{bmatrix}$$

1.3 Explicit Computation of the State Transition Matrix e^{At}

General matrices may have structures other than the diagonal and Jordan canonical forms. However, via similar transformation, we can readily transform a general matrix to a diagonal or Jordan form under a different choice of state vectors.

Principle Concept.

1. Given

$$\dot{x}(t) = Ax(t) + Bu(t), \ x(t_0) = x_0 \in \mathbb{R}^n, \ A \in \mathbb{R}^{n \times n}$$

we will find a nonsingular matrix $T \in \mathbb{R}^{n \times n}$ such that a coordinate transformation defined by $x(t) = Tx^*(t)$ yields

$$\frac{d}{dt} (Tx^{*}(t)) = ATx^{*}(t) + Bu(t)$$
$$\frac{d}{dt}x^{*}(t) = \underbrace{T^{-1}AT}_{\triangleq \Lambda}x^{*}(t) + \underbrace{T^{-1}B}_{B^{*}}u(t), \ x^{*}(0) = T^{-1}x_{0}$$

where Λ is diagonal or in Jordan form.

- 2. Now $x^*(t)$ can be solved easily, and the free response is $x^*(t) = e^{\Lambda t} x^*(0)$. For example, when $\Lambda = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$, we would readily obtain $x^*(t) = \begin{bmatrix} e^{\lambda_1 t} & 0 \\ 0 & e^{\lambda_2 t} \end{bmatrix} \begin{bmatrix} x_1^*(0) \\ x_2^*(0) \end{bmatrix} = \begin{bmatrix} e^{\lambda_1 t} x_1^*(0) \\ e^{\lambda_2 t} x_2^*(0) \end{bmatrix}$.
- 3. As $x(t) = Tx^*(t)$, the above implies

$$x(t) = Te^{\Lambda t}T^{-1}x_0$$

4. From the original state-space description, $x(t) = e^{At}x_0$. Hence

$$e^{At} = T e^{\Lambda t} T^{-1}$$

Existence of Solutions. The solution of T comes from the theory of eigenvalues and eigenvectors in linear algebra.

More generally If two matrices $A, B \in \mathbb{C}^{n \times n}$ are similar: $A = TBT^{-1}, T \in \mathbb{C}^{n \times n}$, then

- their A^n and B^n are also similar: e.g., $AA = TBT^{-1}TBT^{-1} = TB^2T^{-1}$
- $\bullet\,$ their exponential matrices are also similar

$$e^{At} = Te^{Bt}T^{-1}$$

as

$$Te^{Bt}T^{-1} = T(I + Bt + \frac{1}{2}B^{2}t^{2} + \dots)T^{-1} = TIT^{-1} + TBtT^{-1} + \frac{1}{2}TB^{2}t^{2}T^{-1} + \dots$$
$$= I + At + \frac{1}{2}A^{2}t^{2} + \dots = e^{At}$$

Eigenvalues and Eigenvectors. The principle concept of computing e^{At} in this section relies on the similarity transform $\Lambda = T^{-1}AT$, where Λ is structurally simple: i.e., in diagonal or Jordan form. We already observed the resulting convenience in computing $x^*(t) = e^{\Lambda t}x^*(0) \stackrel{\text{e.g.}}{=} \begin{bmatrix} e^{\lambda_1 t}x_1^*(0) \\ e^{\lambda_2 t}x_2^*(0) \end{bmatrix}$. Under the coordinate transformation defined by $x(t) = Tx^*(t)$, we then have

$$x(t) = Te^{\Lambda t}x^*(0) \stackrel{\text{e.g.}}{=} \underbrace{[t_1, t_2]}_{T} \begin{bmatrix} e^{\lambda_1 t}x_1^*(0) \\ e^{\lambda_2 t}x_2^*(0) \end{bmatrix} = e^{\lambda_1 t}x_1^*(0)t_1 + e^{\lambda_2 t}x_2^*(0)t_2$$

in other words, the state trajectory is conveniently decomposed into two modes along the directions defined by t_1 and t_2 , the column vectors of T.

In practice, Λ and T are obtained using the tools of eigenvalues and eigenvectors.

For $A \in \mathbb{R}^{n \times n}$, an eigenvalue $\lambda \in \mathcal{C}$ of A is the solution to the characteristic equation

$$\det\left(A - \lambda I\right) = 0 \tag{14}$$

The corresponding eigenvectors are the nonzero solutions to

$$At = \lambda t \Leftrightarrow (A - \lambda I) t = 0 \tag{15}$$

The case with distinct eigenvalues (diagonalization). When $A \in \mathbb{R}^{n \times n}$ has n distinct eigenvalues such that

$$Ax_1 = \lambda_1 x_1$$
$$Ax_2 = \lambda_2 x_2$$
$$\vdots$$
$$Ax_n = \lambda_n x_n$$

we can write the above as

$$A\underbrace{[x_1, x_2, \dots, x_n]}_{\triangleq T} = [\lambda_1 x_1, \lambda_2 x_2, \dots, \lambda_n x_n] = [x_1, x_2, \dots, x_n] \underbrace{\begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{bmatrix}}_{A}$$

The matrix $[x_1, x_2, \ldots, x_n]$ is square. From linear algebra, the eigenvectors are linearly independent and $[x_1, x_2, \ldots, x_n]$ is invertible. Hence

$$A = T\Lambda T^{-1}, \ \Lambda = T^{-1}AT$$

Example 1. Mechanical system with strong damping

Consider a spring-mass-damper system with m = 1, k = 2, b = 3. Let x_1 and x_2 be the position and velocity of the mass, respectively. We have

- Find eigenvalues: $det(A \lambda I) = det \begin{bmatrix} -\lambda & 1 \\ -2 & -\lambda 3 \end{bmatrix} = (\lambda + 2)(\lambda + 1) \Rightarrow \lambda_1 = -2, \lambda_2 = -1$
- Find associate eigenvectors:

$$-\lambda_1 = -2: (A - \lambda_1 I) t_1 = 0 \Rightarrow t_1 = \begin{bmatrix} 1 \\ -2 \\ -\lambda_1 = -1: (A - \lambda_2 I) t_2 = 0 \Rightarrow t_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

- Define T and A: $T = \begin{bmatrix} t_1 & t_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -2 & -1 \end{bmatrix}$, $\Lambda = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} = \begin{bmatrix} -2 & 0 \\ 0 & -1 \end{bmatrix}$
- Compute $T^{-1} = \begin{bmatrix} 1 & 1 \\ -2 & -1 \end{bmatrix}^{-1} = \begin{bmatrix} -1 & -1 \\ 2 & 1 \end{bmatrix}$
- Compute $e^{At} = Te^{\Lambda t}T^{-1} = T\begin{bmatrix} e^{-2t} & 0\\ 0 & e^{-1t} \end{bmatrix}T^{-1} = \begin{bmatrix} -e^{-2t} + 2e^{-t} & -e^{-2t} + e^{-t}\\ 2e^{-2t} 2e^{-t} & 2e^{-2t} e^{-t} \end{bmatrix}$

Physical interpretations Let us revisit the intuition at the beginning of this subsection:

- $x(t) = e^{\lambda_1 t} x_1^*(0) t_1 + e^{\lambda_2 t} x_2^*(0) t_2$ decomposes the state trajectory into two modes along the direction of the two eigenvectors t_1 and t_2 .
- The two modes are scaled by $x_1^*(0)$ and $x_2^*(0)$ defined from $x(0) = Tx^*(0)$, or more explicitly, $x(0) = [t_1, t_2][x_1^*(0), x_2^*(0)]^T = x_1^*(0)t_1 + x_2^*(0)t_2$. This is nothing but decomposing x(0) into the sum of two vectors along the directions of the eigenvectors; and $x_1^*(0)$ and $x_2^*(0)$ are the coefficients of the decomposition!



• If the initial condition x(0) is aligned with one eigenvector, say, t_1 , then $x_2^*(0) = 0$. The decomposition $x(t) = e^{\lambda_1 t} x_1^*(0) t_1 + e^{\lambda_2 t} x_2^*(0) t_2$ then dictates that x(t) will stay in the direction of t_1 . In other words, if the state initiates along the direction of one eigenvector, then the free response will stay in that direction without "making turns". If $\lambda_1 < 0$, then x(t) will move towards the origin of the state space; if $\lambda_1 = 0$, x(t) will stay at the initial point; and if positive, x(t) will move away from the origin along t_1 . Furthermore, the magnitude of λ_1 determines the speed of response.



The case with complex eigenvalues Consider the undamped spring-mass system

$$\frac{d}{dt} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}}_A \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \ \det(A - \lambda I) = \lambda^2 + 1 = 0 \Rightarrow \lambda_{1,2,} = \pm j.$$

The eigenvectors are

$$\lambda_1 = j: \ (A - jI)t_1 = 0 \Rightarrow t_1 = \begin{bmatrix} 1\\ j \end{bmatrix}$$
$$\lambda_2 = -j: \ (A + jI)t_2 = 0 \Rightarrow t_2 = \begin{bmatrix} 1\\ -j \end{bmatrix} \text{ (complex conjugate of } t_1\text{)}.$$

Hence

$$T = \begin{bmatrix} 1 & 1\\ j & -j \end{bmatrix}, \ T^{-1} = \frac{1}{2} \begin{bmatrix} 1 & -j\\ 1 & j \end{bmatrix}, \ e^{At} = Te^{\Lambda t}T^{-1} = T\begin{bmatrix} e^{jt} & 0\\ 0 & e^{-jt} \end{bmatrix} T^{-1} = \begin{bmatrix} \cos t & \sin t\\ -\sin t & \cos t \end{bmatrix}.$$

As an exercise, for a general $A \in \mathbb{R}^{2\times 2}$ with complex eigenvalues $\sigma \pm j\omega$, you can show that by using $T = [t_R, t_I]$ where t_R and t_I are the real and the imaginary parts of t_1 , an eigenvector associated with $\lambda_1 = \sigma + j\omega$, $x = Tx^*$ transforms $\dot{x} = Ax$ to

$$\dot{x}^*(t) = \begin{bmatrix} \sigma & \omega \\ -\omega & \sigma \end{bmatrix} x^*(t)$$

and

$$e^{\begin{bmatrix} \sigma & \omega \\ -\omega & \sigma \end{bmatrix}^{t}} = \begin{bmatrix} e^{\sigma t} \cos \omega t & e^{\sigma t} \sin \omega t \\ -e^{\sigma t} \sin \omega t & e^{\sigma t} \cos \omega t \end{bmatrix}.$$



The case with repeated eigenvalues, via generalized eigenvectors. Consider $A = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$, which has two repeated eigenvalues $\lambda(A) = 2$ and

$$(A - \lambda I) t_1 = 0 \Rightarrow t_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

No other linearly independent eigenvectors exist. How do we go further? As A is already very similar to the Jordan form, we try instead

$$A\begin{bmatrix} t_1 & t_2 \end{bmatrix} = \begin{bmatrix} t_1 & t_2 \end{bmatrix} \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix},$$

which requires $At_2 = t_1 + \lambda t_2$, i.e.,

$$(A - \lambda I) t_2 = t_1 \Leftrightarrow \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix} t_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$
$$\Rightarrow t_2 = \begin{bmatrix} 0 \\ 0.5 \end{bmatrix}$$

 t_2 is linearly independent from t_1 . Together, t_1 and t_2 span the 2-dimensional vector space. As such, t_2 is called a generalized eigenvector.

For general 3×3 matrices with $\det(\lambda I - A) = (\lambda - \lambda_m)^3$, i.e., $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_m$, we look for T such that

$$A = TJT^{-1}$$

where J has three canonical forms:

$$i), \begin{bmatrix} \lambda_m & 0 & 0\\ 0 & \lambda_m & 0\\ 0 & 0 & \lambda_m \end{bmatrix}, ii), \begin{bmatrix} \lambda_m & 1 & 0\\ 0 & \lambda_m & 0\\ 0 & 0 & \lambda_m \end{bmatrix} \text{ or } \begin{bmatrix} \lambda_m & 0 & 0\\ 0 & \lambda_m & 1\\ 0 & 0 & \lambda_m \end{bmatrix}, iii), \begin{bmatrix} \lambda_m & 1 & 0\\ 0 & \lambda_m & 1\\ 0 & 0 & \lambda_m \end{bmatrix}.$$

- Case i): this happens when A has three linearly independent eigenvectors, i.e., $(A \lambda_m I)t = 0$ yields t_1, t_2 , and t_3 that span the 3-d vector space. This happens when nullity $(A \lambda_m I) = 3$, namely, rank $(A \lambda_m I) = 3 1$ nullity $(A \lambda_m I) = 0$.
- Case ii): this happens when $(A \lambda_m I)t = 0$ yields two linearly independent solutions, i.e., when nullity $(A \lambda_m I) = 2$. We then have, e.g.,

$$A[t_1, t_2, t_3] = [t_1, t_2, t_3] \begin{bmatrix} \lambda_m & 1 & 0\\ 0 & \lambda_m & 0\\ 0 & 0 & \lambda_m \end{bmatrix} \Leftrightarrow [\lambda_m t_1, t_1 + \lambda_m t_2, \lambda_m t_3] = [At_1, At_2, At_3]$$

 t_1 and t_3 are the directly computed eigenvectors. For the generalized eigenvector t_2 , the second column of the equality gives

$$(A - \lambda_m I) t_2 = t_1$$

• Case iii): this is for the case when $(A - \lambda_m I)t = 0$ yields only one linearly independent solution, i.e., when nullity $(A - \lambda_m I) = 1$. We then have,

$$A[t_1, t_2, t_3] = [t_1, t_2, t_3] \begin{bmatrix} \lambda_m & 1 & 0\\ 0 & \lambda_m & 1\\ 0 & 0 & \lambda_m \end{bmatrix} \Leftrightarrow [\lambda_m t_1, t_1 + \lambda_m t_2, t_2 + \lambda_m t_3] = [At_1, At_2, At_3]$$

yielding

$$(A - \lambda_m I) t_1 = 0$$
$$(A - \lambda_m I) t_2 = t_1$$
$$(A - \lambda_m I) t_3 = t_2$$

where t_2 and t_3 are the generalized eigenvectors.

Example 2. Consider

$$A = \begin{bmatrix} -1 & 1\\ -1 & 1 \end{bmatrix}, \text{ det} (A - \lambda I) = (\lambda + 1) (\lambda - 1) - 1 = \lambda^2 \Rightarrow \lambda_1 = \lambda_2 = 0.$$

Two repeated eigenvalues with $rank(A - 0I) = 1 \Rightarrow$ only one linearly independent eigenvector:

$$(A - 0I) t_1 = 0 \Rightarrow t_1 = \begin{bmatrix} 1\\1 \end{bmatrix}$$

Generalized eigenvector:

$$(A - 0I) t_2 = t_1 \Rightarrow t_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Coordinate transform matrix:

$$T = [t_1, t_2] = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \ T^{-1} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix},$$

$$J = T^{-1}AT = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \ e^{At} = Te^{Jt}T^{-1} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 1-t & t \\ -t & 1+t \end{bmatrix}.$$

The first eigenvector implies that if $x_1(0) = x_2(0)$ then the response is characterized by $e^{0t} = 1$, i.e., $x_1(t) = x_1(0) = x_2(0) = x_2(t)$. This makes sense because $\dot{x}_1 = -x_1 + x_2$ from the state equation.

Example 3 (Multiple eigenvectors). Obtain the eigenvalues and eigenvectors of

$$A = \left[\begin{array}{rrr} -2 & 2 & -3 \\ 2 & 1 & -6 \\ -1 & -2 & 0 \end{array} \right]$$

Analogous procedures give that

$$\lambda_1 = 5, \ \lambda_2 = \lambda_3 = -3.$$

So there are repeated eigenvalues. For $\lambda_1 = 5$, $(A - 5I)t_1 = 0$ gives

$$\begin{bmatrix} -7 & 2 & -3\\ 2 & -4 & -6\\ -1 & -2 & -5 \end{bmatrix} t_1 = 0 \Rightarrow \begin{bmatrix} 1 & 0 & 1\\ 0 & 1 & 2\\ 1 & 0 & 1 \end{bmatrix} t_1 = 0 \Rightarrow t_1 = \begin{bmatrix} 1\\ 2\\ -1 \end{bmatrix}.$$

For $\lambda_2 = \lambda_3 = -3$, the characteristic matrix is

$$A + 3I = \left[\begin{array}{rrr} 1 & 2 & -3 \\ 2 & 4 & -6 \\ -1 & -2 & 3 \end{array} \right]$$

The second row is the first row multiplied by 2. The third row is the negative of the first row. So the characteristic matrix has only rank 1. The characteristic equation

$$(A - \lambda_2 I) t = 0$$

has two linearly independent solutions

$$\left[\begin{array}{c} -2\\1\\0\end{array}\right], \left[\begin{array}{c} 3\\0\\1\end{array}\right]$$

Then

$$T = \begin{bmatrix} 1 & -2 & 3 \\ 2 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}, \ J = \begin{bmatrix} 5 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & -3 \end{bmatrix}.$$

Physical interpretation. When $\dot{x} = Ax$, $A = TJT^{-1}$ with $J = \begin{bmatrix} \lambda_m & 1 & 0 \\ 0 & \lambda_m & 0 \\ 0 & 0 & \lambda_m \end{bmatrix}$, we have

$$x(t) = e^{At}x(0) = T \begin{bmatrix} e^{\lambda_m t} & te^{\lambda_m t} & 0\\ 0 & e^{\lambda_m t} & 0\\ 0 & 0 & e^{\lambda_m t} \end{bmatrix} T^{-1}x(0) = T \begin{bmatrix} e^{\lambda_m t} & te^{\lambda_m t} & 0\\ 0 & e^{\lambda_m t} & 0\\ 0 & 0 & e^{\lambda_m t} \end{bmatrix} \mathcal{T}^{-1}T x^I(0)$$

If the initial condition is in the direction of t_1 , i.e., $x^*(0) = [x_1^*(0), 0, 0]^T$ and $x_1^*(0) \neq 0$, the above equation yields $x(t) = x_1^*(0)t_1e^{\lambda_m t}$. If x(0) starts in the direction of t_2 , i.e., $x^*(0) = [0, x_2^*(0), 0]^T$, then $x(t) = x_2^*(0)(t_1te^{\lambda_m t} + t_2e^{\lambda_m t})$. In this case, the response does not remain in the direction of t_2 but is confined in the subspace spanned by t_1 and t_2 .

Exercise 1. Obtain eigenvalues of J and e^{Jt} by inspection:

$$J = \begin{bmatrix} -1 & 0 & 0 & 0 & 0 \\ 0 & -2 & 1 & 0 & 0 \\ 0 & -1 & -2 & 0 & 0 \\ 0 & 0 & 0 & -3 & 1 \\ 0 & 0 & 0 & 0 & -3 \end{bmatrix}$$

1.4 Explicit Computation of the State Transition Matrix A^k

Everything in computing the similarity transform $A = T\Lambda T^{-1}$ or $A = TJT^{-1}$ applies to the discrete-time case. The state transition matrix in this case is

$$A^{k} = T\Lambda^{k}T^{-1} \text{ or } A^{k} = TJ^{k}T^{-1}.$$

You should be able to derive these results:



Exercise 3. Show that

$$J = \begin{bmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{bmatrix} \Rightarrow J^{k} = \begin{bmatrix} \lambda^{k} & k\lambda^{k-1} & \frac{1}{2!}k\left(k-1\right)\lambda^{k-2} & \frac{1}{3!}k\left(k-1\right)\left(k-2\right)\lambda^{k-3} \\ 0 & \lambda^{k} & k\lambda^{k-1} & \frac{1}{2!}k\left(k-1\right)\lambda^{k-2} \\ 0 & 0 & \lambda^{k} & k\lambda^{k-1} \\ 0 & 0 & 0 & \lambda^{k} \end{bmatrix}.$$

1.5 Transition Matrix via Inverse Transformation

We have now

Continuous-time systemDiscrete-time systemstate equation
$$\dot{x}(t) = Ax(t) + Bu(t), x(0) = x_0$$
 $x(k+1) = Ax(k) + Bu(k), x(0) = x_0$ solution $x(t) = \underbrace{e^{At}x(0)}_{\text{free response}} + \underbrace{\int_{0}^{t} e^{A(t-\tau)}Bu(\tau)d\tau}_{\text{forced response}}$ $x(k) = \underbrace{A^{k}x(0)}_{\text{free response}} + \underbrace{\sum_{j=0}^{(k-1)}A^{(k-1-j)}Bu(j)}_{\text{forced response}}$ transition matrix e^{At} A^{k}

We also know from Laplace transform, that

$$\dot{x}(t) = Ax(t) + Bu(t)$$
$$X(s) = \underbrace{(sI - A)^{-1} x(0)}_{\text{free response}} + \underbrace{(sI - A)^{-1} BU(s)}_{\text{free response}}$$

Comparing x(t) and X(s) gives

$$e^{At} = \mathcal{L}^{-1}\left\{ (sI - A)^{-1} \right\}$$
(16)

Example 4. Consider $A = \begin{bmatrix} \sigma & \omega \\ -\omega & \sigma \end{bmatrix}$. We have

$$e^{At} = \mathcal{L}^{-1} \begin{bmatrix} s - \sigma & -\omega \\ \omega & s - \sigma \end{bmatrix}^{-1} = \mathcal{L}^{-1} \left\{ \frac{1}{(s - \sigma)^2 + \omega^2} \begin{bmatrix} s - \sigma & \omega \\ -\omega & s - \sigma \end{bmatrix} \right\}$$
$$= e^{\sigma t} \begin{bmatrix} \cos(\omega t) & \sin(\omega t) \\ -\sin(\omega t) & \cos(\omega t) \end{bmatrix}$$

Similarly, for the discrete time case, we have $X(z) = (zI - A)^{-1} zx(0) + (zI - A)^{-1} BU(s)$ and

$$A^{k} = \mathcal{Z}^{-1}\left\{ (zI - A)^{-1} z \right\}$$
(17)

Example 5. Consider $A = \begin{bmatrix} \sigma & \omega \\ -\omega & \sigma \end{bmatrix}$. We have

$$\begin{split} A^{k} &= \mathcal{Z}^{-1} \left\{ z \begin{bmatrix} z - \sigma & -\omega \\ \omega & z - \sigma \end{bmatrix}^{-1} \right\} = \mathcal{Z}^{-1} \left\{ \frac{z}{(z - \sigma)^{2} + \omega^{2}} \begin{bmatrix} z - \sigma & \omega \\ -\omega & z - \sigma \end{bmatrix} \right\} \\ &= \mathcal{Z}^{-1} \left\{ \frac{z}{z^{2} - 2r\cos\theta z + r^{2}} \begin{bmatrix} z - r\cos\theta & r\sin\theta \\ -r\sin\theta & z - r\cos\theta \end{bmatrix} \right\}, \ r = \sqrt{\sigma^{2} + \omega^{2}}, \ \theta = \tan^{-1} \frac{\omega}{\sigma} \\ &= r^{k} \begin{bmatrix} \cos k\theta & \sin k\theta \\ -\sin k\theta & \cos k\theta \end{bmatrix} \end{split}$$

Example 6. Consider $A = \begin{bmatrix} 0.7 & 0.3 \\ 0.1 & 0.5 \end{bmatrix}$. We have

$$(zI - A)^{-1} z = \begin{bmatrix} \frac{z(z-0.5)}{(z-0.8)(z-0.4)} & \frac{0.3z}{(z-0.8)(z-0.4)} \\ \frac{0.1z}{(z-0.8)(z-0.4)} & \frac{z(z-0.7)}{(z-0.8)(z-0.4)} \end{bmatrix} = \begin{bmatrix} \frac{0.75z}{z-0.8} + \frac{0.25z}{z-0.4} & \frac{0.75z}{z-0.8} - \frac{0.75z}{z-0.4} \\ \frac{0.25z}{z-0.8} - \frac{0.25z}{z-0.4} & \frac{0.25z}{z-0.8} + \frac{0.75z}{z-0.4} \end{bmatrix}$$

$$\Rightarrow A^{k} = \begin{bmatrix} 0.75 (0.8)^{k} + 0.25 (0.4)^{k} & 0.75 (0.8)^{k} - 0.75 (0.4)^{k} \\ 0.25 (0.8)^{k} - 0.25 (0.4)^{k} & 0.25 (0.8)^{k} + 0.75 (0.4)^{k} \end{bmatrix}$$

ME 547: Linear Systems State-Space Discretization

Xu Chen

University of Washington

UW Linear Systems (X. Chen, ME547)

SS discretization

Motivation

- For mechanical systems, physics often give differential-equation models.
- When implementing controls digitally (e.g., on a microcontroller), a continuous-time system must be represented in the discrete-time domain.
- Sampler: converts a time function into a discrete sequence, e.g.,





where $y(k) \triangleq y(t_k) = y(Tk)$

1/8

Signal holding

Zero-order Hold (ZOH): converts a sequence into a "stair-case" time function, e.g.,



where u(t) = u(k) for $k \in [kT, (k+1)T)$.

UW Linear Systems (X. Chen, ME547)

SS discretization

3/8

Problem definition

Consider a continuous time system preceded by a ZOH:



• Goal: to obtain the model between $u(t_k)$ and $x(t_k)$.

Solution

$$\underbrace{u(t_k) \qquad u(t)}_{\text{Zero Order Hold}} \underbrace{\frac{dx}{dt} = Ax + Bu}_{x(t_k)}$$

1. Starting from t_k , the solution of $\dot{x} = Ax + Bu$ at time t_{k+1} is

$$x(t_{k+1}) = e^{A(t_{k+1}-t_k)}x(t_k) + \int_{t_k}^{t_{k+1}} e^{A(t_{k+1}-\tau_o)}Bu(\tau_o)d\tau_o$$

= $e^{A(t_{k+1}-t_k)}x(t_k) + u(t_k) \underbrace{\int_{t_k}^{t_{k+1}} e^{A(t_{k+1}-\tau_o)}Bd\tau_o}_{=\int_T^0 e^{A\eta}Bd(-\eta) = -\int_T^0 e^{A\eta}Bd\eta}$
2. Noting $-\int_T^0 e^{A\eta}Bd\eta = \int_0^T e^{A\tau}Bd\tau$ and denoting t_k as k yield

$$x(k+1) = A_d x(k) + B_d u(k), \ A_d = e^{AT}, \ B_d = \int_0^T e^{A\tau} B d\tau$$

UW Linear Systems (X. Chen, ME547)

SS discretization

5/8

Analysis Mapping of eigenvalues

- e^{At} has the same eigenvalues as e^{Λt}: Λ is in diagonal or Jordan form with diagonal elements formed by eigenvalues of A
- ► ⇒ eigenvalues of $A_d = e^{AT}$ are $e^{\lambda_i T}$'s where λ_i is an eigenvalue of A
- same conclusion can be drawn from the spectral mapping theorem

Theorem (Spectral Mapping Theorem)

Take any $A \in \mathbb{C}^{n \times n}$ and a polynomial function $f(\cdot)$ (more generally, analytic functions). Then

$$\operatorname{eig}(f(A)) = f(\operatorname{eig}(A))$$

Example (Compute the eigenvalues)

$$A = \begin{bmatrix} 99.8 & 2000 \\ -2000 & 99.8 \end{bmatrix}$$
Solution:

$$A = 99.8I + 2000 \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \Rightarrow \lambda(A) = 99.8 + 2000\lambda \left\{ \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \right\}$$
$$= 99.8 \pm 2000i$$

UW Linear Systems (X. Chen, ME547)

SS discretization

7/8

*Analysis

Explicit form of B_d when A is nonsingular

$$x(k+1) = A_d x(k) + B_d u(k), \ A_d = e^{AT}, \ B_d = \int_0^T e^{A\tau} B d\tau$$

• Using
$$e^{A\tau} = I + A\tau + \frac{1}{2!}A^2\tau^2 + \dots$$
 gives

$$B_{d} = \int_{0}^{T} \left(I + A\tau + \frac{1}{2!}A^{2}\tau^{2} + \dots \right) d\tau B$$
$$= \left(IT + \frac{1}{2}AT^{2} + \frac{1}{3!}A^{2}T^{3} + \dots \right) B$$
$$= A^{-1} \left(e^{AT} - I \right) B = \underline{A^{-1} \left(A_{d} - I \right) B}$$

Essentials of Control Systems Discretization of Continuous-time Transfer-function Systems

Xu Chen

University of Washington

UW Linear Systems (X. Chen, ME547)

TF discretization

Overview

Consider the discrete-time controller implementation scheme

$$u[k] \longrightarrow \overline{ZOH} \xrightarrow{u(t)} \overline{G(s)} \xrightarrow{y(t)} \xrightarrow{\Delta T} y[k]$$

where u[k] and y[k] have the same sampling time.

- for this note, we use [k] to distinguish DT signals from their CT counter parts
- Goal: to derive the transfer function from u[k] to y[k].
- Solution concept: let u[k] be a discrete-time impulse (whose Z transform is 1) and obtain the Z transform of y[k].

1/8

Solution

$$u[k] \longrightarrow \overline{ZOH} \xrightarrow{u(t)} \overline{G(s)} \xrightarrow{y(t)} \xrightarrow{\Delta T} y[k]$$

• u[k] is a DT impulse \Rightarrow after ZOH

$$u(t) = \begin{cases} 1, & 0 \le t < \Delta T \\ 0, & \text{otherwise} \end{cases} = 1(t) - 1(t - \Delta T) \Longrightarrow U(s) = \frac{1 - e^{-s\Delta T}}{s}$$

► Hence

$$y(t) = \mathcal{L}^{-1}\left[G(s)\frac{1-e^{-s\Delta T}}{s}\right] = \mathcal{L}^{-1}\left[G(s)\frac{1}{s}\right] - \mathcal{L}^{-1}\left[G(s)\frac{e^{-s\Delta T}}{s}\right]$$

UW Linear Systems (X. Chen, ME547)

TF discretization

3/8

Solution

$$u[k] \longrightarrow \overline{ZOH} \xrightarrow{u(t)} G(s) \xrightarrow{y(t)} \xrightarrow{\Delta T} y[k]$$

$$y(t) = \mathcal{L}^{-1}\left[G(s)\frac{1-e^{-s\Delta T}}{s}\right] = \mathcal{L}^{-1}\left[G(s)\frac{1}{s}\right] - \mathcal{L}^{-1}\left[G(s)\frac{e^{-s\Delta T}}{s}\right]$$

Sampling y(t) at ΔT and performing Z transform give:

$$G(z) = \mathcal{Z} \left\{ \begin{array}{c} \overbrace{\mathcal{Y}(t)} \\ \mathcal{L}^{-1} \left[G(s) \frac{1}{s} \right] \\ \underbrace{\mathcal{L}^{-1} \left[G(s) \frac{1}{s} \right] }_{\underline{t=k\Delta T}} \\ \underbrace{\mathcal{L}^{-1} \left[G(s) \frac{e^{-s\Delta T}}{s} \right] }_{\underline{t=k\Delta T}} \\ \underbrace{\mathcal{L}^{-1} \left[G(s) \frac{1}{s} \right] }_{\underline{t=k\Delta T}}$$

Solution

$$u[k] \longrightarrow \overline{ZOH} \xrightarrow{u(t)} \overline{G(s)} \xrightarrow{y(t)} \xrightarrow{\Delta T} y[k]$$

Fact

The zero order hold equivalent of G(s) is

$$G(z) = (1 - z^{-1})\mathcal{Z}\left\{ \mathcal{L}^{-1}\left[G(s)\frac{1}{s}\right]\Big|_{t=k\Delta T}\right\}$$

where ΔT is the sampling time.

UW Linear Systems (X. Chen, ME547)

TF discretization

Example

Obtain the ZOH equivalent of

$$G(s) = rac{a}{s+a}$$

Following the discretization procedures we have $\frac{G(s)}{s} = \frac{a}{s(s+a)} = \frac{1}{s} - \frac{1}{s+a}$ and hence

$$\mathcal{L}^{-1}\left\{rac{G(s)}{s}
ight\}=1(t)-e^{-at}1(t)$$

Sampling at ΔT gives $1[k] - e^{-ak\Delta T}1[k]$, whose Z transform is

$$\frac{z}{z-1} - \frac{z}{z-e^{-a\Delta T}} = \frac{z(1-e^{-a\Delta T})}{(z-1)(z-e^{-a\Delta T})}$$

Hence the ZOH equivalent is

$$(1-z^{-1})\frac{z(1-e^{-a\Delta T})}{(z-1)(z-e^{-a\Delta T})} = \frac{1-e^{-a\Delta T}}{z-e^{-a\Delta T}}$$

5/8

Matlab command

In MATLAB, the function *c2d.m* computes the ZOH equivalent of a continuous-time transfer function, as well as other discrete equivalents. For

$$G(s)=\frac{1}{s^2}$$

and \Delta T = 1, the following script
T=1;
numG=1; denG=[1 0 0];
G = tf(numG,denG);
Gd = c2d(G,T);
produces the correct ZOH equivalent.

UW Linear Systems (X. Chen, ME547)

TF discretization

Exercise

Exercise

Find the zero order hold equivalent of $G(s) = e^{-Ls}$, $2\Delta T < L < 3\Delta T$, where ΔT is the sampling time.

7/8

ME 547: Linear Systems Stability

Xu Chen

University of Washington

UW Linear Systems (X. Chen, ME547)

Stability

1/73

Outline

- 1. Definitions in Lyapunov stability analysis
- 2. Stability of LTI systems: method of eigenvalue/pole locations
- 3. Lyapunov's approach to stability
 - Relevant tools Lyapunov stability theorems Instability theorem Discrete-time case

4. Recap

1. Definitions in Lyapunov stability analysis

2. Stability of LTI systems: method of eigenvalue/pole locations

3. Lyapunov's approach to stability

Relevant tools Lyapunov stability theorems Instability theorem Discrete-time case

4. Recap

UW Linear Systems (X. Chen, ME547)

Stability

3/73

Finite dimensional vector norms

Let $v \in \mathbb{R}^n$. A norm is:

a metric in vector space: a function that assigns a real-valued length to each vector in a vector space, e.g.,

▶ 2 (Euclidean) norm: $\|v\|_2 = \sqrt{v^T v} = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}$ default in this set of notes: $\|\cdot\| = \|\cdot\|_2$

Equilibrium state

For *n*-th order unforced system

$$\dot{x}=f\left(x,t\right),\ x(t_{0})=x_{0}$$

an equilibrium state/point x_e is one such that

$$f(x_e,t)=0, \ \forall t$$

- the condition must be satisfied by all $t \ge 0$.
- if a system starts at equilibrium state, it stays there
- e.g., (inverted) pendulum resting at the verticle direction
- without loss of generality, we assume the origin is an equilibrium point

UW Linear Systems (X. Chen, ME547)

Stability

Equilibrium state of a linear system

For a linear system

$$\dot{x}(t)=A(t)x(t),\,\,x(t_0)=x_0$$

- origin $x_e = 0$ is always an equilibrium state
- when A(t) is singular, multiple equilibrium states exist

5/73
Continuous function

The function $f : \mathbb{R} \to \mathbb{R}$ is continuous at x_0 if $\forall \epsilon > 0$, there exists a $\delta(x_0, \epsilon) > 0$ such that

$$|x - x_0| < \delta \Longrightarrow |f(x) - f(x_0)| < \epsilon$$

Graphically, continuous functions is a single unbroken curve:



Figure: Continuous functions

e.g.,
$$\sin x$$
, x^2 , $sign(x - 1.5)$

UW Linear Systems (X. Chen, ME547)

Stability

Lyapunov's definition of stability

- Lyapunov invented his stability theory in 1892 in Russia. Unfortunately, the elegant theory remained unknown to the West until approximately 1960.
- The equilibrium state 0 of x = f(x, t) is stable in the sense of Lyapunov (s.i.L) if for all € > 0, and t₀, there exists δ(€, t₀) > 0 such that ||x(t₀)||₂ < δ gives ||x(t)||₂ < € for all t ≥ t₀</p>



Figure: Stable s.i.L: $||x(t_0)|| < \delta \Rightarrow ||x(t)|| < \epsilon \ \forall t \ge t_0$.

Stability

Asymptotic stability

The equilibrium state 0 of $\dot{x} = f(x, t)$ is asymptotically stable if

- it is stable in the sense of Lyapunov, and
- ▶ for all $\epsilon > 0$ and t_0 , there exists $\delta(\epsilon, t_0) > 0$ such that $||x(t_0)||_2 < \delta$ gives $x(t) \rightarrow 0$



Figure: Asymptotically stable i.s.L: $\|x(t_0)\| < \delta \Rightarrow \|x(t)\| \to 0$.



