# Introduction to Dynamic Systems

# (Network Mathematics Graduate Programme)


Martin Corless

School of Aeronautics & Astronautics
Purdue University
West Lafayette, Indiana
corless@purdue.edu


July 14, 2011

# Contents


**I Representation of Dynamical Systems** **vii**

**1 Introduction** **1**

1.1 Ingredients . . . 4
1.2 Some notation . . . 4
1.3 MATLAB . . . 4

**2 State space representation of dynamical systems** **5**

2.1 Linear examples . . . 5
2.1.1 A first example . . . 5
2.1.2 The unattached mass . . . 6
2.1.3 Spring-mass-damper . . . 6
2.1.4 A simple structure . . . 7
2.2 Nonlinear examples . . . 8
2.2.1 A first nonlinear system . . . 8
2.2.2 Planar pendulum . . . 9
2.2.3 Attitude dynamics of a rigid body . . . 9
2.2.4 Body in central force motion . . . 10
2.2.5 Double pendulum on cart . . . 11
2.2.6 Two-link robotic manipulator . . . 12
2.3 Discrete-time examples . . . 13
2.3.1 The discrete unattached mass . . . 13
2.3.2 Additive increase multiplicative decrease (AIMD) algorithm for resource allocation . . . 14
2.4 General representation . . . 16
2.4.1 Continuous-time . . . 16
2.4.2 Discrete-time . . . 19
2.4.3 Exercises . . . 21
2.5 Vectors . . . 23
2.5.1 Vector spaces and $\mathbb{R}^n$ . . . 23
2.5.2 $\mathbb{R}^2$ and pictures . . . 25
2.5.3 Derivatives . . . 26
2.5.4 MATLAB . . . 26
2.6 Vector representation of dynamical systems . . . 27
2.7 Solutions and equilibrium states : continuous-time . . . 30
2.7.1 Equilibrium states . . . 30

2.7.2 Exercises . . . 32
2.7.3 Controlled equilibrium states . . . 33
2.8 Solutions and equilibrium states: discrete-time . . . 33
2.8.1 Equilibrium states . . . 33
2.8.2 Controlled equilibrium states . . . 35
2.9 Numerical simulation . . . 36
2.9.1 MATLAB . . . 36
2.9.2 Simulink . . . 37
2.10 Exercises . . . 38

**3 Linear time-invariant (LTI) systems** **41**
3.1 Matrices . . . 41
3.2 Linear time-invariant systems . . . 46
3.2.1 Continuous-time . . . 46
3.2.2 Discrete-time . . . 47
3.3 Linearization about an equilibrium solution . . . 48
3.3.1 Derivative as a matrix . . . 48
3.3.2 Linearization of continuous-time systems . . . 49
3.3.3 Implicit linearization . . . 52
3.3.4 Linearization of discrete-time systems . . . 55

**4 Transfer functions** **59**
4.1 ⟵ Transfer functions ⟵ . . . 59
4.1.1 The Laplace transform . . . 59
4.1.2 Transfer functions . . . 60
4.1.3 Poles and zeros . . . 67
4.1.4 Discrete-time . . . 71
4.1.5 Exercises . . . 71
4.2 Some transfer function properties . . . 73
4.2.1 Power series expansion and Markov parameters . . . 73
4.2.2 More properties* . . . 73
4.3 State space realization of transfer functions . . . 75
4.4 Realization of SISO transfer functions . . . 76
4.4.1 Controllable canonical form realization . . . 76
4.4.2 Observable canonical form realization . . . 77
4.5 Realization of MIMO systems . . . 80
4.5.1 Controllable realizations . . . 80
4.5.2 Observable realizations . . . 83
4.5.3 Alternative realizations* . . . 84