1. Definitions in Lyapunov stability analysis

2. Stability of LTI systems: method of eigenvalue/pole locations

3. Lyapunov's approach to stability

Relevant tools Lyapunov stability theorems Instability theorem Discrete-time case

4. Recap

Stability of LTI systems: method of eigenvalue/pole locations

the stability of the equilibrium point 0 for $\dot{x} = Ax$ or x(k+1) = Ax(k) can be concluded immediately based on the eigenvalues, λ 's, of A:

- ► the response $e^{At}x(t_0)$ involves modes such as $e^{\lambda t}$, $te^{\lambda t}$, $e^{\sigma t} \cos \omega t$, $e^{\sigma t} \sin \omega t$
- the response A^kx(k₀) involves modes such as λ^k, kλ^{k-1}, r^k cos kθ, r^k sin kθ
- $e^{\sigma t} \rightarrow 0$ if $\sigma < 0$; $e^{\lambda t} \rightarrow 0$ if $\lambda < 0$
- ► $\lambda^k \to 0$ if $|\lambda| < 1$; $r^k \to 0$ if $|r| = \left|\sqrt{\sigma^2 + \omega^2}\right| = |\lambda| < 1$

UW Linear Systems (X. Chen, ME547)

Stability

11 / 73

Stability of the origin for $\dot{x} = Ax$

stability	$\lambda_i(A)$
at O	
unstable	Re $\{\lambda_i\} > 0$ for some λ_i or Re $\{\lambda_i\} \leq 0$ for all λ_i 's but
	for a repeated λ_m on the imaginary axis with
	multiplicity <i>m</i> , nullity $(A - \lambda_m I) < m$ (Jordan form)
stable	Re $\{\lambda_i\} \leq 0$ for all λ_i 's and \forall repeated λ_m on the
i.s.L	imaginary axis with multiplicity <i>m</i> ,
	$nullity\left(A - \lambda_{m} I ight) = m ext{ (diagonal form)}$
asymptotical	If $Re \{\lambda_i\} < 0$ for all λ_i (A is then called Hurwitz)
stable	

Example (Unstable moving mass)

$$\dot{x} = Ax, \ A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

$$\lambda_1 = \lambda_2 = 0, \ m = 2,$$

nullity $(A - \lambda_i I) =$ nullity $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = 1 < m$

► i.e., two repeated eigenvalues but needs a generalized eigenvector ⇒ Jordan form after similarity transform

• verify by checking
$$e^{At} = \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}$$
: *t* grows unbounded

UW Linear Systems (X. Chen, ME547)

Stability

Example (Stable in the sense of Lyapunov)

$$\dot{x} = Ax, \ A = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Routh-Hurwitz criterion

the Routh Test (by E.J. Routh, in 1877): a simple algebraic procedure to determine how many roots a given polynomial

$$A(s)=a_ns^n+a_{n-1}s^{n-1}+\cdots+a_1s+a_0$$

has in the closed right-half complex plane, without the need to explicitly solve for the roots

- German mathematician Adolf Hurwitz independently proposed in 1895 to approach the problem from a matrix perspective
- popular if stability is the only concern and no details on eigenvalues (e.g., speed of response) are needed

UW Linear Systems (X. Chen, ME547)

Stability

Routh-Hurwitz criterion

- the asymptotic stability of the equilibrium point 0 for $\dot{x} = Ax$ can also be concluded based on the Routh-Hurwitz criterion
- simply apply the Routh Test to $A(s) = \det(sI A)$
- ▶ recap: the poles of transfer function $G(s) = C (sI A)^{-1} B + D$ come from det (sI A) in computing the inverse $(sI A)^{-1}$

The Routh Array for $A(s) = a_n s^n + a_{n-1} s^{n-1} + \dots + a_1 s + a_0$, construct $\begin{vmatrix} s^n & a_n & a_{n-2} & a_{n-4} & a_{n-6} & \dots \\ s^{n-1} & a_{n-1} & a_{n-3} & a_{n-5} & a_{n-7} & \dots \\ s^{n-2} & q_{n-2} & q_{n-4} & q_{n-6} & \dots \end{vmatrix}$

- first two rows contain the coefficients of A(s)
- third row constructed from the previous two rows via

The Routh Array

for $A(s) = a_n s^n + a_{n-1} s^{n-1} + \cdots + a_1 s + a_0$, construct

All roots of A(s) are on the left half s-plane if and only if all elements of the first column of the Routh array are positive.

The Routh Array

Example $(A(s) = 2s^4 + s^3 + 3s^2 + 5s + 10)$					
<i>s</i> ⁴	2	3	10		
<i>s</i> ³	1	5	0		
s^2	$3 - \frac{2 \times 5}{1} = -7$	10	0		
s^1	$5 - \frac{1 \times 10}{-7}$	0	0		
<i>s</i> ⁰	10	0	0		

- two sign changes in the first column
- unstable and two roots in the right half side of s-plane



The Routh Array

special cases:

- If the 1st element in any one row of Routh's array is zero, one can replace the zero with a small number e and proceed further.
- If the elements in one row of Routh's array are all zero, then the equation has at least one pair of real roots with equal magnitude but opposite signs, and/or the equation has one or more pairs of imaginary roots, and/or the equation has pairs of complex-conjugate roots forming symmetry about the origin of the *s*-plane.
- There are other possible complications, which we will not pursue further. See, e.g. "Automatic Control Systems", by Kuo, 7th ed., pp. 339-340.

Stability of the origin for x(k+1) = f(x(k), k)

 stability analysis follows analogously for nonlinear time-varying discrete-time systems of the form

$$x(k+1) = f(x(k), k), \ x(k_0) = x_0$$

• equilibrium point x_e :

$$f(x_e,k) = x_e, \ \forall k$$

without loss of generality, 0 is assumed an equilibrium point

UW Linear Systems (X. Chen, ME547)

Stability

21/73

Stability of the origin for x(k+1) = Ax(k)

stability	$\lambda_i(A)$
at O	
unstable	$ \lambda_i > 1$ for some λ_i or $ \lambda_i \le 1$ for all λ_i 's but for a
	repeated λ_m on the unit circle with multiplicity m_i
	$nullity\left(A - \lambda_m I ight) < m ext{ (Jordan form)}$
stable	$ \lambda_i \leq 1$ for all λ_i 's but for any repeated λ_m on the unit
i.s.L	circle with multiplicity m , nullity $(A - \lambda_m I) = m$
	(diagonal form)
asymptotica	ally $ \lambda_i < 1$ for all λ_i (such a matrix is called a Schur
stable	matrix)

Routh-Hurwitz criterion for DT LTI systems

- the stability domain $|\lambda_i| < 1$ is a unit disk
- Routh array validates stability in the left-half plane
- bilinear transformation maps the closed left half s-plane to the closed unit disk in z-plane



Routh-Hurwitz criterion for DT LTI systems

apply bilinear transform

$$A(z)|_{z=\frac{1+s}{1-s}} = \left(\frac{1+s}{1-s}\right)^n + a_1 \left(\frac{1+s}{1-s}\right)^{n-1} + \dots + a_n =: \frac{A^*(s)}{(1-s)^n}$$

apply Routh test to

$$A^*(s) = a_n^* s^n + a_{n-1}^* s^{n-1} + \dots + a_0^* = A(z)|_{z=\frac{1+s}{1-s}} (1-s)^n$$

Routh-Hurwitz criterion for DT LTI systems

Example $(A(z) = z^3 + 0.8z^2 + 0.6z + 0.5)$

- $A^*(s) = A(z)|_{z=\frac{1+s}{1-s}} (1-s)^3 = (1+s)^3 + 0.8 (1+s)^2 (1-s) + 0.6 (1+s) (1-s)^2 + 0.5 (1-s)^3 = 0.3s^3 + 3.1s^2 + 1.7s + 2.9$
- ► Routh array

<i>s</i> ³	0.3	1.7
<i>s</i> ²	3.1	2.9
5	$1.7 - \frac{0.3 \times 2.9}{3.1} = 1.42$	0
s^0	2.9	0

► all elements in first column are positive ⇒ roots of A(z) are all in the unit circle

UW Linear Systems (X. Chen, ME547)

Stability

25 / 73

1. Definitions in Lyapunov stability analysis

2. Stability of LTI systems: method of eigenvalue/pole locations

3. Lyapunov's approach to stability

Relevant tools Lyapunov stability theorems Instability theorem Discrete-time case

4. Recap

Lyapunov's approach to stability

The direct method of Lyapunov to stability problems:

- no need for explicit solutions to system responses
- ► an "energy" perspective
- fit for general dynamic systems (linear/nonlinear, time-invariant/time-varying)

UW Linear Systems (X. Chen, ME547

Stability

27 / 73

Stability from an energy viewpoint: Example

Consider spring-mass-damper systems:

$$\dot{x}_1 = x_2 \qquad (x_1: \text{ position}; \ x_2: \text{ velocity}) \\ \dot{x}_2 = -\frac{k}{m}x_1 - \frac{b}{m}x_2, \ b > 0 \qquad (\text{Newton's law})$$

λ (A)'s are in the left-half s-plane⇒ asymptotically stable
 total energy

 $\mathcal{E}(t) = \text{potential energy} + \text{kinetic energy} = \frac{1}{2}kx_1^2 + \frac{1}{2}mx_2^2$

energy dissipates / is dissipative:

$$\dot{\mathcal{E}}(t) = kx_1\dot{x}_1 + mx_2\dot{x}_2 = -bx_2^2 \le 0$$

E = 0 only when x₂ = 0. Since [x₁, x₂]^T = 0 is the only equilibrium state, the motion will not stop at x₂ = 0, x₁ ≠ 0. Thus the energy will keep decreasing toward 0 which is achieved at the origin.

Stability from an energy viewpoint: Generalization

Consider unforced, time-varying, nonlinear systems

$$\dot{x}(t) = f(x(t), t), \ x(t_0) = x_0 \ x(k+1) = f(x(k), k), \ x(k_0) = x_0$$

- assume the origin is an equilibrium state
- ► energy function ⇒ Lyapunov function: a scalar function of x and t (or x and k in discrete-time case)
- goal is to relate properties of the state through the Lyapunov function
- main tool: matrix formulation, linear algebra, positive definite functions

UW Linear Systems (X. Chen, ME547)

Relevant tools

Quadratic functions

intrinsic in energy-like analysis, e.g.

$$\frac{1}{2}kx_1^2 + \frac{1}{2}mx_2^2 = \frac{1}{2}\begin{bmatrix}x_1\\x_2\end{bmatrix}^T\begin{bmatrix}k&0\\0&m\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}$$

convenience of matrix formulation:

$$\frac{1}{2}kx_1^2 + \frac{1}{2}mx_2^2 + x_1x_2 = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}^T \begin{bmatrix} \frac{k}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{m}{2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$
$$\frac{1}{2}kx_1^2 + \frac{1}{2}mx_2^2 + x_1x_2 + c = \begin{bmatrix} x_1 \\ x_2 \\ 1 \end{bmatrix}^T \begin{bmatrix} \frac{k}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{m}{2} & 0 \\ 0 & 0 & c \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ 1 \end{bmatrix}$$

general quadratic functions in matrix form

$$Q(x) = x^T P x, \ P^T = P$$

Relevant tools

Symmetric matrices

- recall: a real square matrix A is
 - \blacktriangleright symmetric if $A = A^T$
 - \blacktriangleright skew-symmetric if $A = -A^T$
- examples:

$$\left[\begin{array}{rrrr}1 & 2\\2 & 1\end{array}\right], \left[\begin{array}{rrrr}1 & 2\\-2 & 1\end{array}\right], \left[\begin{array}{rrrr}0 & 2\\-2 & 0\end{array}\right]$$

Fact

Any real square matrices can be decomposed as the sum of a symmetric matrix and a skew symmetric matrix:

e.g.
$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 2.5 \\ 2.5 & 4 \end{bmatrix} + \begin{bmatrix} 0 & -0.5 \\ 0.5 & 0 \end{bmatrix}$$

formula: $P = \frac{P + P^{T}}{2} + \frac{P - P^{T}}{2}$
s (X. Chen, ME547) Stability 31/73

UW Linear Systems (X. Chen, ME547)

Stabilit

Relevant tools

Symmetric matrices

- ► A real square matrix $A \in \mathbb{R}^{n \times n}$ is orthogonal if $A^T A = A A^T = I$
- meaning that the columns of A form a orthonormal basis of \mathbb{R}^n
- to see this, writing A in the column-vector notation

$$A = \begin{bmatrix} | & | & | & | & | \\ a_1 & a_2 & \dots & a_n \\ | & | & | & | & | \end{bmatrix}$$

we get

$$A^{T}A = \begin{bmatrix} a_{1}^{T}a_{1} & a_{1}^{T}a_{2} & \dots & a_{1}^{T}a_{n} \\ a_{2}^{T}a_{1} & a_{2}^{T}a_{2} & \dots & a_{2}^{T}a_{n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n}^{T}a_{1} & a_{n}^{T}a_{2} & \dots & a_{n}^{T}a_{n} \end{bmatrix} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{bmatrix}$$

namely, $a_i^T a_j = 1 a_i^T a_m = 0 \ \forall j \neq m$.

Theorem

The eigenvalues of symmetric matrices are all real.

Proof: $\forall : A \in \mathbb{R}^{n \times n}$ with $A^T = A$.

Eigenvalue-eigenvector pair: $Au = \lambda u \Rightarrow \overline{u}^T A u = \lambda \overline{u}^T u$, where \overline{u} is the complex conjugate of u. $\overline{u}^T A u$ is a real number, as

$$\overline{\overline{u}^{T}Au} = u^{T}\overline{A}\overline{u}$$

$$= u^{T}A\overline{u} \quad \because A \in \mathbb{R}^{n \times n}$$

$$= u^{T}A^{T}\overline{u} \quad \because A = A^{T}$$

$$= \lambda u^{T}\overline{u} \quad \because (Au)^{T} = (\lambda u)^{T}$$

$$= \lambda \overline{u}^{T}u \quad \because u^{T}\overline{u} \in \mathbb{R}$$

$$= \overline{u}^{T}Au \quad \because Au = \lambda u$$

Also, $\overline{u}^T u \in \mathbb{R}$. Thus $\lambda = \frac{\overline{u}^T A u}{\overline{u}^T u}$ must also be a real number.

UW Linear Systems (X. Chen, ME547)

Stability

Important properties of symmetric matrices

Theorem

The eigenvalues of symmetric matrices are all real.

Theorem

The eigenvalues of skew-symmetric matrices are all imaginary or zero.

Theorem

All eigenvalues of an orthogonal matrix have a magnitude of 1.

matrix structure	analogy in complex plane
symmetric	real line
skew-symmetric	imaginary line
orthogonal	unit circle

Example

•
$$\begin{bmatrix} 0 & 2 \\ 2 & 0 \end{bmatrix}$$
: eigenvalues $(= \pm 2)$ are all real
• $\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 2 \\ 2 & 0 \end{bmatrix} \Rightarrow$ eigenvalues $(= 1 \pm 2$ by spectral mapping theorem) are all real
• $\begin{bmatrix} 0 & 2 \\ -2 & 0 \end{bmatrix}$: eigenvalues $(= \pm 2j)$ are all imaginary
• $\begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 2 \\ -2 & 0 \end{bmatrix} \Rightarrow$ eigenvalues are $1 \pm 2j$

UW Linear Systems (X. Chen, ME547)

Stability

35 / 73

The spectral theorem for symmetric matrices

When $A \in \mathbb{R}^{n \times n}$ has *n* distinct eigenvalues, we can do diagonalization $A = U \wedge U^{-1}$. The following spectral theorem significantly simplifies the result when A is symmetric.

Theorem (Symmetric eigenvalue decomposition (SED))

 $\forall : A \in \mathbb{R}^{n \times n}, \ A^T = A$, there always exist $\lambda_i \in \mathbb{R}$ and $u_i \in \mathbb{R}^n$, s.t.

$$A = \sum_{i=1}^{n} \lambda_i u_i u_i^T = U \Lambda U^T$$
(1)

- \triangleright λ_i 's: eigenvalues of A
- u_i : eigenvector associated to λ_i , normalized to have unity norms
- $\blacktriangleright U = [u_1, u_2, \cdots, u_n]^T \text{ is orthogonal: } U^T U = UU^T = I$
- $\blacktriangleright \Lambda = diagonal(\lambda_1, \lambda_2, \dots, \lambda_n)$

Elements of proof for SED

Theorem

 $\forall : A \in \mathbb{R}^{n \times n}$ with $A^T = A$, then eigenvectors of A, associated with different eigenvalues, are orthogonal.

Proof.

Let $Au_i = \lambda_i u_i$ and $Au_j = \lambda_j u_j$. Then $u_i^T A u_j = u_i^T \lambda_j u_j = \lambda_j u_i^T u_j$. In the meantime, $u_i^T A u_j = u_i^T A^T u_j = (Au_i)^T u_j = \lambda_i u_i^T u_j$. So $\lambda_i u_i^T u_j = \lambda_j u_i^T u_j$. But $\lambda_i \neq \lambda_j$. It must be that $u_i^T u_j = 0$.

SED now follows:

- ► If A has distinct eigenvalues, then U = [u₁, u₂, ···, u_n]^T is orthogonal if we normalize all the eigenvectors to unity norm.
- If A has r(< n) distinct eigenvalues, it turns out can choose multiple orthogonal eigenvectors for the eigenvalues with none-unity multiplicities. See proof in supplementary notes.

```
UW Linear Systems (X. Chen, ME547)
```

Stability

37 / 73

Rethinking symmetric matrices

With the spectral theorem, next time we see a symmetric matrix A, we immediately know that

- $\blacktriangleright \lambda_i \text{ is real for all } i$
- ▶ associated with λ_i , we can always find a real eigenvector
- ▶ ∃ an orthonormal basis $\{u_i\}_{i=1}^n$, which consists of the eigenvectors
- if A ∈ ℝ^{2×2}, then if you compute first λ₁, λ₂ and u₁, you won't need to go through the regular math to get u₂, but can simply solve for a u₂ that is orthogonal to u₁ with ||u₂|| = 1.

Rethinking symmetric matrices

Example

SED of
$$A = \begin{bmatrix} 5 & \sqrt{3} \\ \sqrt{3} & 7 \end{bmatrix}$$
. Computing the eigenvalues gives
det $\begin{bmatrix} 5-\lambda & \sqrt{3} \\ \sqrt{3} & 7-\lambda \end{bmatrix} = 35 - 12\lambda + \lambda^2 - 3 = (\lambda - 4) (\lambda - 8) = 0$
 $\Rightarrow \lambda_1 = 4, \ \lambda_2 = 8$
First normalized eigenvector:

$$(A - \lambda_1 I) t_1 = 0 \Rightarrow \begin{bmatrix} 1 & \sqrt{3} \\ \sqrt{3} & 3 \end{bmatrix} t_1 = 0 \Rightarrow t_1 = \begin{bmatrix} -\frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{bmatrix}$$

A is symmetric \Rightarrow eigenvectors are orthogonal to each other:

choose $t_2 = \begin{vmatrix} \frac{2}{\sqrt{3}} \\ \frac{\sqrt{3}}{2} \end{vmatrix}$. No need to solve $(A - \lambda_2 I) t_2 = 0!$

Stabilit

39 / 73

Theorem (Eigenvalues of symmetric matrices)

If $A = A^T \in \mathbb{R}^{n \times n}$, then the eigenvalues of A satisfy

$$\lambda_{\max} = \max_{x \in \mathbb{R}^n, \ x \neq 0} \frac{x^T A x}{\|x\|_2^2} \tag{2}$$

$$\lambda_{\min} = \min_{x \in \mathbb{R}^n, \ x \neq 0} \frac{x^T A x}{\|x\|_2^2} \tag{3}$$

Proof.

Perform SED to get $A = \sum_{i=1}^{n} \lambda_i u_i^T u_i$ where $\{u_i\}_{i=1}^{n}$ spans \mathbb{R}^n . Then any vector $x \in \mathbb{R}^n$ can be decomposed as $x = \sum_{i=1}^{n} \alpha_i u_i$. Thus

$$\max_{x \neq 0} \frac{x^T A x}{\|x\|_2^2} = \max_{\alpha_i} \frac{\left(\sum_i \alpha_i u_i\right)^T \sum_i \lambda_i \alpha_i u_i}{\sum_i \alpha_i^2} = \max_{\alpha_i} \frac{\sum_i \lambda_i \alpha_i^2}{\sum_i \alpha_i^2} = \lambda_{\max}$$

UW Linear Systems (X. Chen, ME547)

Positive definite matrices

► eigenvalues of symmetric matrices are real ⇒ we can order the eigenvalues.

Definition

A symmetric matrix P is called positive-definite if all its eigenvalues are positive.

Equivalently,

Definition (Positive Definite Matrices)

A symmetric matrix $P \in \mathbb{R}^{n \times n}$ is called **positive-definite**, written $P \succ 0$, if $x^T P x > 0$ for all $x (\neq 0) \in \mathbb{R}^n$. P is called **positive-semidefinite**, written $P \succeq 0$, if $x^T P x \ge 0$ for all $x \in \mathbb{R}^n$

UW Linear Systems (X. Chen, ME547)

Stability

41 / 73

Negative definite matrices

Definition

A symmetric matrix $Q \in \mathbb{R}^{n \times n}$ is called **negative-definite**, written $Q \prec 0$, if $-Q \succ 0$, i.e., $x^T Q x < 0$ for all $x (\neq 0) \in \mathbb{R}^n$. Q is called **negative-semidefinite**, written $Q \preceq 0$, if $x^T Q x \leq 0$ for all $x \in \mathbb{R}^n$

▶ When A and B have compatible dimensions, $A \succ B$ means $A - B \succ 0$.

Positive definite matrices

Positive-definite matrices can have negative entries:

Example

$$P = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \text{ is positive-definite, as } P = P^T \text{ and take any}$$

$$v = [x, y]^T, \text{ we have}$$

$$v^T P v = \begin{bmatrix} x \\ y \end{bmatrix}^T \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 2x^2 + 2y^2 - 2xy$$

$$= x^2 + y^2 + (x - y)^2 \ge 0$$
and the equality sign holds only when $x = y = 0$.

UW Linear Systems (X. Chen, ME547)

Stability

43 / 73

Positive definite matrices

Conversely, matrices whose entries are all positive are not necessarily positive-definite.

Example

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$$
 is not positive-definite:

 $\begin{bmatrix} 1 \\ -1 \end{bmatrix}^T \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = -2 < 0$

Positive definite matrices

Theorem

For a symmetric matrix P, $P \succ 0$ if and only if all the eigenvalues of P are positive.

Proof.

Since P is symmetric, we have

$$\lambda_{\max}(P) = \max_{x \in \mathbb{R}^n, \ x \neq 0} \frac{x^T A x}{\|x\|_2^2} \tag{4}$$

$$\lambda_{\min}\left(P\right) = \min_{x \in \mathbb{R}^n, \ x \neq 0} \frac{x^{\gamma} A x}{\|x\|_2^2} \tag{5}$$

which gives $x^T A x \in [\lambda_{\min} ||x||_2^2, \ \lambda_{\max} ||x||_2^2]$. Thus $x^T A x > 0, \ x \neq 0 \Leftrightarrow \lambda_{\min} > 0$.

UW Linear Systems (X. Chen, ME547)

Stability

45 / 73

Relevant tools

Checking positive definiteness of a matrix.

We often use the following necessary and sufficient conditions to check if a symmetric matrix P is positive (semi-)definite or not:

- P ≻ 0 (P ≥ 0) ⇔ the leading principle minors defined below are positive (nonnegative)
- P ≻ 0 (P ≥ 0) ⇔ P can be decomposed as P = N^TN where N is nonsingular (singular)

Definition

The leading principle minors of $P = \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix}$ are defined as p_{11} , det $\begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}$, det P.

Relevant tools

Checking positive definiteness of a matrix.

Example

None of the following matrices are positive definite:

$$\left[\begin{array}{cc} -1 & 0 \\ 0 & 1 \end{array}\right], \left[\begin{array}{cc} -1 & 1 \\ 1 & 2 \end{array}\right], \left[\begin{array}{cc} 2 & 1 \\ 1 & -1 \end{array}\right], \left[\begin{array}{cc} 1 & 2 \\ 2 & 1 \end{array}\right]$$

Relevant tools

Definition (Positive Definite Functions)

A continuous time function $W : \mathbb{R}^n \to \mathbb{R}_+$, called to be PD, satisfying

•
$$W(x) > 0$$
 for all $x \neq 0$

•
$$W(0) = 0$$

• $W(x) \to \infty$ as $|x| \to \infty$ uniformly in x

In the three dimensional space, positive definite functions are "bowl-shaped", e.g., $W(x_1, x_2) = x_1^2 + x_2^2$.



Relevant tools

Definition (Locally Positive Definite Functions)

A continuous time function $W: \mathbb{R}^n \to \mathbb{R}_+$, called to be LPD, satisfying

- W(x) > 0 for all $x \neq 0$ and |x| < r
- \blacktriangleright W(0) = 0

In the three dimensional space, locally positive definite functions are "bowl-shaped" locally, e.g., $W(x_1, x_2) = x_1^2 + \sin^2 x_2$ for $x_1 \in \mathbb{R}$ and $|x_2| < \pi$



Relevant tools

Exercise

Let $x = [x_1, x_2, x_3]^T$. Check the positive definiteness of the following functions

1.
$$V(x) = x_1^4 + x_2^2 + x_3^4$$
 (PD)
2. $V(x) = x_1^2 + x_2^2 + 3x_3^2 - x_3^4$ (LPD for $|x_3| < \sqrt{3}$)

- 1. Definitions in Lyapunov stability analysis
- 2. Stability of LTI systems: method of eigenvalue/pole locations
- 3. Lyapunov's approach to stability

Relevant tools Lyapunov stability theorems Instability theorem Discrete-time case

4. Recap

UW Linear Systems (X. Chen, ME547)

Stability

51/73

Lyapunov stability theorems

recall the spring mass damper example in matrix form

$$\frac{d}{dt} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = A \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{b}{m} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

energy function is PD:

 $\mathcal{E}(t) = \text{potential energy} + \text{kinetic energy} = \frac{1}{2}kx_1^2 + \frac{1}{2}mx_2^2$ and its derivative is NSD:

$$\dot{\mathcal{E}}(t) = \left[\frac{\partial \mathcal{E}}{\partial x_1}, \frac{\partial \mathcal{E}}{\partial x_2}\right] \left[\dot{x}_1, \dot{x}_2\right]^T = k_1 x_1 \dot{x}_1 + m x_2 \dot{x}_2 \tag{6}$$

$$= k_1 x_1 x_2 + m x_2 \left(-\frac{k}{m} x_1 - \frac{b}{m} x_2 \right) = \left[\frac{\partial \mathcal{E}}{\partial x_1}, \frac{\partial \mathcal{E}}{\partial x_2} \right] A x \quad (7)$$
$$= -b x_2^2$$

Remark:
\vec{\vec{\vec{\vec{E}}}}(t) is a derivative along the state trajectory: (6) takes the derivative of \$\vec{\vec{E}}\$ w.r.t. \$x = [x_1, x_2]^T\$; (7) is the time derivative along the trajectory of the state.

The notion of derivative along state trajectories

Generalizing the concept to system x = f (x): let V (x) be a general energy function, the energy dissipation w.r.t. time is

$$\frac{dV(x)}{dt} = \begin{bmatrix} \frac{\partial V}{\partial x_1}, \frac{\partial V}{\partial x_2}, \dots, \frac{\partial V}{\partial x_n} \end{bmatrix} \begin{bmatrix} f_1(x) \\ \vdots \\ f_n(x) \end{bmatrix}$$

also denoted as $L_f V(x)$, the Lie derivative of V(x) w.r.t. f(x).

We concluded stability of the system by analyzing how energy will dissipate to zero along the trajectory of the state.



Stability

Theorem

The equilibrium point 0 of $\dot{x}(t) = f(x(t), t)$, $x(t_0) = x_0$ is <u>stable in</u> <u>the sense of Lyapunov</u> if there exists a locally positive definite function V(x, t) such that $\dot{V}(x, t) \leq 0$ for all $t \geq t_0$ and all x in a local region x : |x| < r for some r > 0.

- ▶ such a V(x, t) is called a Lyapunov function
- i.e., V (x) is PD and V(x) is negative semidefinite in a local region |x| < r

Theorem

The equilibrium point 0 of $\dot{x}(t) = f(x(t), t)$, $x(t_0) = x_0$ is <u>locally</u> <u>asymptotically stable</u> if there exists a Lyapunov function V(x) such that $\dot{V}(x)$ is locally negative definite.

Theorem

The equilibrium point 0 of $\dot{x}(t) = f(x(t), t)$, $x(t_0) = x_0$ is globally asymptotically stable if there exists a Lyapunov function V(x) such that V(x) is positive definite and $\dot{V}(x)$ is negative definite.

UW Linear Systems (X. Chen, ME547)

Stability

Lyapunov stability concept for linear systems

- ► for linear system $\dot{x} = Ax$, a good Lyapunov candidate is the quadratic function $V(x) = x^T P x$ where $P = P^T$ and $P \succ 0$
- the derivative along the state trajectory is then

$$\dot{V}(x) = \dot{x}^T P x + x^T P \dot{x}$$

= $(Ax)^T P x + x^T P A x$
= $x^T (A^T P + P A) x$

- ▶ such a $V(x) = x^T P x$ is a Lyapunov function for $\dot{x} = A x$ when $A^T P + P A \leq 0$
- and the origin is stable in the sense of Lyapunov

Theorem (Lyapunov stability theorem for linear systems)

For $\dot{x} = Ax$ with $A \in \mathbb{R}^{n \times n}$, the origin is asymptotically stable if and only if for any symmetric positive definite matrix $Q \succ 0$, <u>the</u> Lyapunov equation

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

has a unique positive definite solution $P \succ 0$, $P^T = P$.

Proof.

$$"\Rightarrow": \frac{\dot{V}}{V} = -\frac{x^{T}Qx}{x^{T}Px} \leq -\underbrace{\frac{(\lambda_{Q})_{\min}}{(\lambda_{P})_{\max}}}_{\text{(}\lambda_{P})_{\max}} \Longrightarrow V(t) \leq e^{-\alpha t}V(0). \text{ Since}$$

 $Q \succ 0 \text{ and } P \succ 0, \ (\lambda_Q)_{\min} > 0 \text{ and } (\lambda_P)_{\max} > 0.$ Thus $\alpha > 0; \ V(t)$ decays exponentially to zero. $V(x) \succ 0 \Rightarrow V(x) = 0$ only at x = 0. Therefore, $x \rightarrow 0$ as $t \rightarrow \infty$, regardless of the initial condition.

UW Linear Systems (X. Chen, ME547)

Stability

58 / 73

Proof.

" \Leftarrow ": if 0 of $\dot{x} = Ax$ is asymptotically stable, then all eigenvalues of A have negative real parts. For any Q, the Lyapunov equation has a unique solution P. Note $x(t) = e^{At}x_0 \to 0$ as $t \to \infty$. We have

$$x^{T}(\infty) Px(\infty) - x^{T}(0) Px(0) = \int_{0}^{\infty} \frac{d}{dt} x^{T}(t) Px(t) dt = \int_{0}^{\infty} x^{T}(t) \left(A^{T}P + PA\right) x(t) dt$$
$$\Rightarrow x(0)^{T} Px(0) = \int_{0}^{\infty} x^{T}(t) Qx(t) dt = \int_{0}^{\infty} x(0) e^{A^{T}t} Q e^{At} x(0) dt$$

If $Q \succ 0$, there exists a nonsingular N matrix: $Q = N^T N$. Thus $x(0)^T P x(0) = \int_0^\infty \|Ne^{At} x(0)\|^2 dt \ge 0$ $x(0)^T P x(0) = 0$ only if $x_0 = 0$

Thus $P \succ 0$. Furthermore

$$P = \int_0^\infty e^{A^T t} Q e^{At} dt$$

Stability

Lyapunov stability theorems

Example

Let
$$\dot{x} = Ax$$
, $A = \begin{bmatrix} -1 & 1 \\ -1 & 0 \end{bmatrix}$. The Lyapunov equation is
$$\begin{bmatrix} -1 & 1 \\ -1 & 0 \end{bmatrix}^{T} \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} + \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} \begin{bmatrix} -1 & 1 \\ -1 & 0 \end{bmatrix} = -\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \underset{Q}{\overset{Q}{}}$$

We need

$$\begin{array}{l} -2p_{11}-2p_{12}=-1\\ -p_{12}-p_{22}+p_{11}=0\\ 2p_{12}=-1 \end{array} \Rightarrow \begin{cases} p_{11}=1\\ p_{22}=3/2\\ p_{12}=-1/2 \end{cases}$$

leading principle minors: $p_{11} > 0$, $p_{11}p_{22} - p_{12}^2 > 0$ $\Rightarrow P \succ 0 \Rightarrow$ asymptotically stable

UW Linear Systems (X. Chen, ME547)

Stability

59 / 73

Essense of the Lyapunov Eq.

Observations:

► $A^T P + PA$ is a linear operation on P: e.g.,

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \ Q = \begin{bmatrix} | & | \\ q_1 & q_2 \\ | & | \end{bmatrix}, \ P = \begin{bmatrix} | & | \\ p_1 & p_2 \\ | & | \end{bmatrix}$$
$$A^{T} \begin{bmatrix} | & | \\ p_1 & p_2 \\ | & | \end{bmatrix} + \begin{bmatrix} | & | \\ p_1 & p_2 \\ | & | \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = -\begin{bmatrix} | & | \\ q_1 & q_2 \\ | & | \end{bmatrix}$$
$$A^{T} p_1 + a_{11}p_1 + a_{21}p_2 = -q_1$$
$$A^{T} p_2 + a_{12}p_1 + a_{22}p_2 = -q_2$$

Essense of the Lyapunov Eq.

Observations: with now

$$A^T P + P A = Q \Leftrightarrow \begin{cases} A^T p_1 + a_{11} p_1 + a_{21} p_2 &= -q_1 \\ A^T p_2 + a_{12} p_1 + a_{22} p_2 &= -q_2 \end{cases}$$

▶ can stack the columns of $A^T P + PA$ and Q to yield, e.g.

$$\begin{bmatrix} A^{T} & 0 \\ 0 & A^{T} \end{bmatrix} \begin{bmatrix} p_{1} \\ p_{2} \end{bmatrix} + \begin{bmatrix} a_{11}I & a_{21}I \\ a_{12}I & a_{22}I \end{bmatrix} \begin{bmatrix} p_{1} \\ p_{2} \end{bmatrix} = -\begin{bmatrix} q_{1} \\ q_{2} \end{bmatrix}$$
$$\underbrace{\left\{ \begin{bmatrix} A^{T} & 0 \\ 0 & A^{T} \end{bmatrix} + \begin{bmatrix} a_{11}I & a_{21}I \\ a_{12}I & a_{22}I \end{bmatrix} \right\}}_{L_{A}} \begin{bmatrix} p_{1} \\ p_{2} \end{bmatrix} = -\begin{bmatrix} q_{1} \\ q_{2} \end{bmatrix}$$

UW Linear Systems (X. Chen, ME547)

Stability

61/73

The Lyapunov Eq.: Existence of solution

$$L_{A} = \begin{bmatrix} A^{T} & 0 \\ 0 & A^{T} \end{bmatrix} + \begin{bmatrix} a_{11}I & a_{21}I \\ a_{12}I & a_{22}I \end{bmatrix}$$

• can simply write $L_A = \underbrace{I \otimes A^T + A^T \otimes I}_{\text{mirror symmetric}}$ using the Kronecker

product notation
$$B \otimes C = \begin{bmatrix} b_{11}C & b_{11}C & \dots & b_{11}C \\ b_{21}C & b_{22}C & \dots & b_{2n}C \\ \vdots & \vdots & \dots & \vdots \\ b_{m1}C & b_{m2}C & \dots & b_{mn}C \end{bmatrix}$$

- can show that L_A is invertible if and only if $\lambda_i + \lambda_j \neq 0$ for all eigenvalues of A.
- ► To check, let $A^T u_i = \lambda_i u_i$ and $A^T u_j = \lambda_j u_j$. Note that $u_i u_j^T A + A^T u_i u_j^T = u_i (\lambda_j u_j)^T + \lambda_i u_i u_j^T = (\lambda_i + \lambda_j) u_i u_j^T$. So $\lambda_i + \lambda_j$ is an eigenvalue of the operator $L_A(P)$. If $\lambda_i + \lambda_j \neq 0$, the operator is invertible.

The Lyapunov Eq.: eigenvalues



UW Linear Systems (X. Chen, ME547)

Stability

63 / 73

Procedures of Lyapunov's direct method

- Given A, select an arbitrary positive-definite symmetric matrix Q (e.g., I).
- 2. Find the solution matrix *P* to the continuous- or discrete-time Lyapunov equation.
- 3. If a solution *P* cannot be found, the origin is not asymptotically stable.
- 4. If a solution is found:
 - if P is positive-definite, then A is Hurwitz and the origin is asymptotically stable;
 - if P is not positive-definite, then A has at least one eigenvalue with a positive real part and the origin is an unstable equilibrium.

It suffices to select Q = I

For linear systems we can let Q = I and check whether the resulting P is positive definite. If it is, then we can assert the asymptotic stability:

► take any $Q \succ 0$. there exists $Q = N^T N$, where N is invertible, yielding

- $\tilde{A} = N^{-1}AN$ and A are similar matrices and have the same eigenvalues
- $\tilde{P} = N^T P N$ and P have the same definiteness. If we can find a positive definite solution P then the \tilde{P} will also be positive definite. Vise versa.

UW Linear Systems (X. Chen, ME547)

Instability theorem

failure to find a Lyapunov function does not imply instability

Stability

 (for nonlinear systems, Lyapunov function can be nontrivial to find)

Theorem

The equilibrium state 0 of $\dot{x} = f(x)$ is unstable if there exists a function W(x) such that

- $\dot{W}(x)$ is PD locally: $\dot{W}(x) > 0 \forall |x| < r$ for some r and $\dot{W}(0) = 0$
- $\blacktriangleright W(0) = 0$
- there exist states x arbitrarily close to the origin such that W(x) > 0

Discrete-time case: key concept of Lyapunov

For the discrete-time system

$$x\left(k+1\right)=Ax\left(k\right)$$

we consider a quadratic Lyapunov function candidate

$$V(x) = x^T P x, \ P = P^T \succ 0$$

and compute $\Delta V(x)$ along the trajectory of the state

$$V(x(k+1)) - V(x(k)) = x^{T}(k) \underbrace{\left[A^{T}PA - P
ight]}_{\triangleq -Q} x(k)$$

The above gives the DT Lyapunov stability theorem for LTI systems.



DT Lyapunov stability theorem for linear systems

Theorem

For system x(k + 1) = Ax(k) with $A \in \mathbb{R}^{n \times n}$, the origin is asymptotically stable if and only if $\exists Q \succ 0$, such that <u>the</u> discrete-time Lyapunov equation

$$A^T P A - P = -Q$$

has a unique positive definite solution $P \succ 0$, $P^T = P$.

The DT Lyapunov Eq.

$$A^T P A - P = -Q$$

Solution to the DT Lyapunov equation, when asymptotic stability holds (A is Schur), comes from the following:

$$V(x(\infty))^{-} V(x(0)) = \sum_{k=0}^{\infty} x^{T}(k) [A^{T}PA - P] x(k)$$
$$= -\sum_{k=0}^{\infty} x^{T}(0) (A^{T})^{k} QA^{k} x(0)$$
$$\Rightarrow P = \sum_{k=0}^{\infty} (A^{T})^{k} QA^{k}$$

• Can show that the discrete-time Lyapunov operator $L_A = A^T P A - P$ is invertible if and only if for all i, j $(\lambda_A)_i (\lambda_A)_j \neq 1$

UW Linear Systems (X. Chen, ME547)

Stability

69 / 73

DT Lyapunov stability: MATLAB command

Example				
$x(k+1) = Ax(k), \ A =$	- 0 0	1 0	0 - 1	
	0.275	-0.225	-0.1	

Matlab Commands: $A=[0\ 1\ 0;\ 0\ 0\ 1;\ 0.275\ -0.225\ -0.1]$ Q=eye(3) P=dlyap(A',Q) % check function definition in Matlab help eig(P)

Recap

- Internal stability
 - Stability in the sense of Lyapunov: ε , δ conditions
 - Asymptotic stability
- Stability analysis of linear time invariant systems (x = Ax or x(k+1) = Ax(k))
 - Based on the eigenvalues of A
 - Time response modes
 - Repeated eigenvalues on the imaginary axis
 - Routh's criterion
 - No need to solve the characteristic equation
 - Discrete time case: bilinear transform $(z = \frac{1+s}{1-s})$

UW Linear Systems (X. Chen, ME547)

Stability

Recap

Lyapunov equations

Theorem: All the eigenvalues of A have negative real parts if and only if for any given $Q \succ 0$, the Lyapunov equation

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

has a unique solution P and $P \succ 0$.

Note: Given Q, the Lyapunov equation $A^T P + PA = -Q$ has a unique solution, when $\lambda_{A,i} + \lambda_{A,j} \neq 0$ for all i and j.

Theorem: All the eigenvalues of A are inside the unit circle if and only if for any given $Q \succ 0$, the Lyapunov equation

$$A^T P A - P = -Q$$

has a unique solution P and $P \succ 0$.

Note: Given Q, the Lyapunov equation $A^T P A - P = -Q$ has a unique solution, when $\lambda_{A,i}\lambda_{A,j} \neq 1$ for all i and j.

Recap

- P is positive definite if and only if any one of the following conditions holds:
 - 1. All the eigenvalues of P are positive.
 - 2. All the leading principle minors of P are positive.
 - 3. There exists a nonsingular matrix N such that $P = N^T N$.

Stability

ME 547: Linear Systems Controllability and Observability

Xu Chen

University of Washington

UW Linear Systems (X. Chen, ME547)

Controllability and Observability

1/48

Outline

- 1. Concepts
- 2. DT controllability

Controllability and controllable canonical form Controllability and Lyapunov Eq.

3. DT observability

Observability and observable canonical form

- 4. CT cases
- 5. The degrees of controllability and observability
- 6. Transforming controllable systems into controllable canonical forms
- 7. Transforming observable systems into observable canonical forms

Recap

General LTI state-space models:

$$\dot{x}(t) = Ax(t) + Bu(t)$$
 or $x(k+1) = Ax(k) + Bu(k)$
 $y = Cx + Du$

	continuous time	discrete time			
Lyapunov Eq.	$A^{T}P + PA = -Q$	$A^T P A - P = -Q$			
unique sol.	$\lambda_i(A) + \lambda_j(A) \neq 0$	$ \lambda_i(A) \lambda_j(A) < 1$			
cond.	$\forall i, j$	$\forall i, j$			
colution	$P = \int_0^\infty e^{A^T t} Q e^{At} dt$	$P = \sum_{k=0}^{\infty} \left(A^T \right)^k Q A^k$			
Solution	(if A is Hurwitz)	(if A is Schur)			

UW Linear Systems (X. Chen, ME547) Controllability and Observability

The concept of controllability and observability

Controllability:

inputs do not act directly on the states but via state dynamics:

$$\dot{x}(t) = Ax(t) + Bu(t)$$
 or $x(k+1) = Ax(k) + Bu(k)$ (1)

can the inputs drive the system to any value in the state space in a finite time?

Observability:

states are not all measured directly but instead impact the output via the output equation:

$$y = Cx + Du$$

can we infer fully the initial state from the outputs and the inputs? (can then reveal the full state trajectory through (1))
In-class demo

Controllability and inverted pendulum on a cart

UW Linear Systems (X. Chen, ME547)

Controllability and Observability

5 / 48

The concept of controllability and observability



- assume x(0) = 0
- because of symmetry, we always have

$$x_{1}(t) = x_{3}(t), \ x_{2}(t) = x_{4}(t), \ \forall t \geq 0$$

• state cannot be arbitrarily steered \Rightarrow uncontrollable

Definition

A discrete-time linear system x(k + 1) = A(k)x(k) + B(k)u(k) is called controllable at k = 0 if there exists a finite time k_1 such that for any initial state x(0) and target state x_1 , there exists a control sequence $\{u(k); k = 0, 1, ..., k_1\}$ that will transfer the system from x(0) at k = 0 to x_1 at $k = k_1$.

UW Linear Systems (X. Chen, ME547)

Controllability and Observability

7 / 48

Controllability of LTI systems

$$x(k+1) = Ax(k) + Bu(k) \Rightarrow x(n) = A^{n}x(0) + \sum_{k=0}^{n-1} A^{n-1-k}Bu(k)$$

$$\Rightarrow x(n) - A^{n}x(0) = \underbrace{\left[B, AB, A^{2}B, \dots, A^{n-1}B\right]}_{P_{d}} \begin{bmatrix} u(n-1) \\ u(n-2) \\ \vdots \\ u(0) \end{bmatrix}$$

- given any x (n) and x (0) in ℝⁿ, [u (n − 1), u (n − 2), ..., u (0)]^T can be solved if the columns of P_d span ℝⁿ
- equivalently, system is controllable if P_d has rank n (full row rank)

Controllability of LTI systems Cont'd

 $x(k+1) = Ax(k) + Bu(k) \Rightarrow$

$$x(n) - A^{n}x(0) = \underbrace{\left[B, AB, A^{2}B, \dots, A^{n-1}B\right]}_{P_{d}} \begin{bmatrix} u(n-1) \\ u(n-2) \\ \vdots \\ u(0) \end{bmatrix}$$

also, no need to go beyond n: adding AⁿB, Aⁿ⁺¹B, ... does not increase the rank of P_d (Cayley Halmilton Theorem):

$$x(k_1) - A^{k_1}x(0) = \underbrace{\begin{bmatrix} B & AB & \dots & A^{n-1}B \mid \dots & A^{k_1-1}B \end{bmatrix}}_{\text{rank}=\text{rank}(P_d)} \begin{bmatrix} u(k_1 - 1) \\ u(k_1 - 2) \\ \vdots \\ u(0) \end{bmatrix}$$

UW Linear Systems (X. Chen, ME547)

Controllability and Observability

Theorem (Cayley Halmilton Theorem)

Let $A \in \mathbb{R}^{n \times n}$. A^n is linearly dependent with $\{I, A, A^2, \cdots A^{n-1}\}$

Proof.

Consider characteristic polynomial

$$p(\lambda) = \lambda^{n} + c_{n-1}\lambda^{n-1} + \dots + c_{1}\lambda + c_{0} = \det(\lambda I - A)$$

= $(\lambda - \lambda_{1})^{m_{1}} \dots (\lambda - \lambda_{p})^{m_{p}}$
 $\Rightarrow p(A) = A^{n} + c_{n-1}A^{n-1} + \dots + c_{1}A + c_{0}I$
= $(A - \lambda_{1}I)^{m_{1}} \dots (A - \lambda_{p}I)^{m_{p}}, \quad m_{1} + m_{2} + \dots + m_{p} = I$

Take any eigenvector or generalized eigenvector t_i , say, associated to λ_i : $p(A) t_i = (A - \lambda_1 I)^{m_1} \dots (A - \lambda_p I)^{m_p} t_i =$ $(A - \lambda_1 I)^{m_1} \dots (A - \lambda_p I)^{m_p - 1} (\lambda_i t_i - \lambda_p t_i) = (\lambda_i - \lambda_1)^{m_1} \dots (\lambda_i - \lambda_p)^{m_p} t_i = 0$ \blacktriangleright Therefore $p(A) [t_1, t_2, \dots, t_n] = 0$. \triangleright But $T = [t_1, t_2, \dots, t_n]$ is invertible. Hence $p(A) = 0 \Rightarrow A^n = -c_0 I - c_1 A - \dots - c_{n-1} A^{n-1}$. Arthur Cayley: 1821-1895, British mathematician

algebraic theory of curves and surfaces, group theory, linear algebra, graph theory, invariant theory, ...

extraordinarily prolific career: ~1,000 math papers
 William Hamilton: 1805-1865, Irish mathematician

- optics and classical mechanics in physics, dynamics, algebra, quaternions, ...
- quaternions: extending complex numbers to higher spatial dimensions: 4D case

$$i^2 = j^2 = k^2 = ijk = -1$$

now used in computer graphics, control theory, orbital mechanics, e.g., spacecraft attitude-control systems

UW Linear Systems (X. Chen, ME547) Controllability and Observability

11/48

Theorem (Controllability Theorem)

The n-dimensional r-input LTI system with $x(k+1) = Ax(k) + Bu(k), A \in \mathbb{R}^{n \times n}, B \in \mathbb{R}^{n \times r}$ is controllable if and only if either one of the following is satisfied

1. The $n \times nr$ controllability matrix

$$P_d = \left[B, AB, A^2B, \dots, A^{n-1}B\right]$$

has rank n. (proved in previous three slides)

2. The controllability gramian

$$W_{cd} = \sum_{k=0}^{k_1} A^k B B^T \left(A^T \right)^k$$

is nonsingular for some finite k_1 .

Proof: from controllability matrix to gramian

Recall

$$x(n) - A^{n}x(0) = \underbrace{\left[B, AB, A^{2}B, \dots, A^{n-1}B\right]}_{P_{d}} \left[u(n-1), u(n-2), \dots, u(0)\right]^{T}$$
(2)

$$P_d \text{ is full row rank} \Rightarrow P_d P_d^T = \underbrace{\sum_{k=0}^n A^k B B^T (A^T)^k}_{W_{cd} \text{ at } k_1 = n} \text{ is nonsingular}$$

• a solution to (2) is

$$\left[u(n-1), u(n-2), \dots, u(0)\right]^{T} = P_{d}^{T} \left(P_{d} P_{d}^{T}\right)^{-1} \left[x(n) - A^{n} x(0)\right]$$

UW Linear Systems (X. Chen, ME547) Controllability and Observability

13/48

Example

$$A = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 1 \\ 0 & 0 & \lambda_2 \end{bmatrix}, B = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$
$$P_d = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & \lambda_2 + \lambda_2 \\ 1 & \lambda_2 & \lambda_2^2 \end{bmatrix} \Rightarrow \operatorname{rank}(P_d) = 2 < 3 \Rightarrow \text{uncontrollable}$$
Intuition: $\dot{x}_1 = \lambda_1 x_1$ is not impacted by the control input at all.



Example



 \Rightarrow uncontrollable

Analysis: controllability and controllable canonical form

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -a_0 & -a_1 & -a_2 \end{bmatrix}, B = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

controllability matrix

$$P_d = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & -a_2 \\ 1 & -a_2 & -a_1 + a_2^2 \end{bmatrix}$$

has full row rank

system in controllable canonical form is controllable

UW Linear Systems (X. Chen, ME547) Controllability and Observability 17/48

Analysis: controllability gramian and Lyapunov Eq.

$$W_{cd} = \sum_{k=0}^{k_1} A^k B B^T \left(A^T
ight)^k$$

• If A is Schur, k_1 can be set to ∞

$$W_{cd} = \sum_{k=0}^{\infty} A^k \underbrace{BB^T}_{Q} \left(A^T \right)^k$$

which can be solved via the Lyapunov Eq.

$$AW_{cd}A^T - W_{cd} = -BB^T$$

Analysis: controllability and similarity transformation

$$\begin{cases} x(k+1) = Ax(k) + Bu(k) & \xrightarrow{x=Tx^*} \\ y(k) = Cx(k) + Du(k) & \xrightarrow{x=Tx^*} \end{cases} \begin{cases} x^*(k+1) = \overbrace{T^{-1}AT}^{\tilde{A}} x^*(k) + \overbrace{T^{-1}B}^{\tilde{B}} u(k) \\ y(k) = CTx^*(k) + Du(k) \end{cases}$$

controllability matrix

$$P_d^* = \begin{bmatrix} \tilde{B}, \tilde{A}\tilde{B}, \dots, \tilde{A}^{n-1}\tilde{B} \end{bmatrix}$$
$$= \begin{bmatrix} T^{-1}B, T^{-1}AB, \dots, T^{-1}A^{n-1}B \end{bmatrix} = T^{-1}P_d$$

hence (A, B) controllable $\Leftrightarrow (T^{-1}AT, T^{-1}B)$ controllable

The controllability property is invariant under any coordinate transformation.

UW Linear Systems (X. Chen, ME547) Controllability and Observability

19/48

* Popov-Belevitch-Hautus (PBH) controllability

test

the full rank condition of the controllability matrix

 $P_d = \begin{bmatrix} B, AB, A^2B, \dots, A^{n-1}B \end{bmatrix}$

is equivalent to: the matrix $[A - \lambda I, B]$ having full row rank at every eigenvalue, λ , of A

► to see this: if [A - λI, B] is not full row rank then there exists nonzero vector (a left eigenvector) such that

$$v^{T}[A - \lambda I B] = 0$$
$$\Leftrightarrow v^{T}A = \lambda v^{T}$$
$$v^{T}B = 0$$

i.e., the input vector B is orthogonal to a left eigenvector of A.

Example



UW Linear Systems (X. Chen, ME547) Controllability and Observability

21/48

- 1. Concepts
- 2. DT controllability

Controllability and controllable canonical form Controllability and Lyapunov Eq.

- 3. DT observability Observability and observable canonical form
- 4. CT cases
- 5. The degrees of controllability and observability
- 6. Transforming controllable systems into controllable canonical forms
- 7. Transforming observable systems into observable canonical forms

Observability of LTI systems

Definition

A discrete-time linear system

$$x(k+1) = A(k)x(k) + B(k)u(k)$$

$$y(k) = C(k)x(k) + D(k)u(k)$$

is called observable at k = 0 if there exists a finite time k_1 such that for any initial state x(0), the knowledge of the input $\{u(k); k = 0, 1, ..., k_1\}$ and $\{y(k); k = 0, 1, ..., k_1\}$ suffice to determine the state x(0). Otherwise, the system is said to be unobservable at time k = 0.

UW Linear Systems (X. Chen, ME547) Controllability and Observability

23 / 48

Observability of LTI systems

let us start with the unforced system

$$egin{aligned} &x\left(k+1
ight)=Ax\left(k
ight), \ A\in\mathbb{R}^n\ &y\left(k
ight)=Cx\left(k
ight), \ y\in\mathbb{R}^m \end{aligned}$$

 $x\left(k
ight)=A^{k}x\left(0
ight)$ and $y\left(k
ight)=Cx\left(k
ight)$ give

$$\begin{bmatrix} y(0) \\ y(1) \\ \vdots \\ y(n-1) \end{bmatrix} = \underbrace{\begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix}}_{Q_d:nm \times n} x(0)$$

if the linear matrix equation has a nonzero solution x (0), the system is observable.

Observability of LTI systems



Observability of LTI systems

$$\underbrace{\begin{bmatrix} y(0) - y_{\text{forced}}(0) \\ y(1) - y_{\text{forced}}(1) \\ \vdots \\ y(n-1) - y_{\text{forced}}(n-1) \end{bmatrix}}_{Y} = \underbrace{\begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix}}_{Q_{d}} x(0)$$

• x(0) can be solved if Q_d has rank *n* (full column rank):

▶ pick *n* linearly independent rows from Q_d to form Q_{d_n} , yielding $Q_{d_n} \times (0) = Y_n$

• Q_d is full column rank $\Rightarrow Q_{d_n}$ is full column rank $\Rightarrow x(0) = Q_{d_n}^{-1} Y_n$

one way to write the solution is

$$x(0) = \left(Q_d^T Q_d\right)^{-1} Q_d^T Y$$

Observability of LTI systems Cont'd

$$\underbrace{\begin{bmatrix} y(0) - y_{\text{forced}}(0) \\ y(1) - y_{\text{forced}}(1) \\ \vdots \\ y(n-1) - y_{\text{forced}}(n-1) \end{bmatrix}}_{Y} = \underbrace{\begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix}}_{Q_{d}} x(0)$$

> also, no need to go beyond n in Q_d : adding CA^n , CA^{n+1} , ... does not increase the column rank of Q_d (Cayley Halmilton Theorem)

UW Linear Systems (X. Chen, ME547) Controllability and Observability

27 / 48

Theorem (Observability Theorem)

System x(k + 1) = Ax(k) + Bu(k), y(k) = Cx(k) + Du(k), $A \in \mathbb{R}^{n \times n}$, $C \in \mathbb{R}^{m \times n}$ is observable if and only if either one of the following is satisfied

1. The observability matrix $Q_d = \begin{bmatrix} C \\ CA \\ \vdots \\ CAB = 1 \end{bmatrix}$ has full column rank

2. The observability gramian

$$W_{od} = \sum_{k=0}^{k_1} \left(A^T \right)^k C^T C A^k$$
 is nonsingular for some finite k_1

3. * PBF test: The matrix $\begin{bmatrix} A - \lambda I \\ C \end{bmatrix}$ has full column rank at every eigenvalue, λ , of A.

Proof: from observability matrix to gramian

$$Q_{d} = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} \quad W_{od} = \sum_{k=0}^{k_{1}} (A^{T})^{k} C^{T} CA^{k}$$

•
$$Q_d$$
 is full column rank $\Rightarrow Q_d^T Q_d = \sum_{\substack{k=0 \ W_{od} \text{ at } k_1 = n}}^n (A^T)^k C^T C A^k$ is

nonsingular

UW Linear Systems (X. Chen, ME547) Controllability and Observability

29 / 48

Observability check

Analogous to the case in controllability, the observability property is invariant under any coordinate transformation:

$$(A, C)$$
 is observable $\iff (T^{-1}AT, CT)$ is observable

• If A is Schur, k_1 can be set to ∞ in the observability gramian

$$W_{od} = \sum_{k=0}^{\infty} \left(A^T \right)^k C^T C A^k$$

and we can compute by solving the Lyapunov equation

$$A^T W_{od} A - W_{od} = -C^T C$$

The solution is nonsingular if and only if the system is observable. In fact, $W_{od} \succeq 0$ by definition \Rightarrow "nonsingular" can be replaced with "positive definite".

Observability and observable canonical form

$$A = \begin{bmatrix} -a_2 & 1 & 0 \\ -a_1 & 0 & 1 \\ -a_0 & 0 & 0 \end{bmatrix}, \ C = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}$$

observability matrix

$$Q_{d} = \begin{bmatrix} C \\ CA \\ CA^{2} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -a_{2} & 1 & 0 \\ a_{2}^{2} - a_{1} & -a_{2} & 1 \end{bmatrix}$$

has full column rank

system in observable canonical form is observable

UW Linear Systems (X. Chen, ME547) Controllability and Observability 31 / 48 * PBH test for observability The matrix $\begin{bmatrix} A - \lambda I \\ C \end{bmatrix}$ has full column rank at every eigenvalue, λ , of A. ▶ if not full rank then there exists a nonzero eigenvector *v*: $\begin{array}{c} Av = \lambda v \\ Cv = 0 \\ \Rightarrow CAv = \lambda Cv = 0 \\ \vdots \end{array} \rightarrow \left[\begin{array}{c} C \\ CA \\ \vdots \\ CA^{n-1} \end{array} \right] v = 0 \Rightarrow \text{unobservable}$ $CA^{n-1}v = 0$ the reverse direction is analogous **interpretation**: some non-zero initial condition $x_0 = v$ will generate zero output, which is not distinguishable from the origin.

1. Concepts

2. DT controllability

Controllability and controllable canonical form Controllability and Lyapunov Eq.

3. DT observability

Observability and observable canonical form

- 4. CT cases
- 5. The degrees of controllability and observability
- 6. Transforming controllable systems into controllable canonical forms
- 7. Transforming observable systems into observable canonical forms

UW Linear Systems (X. Chen, ME547) Controllability and Observability

33 / 48

Theorem (Controllability of continuous-time systems)

The n-dimensional r-input LTI system with $\dot{x} = Ax + Bu$, $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times r}$ is controllable if and only if either one of the following is satisfied

1. The $n \times nr$ controllability matrix

$$P = \left[B, AB, A^2B, \dots, A^{n-1}B\right]$$

has rank n.

2. The controllability gramian

$$W_{cc} = \int_0^t e^{A\tau} B B^T e^{A^T \tau} d\tau$$

is nonsingular for any t > 0.

Theorem (Observability of continuous-time systems)

System $\dot{x} = Ax + Bu$, y = Cx + Du, $A \in \mathbb{R}^{n \times n}$, $C \in \mathbb{R}^{m \times n}$ is observable if and only if either one of the following is satisfied

1. The $(mn) \times n$ observability matrix

 $Q = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix}$ has rank n (full column rank)

The observability gramian 2.

$$W_{oc} = \int_{0}^{t} e^{A^{T}\tau} C^{T} C e^{A\tau} d\tau$$
 is nonsingular for any $t > 0$

reading: Linear System Theory and Design by Chen, Chap 6.

UW Linear Systems (X. Chen, ME547) Controllability and Observability

35 / 48

Summary: computing the gramians

	Controllability Gramian	Observability Gramian
continuous time	$\int_{0}^{t} e^{A\tau} B B^{T} \left(e^{A\tau} \right)^{T} d\tau$	$\int_0^t \left(e^{A\tau} \right)^T C^T C e^{A\tau} d\tau$
Lyapunov eq.		
if $t ightarrow \infty$	$AW_c + W_c A^T = -BB^T$	$A^T W_o + W_o A = -C^T C$
& A is Hurwitz		
discrete time	$\sum_{k=0}^{k_1} A^k B B^T \left(A^T \right)^k$	$\sum_{k=0}^{k_1} (A^T)^k C^T C A^k$
Lyapunov eq.		
if $k_1 o \infty$	$AW_{cd}A^T - W_{cd} = -BB^T$	$A^T W_{od} A - W_{od} = -C^T C$
& A is Schur		

- duality: (A, B) is controllable if and only if $(\overline{A}, \overline{C}) = (A^T, B^T)$ is observable
- prove by comparing the gramians

Exercise

$$A = \begin{bmatrix} -2 & 0 & 0 \\ 1 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}, B = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$
$$C = \begin{bmatrix} 1 & 0 & 1 \end{bmatrix}$$
$$exercise: show that the system is not observable.$$
$$in fact, by similarity transform \overline{x} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} x, we get$$
$$\overline{A} = \begin{bmatrix} -2 & 0 & | & 0 \\ 0 & 0 & 0 \\ 1 & 2 & | & 0 \end{bmatrix}, \ \overline{B} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$
$$\overline{C} = \begin{bmatrix} 1 & 1 & | & 0 \end{bmatrix}$$
where the third state is not observable.

UW Linear Systems (X. Chen, ME547) Controllability and Observability

37 / 48

- 1. Concepts
- 2. DT controllability

Controllability and controllable canonical form Controllability and Lyapunov Eq.

3. DT observability

Observability and observable canonical form

- 4. CT cases
- 5. The degrees of controllability and observability
- 6. Transforming controllable systems into controllable canonical forms
- 7. Transforming observable systems into observable canonical forms

The degree of controllability

consider two systems

$$S_{1}:x(k+1) = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x(k) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(k)$$
$$S_{2}:x(k+1) = \begin{bmatrix} 0 & 0.01 \\ 0 & 1 \end{bmatrix} x(k) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(k)$$

both systems are controllable:

$$P_{d_1} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \qquad P_{d_2} = \begin{bmatrix} 0 & 0.01 \\ 1 & 1 \end{bmatrix}$$

- however, P_{d2} is nearly singular, hinting that S₂ is not "easy" to control
- e.g., to move from $x(0) = [0, 0]^T$ to $x(1) = [1, 1]^T$ in two steps, we need

$$S_1: \{u(0), u(0)\} = \{1, 1\}$$
 $S_2: \{u(0), u(0)\} = \{100, -99\}$

UW Linear Systems (X. Chen, ME547)

Controllability and Observability

39 / 48

The degree of observability

consider two systems

$$S_{1} : x (k+1) = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x (k) \qquad y (k) = \begin{bmatrix} 1 & 0 \end{bmatrix} x (k)$$
$$S_{2} : x (k+1) = \begin{bmatrix} 1 & 0.01 \\ 0 & 0 \end{bmatrix} x (k) \qquad y (k) = \begin{bmatrix} 1 & 0 \end{bmatrix} x (k)$$

both systems are observable:

$$Q_{d_1}=\left[egin{array}{cc} 1 & 0\ 0 & 1 \end{array}
ight], \qquad Q_{d_2}=\left[egin{array}{cc} 1 & 0\ 1 & 0.01 \end{array}
ight]$$

- however, Q_{d2} is nearly singular, hinting that S₂ is not "easy" to observe
- e.g., to infer $x(0) = [2, 1]^T$, the two measurements y(0) = 2 and y(1) = CAx(0) = 2.001 are nearly identical in S_2 !

1. Concepts

2. DT controllability

Controllability and controllable canonical form Controllability and Lyapunov Eq.

- 3. DT observability Observability and observable canonical form
- 4. CT cases
- 5. The degrees of controllability and observability
- 6. Transforming controllable systems into controllable canonical forms
- 7. Transforming observable systems into observable canonical forms

UW Linear Systems (X. Chen, ME547) Controllability and Observability

Transforming single-input controllable system into ccf

Let
$$x = M\tilde{x}$$
, where $M = \begin{bmatrix} | & | & | & | \\ m_1 & m_2 & \dots & m_n \\ | & | & | & | \end{bmatrix}$, then
 $\dot{\tilde{x}} = M^{-1}\dot{x} = M^{-1}(Ax + Bu) = M^{-1}AM\tilde{x} + \underbrace{M^{-1}B}_{\tilde{B}}u$

If system is controllable, we show how to transform the state equation into the controllable canonical form.

▶ goal 1: \tilde{B} be in controllable canonical form \Leftrightarrow

$$M^{-1}B = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \Rightarrow B = \begin{bmatrix} m_1, m_2, \dots, m_n \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} = m_n$$

41 / 48

Transforming SI controllable system into ccf Let $x = M\tilde{x}$, where $M = [m_1, m_2, ..., m_n]$, then

$$\dot{\tilde{x}} = M^{-1}\dot{x} = M^{-1}\left(Ax + Bu\right) = \underbrace{M^{-1}AM}_{\tilde{A}}\tilde{x} + M^{-1}Bu$$

▶ goal 2: \tilde{A} be in controllable canonical form \Leftrightarrow

$$A[m_1, m_2, \dots, m_n] = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ \vdots & 0 & \ddots & 0 & \vdots \\ \vdots & \ddots & \ddots & 1 & 0 \\ 0 & \dots & 0 & 0 & 1 \\ -a_0 & -a_1 & \dots & -a_{n-1} \end{bmatrix}$$

UW Linear Systems (X. Chen, ME547)

Controllability and Observability

43 / 48

Transforming SI controllable system into ccf

Let $x = M\tilde{x}$, where $M = [m_1, m_2, \ldots, m_n]$, then

$$\dot{\tilde{x}} = M^{-1}\dot{x} = M^{-1}(Ax + Bu) = M^{-1}AM\tilde{x} + M^{-1}Bu$$

solving goals 1 and 2 yields

$$m_n = B$$

 $m_{n-1} = Am_n + a_{n-1}m_n$
 $m_{n-2} = Am_{n-1} + a_{n-2}m_n$
 $m_{i-1} = Am_i + a_{i-1}m_n, i = n, ..., 2$
 \vdots

▶ when implementing, obtain a₀, a₁, ..., a_{n-1} first by calculating det (sl − A) = sⁿ + a_{n-1}sⁿ⁻¹ + ··· + a₁s + a₀

Transforming single-output (SO) observable system into ocf Let $x = R^{-1}\tilde{x}$, where $R = [r_1^T, r_2^T, \dots, r_n^T]^T$ (r_i is a row vector). $\ddot{x} = R\dot{x} = R(Ax + Bu) = \underbrace{RAR^{-1}}_{\tilde{A}}\tilde{x} + RBu$ $y = Cx = \underbrace{CR^{-1}}_{\tilde{C}}\tilde{x}$

If system is observable, we show how to transform the state equation into the observable canonical form.

▶ goal 1: \tilde{C} be in observable canonical form \Leftrightarrow

$$CR^{-1} = \begin{bmatrix} 1\\0\\\vdots\\0 \end{bmatrix}^T \Rightarrow C = r_1$$

UW Linear Systems (X. Chen, ME547)

Controllability and Observability

Transforming SO observable system into ocf Let $x = R^{-1}\tilde{x}$, where $R = [r_1^T, r_2^T, \dots, r_n^T]^T$ (r_i is a row vector).

$$\dot{\tilde{x}} = R\dot{x} = R\left(Ax + Bu\right) = \underbrace{RAR^{-1}}_{\tilde{A}}\tilde{x} + RBu$$
$$y = Cx = \underbrace{CR^{-1}}_{\tilde{C}}\tilde{x}$$

▶ goal 2: \tilde{A} be in observable canonical form \Leftrightarrow

$$\begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{bmatrix} A = \begin{bmatrix} -a_{n-1} & 1 & 0 & \dots & 0 \\ \vdots & 0 & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & 0 \\ -a_1 & \vdots & \ddots & \ddots & 1 \\ -a_0 & 0 & \dots & 0 & 0 \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{bmatrix}$$

45 / 48

Transforming SO observable system into ocf
Let
$$x = R^{-1}\tilde{x}$$
, where $R = [r_1^T, r_2^T, \dots, r_n^T]^T$ (r_i is a row vector).
 $\dot{\tilde{x}} = R\dot{x} = R(Ax + Bu) = \underbrace{RAR^{-1}}_{\tilde{A}}\tilde{x} + RBu$
 $y = Cx = \underbrace{CR^{-1}}_{\tilde{C}}\tilde{x}$
• solving goals 1 and 2 yields
 $r_1 = C$
 $r_2 = r_1A + a_{n-1}r_1$
 $r_3 = r_2A + a_{n-2}r_1$
 $r_{i+1} = r_iA + a_{n-i}r_1$, $i = 1, \dots, n-1$
:

▶ when implementing, obtain a₀, a₁, ..., a_{n-1} first by calculating det (sl − A)

UW Linear Systems (X. Chen, ME547) Controllability and Observability

47 / 48

Transforming SO observable system into ocf

xample
$$x(k+1) = \begin{bmatrix} 1 & 0.01 \\ 0 & 0 \end{bmatrix} x(k) \qquad y(k) = \begin{bmatrix} 1 & 0 \end{bmatrix} x(k)$$

$$det (A - \lambda I) = \lambda^{2} - \lambda \Rightarrow a_{1} = -1, \ a_{0} = 0$$

$$r_{1} = C = [1, 0]$$

$$r_{2} = r_{1}C + a_{1}r_{1} = [1, 0]A + (-1)[1, 0]$$

$$R = \begin{bmatrix} 1 & 0 \\ 0 & 0.01 \end{bmatrix}, R^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 100 \end{bmatrix}$$

$$\tilde{C} = CR^{-1} = [1, 0] \iff ocf!$$

$$\tilde{A} = RAR^{-1} = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \iff ocf!$$

Controllability and Observability

ME 547: Linear Systems Controllable and Observable Subspaces Kalman Canonical Decomposition

Xu Chen

University of Washington

UW Linear Systems (X. Chen, ME547)

Kalman decomposition

1/31

1. Controllable subspace

- 2. Observable subspace
- 3. Separating the uncontrollable subspace

Discrete-time version Continuous-time version Stabilizability

4. Separating the unobservable subspace

Discrete-time version Detectability Continuous-time version

- 5. Transfer-function perspective
- 6. Kalman decomposition

Controllable subspace: Introduction

Example

$$\bar{A} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \ \bar{B} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \Leftrightarrow \begin{cases} x_1(k+1) &= x_1(k) + u(k) \\ x_2(k+1) &= 0 \end{cases}$$
$$\bar{A} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \ \bar{B} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \Leftrightarrow \begin{cases} x_1(k+1) &= x_1(k) + x_2(k) + u(k) \\ x_2(k+1) &= x_2(k) \end{cases}$$

- there exists controllable and uncontrollable states: x₁ controllable and x₂ uncontrollable
- how to compute the dimensions of the two for general systems?
- ▶ how to separate them?

UW Linear Systems (X. Chen, ME547)

Kalman decomposition

Controllable subspace: Assumptions

Consider an uncontrollable LTI system

$$\begin{aligned} x\left(k+1\right) &= Ax\left(k\right) + Bu\left(k\right), \ A \in \mathbb{R}^{n \times n} \\ y\left(k\right) &= Cx\left(k\right) + Du\left(k\right) \end{aligned}$$

Let the controllability matrix

$$P = [B, AB, A^2B, \dots, A^{n-1}B]$$

have rank $n_1 < n$.

3/31

Controllable subspace

▶ The controllable subspace χ_C is the set of all vectors $x \in \mathbb{R}^n$ that can be reached from the origin.

From

$$x(n) - A^{n}x(0) = \underbrace{\left[B, AB, A^{2}B, \dots, A^{n-1}B\right]}_{P} \begin{bmatrix} u(n-1) \\ u(n-2) \\ \vdots \\ u(0) \end{bmatrix}$$

 $\chi_{\mathcal{C}}$ is the range space of P: $\chi_{\mathcal{C}} = \mathcal{R}(P)$

UW Linear Systems (X. Chen, ME547)

Kalman decomposition

5/31

1. Controllable subspace

2. Observable subspace

3. Separating the uncontrollable subspace

Discrete-time version Continuous-time version Stabilizability

4. Separating the unobservable subspace

Discrete-time version Detectability Continuous-time version

- 5. Transfer-function perspective
- 6. Kalman decomposition

Observable subspace: Introduction

Example

$$ar{A} = \left[egin{array}{c} 1 & 0 \ 1 & 1 \end{array}
ight], \ ar{B} = \left[egin{array}{c} 1 \ 0 \end{array}
ight], \ \Leftrightarrow egin{array}{c} x_1(k+1) &= x_1(k) + u(k) \ x_2(k+1) &= x_1(k) + x_2(k) \ y(k) &= x_1(k) + x_2(k) \ y(k) &= x_1(k) \end{array}$$

- exists observable and unobservable states: x₁ observable and x₂ unobservable
- how to separate the two?
- how to separate controllable but observable states, controllable but unobservable states, etc?

UW Linear Systems (X. Chen, ME547)

Kalman decomposition

Observable subspace: Assumptions

Consider an unobservable LTI system

$$\begin{aligned} x\left(k+1\right) &= Ax\left(k\right) + Bu\left(k\right), \ A \in \mathbb{R}^{n \times n} \\ y\left(k\right) &= Cx\left(k\right) + Du\left(k\right) \end{aligned}$$

Let the observability matrix

$$Q = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix}$$

have rank $n_2 < n$.

7/31

Unobservable subspace

- The unobservable subspace \(\chi_{uo}\) is the set of all nonzero initial conditions \(x(0) ∈ \mathbb{R}^n\) that produce a zero free response.
- From

$$\underbrace{\begin{bmatrix} y(0) \\ y(1) \\ \vdots \\ y(n-1) \end{bmatrix}}_{Y} = \underbrace{\begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix}}_{Q} x(0)$$

 $\chi_{\it uo}$ is the null space of ${\it Q}$: $\chi_{\it uo}=\mathcal{N}\left({\it Q}
ight)$

UW Linear Systems (X. Chen, ME547)

Kalman decomposition

9/31

1. Controllable subspace

- 2. Observable subspace
- 3. Separating the uncontrollable subspace

Discrete-time version Continuous-time version Stabilizability

4. Separating the unobservable subspace

Discrete-time version Detectability Continuous-time version

- 5. Transfer-function perspective
- 6. Kalman decomposition

Separating the uncontrollable subspace

▶ recall 1: similarity transform $x = Mx^*$ preserves controllability

$$\begin{cases} x(k+1) = Ax(k) + Bu(k) \\ y(k) = Cx(k) + Du(k) \end{cases} \Rightarrow \begin{cases} x^*(k+1) = M^{-1}AMx^*(k) + M^{-1}Bu(k) \\ y(k) = CMx^*(k) + Du(k) \end{cases}$$

recall 2: the uncontrollable system structure at introduction

$$\bar{A} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \ \bar{B} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \Leftrightarrow \begin{cases} x_1(k+1) &= x_1(k) + x_2(k) + u(k) \\ x_2(k+1) &= x_2(k) \end{cases}$$

decoupled structure for generalized systems

$$\begin{bmatrix} \bar{x}_{c}(k+1) \\ \bar{x}_{uc}(k+1) \end{bmatrix} = \begin{bmatrix} \bar{A}_{c} & \bar{A}_{12} \\ 0 & \bar{A}_{uc} \end{bmatrix} \begin{bmatrix} \bar{x}_{c}(k) \\ \bar{x}_{uc}(k) \end{bmatrix} + \begin{bmatrix} \bar{B}_{c} \\ 0 \end{bmatrix} u(k)$$
$$y(k) = \begin{bmatrix} \bar{C}_{c} & \bar{C}_{uc} \end{bmatrix} \begin{bmatrix} \bar{x}_{c}(k) \\ \bar{x}_{uc}(k) \end{bmatrix} + Du(k)$$

 \bar{x}_{uc} impacted by neither u nor \bar{x}_c .

UW Linear Systems (X. Chen, ME547)

Kalman decomposition

11/31

Theorem (Kalman canonical form (controllability))

Let $x \in \mathbb{R}^n$, x(k+1) = Ax(k) + Bu(k), y(k) = Cx(k) + Du(k) be uncontrollable with rank of the controllability matrix, rank(P) = $n_1 < n$. Let $M = \begin{bmatrix} M_c & M_{uc} \end{bmatrix}$, where $M_c = [m_1, \ldots, m_{n_1}]$ consists of n_1 linearly independent columns of P, and $M_{uc} = [m_{n_1+1}, \ldots, m_n]$ are added columns to complete the basis and yield a nonsingular M. Then $x = M\bar{x}$ transforms the system equation to

$$\begin{bmatrix} \bar{x}_{c}(k+1) \\ \bar{x}_{uc}(k+1) \end{bmatrix} = \begin{bmatrix} \bar{A}_{c} & \bar{A}_{12} \\ 0 & \bar{A}_{uc} \end{bmatrix} \begin{bmatrix} \bar{x}_{c}(k) \\ \bar{x}_{uc}(k) \end{bmatrix} + \begin{bmatrix} \bar{B}_{c} \\ 0 \end{bmatrix} u(k)$$
$$y(k) = \begin{bmatrix} \bar{C}_{c} & \bar{C}_{uc} \end{bmatrix} \begin{bmatrix} \bar{x}_{c}(k) \\ \bar{x}_{uc}(k) \end{bmatrix} + Du(k)$$

Furthermore, (\bar{A}_c, \bar{B}_c) is controllable, and $C(zI - A)^{-1}B + D = \bar{C}_c(zI - \bar{A}_c)^{-1}\bar{B}_c + D$ Theorem (Kalman canonical form (controllability))

$$\begin{bmatrix} \bar{x}_c(k+1) \\ \bar{x}_{uc}(k+1) \end{bmatrix} = \begin{bmatrix} \bar{A}_c & \bar{A}_{12} \\ 0 & \bar{A}_{uc} \end{bmatrix} \begin{bmatrix} \bar{x}_c(k) \\ \bar{x}_{uc}(k) \end{bmatrix} + \overbrace{\begin{bmatrix} \bar{B}_c \\ 0 \end{bmatrix}}^{M^{-1}B} u(k)$$

intuition: the "B" matrix after transformation

- ► columns of $B \in$ column space of P, which is equivalent to $\mathcal{R}(M_c)$
- columns of M_{uc} and M_c are linearly independent ⇒ columns of B ∉ R (M_{uc})

thus

$$B = \begin{bmatrix} M_c & M_{uc} \end{bmatrix} \begin{bmatrix} \frac{\text{denote as } \bar{B}_c}{*} \\ 0 \end{bmatrix} \Rightarrow M^{-1}B = \begin{bmatrix} \bar{B}_c \\ 0 \end{bmatrix}$$

UW Linear Systems (X. Chen, ME547)

Kalman decomposition

13/31

Theorem (Kalman canonical form (controllability))

$$\begin{bmatrix} \bar{x}_c(k+1) \\ \bar{x}_{uc}(k+1) \end{bmatrix} = \underbrace{\begin{bmatrix} \bar{A}_c & \bar{A}_{12} \\ 0 & \bar{A}_{uc} \end{bmatrix}}_{M^{-1}AM} \begin{bmatrix} \bar{x}_c(k) \\ \bar{x}_{uc}(k) \end{bmatrix} + \begin{bmatrix} \bar{B}_c \\ 0 \end{bmatrix} u(k)$$

intuition: the "A" matrix after transformation

▶ range space of *M_c* is "*A*-invariant":

columns of $AM_c \in \{AB, A^2B, \ldots, A^nB\} \in \mathcal{R}(M_c)$

where columns of $A^{n}B \in \mathcal{R}(P) = \mathcal{R}(M_{c})$ (:: Cayley Halmilton Thm)

• i.e.,
$$AM_c = M_c \bar{A}_c$$
 for some $\bar{A}_c \Rightarrow$

$$A[M_c, M_{uc}] = [M_c, M_{uc}] \underbrace{\begin{bmatrix} \bar{A}_c & \stackrel{\triangleq \bar{A}_{12}}{& & \\ & \stackrel{\triangleq \bar{A}_{uc}}{& & \\ & 0 & \stackrel{\textstyle{}}{\overset{\textstyle{}}{\overset{\textstyle{}}{\overset{\textstyle{}}{\overset{\textstyle{}}}}} \end{bmatrix}}_{\bar{a}} \Rightarrow M^{-1}AM = \begin{bmatrix} \bar{A}_c & \bar{A}_{12} \\ 0 & \bar{A}_{uc} \end{bmatrix}$$

Theorem (Kalman canonical form (controllability))

$$\begin{bmatrix} \bar{x}_c(k+1) \\ \bar{x}_{uc}(k+1) \end{bmatrix} = \underbrace{\left[\begin{array}{c} \bar{A}_c & \bar{A}_{12} \\ 0 & \bar{A}_{uc} \end{array} \right]}_{\left[\begin{array}{c} \bar{x}_c(k) \\ \bar{x}_{uc}(k) \end{array} \right]} + \underbrace{\left[\begin{array}{c} \bar{B}_c \\ 0 \end{bmatrix}}_{\left[\begin{array}{c} \bar{B}_c \\ 0 \end{bmatrix}} u(k)$$

 (\bar{A}_c, \bar{B}_c) is controllable

controllability matrix after similarity transform

$$\bar{P} = \begin{bmatrix} \bar{B}_c & \bar{A}_c \bar{B}_c & \dots & \bar{A}_c^{n_1-1} \bar{B}_c \\ 0 & 0 & \dots & 0 \end{bmatrix} \dots \quad \bar{A}_c^{n-1} \bar{B}_c \\ = \begin{bmatrix} \bar{P}_c & \bar{A}_c^{n_1} \bar{B}_c & \dots & \bar{A}_c^{n-1} \bar{B}_c \\ 0 & 0 & \dots & 0 \end{bmatrix}$$

▶ similarity transform does not change controllability⇒ $rank(\bar{P}) = rank(P) = n_1$

▶ thus $\operatorname{rank}(\bar{P}_c) = n_1 \Rightarrow (\bar{A}_c, \bar{B}_c)$ is controllable

UW Linear Systems (X. Chen, ME547)

Kalman decomposition

15 / 31

Theorem (Kalman canonical form (controllability))

$$\begin{bmatrix} \bar{x}_{c}(k+1) \\ \bar{x}_{uc}(k+1) \end{bmatrix} = \begin{bmatrix} \bar{A}_{c} & \bar{A}_{12} \\ 0 & \bar{A}_{uc} \end{bmatrix} \begin{bmatrix} \bar{x}_{c}(k) \\ \bar{x}_{uc}(k) \end{bmatrix} + \begin{bmatrix} \bar{B}_{c} \\ 0 \end{bmatrix} u(k)$$
$$y(k) = \begin{bmatrix} \bar{C}_{c} & \bar{C}_{uc} \end{bmatrix} \begin{bmatrix} \bar{x}_{c}(k) \\ \bar{x}_{uc}(k) \end{bmatrix} + Du(k)$$

 $\underline{C(zI - A)^{-1}B + D = \bar{C}_c(zI - \bar{A}_c)^{-1}\bar{B}_c + D}$

we can check that

$$\begin{bmatrix} \bar{C}_{c} & \bar{C}_{uc} \end{bmatrix} \begin{bmatrix} zI - \bar{A}_{c} & -\bar{A}_{12} \\ 0 & zI - \bar{A}_{uc} \end{bmatrix}^{-1} \begin{bmatrix} \bar{B}_{c} \\ 0 \end{bmatrix} + D$$
$$= \begin{bmatrix} \bar{C}_{c} & \bar{C}_{uc} \end{bmatrix} \begin{bmatrix} (zI - \bar{A}_{c})^{-1} & * \\ 0 & (zI - \bar{A}_{uc})^{-1} \end{bmatrix} \begin{bmatrix} \bar{B}_{c} \\ 0 \end{bmatrix} + D$$
$$= \bar{C}_{c} (zI - \bar{A}_{c})^{-1} \bar{B}_{c} + D$$

Matlab commands

$$\begin{bmatrix} \bar{x}_c(k+1) \\ \bar{x}_{uc}(k+1) \end{bmatrix} = \underbrace{\begin{bmatrix} \bar{A}_c & \bar{A}_{12} \\ 0 & \bar{A}_{uc} \end{bmatrix}}_{M^{-1}AM} \begin{bmatrix} \bar{x}_c(k) \\ \bar{x}_{uc}(k) \end{bmatrix} + \underbrace{\begin{bmatrix} \bar{B}_c \\ 0 \end{bmatrix}}_{M^{-1}B} u(k)$$

 $x = M\bar{x}$ where $M = \begin{bmatrix} M_c & M_{uc} \end{bmatrix}$

- *M_c* = [*m*₁,..., *m<sub>n₁*] consists of all the linearly independent columns of *P*: Mc = orth(P)
 </sub>
- *M_{uc}* = [*m_{n1+1}*,...,*m_n*] are added columns to complete the basis and yield a nonsingular *M*
 - from linear algebra: the orthogonal complement of the range space of P is the null space of P^T:

$$\mathbb{R}^{n}=\mathcal{R}\left(P\right)\oplus\mathcal{N}\left(P^{T}\right)$$

hence Muc = null(P') (the transpose is important here)

UW Linear Systems (X. Chen, ME547)

Kalman decomposition

17 / 31

The techniques apply to CT systems

Theorem (Kalman canonical form (controllability))

Let a n-dimensional state-space system $\dot{x} = Ax + Bu$, y = Cx + Dube uncontrollable with the rank of the controllability matrix rank(P) = $n_1 < n$. Let $M = \begin{bmatrix} M_c & M_{uc} \end{bmatrix}$ where $M_c = [m_1, \ldots, m_{n_1}]$ consists of n_1 linearly independent columns of P, $M_{uc} = [m_{n_1+1}, \ldots, m_n]$ are added columns to complete the basis for \mathbb{R}^n and yield a nonsingular M. Then the similarity transformation $x = M\bar{x}$ transforms the system equation to

$$\frac{d}{dt} \begin{bmatrix} \bar{x}_c \\ \bar{x}_{uc} \end{bmatrix} = \begin{bmatrix} \bar{A}_c & \bar{A}_{12} \\ 0 & \bar{A}_{uc} \end{bmatrix} \begin{bmatrix} \bar{x}_c \\ \bar{x}_{uc} \end{bmatrix} + \begin{bmatrix} \bar{B}_c \\ 0 \end{bmatrix} u$$
$$y = \begin{bmatrix} \bar{C}_c & \bar{C}_{uc} \end{bmatrix} \begin{bmatrix} \bar{x}_c \\ \bar{x}_{uc} \end{bmatrix} + Du$$

Example

$$\frac{d}{dt} \begin{bmatrix} \mathbf{v}_m \\ F_{k_1} \\ F_{k_2} \end{bmatrix} = \begin{bmatrix} -b/m & -1/m & -1/m \\ k_1 & 0 & 0 \\ k_2 & 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{v}_m \\ F_{k_1} \\ F_{k_2} \end{bmatrix} + \begin{bmatrix} 1/m \\ 0 \\ 0 \end{bmatrix} F$$
Let $m = 1, b = 1$

$$P = \begin{bmatrix} 1 & -1 & 1-k_1-k_2 \\ 0 & k_1 & -k_1 \\ 0 & k_2 & -k_2 \end{bmatrix}, \ M = \begin{bmatrix} 1 & -1 & 0 \\ 0 & k_1 & 0 \\ 0 & k_2 & 1 \end{bmatrix}, \ M^{-1} = \begin{bmatrix} 1 & 1/k_1 & 0 \\ 0 & 1/k_1 & 0 \\ 0 & -k_2/k_1 & 1 \end{bmatrix}$$
 $\bar{A} = M^{-1}AM = \begin{bmatrix} 0 & -(k_1+k_2) & 1 \\ 1 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \ \bar{B} = M^{-1}B = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$

UW Linear Systems (X. Chen, ME547)

Kalman decomposition

19/31

Stabilizability

$$\begin{bmatrix} \bar{x}_{c}(k+1) \\ \bar{x}_{uc}(k+1) \end{bmatrix} = \begin{bmatrix} \bar{A}_{c} & \bar{A}_{12} \\ 0 & \bar{A}_{uc} \end{bmatrix} \begin{bmatrix} \bar{x}_{c}(k) \\ \bar{x}_{uc}(k) \end{bmatrix} + \begin{bmatrix} \bar{B}_{c} \\ 0 \end{bmatrix} u(k)$$
$$y(k) = \begin{bmatrix} \bar{C}_{c} & \bar{C}_{uc} \end{bmatrix} \begin{bmatrix} \bar{x}_{c}(k) \\ \bar{x}_{uc}(k) \end{bmatrix} + Du(k)$$

The system is *stabilizable* if

- all its unstable modes, if any, are controllable
- i.e., the uncontrollable modes are stable (A
 _{uc} is Schur, namely, all eigenvalues are in the unit circle)

- 1. Controllable subspace
- 2. Observable subspace
- Separating the uncontrollable subspace Discrete-time version Continuous-time version Stabilizability
- Separating the unobservable subspace Discrete-time version Detectability Continuous-time version
- 5. Transfer-function perspective
- 6. Kalman decomposition

UW Linear Systems (X. Chen, ME547)

Kalman decomposition

21/31

Separating the unobservable subspace

Fixed recall 1: similarity transform $x = O^{-1}x^*$ preserves observability

$$\begin{cases} x(k+1) = Ax(k) + Bu(k) \\ y(k) = Cx(k) + Du(k) \end{cases} \Rightarrow \begin{cases} x^*(k+1) = OAO^{-1}x^*(k) + OBu(k) \\ y(k) = CO^{-1}x^*(k) + Du(k) \end{cases}$$

an unobservable system structure

$$\bar{A} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \ \bar{B} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \Leftrightarrow \begin{cases} x_1(k+1) &= x_1(k) + u(k) \\ x_2(k+1) &= x_1(k) + x_2(k) \\ y(k) &= x_1(k) \end{cases}$$
$$\bar{C} = \begin{bmatrix} 1 & 0 \end{bmatrix}$$

decoupled structure for generalized systems

$$\begin{bmatrix} \bar{x}_o(k+1) \\ \bar{x}_{uo}(k+1) \end{bmatrix} = \begin{bmatrix} \bar{A}_o & 0 \\ \bar{A}_{21} & \bar{A}_{uo} \end{bmatrix} \begin{bmatrix} \bar{x}_o(k) \\ \bar{x}_{uo}(k) \end{bmatrix} + \begin{bmatrix} \bar{B}_o \\ \bar{B}_{uo} \end{bmatrix} u(k)$$
$$y(k) = \begin{bmatrix} \bar{C}_o & 0 \end{bmatrix} \begin{bmatrix} \bar{x}_o(k) \\ \bar{x}_{uo}(k) \end{bmatrix} + Du(k)$$

the "observed" \bar{x}_o doesn't reflect \bar{x}_{uc} $(\bar{x}_o(k+1) = \bar{A}_o \bar{x}_o(k) + \bar{B}_o u(k))$

Theorem (Kalman canonical form (observability))

Let $x \in \mathbb{R}^n$, x(k+1) = Ax(k) + Bu(k), y(k) = Cx(k) + Du(k) be unobservable with rank of the observability matrix,

rank(Q) = $n_2 < n$. Let $O = \begin{bmatrix} O_o \\ O_{uo} \end{bmatrix}$ where O_o consists of n_2 linearly independent rows of Q, and $O_{uo} = \begin{bmatrix} o_{n_1+1}^T, \dots, o_n^T \end{bmatrix}^T$ are added rows to complete the basis and yield a nonsingular O. Then $\bar{x} = Ox$ transforms the system equation to

$$\begin{bmatrix} \bar{x}_o(k+1) \\ \bar{x}_{uo}(k+1) \end{bmatrix} = \begin{bmatrix} \bar{A}_o & 0 \\ \bar{A}_{21} & \bar{A}_{uo} \end{bmatrix} \begin{bmatrix} \bar{x}_o(k) \\ \bar{x}_{uo}(k) \end{bmatrix} + \begin{bmatrix} \bar{B}_o \\ \bar{B}_{uo} \end{bmatrix} u(k)$$
$$y(k) = \begin{bmatrix} \bar{C}_o & 0 \end{bmatrix} \begin{bmatrix} \bar{x}_o(k) \\ \bar{x}_{uo}(k) \end{bmatrix} + Du(k)$$

Furthermore, (\bar{A}_o, \bar{O}_o) is observable, and $C(zI - A)^{-1}B + D = \bar{C}_o(zI - \bar{A}_o)^{-1}\bar{B}_o + D$

UW Linear Systems (X. Chen, ME547)

Kalman decomposition

23/31

Theorem (Kalman canonical form)

Case for observability

$$\begin{bmatrix} \bar{x}_o(k+1) \\ \bar{x}_{uo}(k+1) \end{bmatrix} = \begin{bmatrix} \bar{A}_o & 0 \\ \bar{A}_{21} & \bar{A}_{uo} \end{bmatrix} \begin{bmatrix} \bar{x}_o(k) \\ \bar{x}_{uo}(k) \end{bmatrix} + \begin{bmatrix} \bar{B}_o \\ \bar{B}_{uo} \end{bmatrix} u(k)$$
$$y(k) = \begin{bmatrix} \bar{C}_o & 0 \end{bmatrix} \begin{bmatrix} \bar{x}_o(k) \\ \bar{x}_{uo}(k) \end{bmatrix} + Du(k)$$

v.s. case for controllability

$$\begin{bmatrix} \bar{x}_{c}(k+1) \\ \bar{x}_{uc}(k+1) \end{bmatrix} = \begin{bmatrix} \bar{A}_{c} & \bar{A}_{12} \\ 0 & \bar{A}_{uc} \end{bmatrix} \begin{bmatrix} \bar{x}_{c}(k) \\ \bar{x}_{uc}(k) \end{bmatrix} + \begin{bmatrix} \bar{B}_{c} \\ 0 \end{bmatrix} u(k)$$
$$y(k) = \begin{bmatrix} \bar{C}_{c} & \bar{C}_{uc} \end{bmatrix} \begin{bmatrix} \bar{x}_{c}(k) \\ \bar{x}_{uc}(k) \end{bmatrix} + Du(k)$$

Intuition: duality between controllability and observability

(A, B) unconrollable $\Leftrightarrow (A^T, B^T)$ unobservable

Kalman decomposition

Detectability

$$\begin{bmatrix} \bar{x}_o(k+1) \\ \bar{x}_{uo}(k+1) \end{bmatrix} = \begin{bmatrix} \bar{A}_o & 0 \\ \bar{A}_{21} & \bar{A}_{uo} \end{bmatrix} \begin{bmatrix} \bar{x}_o(k) \\ \bar{x}_{uo}(k) \end{bmatrix} + \begin{bmatrix} \bar{B}_o \\ \bar{B}_{uo} \end{bmatrix} u(k)$$
$$y(k) = \begin{bmatrix} \bar{C}_o & 0 \end{bmatrix} \begin{bmatrix} \bar{x}_o(k) \\ \bar{x}_{uo}(k) \end{bmatrix} + Du(k)$$

The system is *detectable* if

- all its unstable modes, if any, are observable
- i.e., the unobservable modes are stable $(\bar{A}_{uo}$ is Schur)

UW Linear Systems (X. Chen, ME547)

Kalman decomposition

25 / 31

Continuout-time version

Theorem (Kalman canonical form (observability))

Let a n-dimensional state-space system $\dot{x} = Ax + Bu$, y = Cx + Dube unobservable with the rank of the observability matrix rank(Q) = $n_2 < n$. Then there exists similarity transform $\bar{x} = Ox$ that transforms the system equation to

$$\frac{d}{dt} \begin{bmatrix} \bar{x}_{o} \\ \bar{x}_{uo} \end{bmatrix} = \begin{bmatrix} \bar{A}_{o} & 0 \\ \bar{A}_{21} & \bar{A}_{uo} \end{bmatrix} \begin{bmatrix} \bar{x}_{o} \\ \bar{x}_{uo} \end{bmatrix} + \begin{bmatrix} \bar{B}_{o} \\ \bar{B}_{uo} \end{bmatrix} u$$
$$y = \begin{bmatrix} \bar{C}_{o} & 0 \end{bmatrix} \begin{bmatrix} \bar{x}_{o} \\ \bar{x}_{uo} \end{bmatrix} + Du$$

Furthermore, (\bar{A}_o, \bar{C}_o) is observable, and $C(sI - A)^{-1}B + D = \bar{C}_o(sI - \bar{A}_o)^{-1}\bar{B}_o + D.$

- 1. Controllable subspace
- 2. Observable subspace
- Separating the uncontrollable subspace Discrete-time version Continuous-time version Stabilizability

4. Separating the unobservable subspace Discrete-time version

Detectability Continuous-time version

5. Transfer-function perspective

6. Kalman decomposition

UW Linear Systems (X. Chen, ME547)

Kalman decomposition

27 / 31

Transfer-function perspective

uncontrollable system: $C(zI - A)^{-1}B + D = \overline{C}_c(zI - \overline{A}_c)^{-1}\overline{B}_c + D$ unobservable system: $C(zI - A)^{-1}B + D = \overline{C}_o(zI - \overline{A}_o)^{-1}\overline{B}_o + D$ where $A \in \mathbb{R}^{n \times n}$, $\overline{A}_c \in \mathbb{R}^{n_1 \times n_1}$, $\overline{A}_o \in \mathbb{R}^{n_2 \times n_2}$

Order reduction exists

$$G(z) = C(zI - A)^{-1}B + D = \frac{B(z)}{A(z)}, \ A(z) = \det(zI - A) \ \text{order}: n$$

$$G(z) = \overline{C}_c(zI - \overline{A}_c)^{-1}\overline{B}_c + D = \frac{\overline{B}_c(z)}{\overline{A}_c(z)}, \ \overline{A}_c(z) = \det\left(zI - \overline{A}_c\right) \text{ order } : n_1$$

► ⇒A(z) and B(z) are not co-prime | pole-zero cancellation exists

same applies to unobservable systems
Example

Consider

$$\frac{d}{dt} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u$$
$$y = \begin{bmatrix} c_1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

The transfer function is

$$G(s) = \frac{s + c_1}{s^2 + 3s + 2} = \frac{s + c_1}{(s + 1)(s + 2)}$$

System is in controllable canonical form and is controllable.

observability matrix

$$Q = \left[egin{array}{ccc} c_1 & 1 \ -2 & c_1 - 3 \end{array}
ight], \; \det Q = (c_1 - 1) \, (c_1 - 2)$$

 \Rightarrow unobservable if $c_1 = 1$ or 2

UW Linear Systems (X. Chen, ME547)

Kalman decomposition

29 / 31

- 1. Controllable subspace
- 2. Observable subspace
- 3. Separating the uncontrollable subspace

Discrete-time version Continuous-time version Stabilizability

4. Separating the unobservable subspace

- Discrete-time version Detectability Continuous-time version
- 5. Transfer-function perspective

6. Kalman decomposition

Kalman decomposition

an extended example:

$$A = \begin{bmatrix} A_{11} & 0 & A_{13} & 0 \\ \hline A_{21} & A_{22} & A_{23} & A_{24} \\ \hline 0 & 0 & A_{33} & 0 \\ 0 & 0 & A_{43} & A_{44} \end{bmatrix}, \ B = \begin{bmatrix} B_1 \\ B_2 \\ \hline 0 \\ 0 \end{bmatrix}$$
$$C = \begin{bmatrix} C_1 & 0 & C_3 & 0 \end{bmatrix}$$

- \blacktriangleright A_{ij} , C_i and B_i are nonzero
- The A₁₁ mode is controllable and observable. The A₂₂ mode is controllable but not observable. The A₃₃ mode is not controllable but observable. The A₄₄ mode is not controllable and not observable.

UW Linear Systems (X. Chen, ME547)

Kalman decomposition

ME 547: Linear Systems State Feedback Control

Xu Chen

University of Washington

UW Linear Systems (X. Chen, ME547)

State Feedback

Motivation

- At the center of designing control systems is the idea of feedback.
- In such transfer-function approaches as lead-lag and root locus methods, the primal goal is to achieve a proper map of closed-loop poles with output feedback.

Key questions:

- How much freedom do we have for state-space systems?
- Are there fundamental system properties that yield higher achievable performance?
- How to implement the design algorithms?

1. Goal and realization of state feedback

2. Closed-loop eigenvalue placement by state feedback

UW Linear Systems (X. Chen, ME547)

State Feedback

Goal

Consider an *n*-dimensional state-space system

$$\Sigma: \begin{cases} \dot{x}(t) = Ax(t) + Bu(t) \\ y(t) = Cx(t) + Du(t) \end{cases} \quad x(t_0) = x_0$$

where $x \in \mathbb{R}^n$, $u \in \mathbb{R}^r$, and $y \in \mathbb{R}^m$.

- Denominators of the transfer function
 G (s) = C (sl A)⁻¹ B + D come from the characteristic polynomial det (sl A) that arises when computing the inverse (sl A)⁻¹.
- We shall investigate the use of feedback to alter the qualitative behavior of the system by changing the eigenvalues of the closed-loop "A" matrix.

Realization

$$v \xrightarrow{+} 0 \xrightarrow{u} \overline{x} = Ax + Bu, y = Cx + Du \longrightarrow y$$

Consider the state-feedback law

$$u = -Kx + v \tag{1}$$

- v: new input which we will deal with later
- ▶ $K \in \mathbb{R}^{m \times n}$: *n*-number of states, *m*-number of inputs
- closed-loop system:

$$\Sigma_{cl}: \begin{cases} \dot{x}(t) = (A - BK)x(t) + Bv(t) \\ y(t) = Cx(t) + Du(t) \end{cases} \qquad x(t_0) = x_0 \quad (2)$$

- \blacktriangleright key closed-loop property: eigenvalues of A BK.
- ▶ How freely can we place the eigenvalues of $A_{cl} = A BK$?

UW Linear Systems (X. Chen, ME547)

State Feedback

5/16

1. Goal and realization of state feedback

2. Closed-loop eigenvalue placement by state feedback

Eigenvalue placement by state feedback

Fact

If $\Sigma = (A, B, C, D)$ is in controllable canonical form, we can completely change all the eigenvalues of A - BK by choice of state-feedback gain matrix K.

Problem setup: single-input single-output system in c.c.f.

$$H(s) = \frac{\beta_{n-1}s^{n-1} + \dots + \beta_{1}s + \beta_{0}}{s^{n} + \alpha_{n-1}s^{n-1} + \dots + \alpha_{1}s + \alpha_{0}} + d, \quad \Sigma = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

$$A = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & \vdots \\ \vdots & \dots & \ddots & \ddots & 0 \\ 0 & \dots & \dots & 0 & 1 \\ -\alpha_{0} & \dots & \dots & -\alpha_{n-2} & -\alpha_{n-1} \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

$$C = \begin{bmatrix} \beta_{0} & \beta_{1} & \dots & \dots & \beta_{n-1} \end{bmatrix}, \quad D = d$$

$$\det(sI - A) = s^{n} + \alpha_{n-1}s^{n-1} + \dots + \alpha_{1}s + \alpha_{0} \qquad (3)$$
WW Linear Systems (X. Chen, ME547) Mathematical State State Mathematical State Mathematical State Mathematical State State Mathematical State Mathematical State Mathematical State Mathematical State Mathmatical State Mathematical S

Eigenvalue placement by state feedback: c.c.f.

Goal: achieve desired closed-loop eigenvalue locations
 p₁, · · · , p_n, i.e.

$$\det(sI - (A - BK)) = (s - p_1)(s - p_2) \cdots (s - p_n)$$
(4)

$$= s^{n} + \gamma_{n-1}s^{n-1} + \dots + \gamma_{1}s + \gamma_{0} \qquad (5)$$

• Let $K = [k_0, k_1, \dots, k_{n-1}]$. The structured A and B give

$$BK = \begin{bmatrix} 0\\0\\\vdots\\0\\1 \end{bmatrix} \begin{bmatrix} k_0, k_1, \dots, k_{n-1} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & \dots & 0\\0 & 0 & 0 & 0 & \vdots\\\vdots\\\vdots\\\vdots\\ k_0 & \dots & \dots & 0\\k_0 & \dots & \dots & k_{n-2} & k_{n-1} \end{bmatrix}$$
$$A - BK = \begin{bmatrix} 0 & 1 & 0 & \dots & 0\\0 & 0 & 1 & 0 & \dots & k_{n-2} & k_{n-1} \end{bmatrix}$$
$$\begin{bmatrix} 0 & 1 & 0 & \dots & 0\\0 & 0 & 1 & 0 & \vdots\\\vdots\\ \vdots\\ \dots & \ddots & \ddots & 0\\0 & \dots & \dots & 0 & 1\\-\alpha_0 - k_0 & \dots & \dots & -\alpha_{n-2} - k_{n-2} & -\alpha_{n-1} - k_{n-1} \end{bmatrix}$$

Eigenvalue placement by state feedback: c.c.f.

- A and A BK have the same structure
- the only difference is the last row:



Eigenvalue placement by state feedback: c.c.f.

Eigenvalue-placement Algorithm

- 1 determine desired eigenvalue locations p_1, \cdots, p_n
- 2 calculate desired closed-loop characteristic polynomial $(s p_i)(s p_i) \dots (s p_i) = s^n + \gamma \dots s^{n-1} + \dots$
- $\begin{vmatrix} (s-p_1)(s-p_2)\cdots(s-p_n) = s^n + \gamma_{n-1}s^{n-1} + \cdots + \gamma_1s + \gamma_0 \\ \text{calculate open-loop characteristic polynomial} \end{vmatrix}$

$$det(sI - A) = s^n + \alpha_{n-1}s^{n-1} + \cdots + \alpha_1s + \alpha_0$$

4 define the matrices:

 $\boldsymbol{K} = [\gamma_0 - \alpha_0, \dots, \gamma_{n-1} - \alpha_{n-1}]$

Powerful result: if the system is in controllable canonical form, we can arbitrarily place the closed-loop eigenvalues by state feedback!

General eigenvalue placement by state feedback

- What if the given state-space realization Σ = (A, B, C, D) is not in the required form?
- We can then transform it to c.c.f. via a similarity transformation (See lecture on controllability and observability).
- Powerful fact: if system Σ = (A, B, C, D) is controllable, then we can arbitrarily place the closed-loop eigenvalues via state feedback.

UW Linear Systems (X. Chen, ME547)

State Feedback

Stabilization

- if a single-input system is uncontrollable, arbitrary closed-loop eigenvalue plaement is not available
- Kalman decomposition gives



applying controll law

$$u = -\left[\bar{K}_{c}, \bar{K}_{uc}\right] \left[\begin{array}{c} \bar{x}_{c} \\ \bar{x}_{uc} \end{array}\right] + v$$

gives

$$\frac{d}{dt} \begin{bmatrix} \bar{x}_c \\ \bar{x}_{uc} \end{bmatrix} = \begin{bmatrix} \bar{A}_c - \bar{B}_c \bar{K}_c & \bar{A}_{12} - \bar{B}_c \bar{K}_{uc} \\ 0 & \bar{A}_{uc} \end{bmatrix} \begin{bmatrix} \bar{x}_c \\ \bar{x}_{uc} \end{bmatrix} + \begin{bmatrix} \bar{B}_c \\ 0 \end{bmatrix} v$$

Stabilization cont'd

closed-loop dynamics

$$\frac{d}{dt} \begin{bmatrix} \bar{x}_{c} \\ \bar{x}_{uc} \end{bmatrix} = \underbrace{\begin{bmatrix} \bar{A}_{c} - \bar{B}_{c}\bar{K}_{c} & \bar{A}_{12} - \bar{B}_{c}\bar{K}_{uc} \\ 0 & \bar{A}_{uc} \end{bmatrix}}_{\bar{A}_{cl}} \begin{bmatrix} \bar{x}_{c} \\ \bar{x}_{uc} \end{bmatrix} + \begin{bmatrix} \bar{B}_{c} \\ 0 \end{bmatrix} v$$

closed-loop eigenvalues come from

$$\det \left(\bar{A}_{cI} - \lambda I\right) = \underbrace{\det \left(\overline{\left(\bar{A}_{c} - \bar{B}_{c}\bar{K}_{c}\right) - \lambda I\right)}}_{\text{from the controllable subsystem}} \cdot \underbrace{\det \left(\bar{A}_{uc} - \lambda I\right)}_{\text{uncontrollable eigenvalues}}$$

► ⇒: single-input systems are stabilizable if and only if the uncontrollable portion of the system does not have any unstable eigenvalue.

State Feedback

Discrete-time case

the eigenvalue assignment of discrete-time systems is analogous:

system dynamics:

$$x (k+1) = Ax (k) + Bu (k)$$
$$y (k) = Cx (k)$$

• controller: u(k) = -Kx(k) + v(k)

closed-loop dynamics:

$$x(k+1) = Ax(k) - BKx(k) + Bv(k) = (A - BK)x(k) + Bv(k)$$

 arbitrary closed-loop eigenvalue assignment if system is controllable

The case with output feedback

 if the full state is not measurable, state feedback control is not feasible

consider output feedback

$$\begin{cases} \dot{x} = Ax + Bu \\ y = Cx \Rightarrow \dot{x} = Ax - BFy + Bv = (A - BFC)x + Bv \\ u = -Fy + v \end{cases}$$

- A BFC not as structured as A BK (exercise: write out the case for SISO systems)
- arbitrary closed-loop eigenvalue assignment not feasible

UW Linear Systems (X. Chen, ME547)

State Feedback

15 / 16

The case with output feedback

Example

Controllable mass-spring-damper system

$$\frac{d}{dt} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{b}{m} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{1}{m} \end{bmatrix} u$$
$$\overset{u^* \triangleq \frac{u}{m}}{=} \begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{b}{m} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u^*$$

- ▶ arbitrary closed-loop eigenvalue assignment if $u^* = -k_1x_1 k_2x_2$, namely $U^*(s) = -k_1X_1(s) - k_2X_2(s) = -(k_1 + k_2s)X_1(s) \Rightarrow$ a proportional plus derivative (PD) control law
- if with only proportional control, u* = -k₁x₁, arbitrary closed-loop eigenvalue assignment is not possible

ME 547: Linear Systems Observers and Observer State Feedback Control

Xu Chen

University of Washington

UW Linear Systems (X. Chen, ME547)

Observers and Observer State FB

1/25

Introduction

- full state feedback is usually not available
- the state estimation problem
 - deterministic case: observer design
 - stochastic case: the most frequent option is Kalman filter

Outline

- 1. Concepts
- 2. Continuous-time Luenberger observer
- 3. Discrete-time observers DT full state observer DT full state observer with predictor
- 4. Observer state feedback

UW Linear Systems (X. Chen, ME547) Observers and Observer State FB

3 / 25

Open-loop observer

$$\frac{d}{dt}x(t) = Ax(t) + Bu(t), \ x(k+1) = Ax(k) + Bu(k)$$

conceptually simplest scheme to estimate x:

$$\frac{d}{dt}\hat{x}(t) = A\hat{x}(t) + Bu(t), \ \hat{x}(k+1) = A\hat{x}(k) + Bu(k)$$

with a best guess of initial estimate $\hat{x}(0) \stackrel{e.g.}{=} 0$.

• error dynamics: $e = x - \hat{x}$:

$$\dot{e}(t) = Ae(t), \ e(k+1) = Ae(k), \ e(0) = x_0 - \hat{x}(0)$$

sensitive to input disturbances

- ▶ if *A* is not Hurwitz/Schur, the error diverges
- open-loop observers look simple but do not work in practice

Luenberger (closed-loop) observer concept

given system dynamics

$$\dot{x} = Ax + Bu, \ x(0) = x_0, \ A \in \mathbb{R}^{n \times n}, B \in \mathbb{R}^{n \times r}$$

 $y = Cx, \ y \in \mathbb{R}^{m \times n}$

 in contrast to open-loop observers, the Luenberger observer adds correction based on output differences





Luenberger (closed-loop) observer algorithm



observer realization:

 $\dot{\hat{x}} = A\hat{x} + Bu + L(y - \hat{y}) = A\hat{x} + Bu + L(y - C\hat{x}), \ \hat{x}(0) = 0$ = (A - LC) $\hat{x} + Ly + Bu$

Luenberger (closed-loop) observer error dynamics

system dynamics

$$\dot{x} = Ax + Bu, \ x(0) = x_0, \ A \in \mathbb{R}^{n \times n}, B \in \mathbb{R}^{n \times n}$$

 $y = Cx, \ y \in \mathbb{R}^{m \times n}$

Luenberger observer with correction:

$$\dot{\hat{x}} = A\hat{x} + Bu + L(y - \hat{y}) = A\hat{x} + Bu + L(y - C\hat{x}), \ \hat{x}(0) = 0$$

= $(A - LC)\hat{x} + Ly + Bu$

• error dynamics: $e = x - \hat{x}$:

$$\dot{e} = Ae - LCe = (A - LC)e, \ e(0) = x(0)$$

if all eigenvalues of A – LC are on the left half plane, then the error dynamics can be made asymptotically stable

 UW Linear Systems (X. Chen, ME547)
 Observers and Observer State FB
 7 / 25

Luenberger (closed-loop) observer

Theorem

If (A, C) is an observable pair, then all the eigenvalues of A - LC can be arbitrarily assigned, provided that they are symmetric with respect to the real axis of the complex plane.

we show the SISO case when A and C are in observable canonical form (if not, a similarity transform can help out):

$$A = \begin{bmatrix} -\alpha_{n-1} & 1 & 0 & \dots \\ \vdots & 0 & \ddots & \ddots \\ -\alpha_1 & \vdots & \ddots & 1 \\ -\alpha_0 & 0 & \dots & 0 \end{bmatrix}, B = \begin{bmatrix} \beta_{n-1} \\ \vdots \\ \beta_1 \\ \beta_0 \end{bmatrix}$$
$$C = \begin{bmatrix} 1 & 0 & \dots & 0 \end{bmatrix}, D = d$$
$$\det(sl - A) = s^n + \alpha_n + s^{n-1} + \dots + \alpha_1 s + \alpha_0$$

Observer eigenvalue placement: o.c.f.

Luenberger observer with correction:

$$\dot{\hat{x}} = A\hat{x} + Bu + L(y - \hat{y}) = A\hat{x} + Bu + L(y - C\hat{x}), \ \hat{x}(0) = 0$$

= $(A - LC)\hat{x} + Ly + Bu$

► Goal: place eigenvalues of the observer at locations $\bar{p}_1, \dots, \bar{p}_n$:

$$\det (sI - (A - LC)) = (s - \overline{p}_1)(s - \overline{p}_2) \cdots (s - \overline{p}_n)$$
$$= s^n + \overline{\gamma}_{n-1}s^{n-1} + \cdots + \overline{\gamma}_1s + \overline{\gamma}_0$$

UW Linear Systems (X. Chen, ME547) Observers and Observer State FB

Observer eigenvalue placement: o.c.f.

► Goal: place eigenvalues of the observer at locations $\bar{p}_1, \dots, \bar{p}_n$:

$$\det (sI - (A - LC)) = (s - \overline{p}_1)(s - \overline{p}_2) \cdots (s - \overline{p}_n)$$

 $= s^n + \overline{\gamma}_{n-1}s^{n-1} + \cdots + \overline{\gamma}_1s + \overline{\gamma}_0$

Let $L = [I_0, I_1, \dots, I_{n-1}]^T$. The unique structures of A and C give

$$LC = \begin{bmatrix} l_0 \\ \vdots \\ l_{n-2} \\ l_{n-1} \end{bmatrix} \begin{bmatrix} 1 & 0 & \dots & 0 \end{bmatrix} = \begin{bmatrix} l_0 & 0 & \dots & 0 \\ \vdots & 0 & \ddots & \vdots \\ l_{n-2} & \ddots & \ddots & 0 \\ l_{n-1} & 0 & \dots & 0 \end{bmatrix}$$
$$A - LC = \begin{bmatrix} -\alpha_{n-1} - l_0 & 1 & 0 & \dots & 0 \\ -\alpha_{n-2} - l_1 & 0 & \ddots & \ddots & \vdots \\ \vdots & 0 & \ddots & \ddots & 0 \\ -\alpha_1 - l_{n-2} & \vdots & \ddots & 0 & 1 \\ -\alpha_0 - l_{n-1} & 0 & \dots & 0 & 0 \end{bmatrix}$$

Observer eigenvalue placement: o.c.f.

 \blacktriangleright A and A - LC have the same structure:

$$A = \begin{bmatrix} -\alpha_{n-1} & 1 & 0 & \dots \\ \vdots & 0 & \ddots & \ddots \\ -\alpha_{1} & \vdots & \ddots & 1 \\ -\alpha_{0} & 0 & \dots & 0 \end{bmatrix}, A - LC = \begin{bmatrix} -\alpha_{n-1} - l_{0} & 1 & 0 & \dots \\ \vdots & 0 & \ddots & \ddots \\ -\alpha_{1} - l_{n-2} & \vdots & \ddots & 1 \\ -\alpha_{0} - l_{n-1} & 0 & \dots & 0 \end{bmatrix}$$

Recall: det $(sl - A) = s^{n} + \alpha_{n-1}s^{n-1} + \dots + \alpha_{1}s + \alpha_{0}.$
Thus
det $(sl - (A - LC)) = s^{n} + (\alpha_{n-1} + l_{0}) \atop target: \overline{\gamma}_{n-1}} s^{n-1} + \dots + (\alpha_{0} + l_{n-1}) \atop target: \overline{\gamma}_{0}$
Hence
$$l_{0} = \overline{\gamma}_{n-1} - \alpha_{n-1}$$

$$\vdots$$

$$l_{n-1} = \overline{\gamma}_{0} - \alpha_{0}$$

UW Linear Systems (X. Chen, ME547) Observers and Observer State FB

11 / 25

General observer eigenvalue placement

- What if (A, B, C, D) is not in the observable canonical form?
- We can transform it to o.c.f. via a similarity transform:

• use previous formulas to design \tilde{L} in:

$$\dot{\hat{x}}_{ob} = \left(A_o - \tilde{L}C_o\right)\hat{x}_{ob} + \tilde{L}y + B_o u$$
 (analysis form)

correspondingly in the original state space (via $\hat{x}_{ob} = R\hat{x}$):

$$R\dot{\hat{x}} = \left(RAR^{-1} - \tilde{L}CR^{-1}\right)R\hat{x} + \tilde{L}y + RBu$$

$$\Rightarrow \dot{\hat{x}} = \left(A - \widetilde{R^{-1}\tilde{L}}C\right)\hat{x} + Ly + Bu \qquad \text{(implementation form)}$$

Powerful fact: if system Σ = (A, B, C, D) is observable, then we can arbitrarily place the observer eigenvalues.

Luenberger observer summary

- observer dynamics: $\dot{\hat{x}} = A\hat{x} + Bu + L(y C\hat{x}), \ \hat{x}(0) = 0$
- block diagram



UW Linear Systems (X. Chen, ME547) Observers and Observer State FB 13/25

Luenberger observer summary

system dynamics

$$\dot{x} = Ax + Bu, \ x(0) = x_0, \ A \in \mathbb{R}^{n \times n}, B \in \mathbb{R}^{n \times n}$$

 $y = Cx, \ y \in \mathbb{R}^{m \times 1}$

observer dynamics

$$\dot{\hat{x}} = A\hat{x} + Bu + L(y - C\hat{x}), \ \hat{x}(0) = 0$$
$$= (A - LC)\hat{x} + LCx + Bu$$

augmented system

$$\begin{bmatrix} \dot{x} \\ \dot{\hat{x}} \end{bmatrix} = \begin{bmatrix} A & 0 \\ LC & A - LC \end{bmatrix} \begin{bmatrix} x \\ \hat{x} \end{bmatrix} + \begin{bmatrix} B \\ B \end{bmatrix} u$$

Luenberger observer summary

augmented system

$$\begin{bmatrix} \dot{x} \\ \dot{\hat{x}} \end{bmatrix} = \begin{bmatrix} A & 0 \\ LC & A - LC \end{bmatrix} \begin{bmatrix} x \\ \hat{x} \end{bmatrix} + \begin{bmatrix} B \\ B \end{bmatrix} u$$
$$y = Cx$$

► to see the distribution of eigenvalues, note the error dynamics $\dot{e} = (A - LC)e \Rightarrow$

$$\begin{bmatrix} \dot{x} \\ \dot{e} \end{bmatrix} = \begin{bmatrix} A & 0 \\ 0 & A - LC \end{bmatrix} \begin{bmatrix} x \\ e \end{bmatrix} + \begin{bmatrix} B \\ 0 \end{bmatrix} u$$

 \Rightarrow eigenvalues are separated into: $\lambda(A)$ and observer eigenvalues

• underlying similarity transform: $\begin{bmatrix} x \\ e \end{bmatrix} = \begin{bmatrix} I_n & 0 \\ I_n & -I_n \end{bmatrix} \begin{bmatrix} x \\ \hat{x} \end{bmatrix}$

UW Linear Systems (X. Chen, ME547) Observers and Observer State FB

15 / 25

Discrete-time observers: Introduction

- full state feedback is usually not available
- often observers are implemented in the discrete-time domain
- the discrete-time observer design
 - basic form: analogous to the continuous-time Luenberger observer
 - predict and correct form:
 - direct DT design
 - leverages discrete-time signal properties

Discrete-time full state observer

standard discrete-time observer:

$$\begin{aligned} x\,(k+1) &= Ax\,(k) + Bu\,(k) \\ \hat{x}\,(k+1) &= A\hat{x}\,(k) + Bu\,(k) + L\,(y\,(k) - C\hat{x}\,(k)) \\ y\,(k) &= Cx\,(k) \end{aligned}$$

error dynamics: $e(k) = x(k) - \hat{x}(k)$, e(k+1) = Ae(k) - LCe(k)

overall dynamics

$$\begin{bmatrix} x (k+1) \\ e (k+1) \end{bmatrix} = \begin{bmatrix} A & 0 \\ 0 & A-LC \end{bmatrix} \begin{bmatrix} x (k) \\ e (k) \end{bmatrix} + \begin{bmatrix} B \\ 0 \end{bmatrix} u (k)$$
$$y (k+1) = \begin{bmatrix} C, & 0 \end{bmatrix} \begin{bmatrix} x (k+1) \\ e (k+1) \end{bmatrix}$$



UW Linear Systems (X. Chen, ME547) Observers and Observer State FB

17 / 25

DT full state observer with predictor

- ► motivation: $\hat{x}(k+1) = A\hat{x}(k) + Bu(k) + L(y(k) C\hat{x}(k))$ doesn't use most recent measurement y(k+1) = Cx(k+1)
- discrete-time observer with predictor:

predictor: $\hat{x}(k+1|k) = A\hat{x}(k|k) + Bu(k)$ corrector: $\hat{x}(k+1|k+1) = \hat{x}(k+1|k) + L(y(k+1) - C\hat{x}(k+1|k))$

- $\hat{x}(k|k)$: estimate of x(k) based on measurements up to time k
- $\hat{x}(k|k-1)$: estimate based on measurements up to time k-1
- $e(k) \triangleq x(k) \hat{x}(k|k)$: estimation error
- error dynamics

$$\begin{aligned} \hat{x}(k+1|k+1) &= (I - LC)\,\hat{x}\,(k+1|k) + Ly\,(k+1) \\ &= (I - LC)\,A\hat{x}\,(k|k) + (I - LC)\,Bu\,(k) + Ly\,(k+1) \\ \Rightarrow e\,(k+1) &= x(k+1) - Ly(k+1) - (I - LC)A\hat{x}(k|k) - (I - LC)Bu(k) \\ &= (A - LCA)\,e\,(k) \end{aligned}$$

DT full state observer with predictor

$$e(k+1) = \left(A - L \underbrace{CA}_{\widetilde{C}}\right) e(k), \ e(0) = (I - LC) x_0$$

• the error dynamics can be arbitrarily assigned if the pair $\begin{pmatrix} A, \ \tilde{C} \end{pmatrix} = \begin{pmatrix} A, \ CA \end{pmatrix}$ is observable

• observability matrix

$$\tilde{Q}_d = \begin{bmatrix} \tilde{C} \\ \tilde{C}A \\ \vdots \\ \tilde{C}A^{n-1} \end{bmatrix} = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} A$$

- ▶ if A is invertible, then \tilde{Q}_d has the same rank as Q_d
- (A, \tilde{C}) is observable if (A, C) is observable and A is nonsingular (guaranteed if discretized from a CT system)

UW Linear Systems (X. Chen, ME547) Observers and Observer State

19 / 25

Example

$$\begin{bmatrix} x_1 (k+1) \\ x_2 (k+1) \\ x_3 (k+1) \end{bmatrix} = \begin{bmatrix} -a_2 & 1 & 0 \\ -a_1 & 0 & 1 \\ -a_0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 (k) \\ x_2 (k) \\ x_3 (k) \end{bmatrix} + \begin{bmatrix} b_2 \\ b_1 \\ b_0 \end{bmatrix} u (k),$$

$$y (k) = x_1(k).$$
 Place all eigenvalues of an observer with predictor at the origin.

$$A - LCA = \begin{bmatrix} -a_2 & 1 & 0 \\ -a_1 & 0 & 1 \\ -a_0 & 0 & 0 \end{bmatrix} - \begin{bmatrix} l_1 \\ l_2 \\ l_3 \end{bmatrix} \begin{bmatrix} -a_2 & 1 & 0 \end{bmatrix}$$
$$= \begin{bmatrix} (l_1 - 1)a_2 & 1 - l_1 & 0 \\ l_2a_2 - a_1 & -l_2 & 1 \\ l_3a_2 - a_0 & -l_3 & 0 \end{bmatrix}$$

 $det (A - LCA - \lambda I) = ((I_1 - 1) a_2 - \lambda) (I_2 + \lambda) \lambda + (1 - I_1) (I_3 a_a - a_0) + I_3 ((I_1 - 1) a_2 - \lambda) + \lambda (1 - I_1) (I_2 a_2 - a_1)$ roots must be all $0 \Rightarrow I_1 = 1, I_2 = I_3 = 0.$

1. Concepts

2. Continuous-time Luenberger observer

- 3. Discrete-time observers
 - DT full state observer
 - DT full state observer with predictor

4. Observer state feedback

UW Linear Systems (X. Chen, ME547) Observers and Observer State FB

21 / 25

Observer state feedback

given system dynamics:

$$\dot{x} = Ax + Bu$$

 $y = Cx$

- state feedback control: arbitrary eigenvalue assignment if system controllable
- observer design: arbitrary observer eigenvalue assignment for state estimation if system observerable
- when full states are not available, what's the performance if we combine both?

$$u = -K\hat{x} + v$$

Closed-loop dynamics

full closed-loop system

$$\dot{x} = Ax + Bu$$

$$y = Cx$$

$$\dot{\hat{x}} = A\hat{x} + Bu + L(y - C\hat{x})$$

$$u = -K\hat{x} + v$$

$$\Rightarrow \frac{d}{dt} \begin{bmatrix} x \\ \hat{x} \end{bmatrix} = \begin{bmatrix} A & -BK \\ LC & A - LC - BK \end{bmatrix} \begin{bmatrix} x \\ \hat{x} \end{bmatrix} + \begin{bmatrix} B \\ B \end{bmatrix} v$$
• using again similarity transform
$$\begin{bmatrix} x \\ e \end{bmatrix} = \begin{bmatrix} I_n & 0 \\ I_n & -I_n \end{bmatrix} \begin{bmatrix} x \\ \hat{x} \end{bmatrix}$$
 gives
$$\frac{d}{dt} \begin{bmatrix} x \\ e \end{bmatrix} = \begin{bmatrix} A - BK & BK \\ 0 & A - LC \end{bmatrix} \begin{bmatrix} x \\ e \end{bmatrix} + \begin{bmatrix} B \\ 0 \end{bmatrix} v$$
21/25

Block diagram



The separation theorem

closed-loop dynamics

$$\frac{d}{dt} \begin{bmatrix} x \\ e \end{bmatrix} = \begin{bmatrix} A - BK & BK \\ 0 & A - LC \end{bmatrix} \begin{bmatrix} x \\ e \end{bmatrix} + \begin{bmatrix} B \\ 0 \end{bmatrix} v$$

- powerful result: separation theorem: closed-loop eigenvalues consist of
 - eigenvalues of A BK from the state feedback control design
 - eigenvalues of A LC from the observer design
- can design K and L separately based on discussed tools
- if system is controllable and observable, we can arbitrarily assign the closed-loop eigenvalues

UW Linear Systems (X. Chen, ME547) Observers and Observer State FB

ME 547: Linear Systems Linear Quadratic Optimal Control

Xu Chen

University of Washington

LQ

Motivation

state feedback control:

- allows to arbitrarily assign the closed-loop eigenvalues for a controllable system
- the eigenvalue assignment has been manual thus far
- performance is implicit: we assign eigenvalues to induce proper error convergence

linear quadratic (LQ) optimal regulation control, aka, LQ regulator (or LQR):

- no need to specify closed-loop poles
- performance is explicit: a performance index is defined ahead of time

1. Problem formulation

- 2. Solution to the finite-horizon LQ problem
- 3. From finite-horizon LQ to stationary LQ

UW Linear Systems (X. Chen, ME547)

Goal

Consider an *n*-dimensional state-space system

$$\dot{x}(t) = Ax(t) + Bu(t), \ x(t_0) = x_0 y(t) = Cx(t)$$
(1)

where $x \in \mathbb{R}^n$, $u \in \mathbb{R}^r$, and $y \in \mathbb{R}^m$.

LQ optimal control aims at minimizing the performance index

$$J = \frac{1}{2}x^{T}(t_{f})Sx(t_{f}) + \frac{1}{2}\int_{t_{0}}^{t_{f}} \left(x^{T}(t)Qx(t) + u^{T}(t)Ru(t)\right) dt$$

LQ

- S ≥ 0, Q ≥ 0, R > 0: for a nonnegative cost and well-posed problem
- $\frac{1}{2}x^T(t_f)Sx(t_f)$ penalizes the deviation of x from the origin at t_f
- ► $x^{T}(t)Qx(t)$ $t \in (t_0, t_f)$ penalizes the transient
- often, $Q = C^T C \Rightarrow x^T(t)Qx(t) = y(t)^T y(t)$
- \blacktriangleright $u^{T}(t)Ru(t)$ penalizes large control efforts

Observations

$$J = \frac{1}{2}x^{T}(t_{f})Sx(t_{f}) + \frac{1}{2}\int_{t_{0}}^{t_{f}} (x^{T}(t)Qx(t) + u^{T}(t)Ru(t)) dt$$

▶ when the control horizon is made to be infinitely long, i.e., $t_f \rightarrow \infty$, the problem reduces to the infinite-horizon LQ problem

$$J = \frac{1}{2} \int_{t_0}^{\infty} \left(x^{\mathsf{T}}(t) Q x(t) + u^{\mathsf{T}}(t) R u(t) \right) dt$$

- terminal cost is not needed, as it will turn out, that the control will have to drive x to the origin. Otherwise J will go unbounded.
- often, we have $t_0 = 0$ and

$$J = \frac{1}{2} \int_0^\infty \left(x^T(t) Q x(t) + u^T(t) R u(t) \right) dt$$

UW Linear Systems (X. Chen, ME547)

LQ

5 / 32

1. Problem formulation

2. Solution to the finite-horizon LQ problem

3. From finite-horizon LQ to stationary LQ

Solution to the finite-horizon LQ

Consider the performance index

$$J = \frac{1}{2}x^{T}(t_{f})Sx(t_{f}) + \frac{1}{2}\int_{t_{0}}^{t_{f}} (x^{T}(t)Qx(t) + u^{T}(t)Ru(t)) dt$$

with $\dot{x} = Ax + Bu$, $x(t_0) = x_0$, $S \succeq 0$, $R \succ 0$, and $Q = C^T C$.

• do a Lyapunov-like construction: $V(t) \triangleq \frac{1}{2}x^{T}(t)P(t)x(t)$

then

$$\frac{d}{dt}V(t) = \frac{1}{2}\dot{x}^{T}(t)P(t)x(t) + \frac{1}{2}x^{T}(t)\dot{P}(t)x(t) + \frac{1}{2}x^{T}(t)P(t)\dot{x}(t) = \frac{1}{2}(Ax + Bu)^{T}Px + \frac{1}{2}x^{T}\frac{dP}{dt}x + \frac{1}{2}x^{T}P(Ax + Bu) = \frac{1}{2}\left\{x^{T}(t)\left(A^{T}P + \frac{dP}{dt} + PA\right)x(t) + u^{T}B^{T}Px + x^{T}PBu\right\}$$

UW Linear Systems (X. Chen, ME547)

LQ

7/32

Solution to the finite-horizon LQ with $\frac{d}{dt}V(t)$ from the last slide, we have

$$V(t_f) - V(t_0) = \int_{t_0}^{t_f} \dot{V} dt$$
$$= \frac{1}{2} \int_{t_0}^{t_f} \left(x^T \left(A^T P + PA + \frac{dP}{dt} \right) x + u^T B^T P x + x^T P B u \right) dt$$

adding

$$J = \frac{1}{2}x^{T}(t_{f})Sx(t_{f}) + \frac{1}{2}\int_{t_{0}}^{t_{f}} (x^{T}(t)Qx(t) + u^{T}(t)Ru(t)) dt$$

yields

$$J + V(t_f) - V(t_0) = \frac{1}{2} x^T(t_f) S_X(t_f) + \frac{1}{2} \int_{t_0}^{t_f} \left(x^T \left(A^T P + PA + Q + \frac{dP}{dt} \right) x + \underbrace{u^T B^T P x + x^T P B u}_{\text{products of x and } u} + \underbrace{u^T R u}_{\text{quadratic}} \right) dt$$

Solution to the finite-horizon LQ

• "complete the squares" in $\underbrace{u^T B^T P x + x^T P B u}_{\text{products of } x \text{ and } u} + \underbrace{u^T R u}_{\text{quadratic}}$ (scalar case):

$$u^{T}B^{T}Px + x^{T}PBu + u^{T}Ru \stackrel{\text{scalar case}}{=} Ru^{2} + 2xPBu$$
$$= Ru^{2} + 2\left(xPBR^{-1/2}\right)\underbrace{R^{1/2}u}_{\sqrt{Ru^{2}}} + \left(R^{-1/2}BPx\right)^{2} - \left(R^{-1/2}BPx\right)^{2}$$
$$= \left(R^{1/2}u + R^{-1/2}BPx\right)^{2} - \left(R^{-1/2}BPx\right)^{2}$$

extending the concept to the general vector case:

$$u^{T}B^{T}Px + x^{T}PBu + u^{T}Ru = \underbrace{\|R^{\frac{1}{2}}u + R^{-\frac{1}{2}}B^{T}Px\|_{2}^{2}}_{\text{recall } \|\overrightarrow{a}\|_{2}^{2} = \overrightarrow{a}^{T}\overrightarrow{a}} - x^{T}PBR^{-1}B^{T}Px$$

UW Linear Systems (X. Chen, ME547)

LQ

9 / 32

Solution to the finite-horizon LQ

$$J + V(t_{f}) - V(t_{0}) = \frac{1}{2}x^{T}(t_{f})Sx(t_{f}) + \frac{1}{2}\int_{t_{0}}^{t_{f}} \left(x^{T}\left(A^{T}P + PA + Q + \frac{dP}{dt}\right)x + \underbrace{u^{T}B^{T}Px + x^{T}PBu + u^{T}Ru}_{\|R^{\frac{1}{2}}u + R^{\frac{-1}{2}}B^{T}Px\|_{2}^{2} - x^{T}PBR^{-1}B^{T}Px}\right)dt$$

 \Downarrow "completing the squares"

$$J + \frac{1}{2} x^{T}(t_{f}) P(t_{f}) x(t_{f}) - \frac{1}{2} x^{T}(t_{0}) P(t_{0}) x(t_{0}) = \frac{1}{2} x^{T}(t_{f}) Sx(t_{f}) + \frac{1}{2} \int_{t_{0}}^{t_{f}} \left(x^{T} \left(\frac{dP}{dt} + A^{T} P + PA + Q - PBR^{-1}B^{T} P \right) x + \|R^{\frac{1}{2}}u + R^{\frac{-1}{2}}B^{T} Px\|_{2}^{2} \right) dt$$

Solution to the finite-horizon LQ

$$J + V(t_f) - V(t_0) = \frac{1}{2}x^T(t_f)Sx(t_f) + \frac{1}{2}\int_{t_0}^{t_f} \left(x^T \left(A^T P + PA + Q + \frac{dP}{dt}\right)x + u^T B^T Px + x^T PBu + u^T Ru\right) dt$$

 \Downarrow "completing the squares"

$$J + \frac{1}{2} x^{T}(t_{f}) P(t_{f}) x(t_{f}) - \frac{1}{2} x^{T}(t_{0}) P(t_{0}) x(t_{0}) = \frac{1}{2} x^{T}(t_{f}) Sx(t_{f}) + \frac{1}{2} \int_{t_{0}}^{t_{f}} \left(x^{T} \underbrace{\left(\frac{dP}{dt} + A^{T}P + PA + Q - PBR^{-1}B^{T}P\right)}_{\underline{t_{0}}} x + \|\underline{R^{\frac{1}{2}}u + R^{\frac{-1}{2}}B^{T}Px}\|_{2}^{2} \right) dt$$

the best that the control can do in minimizing the cost is to have

$$u(t) = -K(t)x(t) = \underbrace{-R^{-1}B^{T}P(t)x(t)}_{\underline{\qquad}}$$
$$-\frac{dP}{dt} = \underbrace{A^{T}P + PA - PBR^{-1}B^{T}P + Q}_{\underline{\qquad}}, \ \underline{P(t_{f})} = S$$

to yield the optimal cost $J^0 = \frac{1}{2} x_0^T P(t_0) x_0$

11/32

Observation 1

UW Linear Systems (X. Chen, ME547)

$$u(t) = -K(t) \times (t) = -R^{-1}B^{T}P(t) \times (t)$$
optimal control law
$$-\frac{dP}{dt} = A^{T}P + PA - PBR^{-1}B^{T}P + Q, P(t_{f}) = S$$
the Riccati differential equation

- ► the control u(t) = -R⁻¹B^TP(t)x(t) is a state feedback law (the power of state feedback!)
- the state feedback law is time-varying because of P(t)
- the closed-loop dynamics becomes

$$\dot{x}(t) = Ax(t) + Bu(t) = (A - BR^{-1}B^{T}P(t)) x(t)$$

time-varying closed-loop dynamics

Observation 2

$$u(t) = -K(t) \times (t) = -R^{-1}B^{T}P(t) \times (t)$$
optimal state feedback control
$$-\frac{dP}{dt} = A^{T}P + PA - PBR^{-1}B^{T}P + Q, P(t_{f}) = S$$
the Riccati differential equation

▶ boundary condition of the Riccati equation is given at the final time t_f ⇒ the equation must be integrated backward in time
 ▶ backward integration of

$$-\frac{dP}{dt} = A^{T}P + PA + Q - PBR^{-1}B^{T}P, P(t_{f}) = S$$

is equivalent to the forward integration of

$$\frac{dP^*}{dt} = A^T P^* + P^* A + Q - P^* B R^{-1} B^T P^*, \ P^*(0) = S$$
(2)

by letting $P(t) = P^*(t_f - t)$

 Eq. (2) can be solved by numerical integration, e.g., ODE45 in Matlab

LQ

UW Linear Systems (X. Chen, ME547)

Observation 3

$$J = \frac{1}{2}x^{T}(t_{f})Sx(t_{f}) + \frac{1}{2}\int_{t_{0}}^{t_{f}} (x^{T}(t)Qx(t) + u^{T}(t)Ru(t)) dt$$
$$J^{0} = \frac{1}{2}x_{0}^{T}P(t_{0})x_{0}$$

- the minimum value J^0 is a function of the initial state $x(t_0)$
- ▶ J (and hence J^0) is nonnegative $\Rightarrow P(t_0)$ is at least positive semidefinite
- t₀ can be taken anywhere in (0, t_f) ⇒ P(t) is at least positive semidefinite for any t

Example: LQR of a pure inertia system

Consider

$$\dot{x} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u, \ J = \frac{1}{2} x^{T} (t_{f}) Sx (t_{f}) + \frac{1}{2} \int_{0}^{t_{f}} \left(x^{T} Qx + Ru^{2} \right) dt$$
where $S = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \ Q = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \ R > 0$

• we let $P(t) = P^{*} (t_{f} - t)$ and solve
$$\frac{dP^{*}}{dt} = A^{T} P^{*} + P^{*} A + Q - P^{*} BR^{-1} B^{T} P^{*}, \ P^{*} (0) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\Leftrightarrow \frac{dP^{*}}{dt} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} P^{*} + P^{*} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} - P^{*} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \frac{1}{R} \begin{bmatrix} 0 & 1 \end{bmatrix} P^{*}$$

• letting
$$P^{*} = \begin{bmatrix} P^{*}_{11} & P^{*}_{12} \\ P^{*}_{12} & P^{*}_{22} \end{bmatrix} \Rightarrow \begin{cases} \frac{d}{dt} p^{*}_{11} = 1 - \frac{1}{R} (p^{*}_{12})^{2} \\ \frac{d}{dt} p^{*}_{12} = p^{*}_{11} - \frac{1}{R} P^{*}_{12} P^{*}_{22} \\ \frac{d}{dt} p^{*}_{22} = 2p^{*}_{12} - \frac{1}{R} (p^{*}_{22})^{2} \end{cases} \Rightarrow P^{*}_{12} (0) = 0$$

Example: LQR of a pure inertia system: analysis



Figure: LQ example: $P^*(0) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $P(t) = P^*(t_f - t)$

- if the final time t_f is large, P^{*}(t) forward converges to a stationary value
- ▶ i.e., P(t) backward converges to a stationary value at P(0)

Example: LQR of a pure inertia system: analysis



- a larger R results in a longer transient
- i.e., a larger penalty on the control input yields a longer time to settle

UW Linear Systems (X. Chen, ME547)

Example: LQR of a pure inertia system: analysis



Figure: LQ with different boundary values in Riccati difference Eq.

For the same R, the initial value P(t_f) = S becomes irrelevant as t_f → ∞

- 1. Problem formulation
- 2. Solution to the finite-horizon LQ problem
- 3. From finite-horizon LQ to stationary LQ

 UW Linear Systems (X. Chen, ME547)
 LQ
 19/32

From LQ to stationary LQ



- in the example, we see that P in the Riccati differential Eq. converges to a stationary value given sufficient time
- ▶ when $t_f \rightarrow \infty$, LQ becomes the stationary LQ problem, under two additional conditions that we now discuss in details:
 - (A, B) is controllable/stabilizable
 - (A, C) is observable/detectable

Need for controllability/stabilizability

$$J = \frac{1}{2} x^{\mathsf{T}}(t_f) S x(t_f) + \frac{1}{2} \int_{t_0}^{t_f} \left(x^{\mathsf{T}}(t) Q x(t) + u^{\mathsf{T}}(t) R u(t) \right) dt$$

$$-\frac{dP}{dt} = A^T P + PA - PBR^{-1}B^T P + Q, \ P(t_f) = S$$
 the Riccati differential equation

$$J^{0} = \frac{1}{2} x_{0}^{T} P(t_{0}) x_{0}$$

if (A, B) is controllable or stabilizable, then P(t) is guaranteed to converge to a bounded and stationary value

- for uncontrollable or unstabilizable systems, there can be unstable uncontrollable modes that cause J to be unbounded
- ▶ then if $J^0 = \frac{1}{2}x_0^T P(0) x_0$ is unbounded, we will have $||P(0)|| = \infty$

UW Linear Systems (X. Chen, ME547)

Need for controllability/stabilizability

if (A, B) is controllable or stabilizable, then P(t) is guaranteed to converge to a bounded and stationary value

- e.g.: $\dot{x} = x + 0 \cdot u$, x(0) = 1, Q = 1 and R be any positive value
 - system is uncontrollable and the uncontrollable mode is unstable
 - x(t) will keep increasing to infinity
 - ► $\Rightarrow J = \frac{1}{2} \int_0^\infty (x^T Q x + u^T R u) dt$ unbounded regardless of u(t)
 - ▶ in this case, the Riccati equation is

$$-rac{dP}{dt}=P+P+1=2P+1\Leftrightarrowrac{dP^{*}}{dt}=2P^{*}+1$$

forward integration of P^* (backward integration of P), will drive $P^*(\infty)$ and P(0) to infinity

Need for observability/detectability

$$J = \frac{1}{2} \int_{t_0}^{\infty} \left(x^{\mathsf{T}}(t) Q x(t) + u^{\mathsf{T}}(t) R u(t) \right) dt$$

with $\dot{x} = Ax + Bu$, $x(t_0) = x_0$, $R \succ 0$, and $Q = C^T C$. if (A, C) is observable or detectable, the optimal state feedback control system will be asymptotically stable

- ▶ *intuition*: if the system is observable, y = Cx will relate to all states \Rightarrow regulating $x^T Qx = x^T C^T Cx$ will regulate all states
- formally: if (A, C) is observable (detectable), the solution of the Riccati equation will converge to a positive (semi)definite value P₊ (proof in course notes)

UW Linear Systems (X. Chen, ME547)

From LQ to stationary LQ

	LQ		stationary LQ
Cost	$J = \frac{1}{2} x^{T}(t_f) S x(t_f) + \frac{1}{2} \int_{t_0}^{t_f} \left(x^{T}(t) Q x(t) + u^{T}(t) R u(t) \right) dt$	⇒	$J = \frac{1}{2} \int_{t_0}^{\infty} \left(x^T Q x + u^T R u \right) dt$
Syst.	$\dot{x} = Ax + Bu$	\Rightarrow	$\dot{x} = Ax + Bu$ (A, B) controllable/stabilizable (A, C) observable/detectable
Key Eq.	Riccati Eq. (RE) $-\frac{dP}{dt} = A^T P + PA - PBR^{-1}B^T P$ $+Q, P(t_f) = S$	\Rightarrow	Algebraic RE (ARE) $A^T P + PA - PBR^{-1}B^T P + Q = 0$
Opt. control & cost	$u(t) = -R^{-1}B^{T}P(t)x(t)$ $J^{0} = \frac{1}{2}x_{0}^{T}P(t_{0})x_{0}$	\Rightarrow	$u(t) = -R^{-1}B^{T}P_{+}x(t)$ $J^{0} = \frac{1}{2}x_{0}^{T}P_{+}x_{0}$

More formally: Solution of the infinite-horizon LQ For

$$J = \frac{1}{2} \int_{t_0}^{\infty} \left(x(t)^T Q x(t) + u(t)^T R u(t) \right) dt, \ Q = C^T C$$

with $\dot{x}(t) = Ax(t) + Bu(t)$, $x(t_0) = x_0$ and $R \succ 0$:

- if (A, B) is controllable (stabilizable) and (A, C) is observable (detectable)
- then the optimal control input is given by

$$u(t) = -R^{-1}B^T P_+ x(t)$$

where P₊ (= P₊^T) is the positive (semi)definite solution of the algebraic Riccati equation (ARE)

 $A^T P + P A - P B R^{-1} B^T P + Q = 0$

and the closed-loop system is asymptotically stable, with

$$J_{\min} = J^{0} = \frac{1}{2} x(t_{0})^{T} P_{+} x(t_{0})$$

UW Linear Systems (X. Chen, ME547)

Observations

• the control $u(t) = -R^{-1}B^T Px(t)$ is a *constant* state feedback law

LQ

• under the optimal control, the closed loop is given by

$$\dot{x} = Ax - BR^{-1}B^T Px = (A - BR^{-1}B^T P) \times \text{and } J =$$

 $\frac{1}{2} \int_{t_0}^{\infty} (x^T Qx + u^T Ru) dt = \frac{1}{2} \int_{t_0}^{\infty} x^T (Q + PBR^{-1}B^T P) \times dt$

for the above closed-loop system, the Lyapunov Eq. is

$$A_{c}^{\prime}P + PA_{c} = -Q_{c}$$

$$\Leftrightarrow (A - BR^{-1}B^{T}P)^{T}P + P(A - BR^{-1}B^{T}P) = -Q - PBR^{-1}B^{T}P$$

$$\Leftrightarrow A^{T}P + PA - PBR^{-1}B^{T}P = -Q \text{ (the ARE!)}$$

 when the ARE solution P₊ is positive definite, ¹/₂x^TP₊x is a Lyapunov function for the closed-loop system
Observations

Lyapunov E	Eq. and the ARE:	
Cost	$\overline{J} = \frac{1}{2} \int_0^\infty x^T Q_c x dt$	$J = \frac{1}{2} \int_{t_0}^{\infty} \left(x^T Q x + u^T R u \right) dt$
Syst. dynamics	$\dot{x} = A_c x$	$\dot{x} = Ax + Bu$ (A, B) controllable/stabilizable
Key Eq. Optimal control Opt. cost	$egin{aligned} & A_c^T P + P A_c + Q_c = 0 \ & { m N/A} \ & \overline{J}^0 = rac{1}{2} x^T \left(0 ight) P_+ x \left(0 ight) \end{aligned}$	$(A, C) \text{ observable/detectable}$ $A^{T}P + PA - PBR^{-1}B^{T}P + Q = 0$ $u(t) = -R^{-1}B^{T}P_{+}x(t)$ $J^{0} = \frac{1}{2}x(t_{0})^{T}P_{+}x(t_{0})$

the guaranteed closed-loop stability is an attractive feature
more nice properties will show up later

UW Linear Systems (X. Chen, ME547)

LQ

Example: Stationary LQR of a pure inertia system

Consider

$$\dot{x} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u, \ J = \frac{1}{2} \int_0^\infty \left(x^T \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} x + Ru^2 \right) dt, \ R > 0$$

the ARE is

$$0 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} P + P \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} - P \begin{bmatrix} 0 \\ 1 \end{bmatrix} \frac{1}{R} \begin{bmatrix} 0 & 1 \end{bmatrix} P \Rightarrow P_{+} = \begin{bmatrix} \sqrt{2}R^{1/4} & R^{1/2} \\ R^{1/2} & \sqrt{2}R^{3/4} \end{bmatrix}$$

the closed-loop A matrix can be computed to be

$$A_{c} = A - BR^{-1}B^{T}P_{+} = \begin{bmatrix} 0 & 1 \\ -R^{-1/2} & -\sqrt{2}R^{-1/4} \end{bmatrix}$$

 \blacktriangleright \Rightarrow closed-loop eigenvalues:

$$\lambda_{1,2} = -\frac{1}{\sqrt{2}R^{1/4}} \pm \frac{1}{\sqrt{2}R^{1/4}}j$$

27 / 32

$$\dot{x} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u, \ J = \frac{1}{2} \int_0^\infty \left(x^T \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} x + Ru^2 \right) dt$$



Figure: Eigenvalue $\lambda_{1,2} = -\frac{1}{\sqrt{2}R^{1/4}} \pm \frac{1}{\sqrt{2}R^{1/4}}j$ evolution (root locus)

- R ↑ (more penalty on the control input) ⇒ λ_{1,2} move closer to the origin ⇒ slower state convergence to zero
- ▶ $R \downarrow$ (allow for large control efforts) $\Rightarrow \lambda_{1,2}$ move further to the left of the complex plane \Rightarrow faster speed of closed-loop dynamics

MATLAB commands

care: solves the ARE for a continuous-time system:

$$[P,\Lambda,K] = \operatorname{care}\left(A,B,C^{T}C,R\right)$$

where $K = R^{-1}B^T P$ and Λ is a diagonal matrix with the closed-loop eigenvalues, i.e., the eigenvalues of A - BK, in the diagonal entries.

Iqr and Iqry: provide the LQ regulator with

$$[K, P, \Lambda] = \operatorname{lqr} (A, B, C^T C, R)$$
$$[K, P, \Lambda] = \operatorname{lqry} (\operatorname{sys}, Q_y, R)$$

where sys is defined by $\dot{x} = Ax + Bu$, y = Cx + Du, and

$$J = \frac{1}{2} \int_0^\infty \left(y^T Q_y y + u^T R u \right) dt$$

Additional excellent properties of stationary LQ

we know stationary LQR yields guaranteed closed-loop stability for controllable (stabilizable) and observable (detectable) systems

It turns out that LQ regulators with full state feedback has excellent additional properties of:

- at least a 60 degree phase margin
- infinite gain margin
- stability is guaranteed up to a 50% reduction in the gain



LQ

31/32

Applications and practice

choosing R and Q:

- if there is not a good idea for the structure for Q and R, start with diagonal matrices;
- gain an idea of the magnitude of each state variable and input variable
- ▶ call them $x_{i,\max}$ (i = 1, ..., n) and $u_{i,\max}$ (i = 1, ..., r)
- ▶ make the diagonal elements of Q and R inversely proportional to $||x_{i,\max}||^2$ and $||u_{i,\max}||^2$, respectively.

University of Washington Lecture Notes Linear Algebra for Controls

Xu Chen Bryan T. McMinn Endowed Research Professorship Associate Professor Department of Mechanical Engineering University of Washington chx [AT] uw [DOT] edu

Contents

1	Basic concepts of matrices and vectors1.1Matrix addition and multiplication1.2Matrix transposition	1 2 3
2	Linear systems of equations	5
3	Vector space, linear independence, basis, and span	9
4	Matrix properties 4.1 Rank 4.2 Range and null spaces 4.3 Determinants	10 10 10 10
5	Matrix and linear equations	11
6	Eigenvector and eigenvalue6.1Matrix, mappings, and eigenvectors6.2Computation of eigenvalue and eigenvectors6.3Eigenbases and diagonalization	13 13 14 17
7	Similarity transformation	19
8	Matrix inversion 8.1 Block matrix decomposition and inversion 8.2 *LU and Cholesky decomposition 8.3 Determinant and matrix inverse identity 8.4 Matrix inversion lemma 8.5 Special inverse equalities	20 23 24 25 26 28
9	Spectral mapping theorem	29
10	Matrix exponentials	30
11	Inner product11.1 Inner product spaces11.2 Trace (standard matrix inner product)	31 31 32
12	Norms 12.1 Vector norm 12.2 Induced matrix norm 12.3 Frobenius norm and general matrix norms 12.4 Norm inequalities 12.5 Exercises	33 33 34 34 36 37

13 Symmetric, skew-symmetric, and orthogonal matrices	37
13.1 Definitions and basic properties	37
13.2 Symmetric eigenvalue decomposition (SED)	39
13.3 Symmetric positive-definite matrices	42
13.4 General positive-definite matrices	44
13.5 *Positive-definite functions and non-constant matrices	45
14 Singular value and singular value decomposition (SVD)	46
14.1 Motivation	46
14.2 SVD	46
14.3 Properties of singular values	49
14.4 Exercises	49

1 Basic concepts of matrices and vectors

A linear equation set

$$3x_1 + 4x_2 + 10x_3 = 6$$

$$x_1 + 4x_2 - 10x_3 = 5$$

$$4x_2 + 10x_3 = -1,$$
(1)

can be simply written as

$$\begin{bmatrix} 3 & 4 & 10 \\ 1 & 4 & -10 \\ 0 & 4 & 10 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 6 \\ 5 \\ -1 \end{bmatrix}.$$
 (2)

Equation (2) wrote x_1 , x_2 , and x_3 just once rather than two or three times in (1). There are only three unknowns in the above linear equation set. The notational simplicity and many algebraic convenience that will arise, however, are significant when we have thousands of unknowns...

Formally, we write an $m \times n$ matrix A as

$$A = [a_{jk}] = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & \dots & \dots & a_{2n} \\ \vdots & \dots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}.$$

Here,

- $m \times n$ (reads m by n) is the dimension/size of the matrix. It means that A has m rows and n columns.
- Each element a_{jk} is an entry of the matrix. For two matrices A and B to be equal, it must be that $a_{jk} = b_{jk}$ for any j and k.
- If m = n, A belongs to the class of square matrices. The entries a_{11} , a_{22} , ..., a_{nn} are then called the diagonal entries of A.
 - Upper triangular matrices : square matrices with nonzero entries only on and above the main diagonal.
 - Lower triangular matrices : nonzero entries only on and below the main diagonal.
 - Diagonal matrices : nonzero entries only on the main diagonal.
 - Identity matrice : diagonal and all diagonal entries are 1.
- Vectors: special matrices whose row or column number is one.

- A row vector: $a = [a_1, a_2, \dots, a_n]$; its dimension is $1 \times n$.

– A $m \times 1$ column vector:

$$b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

Example (Matrix and quadratic forms). We can use matrices to express general quadratic functions of vectors. For instance

$$f(x) = x^T A x + 2bx + c$$

is equivalent to

$$f(x) = \begin{bmatrix} x \\ 1 \end{bmatrix}^T \begin{bmatrix} A & b \\ b^T & c \end{bmatrix} \begin{bmatrix} x \\ 1 \end{bmatrix}.$$

1.1 Matrix addition and multiplication

The sum of two matrices A and B (of the same size) is

$$A + B = [a_{jk} + b_{jk}].$$

The **product** between a $m \times n$ matrix A and a scalar c is

$$cA = \left[ca_{jk} \right],$$

i.e. each entry of A is multiplied by c to generate the corresponding entry of cA.

The matrix product C = AB is meaningful only if the column number of A equals the row number of B. The computation is done as shown in the following example:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ \hline a_{21} & \hline a_{22} & \hline a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ \hline b_{21} & b_{22} \\ \hline b_{31} & b_{32} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} \\ \hline c_{21} & c_{22} \\ c_{31} & c_{32} \\ c_{41} & c_{42} \end{bmatrix},$$

where

$$\begin{aligned} c_{21} &= a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31} \\ &= [a_{21}, a_{22}, a_{23}] \begin{bmatrix} b_{11} \\ b_{21} \\ b_{31} \end{bmatrix} \\ &= \text{"second row of } A^{\text{"}} \times \text{"first column of } B^{\text{"}}. \end{aligned}$$

More generally:

$$c_{jk} = a_{j1}b_{1k} + a_{j2}b_{2k} + \dots + a_{jn}b_{nk}$$

= $[a_{j1}, a_{j2}, \dots, a_{jn}] \begin{bmatrix} b_{1k} \\ b_{2k} \\ \vdots \\ b_{nk} \end{bmatrix},$ (3)

namely, the jk entry of C is obtained by multiplying each entry in the jth row of A by the corresponding entry in the kth column of B and then adding these n products. This is called a multiplication of rows into columns.

Matrix multiplication is not commutative: It is a good habit to always check the matrix dimensions when doing matrix products:

$$\begin{array}{ccc} A & B & = & C \\ [m \times n] & [n \times p] & & [m \times p] \end{array}$$

This way it is clear that AB in general does not equal to BA, e.g.,

$$ABC = (AB) C = A (BC) \neq BCA.$$

Matrices as combination of vectors: The matrix-vector product

$$Ax = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 \\ a_{41}x_1 + a_{42}x_2 + a_{43}x_3 \end{bmatrix}$$

is nothing but the weighted sum of the columns of A:

$$Ax = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \\ a_{41} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ a_{32} \\ a_{42} \end{bmatrix} + x_3 \begin{bmatrix} a_{13} \\ a_{23} \\ a_{33} \\ a_{43} \end{bmatrix}$$

1.2 Matrix transposition

Definition 1 (Transpose). The transpose of an $m \times n$ matrix

$$A = [a_{jk}] = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & \dots & \dots & a_{2n} \\ \vdots & \dots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$$

is the $n \times m$ matrix A^T (reads "A transpose") defined as

$$A^{T} = [a_{kj}] = \begin{bmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & \dots & \dots & a_{m2} \\ \vdots & \dots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{bmatrix}.$$

Transposition has the following rules:

- $(A^T)^T = A$
- $(A+B)^T = A^T + B^T$
- $(cA)^T = cA^T$
- $(AB)^T = B^T A^T$

If $A = A^T$, then A is called symmetric. If $A = -A^T$ then A is called skew-symmetric.

2 Linear systems of equations

A linear system of m equations in n unknowns x_1, \ldots, x_n is a set of equations of the form

$$a_{11}x_1 + a_{12}x_2 + \dots a_{1n}x_n = b_1$$

$$a_{21}x_1 + a_{22}x_2 + \dots a_{2n}x_n = b_2$$

$$\dots$$

$$a_{m1}x_1 + a_{m2}x_2 + \dots a_{mn}x_n = b_m$$
(4)

Here,

- The equation set is linear: each variable x_i appears in the first power only.
- If all the b_j are zero, then the linear equation is called a homogeneous system. Otherwise, it is a nonhomogeneous system.
- Homogeneous systems always have at least the trivial solution $x_1 = x_2 = \cdots = x_n = 0$.

The m equations (4) can be written as a single vector equation

$$Ax = b$$
,

where

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \dots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}, \ x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ \vdots \\ x_n \end{bmatrix}, \ b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

Gauss¹ elimination is a systematic method to solve linear equations. Consider

Γ	1	-1	1	Г., Л	[0]	
	-1	1	-1	x_1	0	
İ	0	10	25	$ x_2 =$	90	•
	20	10	0	$\begin{bmatrix} x_3 \end{bmatrix}$	80	
-		~~			<u>تے ہ</u> ت	
		A			b	

The Gauss elimination process is as follows:

¹Johann Carl Friedrich Gauss, 1777-1855, German mathematician: contributed significantly to many fields, including number theory, algebra, statistics, analysis, differential geometry, geodesy, geophysics, electrostatics, astronomy, Matrix theory, and optics.

Gauss was an ardent perfectionist. He was never a prolific writer, refusing to publish work which he did not consider complete and above criticism. Mathematical historian Eric Temple Bell estimated that, had Gauss published all of his discoveries in a timely manner, he would have advanced mathematics by fifty years.

1. Obtain the augmented matrix of the system

$$\begin{bmatrix} A \mid b \end{bmatrix} = \begin{bmatrix} 1 & -1 & 1 & 0 \\ -1 & 1 & -1 & 0 \\ 0 & 10 & 25 & 90 \\ 20 & 10 & 0 & 80 \end{bmatrix}.$$

2. Perform elementary row operation on the augmented matrix, to obtain the Row Echelon Form. Adding the first row to the second row gives

pivot role : $\begin{bmatrix} 1 & -1 & 1 & 0 \\ \hline -1 & 1 & -1 & 0 \\ 0 & 10 & 25 & 90 \\ 20 & 10 & 0 & 80 \end{bmatrix} \xrightarrow{\text{row 2}}_{\text{add pivot role}}$	$\begin{bmatrix} 1 & - \\ 0 & 0 \\ 0 & 1 \\ 20 & 1 \end{bmatrix}$	$ \begin{array}{cccc} -1 & 1 \\ \hline 0 & 0 \\ .0 & 25 \\ .0 & 0 \\ \hline \end{array} $	$\begin{vmatrix} 0 \\ 0 \\ 90 \\ 80 \end{vmatrix}$	
add $-20 \times \text{pivot role}$	$\begin{array}{rrrr} 1 & -1 \\ 0 & 0 \\ 0 & 10 \\ 0 & 30 \end{array}$	$\begin{array}{c c} 1 \\ 0 \\ 25 \\ -20 \end{array}$	0 0 90 80	

What we have done is using the pivot row to eliminate x_1 in the other equations. At this stage, the linear equations look like

 $x_1 - x_2 + x_3 = 0 \tag{5}$

$$0 = 0 \tag{6}$$

$$10x_2 + 25x_3 = 90\tag{7}$$

$$30x_2 - 20x_3 = 80. \tag{8}$$

Re-arranging yields

$$x_1 - x_2 + x_3 = 0 \tag{9}$$

$$10x_2 + 25x_3 = 90 \tag{10}$$

$$30x_2 - 20x_3 = 80\tag{11}$$

$$0 = 0. \tag{12}$$

Moving on, we can get ride of x_2 in the third equation, by adding to it -3 times the second equation. Correspondingly in the augmented matrix, we have

$$\begin{bmatrix} 1 & -1 & 1 & | & 0 \\ 0 & 10 & 25 & | & 90 \\ 0 & 30 & -20 & | & 80 \\ 0 & 0 & 0 & | & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & -1 & 1 & | & 0 \\ 0 & 10 & 25 & | & 90 \\ 0 & 0 & -95 & | & -190 \\ 0 & 0 & 0 & | & 0 \end{bmatrix} \xrightarrow{\text{normalizing}} \underbrace{\begin{bmatrix} 1 & -1 & 1 & | & 0 \\ 0 & 1 & 5/2 & | & 9 \\ 0 & 0 & 1 & | & 38/19 \\ 0 & 0 & 0 & | & 0 \end{bmatrix}}_{\text{normalizing}},$$

the row echelon form

namely

$$x_3 = 38/19$$
$$x_2 + x_3 = 9$$
$$x_1 - x_2 + x_3 = 0.$$

The unknowns can now be readily obtained by back substitution: $x_3=38/19,\ x_2=9-x_3$, $x_1=x_2-x_3.$

Elementary Row Operations for Matrices What we have done can be summarized by the following elementary matrix row operations:

- Interchange of two rows
- Addition of a constant multiple of one row to another row
- Multiplication of a row by a nonzero constant c

Let the final row echelon form be denoted by

$$\left[\begin{array}{c|c} R & f \end{array}\right].$$

We have:

- 1. The two systems Ax = b and Rx = f are equivalent.
- 2. At the end of the Gauss elimination (before the back substitution), the row echelon form of the augmented matrix will be

$$\begin{bmatrix} r_{11} & r_{12} & \dots & \dots & r_{1n} & f_1 \\ r_{22} & \dots & \dots & r_{2n} & f_2 \\ & \ddots & \dots & \ddots & \vdots & \vdots \\ & & r_{rr} & \dots & r_{rn} & f_r \\ & & & & f_{r+1} \\ \vdots \\ & & & & & f_m \end{bmatrix},$$

where all unfilled entries are zero.

- 3. The number of nonzero rows, r, in the row-reduced coefficient matrix R is called the rank of R and also the rank of A.
- 4. Solution concepts:
 - (a) No solution / system is inconsistent: r is less than m and f_{r+1} , f_{r+2} , ..., f_m are not all zero.

- (b) Unique solution: if the system is consistent and r = n, there is exactly one solution, which can be found by back substitution.
- (c) Infinitely many solutions: if $f_{r+1} = f_{r+2} = \ldots = f_m = 0$. To obtain any of these solutions, choose values of x_{r+1}, \ldots, x_n arbitrarily. Then solve the r-th equation for x_r (in terms of those arbitrary values), then the (r-1)-st equation for x_{r-1} , and so on up the line.

3 Vector space, linear independence, basis, and span

Given a set of m vectors $a_1, a_2, ..., a_m$ with the same size,

$$k_1a_1 + k_2a_2 + \dots + k_ma_m$$

is called a linear combination of the vectors. If

$$a_1 = k_2 a_2 + k_3 a_3 + \dots + k_m a_m$$

then a_1 is said to be *linearly dependent* on a_2 , a_3 , ..., a_m . The set

$$\{a_1, a_2, \dots, a_m\}\tag{13}$$

is then a linearly dependent set. The same idea holds if a_2 or any vector in the set (13) is linearly dependent on others.

Generalizing, if

$$k_1 a_1 + k_2 a_2 + \dots + k_m a_m = 0$$

holds if and only if

$$k_1 = k_2 = \dots = k_m = 0,$$

then the vectors in (13) are linearly dependent. This is saying that at least one of the vectors can be expressed as a linear combination of the other vectors.

Why is linear independence important? If a set of vectors is linearly dependent, then we can get rid of one or perhaps more of the vectors until we get a linearly independent set. This set is then the smallest "truly essential" set with which we can work.

Consider a set of n linearly independent vectors, a_1 , a_2 , ..., a_n , each with n components. All the possible linear combinations of a_1 , a_2 , ..., a_n form the vector space \mathbb{R}^n . This is the *span* of the n vectors.

Definition 2 (Basis). A *basis* of V is a set B of vectors in V, such that any $v \in V$ can be uniquely expressed as a finite linear combination of vectors in B.

Example 3. In \mathbb{R}^2

$$v_1 = \begin{bmatrix} 1\\1 \end{bmatrix}, \ v_2 = \begin{bmatrix} 1\\0 \end{bmatrix}$$

is a linearly independent set and forms a basis.

$$v_1 = \begin{bmatrix} 1\\1 \end{bmatrix}, v_2 = \begin{bmatrix} 1\\0 \end{bmatrix}, v_3 = \begin{bmatrix} 3\\4 \end{bmatrix}$$

is not a linearly independent set.

4 Matrix properties

4.1 Rank

Definition 4 (Rank). The rank of a matrix A is the maximum number of linearly independent row or column vectors.

Theorem. Row or column operations do not change the rank of a matrix.

With the concept of linear dependence, many matrix-matrix operations can be understood from the view point of vector manipulations.

Example (Dyad). $A = uv^T$ is called a dyad, where u and v are vectors of proper dimensions. It is a rank 1 matrix, as can be seen that $A = uv^T$ is formed by linear combinations of the vector u, where the weights of the combinations are coefficients of v.

Fact. For $A, B \in \mathbb{R}^{n \times n}$, if rank(A) = n then AB = 0 implies B = 0. If AB = 0 but $A \neq 0$ and $B \neq 0$, then rank(A) < n and rank(B) < n.

4.2 Range and null spaces

Definition 5 (Range space). The range space of a matrix A, denoted as $\mathcal{R}(A)$, is the span of all the column vectors of A.

Definition 6 (Null space). The null space of a matrix $A \in \mathbb{R}^{n \times n}$, denoted as $\mathcal{N}(A)$, is the vector space

$$\{x \in \mathbb{R}^n : Ax = 0\}$$

The dimension of the null space is called *nullity* of the matrix.

Fact 7. The following is true:

$$\mathcal{N}(AA^{T}) = \mathcal{N}(A^{T}); \ \mathcal{R}(AA^{T}) = \mathcal{R}(A).$$

4.3 Determinants

Determinants were originally introduced for solving linear equations in the form of Ax = y, with a square A. They are cumbersome to compute for high-order matrices, but their definitions and concepts are partially very important.

We review only the computations of second- and third-order matrices:

• 2×2 matrices:

$$\det \left[\begin{array}{cc} a & b \\ c & d \end{array} \right] = ad - bc.$$

• 3×3 matrices:

$$\det \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & k \end{bmatrix} = a \det \begin{bmatrix} e & f \\ h & k \end{bmatrix} - b \det \begin{bmatrix} d & f \\ g & k \end{bmatrix} + c \det \begin{bmatrix} d & e \\ g & h \end{bmatrix}$$
$$= aek + bfg + cdh - gec - bdk - ahf,$$

where det
$$\begin{bmatrix} e & f \\ h & k \end{bmatrix}$$
, det $\begin{bmatrix} d & f \\ g & k \end{bmatrix}$, and det $\begin{bmatrix} d & e \\ g & h \end{bmatrix}$ are called the minors of det $\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & k \end{bmatrix}$

Caution: det $(cA) = c^n \det (A)$ (not $c \det (A)$!)

Theorem 8. The determinant of A is nonzero if and only if A is full rank.

You should be able to verify the theorem for 2×2 matrices. The proof will be immediate after introducing the concept of eigenvalues.

Definition 9. A linear transformation is called singular if the determinant of the corresponding transformation matrix is zero.

Fact 10. Determinant facts:

• If A and B are square matrices, then

$$det (AB) = det (BA) = det A det B$$
$$det (A^{T}) = det (A)$$
$$det (A^{*}) = det (A).$$

• If X and Z are square, Y with compatible dimensions, then

 $\det\left(\left[\begin{array}{cc} X & Y \\ 0 & Z \end{array}\right]\right) = \det X \det Z.$

5 Matrix and linear equations

Consider again, using now concepts in range and null spaces of matrices, the linear equations

$$Ax = y. \tag{14}$$

- *Existence* of solutions requires that $y \in \mathcal{R}(A)$.
- The linear equation is called *overdetermined* if it has more equations than unknowns (i.e. A is a tall skinny matrix), *determined* if A is square, *undetermined* if it has fewer equations than unknowns (A is a wide matrix).

• Solutions of the above equation, provided that they exist, is constructed from

$$x = x_o + z$$
: $Az = 0,$ (15)

where x_0 is any (fixed) solution of (14) and z runs through all the homogeneous solutions of Az = 0, namely, z runs through all vectors in the null space of A.

• Uniqueness of a solution: if the null space of A is zero, the solution is unique.

You should be familiar with solving 2nd or 3rd-order linear equations by hand.

6 Eigenvector and eigenvalue

6.1 Matrix, mappings, and eigenvectors

Think of Ax this way: A defines a linear operator; Ax is a vector produced by feeding the vector x to this linear operator. In the two-dimensional case, we can look at Fig. 1. Certainly, Ax does not (at all) need to be in the same direction as x. An example is

$$A_0 = \left[\begin{array}{cc} 1 & 0\\ 0 & 0 \end{array} \right],$$

which gives that

$$A_0 \left[\begin{array}{c} x_1 \\ x_2 \end{array} \right] = \left[\begin{array}{c} x_1 \\ 0 \end{array} \right],$$

namely, Ax is x projected on the first axis in the two-dimensional vector space, which will not be in the same direction as x as long as $x_2 \neq 0$.



Figure 1: Example relationship between x and Ax.

From here comes the concept of eigenvectors and eigenvalues. It says that there are certain "special directions/vectors" (denoted as v_1 and v_2 in our two-dimensional example) for A such that $Av_i = \lambda_i v_i$. Thus Av_i is on the same line as the original vector v_i , just scaled by the eigenvalue λ_i . It can be shown that if $\lambda_1 \neq \lambda_2$, then v_1 and v_2 are linearly independent (your homework). This is saying that any vector in \mathbb{R}^2 can be decomposed as

$$x = a_1 v_1 + a_2 v_2.$$

Therefore

$$Ax = a_1 A v_1 + a_2 A v_2 = a_1 \lambda_1 v_1 + a_2 \lambda_2 v_2$$

Knowing λ_i and v_i thus can directly tell us how Ax looks like. More important, we have decomposed Ax into small modules that are from time to time more handy for analyzing the system properties. Figs. 2 and 3 demonstrate the above idea graphically.

Remark 11. The above geometric interpretations are for matrices with distinct real eigenvalues.







Figure 3: Construction of Ax.

The geometric interpretation above makes eigenvalue a very important concept. *Eigenvalues* are also called *characteristic values* of a matrix. The set of all the eigenvalues of A is called the *spectrum* of A. The largest of the *absolute* values of the eigenvalues of A is called the *spectral radius* of A.

6.2 Computation of eigenvalue and eigenvectors

Formally, eigenvalue and eigenvector are defined as follows. For $A \in \mathbb{R}^{n \times n}$, an eigenvalue λ of A is one for which

$$Ax = \lambda x \tag{16}$$

has a nonzero solution $x \neq 0$. The corresponding solutions are called eigenvectors of A.

Equation (16) is equivalent to

$$(A - \lambda I) x = 0. \tag{17}$$

As $x \neq 0$, the matrix $A - \lambda I$ must be singular, so

$$\det\left(A - \lambda I\right) = 0. \tag{18}$$

det $(A - \lambda I)$ is a polynomial of λ , called the characteristic polynomial. Correspondingly, (18) is called the characteristic equation. So eigenvalues are roots of the characteristic equation. If an $n \times n$ matrix A has n eigenvalues $\lambda_1, \ldots, \lambda_n$, it must be that

$$\det (A - \lambda I) = (\lambda_1 - \lambda) \cdots (\lambda_n - \lambda).$$

After obtaining an eigenvalue λ , we can find the associated eigenvector by solving (17). This is nothing but solving a homogeneous system.

Example 12. Consider

$$A = \left[\begin{array}{cc} -5 & 2\\ 2 & -2 \end{array} \right].$$

Then

$$\det (A - \lambda I) = 0 \Rightarrow \det \left(\begin{bmatrix} -5 - \lambda & 2\\ 2 & -2 - \lambda \end{bmatrix} \right) = 0$$
$$\Rightarrow (5 + \lambda) (2 + \lambda) - 4 = 0$$
$$\Rightarrow \lambda = -1 \text{ or } -6.$$

So A has two eigenvalues: -1 and -6. The characteristic polynomial of A is $\lambda^2 + 7\lambda + 6$.

To obtain the eigenvector associated to $\lambda = -1$, we solve

$$(A - \lambda I) x = 0 \Leftrightarrow \left(\begin{bmatrix} -5 & 2\\ 2 & -2 \end{bmatrix} + 1 \begin{bmatrix} 1 & 0\\ 0 & 1 \end{bmatrix} \right) x = \begin{bmatrix} -4 & 2\\ 2 & -1 \end{bmatrix} x = 0.$$

One solution is

$$x = \left[\begin{array}{c} 1\\2 \end{array} \right].$$

As an exercise, show that an eigenvector associated to $\lambda = -6$ is $\begin{bmatrix} 2 & -1 \end{bmatrix}^T$.

Example 13 (Multiple eigenvectors). Obtain the eigenvalues and eigenvectors of

$$A = \begin{bmatrix} -2 & 2 & -3\\ 2 & 1 & -6\\ -1 & -2 & 0 \end{bmatrix}$$

Analogous procedures give that

$$\lambda_1 = 5, \ \lambda_2 = \lambda_3 = -3.$$

So there are repeated eigenvalues. For $\lambda_2 = \lambda_3 = -3$, the characteristic matrix is

$$A + 3I = \begin{bmatrix} 1 & 2 & -3 \\ 2 & 4 & -6 \\ -1 & -2 & 3 \end{bmatrix}.$$

The second row is the first row multiplied by 2. The third row is the negative of the first row. So the characteristic matrix has only rank 1. The characteristic equation

$$(A - \lambda_2 I) x = 0$$

has two linearly independent solutions

$$\left[\begin{array}{c} -2\\1\\0\end{array}\right], \left[\begin{array}{c} 3\\0\\1\end{array}\right].$$

Theorem 14 (Eigenvalue and determinant). Let $A \in \mathbb{R}^{n \times n}$. Then

$$\det A = \prod_{i=1}^{n} \lambda_i.$$

Proof. Letting $\lambda = 0$ in the characteristic polynomial

$$p(\lambda) = \det (A - \lambda I) = (\lambda_1 - \lambda) (\lambda_2 - \lambda) \dots$$

gives

$$\det (A) = p(0) = \prod_{i=1}^{n} \lambda_i.$$

Example 15. For the two-dimensional	case
--	------

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \Rightarrow p(\lambda) = \det(A - \lambda I) = (a_{11} - \lambda)(a_{22} - \lambda) - a_{12}a_{21}.$$

On the other hand

$$p(\lambda) = (\lambda_1 - \lambda) (\lambda_2 - \lambda).$$

Matching the coefficients we get

$$\lambda_1 + \lambda_2 = a_{11} + a_{22}$$
$$\lambda_1 \lambda_2 = a_{11} a_{22} - a_{12} a_{21}.$$

6.3 Eigenbases and diagonalization

Eigenvectors of an $n \times n$ matrix A may (or may not!) form a basis for \mathbb{R}^n . If we are interested in a transformation y = Ax, such an "eigenbasis" (basis of eigenvectors), if exists, is of great advantage because then we can represent any x in \mathbb{R}^n uniquely as a linear combination of the eigenvectors x_1, \ldots , x_n , say, $x = c_1x_1 + c_2x_2 + \ldots + c_nx_n$. And, denoting the corresponding (not necessarily distinct) eigenvalues of the matrix A by $\lambda_1, \ldots, \lambda_n$, we have $Ax_i = \lambda_i x_i$, so that we simply obtain

$$y = Ax = A (c_1x_1 + c_2x_2 + \dots + c_nx_n)$$
$$= c_1Ax_1 + c_2Ax_2 + \dots + c_nAx_n$$
$$= c_1\lambda_1x_1 + \dots + c_n\lambda_nx_n.$$

This shows that we have decomposed the complicated action of A on an arbitrary vector x into a sum of simple actions (multiplication by scalars) on the eigenvectors of A.

Theorem 16 (Basis of Eigenvectors). If an $n \times n$ matrix A has n distinct eigenvalues, then A has a basis of eigenvectors x_1, \ldots, x_n for \mathbb{R}^n .

Proof. We just need to prove that the n eigenvectors are linearly independent. If not, reorder the eigenvectors and suppose r of them, $\{x_1, x_2, \ldots, x_r\}$, are linearly independent and x_{r+1}, \ldots, x_n are linearly dependent on $\{x_1, x_2, \ldots, x_r\}$. Consider x_{r+1} . There must exist c_1, \ldots, c_{n+1} , not all zero, such that

$$c_1 x_1 + \dots c_{r+1} x_{r+1} = 0. \tag{19}$$

Multiplying A on both sides yields

 $c_1 A x_1 + \dots + c_{r+1} A x_{r+1} = 0.$

Using $Ax_i = \lambda_i x_i$, we have

 $c_1\lambda_1x_1 + \dots + c_{r+1}\lambda_{r+1}x_{r+1} = 0.$

But from (19), we know that

 $c_1\lambda_{r+1}x_1+\ldots c_{r+1}\lambda_{r+1}x_{r+1}=0.$

Subtracting the last two equations gives

$$c_1 \left(\lambda_1 - \lambda_{r+1}\right) x_1 + \dots + c_r \left(\lambda_r - \lambda_{r+1}\right) x_r = 0.$$

None of $\lambda_1 - \lambda_{r+1}, \ldots, \lambda_r - \lambda_{r+1}$ are zero, as the eigenvalues are distinct. Hence not all coefficients $c_1 (\lambda_1 - \lambda_{r+1}), \ldots, c_r (\lambda_r - \lambda_{r+1})$ are zero. Thus $\{x_1, x_2, \ldots, x_r\}$ is not linearly independent–a contradiction with the assumption at the beginning of the proof.

Theorem 16 provides an important decomposition-called diagonalization-of matrices. To show that, we briefly review the concept of matrix inverses first.

Definition 17 (Matrix Inverse). The inverse A^{-1} of a square matrix A satisfies

$$AA^{-1} = A^{-1}A = I.$$

If A^{-1} exists, A is called nonsingular; otherwise, A is singular.

Theorem 18 (Diagonalization of a Matrix). Let an $n \times n$ matrix A have a basis of eigenvectors $\{x_1, x_2, \ldots, x_n\}$, associated to its n distinct eigenvectors $\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$, respectively. Then

$$A = XDX^{-1} = [x_1, x_2, \dots, x_n] \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{bmatrix} [x_1, x_2, \dots, x_n]^{-1}.$$
 (20)

Also,

$$A^m = X D^m X^{-1}, \quad (m = 2, 3, \dots).$$
 (21)

Remark 19. From (21), you can find some intuition about the benefit of (20): A^m can be tedious to compute while D^m is very simple!

Proof. From Theorem 16, the n linearly independent eigenvectors of A form a basis. Write

$$Ax_1 = \lambda_1 x_1$$
$$Ax_2 = \lambda_2 x_2$$
$$\vdots$$
$$Ax_n = \lambda_n x_n$$

as

$$A[x_1, x_2, \dots, x_n] = [x_1, x_2, \dots, x_n] \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{bmatrix}.$$

The matrix $[x_1, x_2, \ldots, x_n]$ is square. Linear independence of the eigenvectors implies that $[x_1, x_2, \ldots, x_n]$ is invertible. Multiplying $[x_1, x_2, \ldots, x_n]^{-1}$ on both sides gives (20).

(21) then immediately follows, as

$$A^{m} = (XDX^{-1})^{m} = XDX^{-1}XDX \dots XDX^{-1} = XD^{m}X^{-1}.$$

Example 20. Let

$$A = \left[\begin{array}{cc} 2 & -3 \\ 1 & -2 \end{array} \right].$$

The matrix has eigenvalues at 1 and -1, with associated eigenvectors

$$\left[\begin{array}{c}3\\1\end{array}\right], \left[\begin{array}{c}1\\1\end{array}\right].$$

Then

$$X = \begin{bmatrix} 3 & 1 \\ 1 & 1 \end{bmatrix}, \ A = X \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} X^{-1}.$$

Now if we are to compute A^{3000} . We just need to do

$$A^{3000} = X \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}^{3000} X^{-1} = I.$$

7 Similarity transformation

Definition 21 (Similar Matrices. Similarity Transformation). An $n \times n$ matrix \hat{A} is called similar to an $n \times n$ matrix A if

$$\hat{A} = T^{-1}AT$$

for some **nonsingular** $n \times n$ matrix T. This transformation, which gives \hat{A} from A, is called a similarity transformation.

Let S_1 and S_2 be two vector spaces of the same dimension. Take the same point P. Let u be its coordinate in S_1 and \hat{u} be its coordinate in S_2 . These coordinates in the two vector spaces are related by some linear transformation T:

$$u = T\hat{u}, \ \hat{u} = T^{-1}u$$

Consider Fig. 4. Let the point P go through a linear transformation A in the vector space S_1 to generate an output point P_o . P_o is physically the same point in both S_1 and S_2 . However, the coordinates of P_o are different: if we see it from "standing inside" S_1 , then

y = Au

If we see it in S_2 , then the coordinate is some other value \hat{y} .



Figure 4: Same points in different vector spaces

How does the linear transformation A mathematically "look like" in S_2 ? Result:

$$\hat{y} = T^{-1}y = T^{-1}Au = (T^{-1}AT)\hat{u}$$

namely, the linear transformation, viewed from \mathcal{S}_2 , is

$$\hat{A} = T^{-1}AT$$

It is central to recognize that the physical operation is the same: P goes to another point P_o . Different is our perspective of viewing this transformation. \hat{A} and A are in this sense called similar.

Purpose of doing similarity transformation: \hat{A} can be simpler! Consider, for instance, the following example



In S_1 , the transformation changes both coordinates of P while in S_2 , only the first coordinate of P is changed.

Theorem 22 (Eigenvalues and Eigenvectors of Similar Matrices). If \hat{A} is similar to A, then \hat{A} has the same eigenvalues as A. Furthermore, if x is an eigenvector of A, then $y = T^{-1}x$ is an eigenvector of \hat{A} corresponding to the same eigenvalue.

8 Matrix inversion

This section provides a more detailed description of matrix inversion. Recall that the inverse A^{-1} of a square nonsingular matrix A satisfies

$$AA^{-1} = A^{-1}A = I.$$

Theorem 23 (Inverse is unique). If A has an inverse, the inverse is unique.

Concepts only. If both B and C are inverses of A, then BA = AB = I and CA = AC = I so that

$$B = IB = (CA)B = CAB = C(AB) = CI = C.$$

Connection with previous topics: The set of all $n \times n$ matrices is not a field. Multiplicative inverse is unique.

Definition 24 (Existence of a matrix inverse). The inverse A^{-1} of an $n \times n$ matrix A exists if and only if the rank of A is n. Hence A is nonsingular if rank(A) = n, and singular if rank(A) < n.

Proof. Let $A \in \mathbb{R}^{n \times n}$ and consider the linear equation

Ax = b.

If A^{-1} exists, then

$$A^{-1}Ax = x = A^{-1}b.$$

Hence $A^{-1}b$ is a solution to the linear equation. It is also unique. If not, then take another solution u; we should have Au = b and $u = A^{-1}b$. Since A^{-1} is unique, it must be that u = x.

Conversely, if A has rank n. Then we can solve Ax = b uniquely by Gauss elimination, to get

x = Bb,

where B is the backward substitution linear transformation in Gauss elimination. Hence

$$Ax = A(Bb) = (AB)b = Ib$$

for any b. Hence

AB = I.

Similarly, substituting Ax = b into x = Bb gives

$$x = B(Ax) = (BA)x = Ix,$$

BA = I.

and hence

Together $B = A^{-1}$ exists.

There are several ways to compute the inverse of a matrix. One approach for low-order matrices is the method of using adjugate matrix (sometimes also called adjoint matrix):

$$A^{-1} = \frac{1}{\det\left(A\right)} \operatorname{\mathsf{adj}}\left(A\right)^{T}.$$

We explain the computation by two examples. You can find additional details in your undergraduate linear algebra course.

• 2×2 example:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} (-1)^{1+1} d & (-1)^{1+2} b \\ (-1)^{2+1} c & (-1)^{2+2} a \end{bmatrix},$$

where b in $(-1)^{1+2} b$ is obtained by:

- noticing b is at row 1 column 2 of A;
- looking at the element at row 2 column 1 of A (notice the transpose in $\operatorname{adj}(A)^{T}$);

- constructing a submatrix of A by removing row 2 and column 1 from it, i.e., [b] in this 2×2 example;
- computing the determinant of this submatrix.
- adding $(-1)^{1+2}$ as a scalar
- 3×3 example:

$$A^{-1} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & k \end{bmatrix}^{-1} = \frac{1}{\det A} \begin{bmatrix} \begin{vmatrix} e & f \\ h & k \end{vmatrix} - \begin{vmatrix} b & c \\ h & k \end{vmatrix} \begin{vmatrix} b & c \\ e & f \end{vmatrix} \\ - \begin{vmatrix} d & f \\ g & k \end{vmatrix} + \begin{vmatrix} a & c \\ g & k \end{vmatrix} - \begin{vmatrix} a & c \\ d & f \end{vmatrix} \\ \begin{vmatrix} d & e \\ g & h \end{vmatrix} - \begin{vmatrix} a & b \\ g & h \end{vmatrix} + \begin{vmatrix} a & b \\ d & e \end{vmatrix}$$

where $|\cdot|$ denotes the determinant of a matrix. Similar as before, the row 1 column 2 element $-\begin{vmatrix} b & c \\ h & k \end{vmatrix}$ is obtained via

$$(-1)^{2+1} \det \left(A \text{ with } [d, e, f], \begin{bmatrix} a \\ d \\ g \end{bmatrix} \text{ removed} \right).$$

Example 25. Find the inverse matrices of

$$A = \begin{bmatrix} 3 & 1 \\ 2 & 4 \end{bmatrix}, B = \begin{bmatrix} -1 & 1 & 2 \\ 3 & -1 & 1 \\ -1 & 3 & 4 \end{bmatrix}, C = \begin{bmatrix} -0.5 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

The answers are:

$$A^{-1} = \begin{bmatrix} 0.4 & -0.1 \\ -0.2 & 0.3 \end{bmatrix}, B^{-1} = \begin{bmatrix} -0.7 & 0.2 & 0.3 \\ -1.3 & -0.2 & 0.7 \\ -1 & 3 & 4 \end{bmatrix}, C^{-1} = \begin{bmatrix} -2 & 0 & 0 \\ 0 & 0.25 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

The related MATLAB command for matrix inversion is inv().

Theorem 26. Inverse of products of matrices can be obtained from inverses of each factor:

$$(AB)^{-1} = B^{-1}A^{-1},$$

and more generally

$$(AB \dots YZ)^{-1} = Z^{-1}Y^{-1} \dots B^{-1}A^{-1}.$$
(22)

Proof. By definition $(AB)(AB)^{-1} = I$. Multiplying A^{-1} on both sides from the left gives

$$B\left(AB\right)^{-1} = A^{-1}$$

Now multiplying the result by B^{-1} on both sides from the left, we get

$$(AB)^{-1} = B^{-1}A^{-1}.$$

The general case (22) follows by induction.

Fact 27. *Inverse of upper (lower) triangular matrices are upper (lower) triangular

Proof. (main idea) We can either use the adjoint matrix method or use the following decomposition of upper(lower) triangular matrices

$$A = D\left(I + N\right)$$

where D is diagonal and N is strictly upper (lower) triangular with zeros diagonal elements. Then using matrix Taylor expansion we have

$$A^{-1} = (I+N)^{-1} D^{-1}$$

= $(I-N+N^2-N^3+N^4-\dots) D^{-1}.$

N is nilpotent: N^k are upper (lower) triangular and $N^n = 0$ for n larger than the row dimension of A. D^{-1} is diagonal. Hence A^{-1} is upper (lower) triangular.

8.1 Block matrix decomposition and inversion

Consider

$$A = \left[\begin{array}{cc} 3 & 4 \\ 1 & 2 \end{array} \right].$$

Recall the key step in performing row operations on matrices in Gauss elimination:

$$\left[\begin{array}{cc} 3 & 4 \\ 1 & 2 \end{array}\right] \rightarrow \left[\begin{array}{cc} 3 & 4 \\ 0 & 2/3 \end{array}\right],$$

where we had substracted one third of the first row in the second row. In matrix representations, the above looks like

$$\begin{bmatrix} 1 & 0 \\ -1/3 & 1 \end{bmatrix} \begin{bmatrix} 3 & 4 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 4 \\ 0 & 2/3 \end{bmatrix}.$$

For more general two by two matrices, we have

$$\begin{bmatrix} 1 & 0 \\ -ca^{-1} & 1 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & b \\ 0 & d - ca^{-1}b \end{bmatrix}.$$

If we want to keep the second row unchanged and simplify the first row, we can do

$$\begin{bmatrix} 1 & -bd^{-1} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a - bd^{-1}c & 0 \\ c & d \end{bmatrix}.$$

Generalizing the concept to blok matrices (with compatible dimensions), we have

$$\begin{bmatrix} I & 0 \\ -B^{T}A^{-1} & I \end{bmatrix} \begin{bmatrix} A & B \\ B^{T} & C \end{bmatrix} = \begin{bmatrix} A & B \\ 0 & C - B^{T}AB \end{bmatrix},$$
$$\begin{bmatrix} A & B \\ 0 & C - B^{T}AB \end{bmatrix} \begin{bmatrix} I & -A^{-1}B \\ 0 & I \end{bmatrix} = \begin{bmatrix} A & 0 \\ 0 & C - B^{T}AB \end{bmatrix}.$$

Thus

and

$$\begin{bmatrix} I & 0 \\ -B^T A^{-1} & I \end{bmatrix} \begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \begin{bmatrix} I & -A^{-1}B \\ 0 & I \end{bmatrix} = \begin{bmatrix} A & 0 \\ 0 & C - B^T AB \end{bmatrix}.$$

Inversion is now very easy:

$$\left\{ \begin{bmatrix} I & 0 \\ -B^{T}A^{-1} & I \end{bmatrix} \begin{bmatrix} A & B \\ B^{T} & C \end{bmatrix} \begin{bmatrix} I & -A^{-1}B \\ 0 & I \end{bmatrix} \right\}^{-1} = \begin{bmatrix} A & 0 \\ 0 & C - B^{T}AB \end{bmatrix}^{-1}$$
$$\implies \begin{bmatrix} I & -A^{-1}B \\ 0 & I \end{bmatrix}^{-1} \begin{bmatrix} A & B \\ B^{T} & C \end{bmatrix}^{-1} \begin{bmatrix} I & 0 \\ -B^{T}A^{-1} & I \end{bmatrix}^{-1} = \begin{bmatrix} A & 0 \\ 0 & C - B^{T}AB \end{bmatrix}^{-1},$$

and hence

$$\begin{bmatrix} A & B \\ B^T & C \end{bmatrix}^{-1} = \begin{bmatrix} I & -A^{-1}B \\ 0 & I \end{bmatrix} \begin{bmatrix} A & 0 \\ 0 & C - B^T A B \end{bmatrix}^{-1} \begin{bmatrix} I & 0 \\ -B^T A^{-1} & I \end{bmatrix}$$
$$= \begin{bmatrix} I & -A^{-1}B \\ 0 & I \end{bmatrix} \begin{bmatrix} A^{-1} & 0 \\ 0 & (C - B^T A B)^{-1} \end{bmatrix} \begin{bmatrix} I & 0 \\ -B^T A^{-1} & I \end{bmatrix}.$$

The above steps work for general partitioned 2 by 2 matrices as well. The result is as follows

$$\begin{bmatrix} I & 0 \\ -CA^{-1} & I \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} I & -BA^{-1} \\ 0 & I \end{bmatrix} = \begin{bmatrix} A & 0 \\ 0 & D - CA^{-1}B \end{bmatrix}$$
$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} = \begin{bmatrix} I & -BA^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} A & 0 \\ 0 & D - CA^{-1}B \end{bmatrix}^{-1} \begin{bmatrix} I & 0 \\ -CA^{-1} & I \end{bmatrix},$$
or

$$\begin{bmatrix} I & -BD^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} I & 0 \\ -D^{-1}C & I \end{bmatrix} = \begin{bmatrix} A - BD^{-1}C & 0 \\ 0 & D \end{bmatrix}$$
$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} = \begin{bmatrix} I & -BD^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} A - BD^{-1}C & 0 \\ 0 & D \end{bmatrix}^{-1} \begin{bmatrix} I & 0 \\ -D^{-1}C & I \end{bmatrix}.$$

8.2 *LU and Cholesky decomposition

Fact 28. The following is true for upper and lower triangular matrices:

$$\begin{bmatrix} I & 0 \\ M & I \end{bmatrix}^{-1} = \begin{bmatrix} I & 0 \\ -M & I \end{bmatrix}$$
$$\begin{bmatrix} I & M \\ 0 & I \end{bmatrix}^{-1} = \begin{bmatrix} I & -M \\ 0 & I \end{bmatrix}.$$

From the last section

$$\begin{bmatrix} I & 0 \\ -CA^{-1} & I \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} I & -BA^{-1} \\ 0 & I \end{bmatrix} = \begin{bmatrix} A & 0 \\ 0 & D - CA^{-1}B \end{bmatrix}.$$

Applying Fact 28 to the last equation gives the *block LU decomposition*:

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} I & 0 \\ CA^{-1} & I \end{bmatrix} \begin{bmatrix} A & 0 \\ 0 & D - CA^{-1}B \end{bmatrix} \begin{bmatrix} I & A^{-1}B \\ 0 & I \end{bmatrix}$$
$$= \begin{bmatrix} I & 0 \\ CA^{-1} & I \end{bmatrix} \begin{bmatrix} A & B \\ 0 & D - CA^{-1}B \end{bmatrix},$$

which shows any square matrix can be decomposed into the product of a lower triangular matrix and an upper triangular matrix.

There is also block Cholesky decomposition

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} I \\ CA^{-1} \end{bmatrix} A \begin{bmatrix} I & A^{-1}B \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & D - CA^{-1}B \end{bmatrix},$$

or using half matrices

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} A^{\frac{1}{2}} \\ CA^{-\frac{*}{2}} \end{bmatrix} \begin{bmatrix} A^{\frac{*}{2}} & A^{-\frac{1}{2}}B \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & Q^{\frac{1}{2}} \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & Q^{\frac{*}{2}} \end{bmatrix}$$
$$Q = D - CA^{-1}B,$$

where

$$A^{\frac{1}{2}}A^{\frac{*}{2}} = A, \ Q^{\frac{1}{2}}Q^{\frac{*}{2}} = Q.$$

 $\left[\begin{array}{cc} A & B \\ C & D \end{array}\right] = LU,$

Hence

where

$$LU = \begin{bmatrix} A^{\frac{1}{2}} & 0\\ CA^{-\frac{*}{2}} & 0 \end{bmatrix} \begin{bmatrix} A^{\frac{*}{2}} & A^{-\frac{1}{2}}B\\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0\\ 0 & Q^{\frac{1}{2}} \end{bmatrix} \begin{bmatrix} 0 & 0\\ 0 & Q^{\frac{*}{2}} \end{bmatrix} = \begin{bmatrix} A^{\frac{1}{2}} & 0\\ CA^{-\frac{*}{2}} & Q^{\frac{1}{2}} \end{bmatrix} \begin{bmatrix} A^{\frac{*}{2}} & A^{-\frac{1}{2}}B\\ 0 & Q^{\frac{*}{2}} \end{bmatrix}.$$

8.3 Determinant and matrix inverse identity

Although $AB \neq BA$ in general, the determinants of products have the following property:

$$\det (AB) = \det (BA) = \det A \det B,$$

where A and B should be square to start with.

Theorem 29 (Sylvester's determinant theorem). For $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{n \times m}$,

$$\det\left(I_m + AB\right) = \det\left(I_n + BA\right),\,$$

where I_m and I_n are the $m \times m$ and $n \times n$ identity matrices, respectively.

Proof. Construct

$$M = \left[\begin{array}{cc} I_m & -A \\ B & I_n \end{array} \right].$$

From the decomposition

$$M = \begin{bmatrix} I_m & 0\\ B & I_n \end{bmatrix} \begin{bmatrix} I_m & -A\\ 0 & I_n + BA \end{bmatrix},$$

we have

$$\det M = \det \left(I_n + BA \right).$$

Alternatively

$$M = \left[\begin{array}{cc} I_m + AB & -A \\ 0 & I_n \end{array} \right] \left[\begin{array}{cc} I_m & 0 \\ B & I_n \end{array} \right].$$

Hence

$$\det M = \det \left(I_m + AB \right).$$

Therefore

$$\det (I_m + AB) = \det M = \det (I_n + BA)$$

r	-	-	-	

More generally, for any invertible $m \times m$ matrix X

x

$$\det (X + AB) = \det (X) \det \left(I_n + BX^{-1}A \right),$$

which comes from

$$X + AB = X \left(I + X^{-1}AB \right)$$

$$\Rightarrow \det \left(X + AB \right) = \det \left[X \left(I + X^{-1}AB \right) \right] = \det X \det \left(I + X^{-1}AB \right).$$

8.4 Matrix inversion lemma

Fact 30 (Matrix inversion lemma). Assume A is nonsingular and $(A + BC)^{-1}$ exists. The following is true

$$(A+BC)^{-1} = A^{-1} \left(I - B \left(CA^{-1}B + I \right)^{-1} CA^{-1} \right).$$
(23)

Proof. Consider

$$(A+BC)x = y. \tag{24}$$

We aim at getting

$$= (*) y$$
, where $(*)$ will be our $(A + BC)^{-1}$. (25)

First, let

$$Cx = d. \tag{26}$$

Equation (24) can be written as

$$Ax + Bd = y$$
$$Cx - d = 0.$$

Solving the first equation yields

$$x = A^{-1} (y - Bd).$$
 (27)

Then (26) becomes

 $CA^{-1}\left(y-Bd\right) = d.$

Combining the terms about d and applying matrix inversion yield

$$d = (CA^{-1}B + I)^{-1} CA^{-1}y.$$

Putting the result in (27) yields

$$x = A^{-1} \left(y - B \left(CA^{-1}B + I \right)^{-1} CA^{-1}y \right)$$

= $A^{-1} \left(I - B \left(CA^{-1}B + I \right)^{-1} CA^{-1} \right) y.$

Comparing the above with (25), we obtain (23).

Exercise 31. The matrix inversion lemma is a powerful tool useful for many applications. One application in adaptive control and system identification uses

$$(A + \phi \phi^T)^{-1} = A^{-1} \left(I - \frac{\phi \phi^T A^{-1}}{\phi^T A^{-1} \phi + 1} \right).$$

Prove the above result. Prove also the general case (called rank one update):

$$(A + bc^{T}) = A^{-1} - \frac{1}{1 + c^{T} A^{-1} b} (A^{-1} b) (c^{T} A^{-1}).$$

Fact 32 (More extended matrix inversion lemma). Assume A, C, and $A + BCB^T$ are nonsingular. The following is true

$$(A + BCB^{T})^{-1} = A^{-1} \left(I - B \left(CB^{T}A^{-1}B + I \right)^{-1} CB^{T}A^{-1} \right)$$
(28)

$$= A^{-1} - A^{-1}B \left(CB^{T}A^{-1}B + I \right)^{-1} CB^{T}A^{-1}$$
(29)

$$= A^{-1} - A^{-1}B \left(B^{T} A^{-1}B + C^{-1} \right)^{-1} B^{T} A^{-1}.$$
 (30)

Proof. Consider

$$\left(A + BCB^T\right)x = y.$$

We aim at getting x = (*) y, where (*) will be our $(A + BCB^T)^{-1}$. First, let

$$CB^T x = d.$$

We have

$$Ax + Bd = y$$
$$CB^T x - d = 0.$$

Solving the first equation yields

Then

$$CB^T A^{-1} \left(y - Bd \right) = d$$

 $x = A^{-1} \left(y - Bd \right).$

gives

$$d = (CB^{T}A^{-1}B + I)^{-1}CB^{T}A^{-1}y.$$

Hence

$$x = A^{-1} \left(y - B \left(CB^T A^{-1} B + I \right)^{-1} CB^T A^{-1} y \right)$$

= $A^{-1} \left(I - B \left(CB^T A^{-1} B + I \right)^{-1} CB^T A^{-1} \right) y$

and (28) follows.

The extended matrix inversion lemma is key in transforming the Kalman filter to the information filter when inverting the innovation of covariance matrices.

8.5 Special inverse equalities

Fact 33. The following matrix equalities are true

• $(I + GK)^{-1}G = G(I + KG)^{-1}$

to prove the result, start with G(I + KG) = (I + GK)G

- $GK(I+GK)^{-1} = G(I+KG)^{-1}K = (I+GK)^{-1}GK$ (the proof uses the first equality twice)
- generalization 1: $(\sigma^2 I + GK)^{-1} G = G (\sigma^2 I + KG)^{-1}$
- generalization 2: if M is invertible then $(M + GK)^{-1}G = M^{-1}G(I + KM^{-1}G)^{-1}$

Exercise 34. Check validity of the following equality, assuming proper dimensions and invertibility of matrices:

•
$$Z(I+Z)^{-1} = I - (I+Z)^{-1}$$

•
$$(I + XY)^{-1} = I - XY (I + XY)^{-1} = I - X (I + YX)^{-1} Y$$

• extension

$$(I + XZ^{-1}Y)^{-1} = I - XZ^{-1}Y(I + XZ^{-1}Y)^{-1} = I - XZ^{-1}(I + YXZ^{-1})^{-1}Y$$
$$= I - X(Z + YX)^{-1}Y$$

9 Spectral mapping theorem

Theorem 35 (Spectral Mapping Theorem). Take any $A \in \mathbb{C}^{n \times n}$ and a polynomial (in s) f(s) (more generally, analytic functions). Then

$$\operatorname{eig}\left(f\left(A\right)\right) = f\left(\operatorname{eig}\left(A\right)\right).$$

Proof. Let

$$f(A) = x_0 I + x_1 A + x_2 A^2 + \dots$$

Let λ be an eigenvalue of A. We first observe that λ^n is an eigenvalue of A^n . This can be seen from $\det(\lambda^n I - A^n) = \det[(\lambda I - A) p(A)] = \det(\lambda I - A) \det(p(A))$ where p(A) is a polynomial of A. Now consider $f(\lambda) = x_0 + x_1\lambda + x_2\lambda^2 + \dots$ We have

$$\det (f(\lambda) I - f(A)) = \det \left[x_1 (\lambda I - A) + x_2 (\lambda^2 I - A^2) + x_3 (\lambda^3 I - A^3) + \dots \right]$$
$$= \det \left[(\lambda I - A) q(A) \right]$$
$$= \det (\lambda I - A) \det (q(A)).$$

Hence $f(\lambda)$ is an eigenvalue of f(A).

Conversely, if γ is an eigenvalue of f(A), we need to prove that γ is in the form of $f(\lambda)$. Factorize the polynomial

$$f(\lambda) - \gamma = a_0 (\lambda - \alpha_1) (\lambda - \alpha_2) \dots (\lambda - \alpha_n)$$

On the other hand, we note that as a matrix polynomial with the same coefficients:

$$f(A) - \gamma I = a_0 \left(A - \alpha_1 I \right) \left(A - \alpha_2 I \right) \dots \left(A - \alpha_n I \right).$$

But det $(f(A) - \gamma I) = 0$, which means that there is at least one α_i such that

$$\det\left(A - \alpha_i I\right) = 0,$$

which says that α_i is an eigenvalue of A. Hence

$$f(\lambda) - \gamma = a_0 (\lambda - \alpha_i) \prod_{k \neq i} (\lambda - \alpha_k) = 0,$$

i.e.

$$\gamma = f\left(\lambda\right),$$

where λ is an eigenvalue of A.

Example 36. Compute the eigenvalues of

$$A = \left[\begin{array}{cc} 99.8 & 2000 \\ -2000 & 99.8 \end{array} \right].$$

Solution:

$$A = 99.8I + 2000 \left[\begin{array}{cc} 0 & 1 \\ -1 & 0 \end{array} \right].$$

So

$$eig(A) = 99.8 + 2000 eig \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} = 99.8 \pm 2000i$$

10 Matrix exponentials

Since the Taylor series

$$e^{st} = 1 + st + \frac{s^2t^2}{2!} + \frac{s^3t^3}{3!} + \dots$$

converges everywhere, we can define the exponential of a matrix $A \in \mathcal{C}^{n \times n}$ by

$$e^{At} = I + At + \frac{A^2t^2}{2!} + \frac{A^3t^3}{3!} + \dots$$

Fact 37. Properties of matrix exponentials

- 1. $e^{A0} = I$
- 2. $e^{A(t+s)} = e^{At}e^{As}$
- 3. If AB = BA then $e^{(A+B)t} = e^{At}e^{Bt} = e^{Bt}e^{At}$
- 4. det $(e^{At}) = e^{trace(A)t}$
- 5. e^{At} is nonsingular for all $t \in \mathcal{R}$ and $(e^{At})^{-1} = e^{-At}$
- 6. e^{At} is the unique solution X of the linear system of ordinary differential equations

 $\dot{X} = AX$, subject to X(0) = I
11 Inner product

11.1 Inner product spaces

Basics: Inner product, or dot product, brings a metric for vector lengths. It takes two vectors and generates a number. In \mathbb{R}^n , we have

$$\langle a, b \rangle \triangleq a^T b = [a_1, a_2, \dots, a_n] \begin{vmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{vmatrix}.$$

Clearly, $\langle a, b \rangle \triangleq a^T b = \langle b, a \rangle$. Letting b = a above, we get the square of the length of a:

$$||a|| = \sqrt{a_1^2 + a_2^2 + \dots + a_n^2}.$$

Formal definitions:

Definition 38. A real vector space V is called a real inner product space, if for any vectors a and b in V there is an associated real number $\langle a, b \rangle$, called the inner product of a and b, such that the following axioms hold:

• (linearity) For all scalars q_1 and q_2 and all vectors $a, b, c \in \mathbf{V}$

$$\langle q_1 a + q_2 b, c \rangle = q_1 \langle a, b \rangle + q_2 \langle b, c \rangle$$

• (symmetry) $\forall a, b \in \mathbf{V}$

 $\langle a, b \rangle = \langle b, a \rangle$

• (positive definiteness) $\forall a \in \mathbf{V}$

 $\langle a, a \rangle \ge 0$

where $\langle a, a \rangle = 0$ if and only if a = 0.

Definition 39 (2-norm of vectors). The length of a vector in V is defined by

$$||a|| = \sqrt{\langle a, a \rangle} \ge 0.$$

For \mathbb{R}^n ,

$$||a|| = \sqrt{a^T a} = \sqrt{a_1^2 + a_2^2 + \dots + a_n^2}.$$

This is the Euclidean norm or 2-norm of the vector. \mathbb{R}^n equiped with the inner product $\langle a, b \rangle = \sqrt{a^T b}$ is called the *n*-dimensional Euclidean space.

Example 40 (Inner product for functions, function spaces). The set of all real-valued continuous functions f(x), g(x), ... $x \in [\alpha, \beta]$ is a real vector space under the usual addition of functions and multiplication by scalars. An inner product on this function space is

$$\langle f,g\rangle = \int_{\alpha}^{\beta} f(x) g(x) dx$$

and the norm of f is

$$\left|\left|f\left(x\right)\right|\right| = \sqrt{\int_{\alpha}^{\beta} f\left(x\right)^{2} dx}.$$

Inner products is also closely related to the geometric relationships between vectors. For the twodimensional case, it is readily seen that

$$v_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \ v_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

is a basis of the vector space. The two vectors are additionally orthogonal, by direct observation. More generally, we have:

Definition 41 (Orthogonal vectors). Vectors whose inner product is zero are called orthogonal.

Definition 42 (Orthonormal vectors). Orthogonal vectors with unity norm is called orthonormal.

Definition 43. The angle between two vectors is defined by

$$\cos \angle (a,b) = \frac{\langle a,b \rangle}{||a|| \cdot ||b||} = \frac{\langle a,b \rangle}{\sqrt{\langle a,a \rangle} \cdot \sqrt{\langle b,b \rangle}}.$$

11.2 Trace (standard matrix inner product)

The trace of an $n \times n$ matrix $A = [a_{ik}]$ is given by

$$\operatorname{Tr}\left(A\right) = \sum_{i=1}^{n} a_{ii}.$$
(31)

Trace defines the so-called matrix inner product:

$$\langle A, B \rangle = \operatorname{Tr} \left(A^T B \right) = \operatorname{Tr} \left(B^T A \right) = \langle B, A \rangle,$$
 (32)

which is closely related to vector inner products. Take an example in $\mathbb{R}^{3\times3}$: write the matrices in the column-vector form $B = [\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3]$, $A = [\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3]$, then

$$A^{T}B = \begin{bmatrix} \mathbf{a}_{1}^{T}\mathbf{b}_{1} & * & * \\ * & \mathbf{a}_{2}^{T}\mathbf{b}_{2} & * \\ * & * & \mathbf{a}_{3}^{T}\mathbf{b}_{3} \end{bmatrix}.$$
 (33)

So

$$\operatorname{Tr} \left(A^{T}B \right) = \mathbf{a}_{1}^{T}\mathbf{b}_{1} + \mathbf{a}_{2}^{T}\mathbf{b}_{2} + \mathbf{a}_{3}^{T}\mathbf{b}_{3},$$

which is nothing but the inner product of the two "stacked" vectors $\begin{bmatrix} \mathbf{a}_{1} \\ \mathbf{a}_{2} \\ \mathbf{a}_{3} \end{bmatrix}$ and $\begin{bmatrix} \mathbf{b}_{1} \\ \mathbf{b}_{2} \\ \mathbf{b}_{3} \end{bmatrix}$. Hence
 $\langle A, B \rangle = \operatorname{Tr} \left(A^{T}B \right) = \left\langle \begin{bmatrix} \mathbf{a}_{1} \\ \mathbf{a}_{2} \\ \mathbf{a}_{3} \end{bmatrix}, \begin{bmatrix} \mathbf{b}_{1} \\ \mathbf{b}_{2} \\ \mathbf{b}_{3} \end{bmatrix} \right\rangle.$

Exercise 44. If x is a vector, show that

$$\operatorname{Tr}(xx^T) = x^T x.$$

12 Norms

Previously we have used $|| \cdot ||$ to denote the Euclidean length function. At different times, it is useful to have more general notions of size and distance in vector spaces. This section is devoted to such generalizations.

12.1 Vector norm

Definition 45. A *norm* is a function that assigns a real-valued length to each vector in a vector space \mathbb{C}^m . To develop a reasonable notion of length, a norm must satisfy the following properties: for any vectors a, b and scalars $\alpha \in \mathbb{C}$,

- the norm of a nonzero vector is positive: $||a|| \ge 0$, and ||a|| = 0 if and only if a = 0
- scaling a vector scales its norm by the same amount: $||\alpha a|| = |\alpha| ||a||$
- triangle inequality: $||a + b|| \le ||a|| + ||b||$

Let w_1 be a $n \times 1$ vector. The most important class of vector norms, the p norms, of w are defined by

$$||w||_p = \left(\sum_{i=1}^n |w_i|^p\right)^{1/p}, \ 1 \le p < \infty.$$

Specifically, we have

 $\|w\|_1 = \sum_{i=1}^n |w_j|$ (absolute column sum)

$$||w||_{\infty} = \max_{i} |w_{i}|$$

 $||w||_2 = \sqrt{w^H w}$ (Euclidean norm)

Remark 46. When unspecified, $|| \cdot ||$ refers to 2 norm in this set of notes.

Intuitions for the infinity norm By definition

$$||w||_{\infty} = \lim_{p \to \infty} \left(\sum_{i=1}^{n} |w_i|^p \right)^{1/p}.$$

Intuitively, as p increases, $\max_i |w_i|$ takes more and more weighting in $\sum_{i=1}^n |w_i|^p$. More rigorously, we have

$$\lim_{p \to \infty} \left((\max |w_i|)^p \right)^{1/p} \le \lim_{p \to \infty} \left(\sum_{i=1}^n |w_i|^p \right)^{1/p} \le \lim_{p \to \infty} \left(\sum_{i=1}^n \left(\max |w_i| \right)^p \right)^{1/p}$$

Both $\lim_{p\to\infty} \left(\left(\max |w_i| \right)^p \right)^{1/p}$ and $\lim_{p\to\infty} \left(\sum_{i=1}^n \left(\max |w_i| \right)^p \right)^{1/p}$ equals $\max_i |w_i|$. Hence $||w||_{\infty} = \max |w_i|$.

12.2 Induced matrix norm

As matrices define linear transformations between vector spaces, it makes sense to have a measure of the "size" of the transformation. Induced matrix norms² are defined by

$$||M||_{p \leftarrow q} = \max_{x \neq 0} \frac{||Mx||_p}{||x||_q}.$$
(34)

In other words, $||M||_{q \leftarrow q}$ is the maximum factor by which M can "stretch" a vector x.

In particular, the following matrix norms are common:

 $\|M\|_{1\leftarrow 1} = \max_j \sum_{i=1}^n |M_{ij}|$ maximum absolute column sum

 $\|M\|_{\infty \leftarrow \infty} = \max_i \sum_{j=1}^m |M_{ij}|$ maximum absolute row sum

 $\|M\|_{2\leftarrow 2} = \sqrt{\lambda_{\max}\left(M^*M\right)}$ maximum singular value

The induced 2 norm can be understood as follows:

$$||M||_{2\leftarrow 2} = \max_{x\neq 0} \frac{||Mx||_2}{||x||_2} = \max_{x\neq 0} \sqrt{\frac{x^*M^*Mx}{\langle x,x\rangle^2}} = \sqrt{\lambda_{\max}(M^*M)}$$

Remark 47. When p = q in (34), often the induced matrix norm is simply written as $||M||_p$.

12.3 Frobenius norm and general matrix norms

Matrix norms do not have to be induced by vector norms.

 $^{^2\}mbox{It}$ is 'induced' from other vector norms as shown in the definition.

Formal definition: Let \mathcal{M}_n be the set of all $n \times n$ real- or complex-valued matrices. We call a function $|| \cdot || : \mathcal{M}_n \to \mathbb{R}$ a matrix norm if for all $A, B \in \mathcal{M}_n$ it satisfies the following axioms:

- 1. $||A|| \ge 0$
- 2. ||A|| = 0 if and only if A = 0
- 3. ||cA|| = |c|||A|| for all complex scalars c
- 4. $||A + B|| \le ||A|| + ||B||$
- 5. $||AB|| \le ||A|| ||B||$

The formal definition of matrix norms is slightly amended from vector norms. This is because although \mathcal{M}_n is itself a vector space of dimension n^2 , it has a natural multiplication operation that is obsent in regular vector spaces. A vector norm on matrices that satisfies the first four axioms and not necessarily axiom 5 is often called a generalized matrix norm.

Frobenius norm: The most important matrix norm which is not induced by a vector norm is the Frobenius norm, defined by

$$||A||_F \triangleq \sqrt{\operatorname{Tr}(A^*A)} = \sqrt{\langle A, A \rangle} = \sqrt{\sum_{i,j} |a_{i,j}|^2}.$$

Frobenius norm is just the Euclidean norm of the matrix, written out as a long column vector:

$$||A||_F = (\operatorname{Tr} (A^*A))^{\frac{1}{2}} = \left(\sum_{i=1}^m \sum_{j=1}^m |a_{i,j}|^2\right)^{\frac{1}{2}}.$$

We also have bounds for Frobenius norms:

$$||AB||_F^2 \le ||A||_F^2 ||B||_F^2.$$

Transforming from one matrix norm to another:

Theorem 48. If $|| \cdot ||$ is a matrix norm on \mathcal{M}_n and if $S \in \mathcal{M}_n$ is nonsingular, then

$$||A||_S = ||S^{-1}AS|| \ \forall A \in \mathcal{M}_n$$

is a matrix norm.

Exercise 49. Prove Theorem 48.

12.4 Norm inequalities

1. Cauchy-Schwartz Inequality:

$$|\langle x, y \rangle| \le ||x||_2 ||y||_2,$$

which is the special case of the Holder inequality

$$\langle x, y \rangle | \le ||x||_p ||y||_q, \ \frac{1}{p} + \frac{1}{q} = 1, \ 1 \le p, q \le \infty.$$
 (35)

Both bounds are tight: for certain choices of x and y, the inequalities become equalities.

2. Bounding induced matrix norms:

$$||AB||_{l\leftarrow n} \le ||A||_{l\leftarrow m} ||B||_{m\leftarrow n},\tag{36}$$

which comes from

$$||ABx||_{l} \le ||A||_{l \leftarrow m} ||Bx||_{m} \le ||A||_{l \leftarrow m} ||B||_{m \leftarrow n} ||x||_{n}$$

In general, the bound is not tight. For instance, $||A^n|| = ||A||^n$ does not hold for $n \ge 2$ unless A has special structures.

3. (35) and (36) are useful for computing bounds of difficult-to-compute norms. For instance, $||A||_2^2$ is expensive to compute but $||A||_1$ and $||A||_{\infty}$ are not. As a special case of (36), we have

$$||A||_2^2 \le ||A||_1 ||A||_{\infty}.$$

We can obtain an upper bound of $||A||_2^2$ by computing $||A||_1||A||_{\infty}$.

4. Any matrix induced norms of A are larger than the absolute eigenvalues of A:

$$|\lambda(A)| \le ||A||_p$$

Hence as a special case, the spectral radius is upper bounded by any matrix norms:

$$\rho\left(A\right) \le \|A\|.$$

5. Let $A \in \mathcal{M}_n$ and $\epsilon > 0$ be given. There is a matrix norm such that

$$\rho(A) \le \|A\| \le \rho(A) + \epsilon.$$

Hint: A can be decomposed as $A = U^* \Delta U$ where U is unitary and Δ is upper triangular [Schur triangularization theorem]. Let $D_t = \text{diag}(t, t^2, \dots, t^n)$ and compute

$$D_t \Delta D_t^{-1} = \begin{bmatrix} \lambda_1 & t^{-1} d_{12} & \dots & t^{-n+1} d_{1n} \\ 0 & \lambda_2 & t^{-1} d_{23} & \dots & t^{-n+2} d_{1n} \\ \vdots & \ddots & \lambda_3 & \ddots & \vdots \\ & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & t^{-1} d_{n-1,n} \\ 0 & \dots & & \dots & 0 & \lambda_n \end{bmatrix}$$

For t large enough, the sum of the absolute values of the off-diagonal entries is less than ϵ and in particular

$$\|D_t \Delta D_t^{-1}\|_1 \le \rho(A) + \epsilon.$$

12.5 Exercises

- 1. Let x be an m vector and A be an $m \times n$ matrix. Verify each of the following inequalities, and give an example when the equality is achieved.
 - (a) $||x||_{\infty} \leq ||x||_{2}$
 - (b) $||x||_2 < \sqrt{m} ||x||_\infty$
 - (c) $||A||_{\infty} \leq \sqrt{n} ||A||_2$
 - (d) $||A||_2 \le \sqrt{m} ||A||_{\infty}$
- 2. Let x be a random vector with mean E[x] = 0 and covariance $E(xx^T) = I$, then

$$||A||_F^2 = \mathbf{E}\left[||Ax||_2^2\right].$$

Hint: use Exercise 44.

13 Symmetric, skew-symmetric, and orthogonal matrices

13.1 Definitions and basic properties

A real square matrix A is called symmetric if $A = A^T$, skew-symmetric if $A = -A^T$.

Fact 50. Any real square matrix A may be written as the sum of a symmetric matrix R and a skew-symmetric matrix S, where

$$R = \frac{1}{2} (A + A^T), \ S = \frac{1}{2} (A - A^T).$$

If $A = [a_{jk}]$, then the complex conjugate of A is denoted as $\overline{A} = [\overline{a}_{jk}]$, i.e., each element $a_{jk} = \alpha + i\beta$ is replaced with its complex conjugate $\overline{a}_{jk} = \alpha - i\beta$.

A square matrix A is called Hermitian if $A^T = \overline{A}$; skew-Hermitian if $A^T = -\overline{A}$.

Example 51. Find the symmetric, skew-symmetric, Hermitian, and skew-Hermitian matrices in the set:

$$\left\{ \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 2i \\ 2i & 1 \end{bmatrix}, \begin{bmatrix} 1 & 2i \\ -2i & 1 \end{bmatrix}, \begin{bmatrix} 0 & 2 \\ -2i & 0 \end{bmatrix}, \begin{bmatrix} 0 & 2+2i \\ 2-2i & 0 \end{bmatrix} \right\}.$$

We introduce one more class of important matrices: a real square matrix A is called **orthogonal**³ if

$$A^T A = A A^T = I. ag{37}$$

Writing A in the column-vector notation

$$A = \left[a_1, a_2, \ldots, a_n\right],$$

³Some people also call use the notion of orthonormal matrix. But orthogonal matrix is more often used (we can say orthonormal basis).

we get the equivalent form of (37):

$$A^{T}A = \begin{bmatrix} a_{1}^{T} \\ a_{2}^{T} \\ \vdots \\ a_{n}^{T} \end{bmatrix} \begin{bmatrix} a_{1}, a_{2}, \dots, a_{n} \end{bmatrix} = \begin{bmatrix} a_{1}^{T}a_{1} & a_{1}^{T}a_{2} & \dots & a_{1}^{T}a_{n} \\ a_{2}^{T}a_{1} & a_{2}^{T}a_{2} & \dots & a_{2}^{T}a_{n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n}^{T}a_{1} & a_{n}^{T}a_{2} & \dots & a_{n}^{T}a_{n} \end{bmatrix} = I.$$

Hence it must be that

$$a_j^T a_j = 1$$
$$a_j^T a_m = 0 \ \forall j \neq m$$

namely, a_j and a_m are orthonormal for any $j \neq m$.

The complex version of an orthogonal matrix is the **unitary matrix**. A square matrix A is called unitary if $A\overline{A}^T = \overline{A}^T A = I$, namely $A^{-1} = \overline{A}^T$.

Remark 52. Often the complex conjugate transpose \overline{A}^T is written as A^* .

Theorem 53. The eigenvalues of symmetric matrices are all real.

Proof. $\forall : A \in \mathbb{R}^{n \times n}$ with $A^T = A$. $Au = \lambda u \Rightarrow \overline{u}^T Au = \lambda \overline{u}^T u$, where \overline{u} is the complex conjugate of u. $\overline{u}^T Au$ is a real number, as

$$\overline{u}^T A \overline{u} = u^T \overline{A} \overline{u}$$

$$= u^T A \overline{u} \quad \because A \in \mathbb{R}^{n \times n}$$

$$= u^T A^T \overline{u} \quad \because A = A^T$$

$$= \lambda u^T \overline{u} \quad \because (Au)^T = (\lambda u)^T$$

$$= \lambda \overline{u}^T u \quad \because u^T \overline{u} \in \mathbb{R}$$

$$= \overline{u}^T A u \quad \because Au = \lambda u.$$

By definition of complex conjugate numbers, $\overline{u}^T u \in \mathbb{R}$. So $\lambda = \frac{\overline{u}^T A u}{\overline{u}^T u}$ is also a real number. \Box

Theorem 54. The eigenvalues of skew-symmetric matrices are all imaginary or zero.

The proof is left as an exercise.

Fact 55. An orthogonal transformation preserves the value of the inner product of vectors a and b in \mathbb{R}^n . That is, for any a and b in \mathbb{R}^n , orthogonal $n \times n$ matrix A, and u = Aa, v = Ab we have $\langle u, v \rangle = \langle a, b \rangle$, as

$$u^T v = a^T A^T A b = a^T b.$$

Hence the transformation also preserves the length or 2-norm of any vector a in \mathbb{R}^n given by $||a||_2 = \sqrt{\langle a, a \rangle}$.

Theorem 56. The determinant of an orthogonal matrix is either 1 or -1.

Proof. $UU^T = I \Rightarrow \det U \det U^T = (\det U)^2 = 1.$

Theorem 57. The eigenvalues of an orthogonal matrix A are real or complex conjugates in pairs and have absolute value 1.

Proof. $Au = \lambda u \Rightarrow A^T Au = \lambda A^T u \Rightarrow u = \lambda A^T u \Rightarrow \overline{u}^T u = \lambda \overline{u}^T A^T u \Rightarrow \overline{u}^T u = \lambda \overline{u}^T \overline{A}^T u = \lambda \overline{\lambda} \overline{\lambda} \overline{u}^T u \Rightarrow (|\lambda|^2 - 1) \overline{u}^T u = 0.$

Properties of the special matrices From the above results, we have the following table:

	1	. •
real matrix	complex matrix	properties
symmetric ($A = A^T$)	Hermitian $(A^* = A)$	eigenvalues are all real
orthogonal	unitary	eigenvalues have unity magnitude; Ax
$(A^T A = A A^T = I)$	$(A^*A = AA^* = I)$	maintains the 2-norm of x
skew-symmetric	skew-Hermitian eigenvalues are all imaginary or zero	
$(A^T = -A)$	$(A^* = -A)$	

Based on the eigenvalue characteristics, we have:

- symmetric and Hermitian matrices are like the real line in the complex domain
- skew-symmetric/Hermitian matrices are like the imaginary line
- orthogonal/unitary matrices are like the unit circle

Exercise 58 (Representation of matrices using special matrices). Many unitary matrices can be mapped as functions of skew-Hermitian matrices as follows

$$U = (I - S)^{-1} (I + S),$$

where $S \neq I$. Show that if S is skew-Hermitian, then U is unitary.

13.2 Symmetric eigenvalue decomposition (SED)

When $A \in \mathbb{R}^{n \times n}$ has n distinct eigenvalues, we have seen the useful result of matrix diagonalization:

$$A = U\Lambda U^{-1} = \begin{bmatrix} u_1, \dots, u_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \begin{bmatrix} u_1, \dots, u_n \end{bmatrix}^{-1},$$
(38)

where λ_i 's are the distinct eigenvalues associated to the eigenvector u_i 's.

The inverse matrix in (38) can be cumbersome to compute though.

The spectral theorem, aka symmetric eigenvalue decomposition theorem,⁴ significantly simplifies the result when A is symmetric.

⁴Recall that the set of all the eigenvalues of A is called the spectrum of A. The largest of the absolute values of the eigenvalues of A is called the spectral radius of A.

Theorem 59. $\forall : A \in \mathbb{R}^{n \times n}, A^T = A$, there always exist λ_i and u_i , such that

$$A = \sum_{i=1}^{n} \lambda_i u_i u_i^T = U \Lambda U^T,$$
(39)

where:5

- λ_i 's: eigenvalues of A
- u_i : eigenvector associated to λ_i , normalized to have unity norms
- $U = [u_1, u_2, \cdots, u_n]^T$ is an orthogonal matrix, i.e., $U^T U = U U^T = I$
- $\{u_1, u_2, \cdots, u_n\}$ forms an orthonormal basis

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

To understand the result, we show an important theorem first.

Theorem 60. $\forall : A \in \mathbb{R}^{n \times n}$ with $A^T = A$, then eigenvectors of A, associated with different eigenvalues, are orthogonal.

Proof. Let $Au_i = \lambda_i u_i$ and $Au_j = \lambda_j u_j$. Then $u_i^T Au_j = u_i^T \lambda_j u_j = \lambda_j u_i^T u_j$. In the meantime, $u_i^T Au_j = u_i^T A^T u_j = (Au_i)^T u_j = \lambda_i u_i^T u_j$. So $\lambda_i u_i^T u_j = \lambda_j u_i^T u_j$. But $\lambda_i \neq \lambda_j$. It must be that $u_i^T u_j = 0$.

Theorem 59 now follows. If A has distinct eigenvalues, then $U = [u_1, u_2, \dots, u_n]^T$ is orthogonal if we normalize all the eigenvectors to unity norm. If A has r(< n) distinct eigenvalues, we can *choose* multiple orthogonal eigenvectors for the eigenvalues with none-unity multiplicities.

Observations:

• If we "walk along" u_i , then

$$Au_{j} = \left(\sum_{i} \lambda_{i} u_{i} u_{i}^{T}\right) u_{j} = \lambda_{j} u_{j} u_{j}^{T} u_{j} = \lambda_{j} u_{j},$$
(40)

where we used the orthonormal condition of $u_i^T u_j = 0$ if $i \neq j$. This confirms that u_j is an eigenvector.

 $\overline{{}^{5}u_{i}u_{i}^{T}} \in \mathbb{R}^{n \times n}$ is a symmetric dyad, the so-called outerproduct of u_{i} and u_{i} . It has the following properties:

- $\forall v \in \mathbb{R}^{n \times 1}$, $(vv^T)_{ij} = v_i v_j$. (Proof: $(vv^T)_{ij} = e_i^T (vv^T) e_j = v_i v_j$, where e_i is the unit vector with all but the i_{th} elements being zero.)
- link with quadratic functions: $q(x) = x^T (vv^T) x = (v^T x)^2$

• $\{u_i\}_{i=1}^n$ is a orthonormal basis $\Rightarrow \forall x \in \mathbb{R}^n$, $\exists x = \sum_i \alpha_i u_i$, where $\alpha_i = \langle x, u_i \rangle$. And we have

$$Ax = A\sum_{i} \alpha_{i} u_{i} = \sum_{i} \alpha_{i} A u_{i} = \sum_{i} \alpha_{i} \lambda_{i} u_{i} = \sum_{i} (\alpha_{i} \lambda_{i}) u_{i},$$
(41)

which gives the (intuitive) picture of the geometric meaning of Ax: decompose first x to the space spanned by the eigenvectors of A, scale each components by the corresponding eigenvalue, sum the results up, then we will get the vector Ax.

With the spectral theorem, next time we see a symmetric matrix A, we immediately know that

- λ_i is real for all i
- associated with λ_i , we can always find one or more real eigenvectors
- \exists an orthonormal basis $\{u_i\}_{i=1}^n$, which consists of the eigenvectors
- if $A \in \mathbb{R}^{2 \times 2}$, then if you compute first λ_1 , λ_2 and u_1 , you won't need to go through the regular math to get u_2 , but can simply solve for a u_2 that is orthogonal to u_1 with $||u_2|| = 1$.

Example 61. Consider the matrix $A = \begin{bmatrix} 5 & \sqrt{3} \\ \sqrt{3} & 7 \end{bmatrix}$. Computing the eigenvalues gives

$$\det \begin{bmatrix} 5-\lambda & \sqrt{3} \\ \sqrt{3} & 7-\lambda \end{bmatrix} = 35 - 12\lambda + \lambda^2 - 3 = (\lambda - 4) (\lambda - 8) = 0$$
$$\Rightarrow \lambda_1 = 4, \ \lambda_2 = 8.$$

We can know one of the eigenvectors from

$$(A - \lambda_1 I) t_1 = 0 \Rightarrow \begin{bmatrix} 1 & \sqrt{3} \\ \sqrt{3} & 3 \end{bmatrix} t_1 = 0 \Rightarrow t_1 = \begin{bmatrix} -\frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{bmatrix}.$$

Note here we normalized t_1 such that $||t_1||_2 = 1$. With the above computation, we no more need to do $(A - \lambda_2 I) t_2 = 0$ for getting t_2 . Keep in mind that A here is symmetric, so has eigenvectors orthogonal to each other. By direct observation, we can see that

$$x = \left[\begin{array}{c} \frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{array}\right]$$

is orthogonal to t_1 and $||x||_2 = 1$. So $t_2 = x$.

Theorem 62 (Eigenvalues of symmetric matrices). If $A = A^T \in \mathbb{R}^{n \times n}$, then the eigenvalues of A satisfy

$$\lambda_{\max} = \max_{x \in \mathbb{R}^n, \ x \neq 0} \frac{x^T A x}{\|x\|_2^2} \tag{42}$$

$$\lambda_{\min} = \min_{x \in \mathbb{R}^n, \ x \neq 0} \frac{x^T A x}{\|x\|_2^2}.$$
(43)

Proof. Perform SED to get

$$A = \sum_{i=1}^{n} \lambda_i u_i^T u_i;$$

where $\{u_i\}_{i=1}^n$ form a basis of \mathbb{R}^n . Then any vector $x \in \mathbb{R}^n$ can be decomposed as

$$x = \sum_{i=1}^{n} \alpha_i u_i.$$

Thus

$$\max_{x \neq 0} \frac{x^T A x}{\|x\|_2^2} = \max_{\alpha_i} \frac{\left(\sum_i \alpha_i u_i\right)^T \sum_i \lambda_i \alpha_i u_i}{\sum_i \alpha_i^2} = \max_{\alpha_i} \frac{\sum_i \lambda_i \alpha_i^2}{\sum_i \alpha_i^2} = \lambda_{\max}.$$

The proof for (43) is analogous and omitted.

13.3 Symmetric positive-definite matrices

Definition 63. A symmetric matrix $P \in \mathbb{R}^{n \times n}$ is called **positive-definite**, written $P \succ 0$, if $x^T P x > 0$ for all $x \neq 0 \in \mathbb{R}^n$. P is called **positive-semidefinite**, written $P \succeq 0$, if $x^T P x \ge 0$ for all $x \in \mathbb{R}^n$

Definition 64. A symmetric matrix $P \in \mathbb{R}^{n \times n}$ is called **negative-definite**, written $P \prec 0$, if $-P \succ 0$, i.e., $x^T P x < 0$ for all $x (\neq 0) \in \mathbb{R}^n$. P is called **negative-semidefinite**, written $P \preceq 0$, if $x^T P x \leq 0$ for all $x \in \mathbb{R}^n$

When A and B have compatible dimensions, $A \succ B$ means $A - B \succ 0$.

Positive-definite matrices can have negative entries, as shown in the next example.

Example 65. The following matrix is positive-definite

$$P = \left[\begin{array}{cc} 2 & -1 \\ -1 & 2 \end{array} \right],$$

as $P = P^T$ and take any $v = [x, y]^T$, we have

$$v^{T}Pv = \begin{bmatrix} x \\ y \end{bmatrix}^{T} \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 2x^{2} + 2y^{2} - 2xy = x^{2} + y^{2} + (x+y)^{2} \ge 0,$$

and the equality sign holds only when x = y = 0.

Conversely, matrices whose entries are all positive are not necessarily positive-definite.

Example 66. The following matrix is not positive-definite

$$A = \left[\begin{array}{cc} 1 & 2 \\ 2 & 1 \end{array} \right],$$

as

$$\begin{bmatrix} 1\\-1 \end{bmatrix}^T \begin{bmatrix} 1 & 2\\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1\\-1 \end{bmatrix} = -2 < 0.$$

Theorem 67. For a symmetric matrix P, $P \succ 0$ if and only if all the eigenvalues of P are positive. *Proof.* Since P is symmetric, we have

$$\lambda_{\max}\left(P\right) = \max_{x \in \mathbb{R}^n, \ x \neq 0} \frac{x^T A x}{\|x\|_2^2} \tag{44}$$

$$\lambda_{\min}(P) = \min_{x \in \mathbb{R}^n, \ x \neq 0} \frac{x^T A x}{\|x\|_2^2},$$
(45)

which gives

$$x^{T}Ax \in [\lambda_{\min} ||x||_{2}^{2}, \ \lambda_{\max} ||x||_{2}^{2}].$$

For $x \neq 0$, $||x||_2^2$ is always positive. It can thus be seen that $x^T A x > 0$, $x \neq 0 \Leftrightarrow \lambda_{\min} > 0$.

Lemma. For a symmetric matrix P, $P \succeq 0$ if and only if all the eigenvalues of P are none-negative.

Theorem. If A is symmetric positive definite, X is full column rank, then X^TAX is positive definite.

Proof. Consider $y(X^TAX) y = x^TAx$, which is always positive unless x = 0. But X is full rank so Xy = x = 0 if and only if y = 0. This proves X^TAX is positive definite.

Fact. All principle submatrices of A are positive definite.

Proof. Use the last theorem. Take $X = e_1$, $X = [e_1, e_2]$, etc. Here e_i is a column vector whose *i*th-entry is 1 and all other entries are zero.

Example 68. The following matrices are all not positive-definite:

 $\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} -1 & 1 \\ 1 & 2 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ 1 & -1 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}.$

Positive-definite matrices are like positive real numbers. We can have the concept of *square root* of positive-definite matrices.

Definition 69. Let $P \succeq 0$. We can perform symmetric eigenvalue decomposition to obtain $P = UDU^T$ where U is orthogonal with $UU^T = I$ and D is diagonal with all diagonal elements being none negative

$$D = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{bmatrix} \succeq 0$$

. Then the square root of P, written $P^{\frac{1}{2}},$ is defined as

$$P^{\frac{1}{2}} = UD^{\frac{1}{2}}U^{T}$$

,where

$$D^{\frac{1}{2}} = \begin{bmatrix} \sqrt{\lambda_1} & 0 & \dots & 0 \\ 0 & \sqrt{\lambda_2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \sqrt{\lambda_n} \end{bmatrix}$$

General positive-definite matrices 13.4

Definition 70. A general square matrix $Q \in \mathbb{R}^{n \times n}$ is called positive-definite, written as $Q \succ 0$, if $x^T Q x > 0 \ \forall x \neq 0.$

We have discussed the case when Q is symmetric. If not, recall that any real square matrix can be decomposed as the sum of a symmetric matrix and a skew symmetric matrix:

$$Q = \frac{Q + Q^T}{2} + \frac{Q - Q^T}{2},$$

where $\frac{Q+Q^T}{2}$ is symmetric. Notice that $x^T \frac{Q-Q^T}{2} x = x^T Q x - (x^T Q x)^T = 0$. So for a general square real matrix:

$$Q \succ 0 \Leftrightarrow Q + Q^T \succ 0.$$

Example 71. The following matrices are positive definite but not symmetric

$$\left[\begin{array}{rrr}1 & 1\\0 & 1\end{array}\right], \left[\begin{array}{rrr}1 & 0\\1 & 1\end{array}\right].$$

For complex matrices with $Q = Q^* = Q_R + jQ_I$, we have

$$Q \succ 0 \Leftrightarrow x^*Qx > 0, \ \forall x \neq 0$$

$$\Leftrightarrow \left(x_R^T - jx_I^T\right) \left(Q_R + jQ_I\right) \left(x_R + jx_I\right) > 0$$

$$\Leftrightarrow \left(\begin{array}{c}x_R\\x_I\end{array}\right)^T \left(\begin{array}{c}1\\j\end{array}\right) \left(\begin{array}{c}Q_R\\Q_I\end{array}\right) \left(\begin{array}{c}1\\j\end{array}\right) \left(\begin{array}{c}1\\j\end{array}\right)^T \left(\begin{array}{c}x_R\\x_I\end{array}\right)$$

$$\Leftrightarrow \left(\begin{array}{c}x_R\\x_I\end{array}\right)^T \left(\begin{array}{c}Q_R\\Q_I\\-Q_I\\Q_R\end{array}\right) \left(\begin{array}{c}x_R\\x_I\end{array}\right) > 0$$

$$\Leftrightarrow x_R^TQ_Rx_R - x_I^TQ_Ix_R + x_R^TQ_Ix_I + x_I^TQ_Rx_I > 0.$$

13.5 *Positive-definite functions and non-constant matrices

We can further extend the concept of positive definiteness to general and even time-varying functions, by placing upper and/or lower bounds that are "positive-definite like".

Define first two special functions:

- 1. class-K function: $\psi \in C^0 : [0, a] \to [0, \infty)$ with $\psi(0) = 0$ and ψ strictly increasing,
- 2. class- K_{∞} function: if the domain $a = \infty$ and $\psi(r) \to \infty$ as $r \to \infty$.

Note: ψ is continuous but does not need to be continuously differentiable, e.g.

$$\psi = \min\left\{x, x^2\right\}$$

is a class-K function.

Lemma 72. Let $V : D \to \mathbb{R}$ be a continuous, positive definite function. Let $B_r \subset D$ for some r > 0. Then there exist class-K functions ψ and ϕ defined on [0, r] such that

$$\phi\left(\|x\|\right) \le V\left(x\right) \le \psi\left(\|x\|\right)$$

for all $x \in B_r$.

- if the domain $D=\mathbb{R}^n$ then $r=\infty$,
- if V(x) is radially unbounded, then ψ and ϕ can be class- K_{∞} .

Definition 73. A time-dependent function V(t, x) is positive-semidefinite if

$$V(t, x) \ge \phi(\|x\|),$$

where ϕ is class-K.

Definition 74. A time-varying matrix P(t) is positive definite if there exists a lower-bounding positive definite matrix such that

$$P(t) \succeq c_3 I \succ 0, \ \forall t \ge 0.$$

14 Singular value and singular value decomposition (SVD)

14.1 Motivation

Symmetric eigenvalue decomposition is great but many matrices are not symmetric. A general matrix A may actually not even be square. Singular value decomposition is an important matrix decomposition technique that works for arbitrary matrices.⁶

For a general none-square matrix $A \in \mathbb{C}^{m \times n}$, eigenvalues and eigenvectors are generalized to

$$Av_j = \sigma_j u_j \tag{46}$$

Be careful about the dimensions: if m > n, we have



It turns out that, if A has full column rank n, then we can find a V that is unitary ($VV^* = V^*V = I$) and a \hat{U} that has orthonormal columns. Hence

$$A = \hat{U}\hat{\Sigma}V^*.$$
(47)

14.2 SVD

(47) forms the so-called reduced singular value decomposition (SVD). The idea of a "full" SVD is as follows. The columns of \hat{U} are n orthonormal vectors in the m-dimensional space \mathbb{C}^m . They do not form a basis for \mathbb{C}^m unless m = n. We can add additional m - n orthonormal columns to \hat{U} and augment it to a unitary matrix U. Now the matrix dimension has changed, $\hat{\Sigma}$ needs to be augmented to compatible dimensions as well. To maintain the equality (47), the newly added elements to $\hat{\Sigma}$ are set to zero.

Theorem 75. Let $A \in \mathbb{C}^{m \times n}$ with rank r. Then we can find orthogonal matrices $U \in \mathbb{C}^{m \times m}$ and $V \in \mathbb{C}^{n \times n}$ such that

$$A = U\Sigma V^*,$$

⁶History of SVD: discovered between 1873 and 1889, independently by several pioneers; did not became widely known in applied mathematics until the late 1960s, when it was shown that SVD can be computed effectively and used as the basis for solving many problems.

where

 $\Sigma \in \mathbb{R}^{m \times n}$ is diagonal $U \in \mathbb{C}^{m \times m}$ is unitary $V \in \mathbb{C}^{n \times n}$ is unitary.

In addition, the diagonal entries σ_j of Σ are nonnegative and in nonincreasing order; that is, $\sigma_1 \ge \sigma_2 \ge \cdots \ge \sigma_r > 0$.

Proof. Notice that A^*A is positive semi-definite. Hence, A^*A has a full set of orthonormal eigenvectors; its eigenvalues are real and nonnegative. Order these eigenvalues as

$$\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_r > \lambda_{r+1} = \lambda_{r+2} = \cdots = \lambda_n = 0.$$

⁷Let $\{v_1, \ldots, v_n\}$ be an orthonormal choice of eigenvectors of A^*A corresponding to these eigenvalues:

$$A^*Av_i = \lambda_i v_i.$$

Then,

$$||Av_i||^2 = v_i^* A^* A v_i = \lambda_i v_i^* v_i = \lambda_i.$$

For i > r, it follows that $Av_i = 0$. For $1 \le i \le r$, we have

$$A^*Av_i = \lambda_i v_i.$$

Recall (46), we define $\sigma_i = \sqrt{\lambda_i}$ and get

$$Av_i = \sigma_i u_i$$
$$A^* u_i = \sigma_i v_i.$$

For $1 \leq i, j \leq r$, we have

$$\langle u_i, u_j \rangle = u_i^* u_j = \frac{1}{\sigma_i \sigma_j} v_i^* A^* A v_j = \frac{1}{\sigma_i \sigma_j} \lambda_j v_i^* v_j = \begin{cases} 1 & i = j, \\ 0 & i \neq j. \end{cases}$$

Hence $\{u_1, \ldots, u_r\}$ is an orthonormal set of eigenvectors. Extending this set to form an orthonormal basis for \mathbb{C}^m gives

$$U = \left[\begin{array}{ccc} u_1, & \dots, & u_r \mid u_{r+1}, & \dots, & u_m \end{array} \right].$$

For $i \leq r$, we already have

$$Av_i = \sigma_i u_i$$

⁷Fact: rank $(A) = \operatorname{rank}(A^*A)$. To see this, notice first, that rank $(A) \ge \operatorname{rank}(A^*A)$ by definition of rank. Second, $A^*Ax = 0 \Rightarrow x^*A^*Ax = 0 \Rightarrow ||Ax|| = 0 \Rightarrow Ax = 0$, hence rank $(A) \le \operatorname{rank}(A^*A)$.

namely

$$A[v_1, \dots, v_r] = [u_1, \dots, u_r] \begin{bmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & \ddots & \\ & & & \sigma_r \end{bmatrix}$$
$$= \begin{bmatrix} u_1, & \dots, & u_r \mid u_{r+1}, & \dots, & u_m \end{bmatrix} \begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_r \\ & & & 0 \\ & & & \vdots \\ & & & 0 \end{bmatrix}$$

For v_{r+1}, \ldots , we have already seen that $Av_{r+1} = Av_{r+2} = \cdots = 0$, hence



Theorem 76. The range space of A is spanned by $\{u_1, \ldots, u_r\}$. The null space of A is spanned by $\{v_{r+1}, \ldots, v_n\}$.

Theorem 77. The nonzero singular values of A are the square roots of the nonzero eigenvalues of A^*A or AA^* .

Theorem 78. $||A||_2 = \sigma_1$, *i.e.*, the induced two norm of A is the maximum singular value of A.

The next important theorem can be easily proved via SVD.

Theorem (Fundermental theory of linear algebra). Let $A \in \mathbb{R}^{m \times n}$. Then

$$\mathcal{R}(A) + \mathcal{N}(A^T) = \mathbb{R}^m,$$

and

$$\mathcal{R}\left(A
ight)\perp\mathcal{N}\left(A^{T}
ight)$$
 .

Proof. By singular value decomposition, we have

$$A = U\Sigma V^T$$
$$A^T = V\Sigma U^T.$$

The range space of A is the first r columns of U, from the first equation. The null space of A^T is the last m - r columns of U, from the second equation.

New intuition of matrix vector operation With $A = U\Sigma V^*$, a new intuition for $Ax = U\Sigma V^*x$ is formed. Since V is unitary, it is norm-preserving, in the sense that V^*x does not change the 2-norm of the vector x. In other words, V^*x only rotates x in \mathbb{C}^n . The diagonal matrix Σ then functions to scale (by its diagonal values) the rotated vector. Finally, U is another rotation in \mathbb{C}^m .

14.3 Properties of singular values

Fact. Let A and B be matrices with compatible dimensions. The following are true

 $\overline{\sigma} (A+B) \leq \overline{\sigma} (A) + \overline{\sigma} (B),$ $\overline{\sigma} (AB) \leq \overline{\sigma} (A) \overline{\sigma} (B).$

Proof. The first inequality comes from

$$\overline{\sigma} \left(A + B \right) = \max_{v \neq 0} \frac{||Av + Bv||_2}{||v||_2} \le \max_{v \neq 0} \frac{||Av||_2 + ||Bv||_2}{||v||_2}.$$

The second inequality uses

$$\overline{\sigma}(AB) = \max_{v \neq 0} \frac{||ABv||_2}{||v||_2} \le \max_{v \neq 0} \frac{||A||_2||Bv||_2}{||v||_2}.$$

14.4 Exercises

1. Compute the singular values of the following matrices

$$(a) \begin{bmatrix} 3 \\ -2 \end{bmatrix}, (b) \begin{bmatrix} 2 \\ 3 \end{bmatrix}, (c) \begin{bmatrix} 0 & 2 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, (d) \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, (e) \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.$$

- 2. Show that if A is real, then it has a real SVD (i.e., U and V are both real).
- 3. For any matrix $A \in \mathbb{R}^{n \times m}$, construct

$$M = \begin{bmatrix} \overbrace{0}^{n \times n} & \overbrace{A}^{n \times m} \\ \overbrace{m \times n}^{T} & \overbrace{m \times m}^{0} \end{bmatrix} \in \mathbb{R}^{(n+m) \times (n+m)},$$

which satisfies

$$M^T = M.$$

 ${\cal M}$ is Hermitian, and hence has real eigenvalues and eigenvectors:

$$\begin{bmatrix} 0 & A \\ A^T & 0 \end{bmatrix} \begin{bmatrix} u_j \\ v_j \end{bmatrix} = \sigma_j \begin{bmatrix} u_j \\ v_j \end{bmatrix}.$$
(48)

- (a) Show that
 - i. v_j is in the co-kernal (perpendicular to kernal/null space) of A and u_j is in the range of A.
 - ii. if σ_j and $\begin{bmatrix} u_j \\ v_j \end{bmatrix}$ form a eigen pair for M, then $-\sigma_j$ and $\begin{bmatrix} u_j^T, -v_j^T \end{bmatrix}^T$ also form an eigen pair for M

iii. eigenvalues of M always appear in pairs that are symmetric to the imaginary axis.

(b) Use the results to show that, if

$$A = \left[\begin{array}{rrr} 1 & 2 & 4 \\ 1 & 4 & 32 \end{array} \right],$$

then M must have eigenvalues that are equal to 0.

4. Suppose $A \in \mathbb{C}^{m \times m}$ and has an SVD $A = U\Sigma V^*$. Find an eigenvalue decomposition of

0	A^*]
A	0	.

 Worst input direction in matrix vector multiplications. Recall that any matrix defines a linear transformation:

$$Mw = z$$

What is the worst input direction for the vector w? Here *worst* means: if we fix the input norm, say ||w|| = 1, ||z|| will reach a maximum value (the worst case) for a specific input direction in w.

- (a) Show that the worst ||z|| is ||M|| when ||w|| = 1.
- (b) Provide procedures to obtain the w that gives the maximum ||z||, for the cases of 1 norm, ∞ norm, and 2 norm.

Index

Α

adjoint matrix, 21 adjugate matrix, 21

В

basis, 9

С

characteristic equation, 14 characteristic polynomial, 14 *characteristic values* of a matrix, 14

D

Determinants, 10 Diagonal matrices, 1 dyad, 10

Е

eigenbasis, 17 eigenvalue, 13 eigenvectors, 13 Elementary Row Operations, 7

Н

homogeneous system, 5

L

Identity matrix, 1 inverse, 17

L

linear combination, 9 linear equation set, 1 linearly independent, 9 lower triangular matrices, 1

М

Matrix inversion lemma, 26 matrix product, 2

Ν

nonhomogeneous system, 5 null space, 10 nullity, 10

O overdetermined, 11

R

range space, 10 rank, 10 row echelon form, 7

S

singular, 11 skew-symmetric, 4 span, 9 spectral radius, 14 spectrum, 14 symmetric, 4

Т

transpose, 3

U

undetermined, **11** Upper triangular matrices, **1**

V

Vectors, $\mathbf{1}$