**II System Behavior and Stability** **89**

**5 Behavior of (LTI) systems : I** **91**
5.1 Initial stuff . . . 91

5.1.1 Continuous-time . . . 91
5.1.2 Discrete-time . . . 92
5.2 Scalar systems . . . 93
5.2.1 Continuous-time . . . 93
5.2.2 First order complex and second order real. . . . 94
5.2.3 DT . . . 96
5.3 Eigenvalues and eigenvectors . . . 97
5.3.1 Real $A$ . . . 100
5.3.2 MATLAB . . . 101
5.3.3 Companion matrices . . . 102
5.4 Behavior of continuous-time systems . . . 103
5.4.1 System significance of eigenvectors and eigenvalues . . . 103
5.4.2 Solutions for nondefective matrices . . . 109
5.4.3 Defective matrices and generalized eigenvectors . . . 110
5.5 Behavior of discrete-time systems . . . 115
5.5.1 System significance of eigenvectors and eigenvalues . . . 115
5.5.2 All solutions for nondefective matrices . . . 116
5.5.3 Some solutions for defective matrices . . . 116
5.6 Similarity transformations . . . 118
5.7 Hotel California . . . 123
5.7.1 Invariant subspaces . . . 123
5.7.2 More on invariant subspaces* . . . 124
5.8 Diagonalizable systems . . . 126
5.8.1 Diagonalizable matrices . . . 126
5.8.2 Diagonalizable continuous-time systems . . . 128
5.8.3 Diagonalizable DT systems . . . 130
5.9 Real Jordan form for real diagonalizable systems* . . . 132
5.9.1 CT systems . . . 133
5.10 Diagonalizable second order systems . . . 133
5.10.1 State plane portraits for CT systems . . . 134
5.11 Exercises . . . 134

**6 $e^{At}$ 141**

6.1 State transition matrix . . . 141
6.1.1 Continuous time . . . 141
6.1.2 Discrete time . . . 143
6.2 Polynomials of a square matrix . . . 144
6.2.1 Polynomials of a matrix . . . 144
6.2.2 Cayley-Hamilton Theorem . . . 145
6.3 Functions of a square matrix . . . 147
6.3.1 The matrix exponential: $e^A$ . . . 150
6.3.2 Other matrix functions . . . 150
6.4 The state transition matrix: $e^{At}$ . . . 151
6.5 Computation of $e^{At}$ . . . 152
6.5.1 MATLAB . . . 152

6.5.2 Numerical simulation . . . 152
6.5.3 Jordan form . . . 152
6.5.4 Laplace style . . . 152
6.6 Sampled-data systems . . . 154
6.7 Exercises . . . 155

**7 Stability and boundedness** **159**
7.1 Boundedness of solutions . . . 159
7.2 Stability of equilibrium states . . . 162
7.3 Asymptotic stability . . . 164
7.3.1 Global asymptotic stability . . . 164
7.3.2 Asymptotic stability . . . 164
7.4 Exponential stability . . . 166
7.5 LTI systems . . . 167
7.6 Linearization and stability . . . 170

**8 Stability and boundedness: discrete time** **173**
8.1 Boundedness of solutions . . . 173
8.2 Stability of equilibrium states . . . 174
8.3 Asymptotic stability . . . 175
8.3.1 Global asymptotic stability . . . 175
8.3.2 Asymptotic stability . . . 175
8.4 Exponential stability . . . 176
8.5 LTI systems . . . 177
8.6 Linearization and stability . . . 178

**9 Basic Lyapunov theory** **179**
9.1 Stability . . . 181
9.1.1 Locally positive definite functions . . . 181
9.1.2 A stability result . . . 183
9.2 Asymptotic stability . . . 187
9.3 Boundedness . . . 189
9.3.1 Radially unbounded functions . . . 189
9.3.2 A boundedness result . . . 189
9.4 Global asymptotic stability . . . 192
9.4.1 Positive definite functions . . . 192
9.4.2 A result on global asymptotic stability . . . 193
9.5 Exponential stability . . . 196
9.5.1 Global exponential stability . . . 196
9.5.2 Proof of theorem **??** . . . 197
9.5.3 Exponential stability . . . 197
9.5.4 A special class of GES systems . . . 199
9.5.5 Summary . . . 201

**10 Basic Lyapunov theory: discrete time*** **203**
10.1 Stability . . . 204
10.2 Asymptotic stability . . . 205
10.3 Boundedness . . . 206
10.4 Global asymptotic stability . . . 207
10.5 Exponential stability . . . 208

**11 Lyapunov theory for linear time-invariant systems** **211**
11.1 Positive and negative (semi)definite matrices . . . 211
11.1.1 Definite matrices . . . 211
11.1.2 Semi-definite matrices* . . . 212
11.2 Lyapunov theory . . . 214
11.2.1 Asymptotic stability results . . . 214
11.2.2 MATLAB. . . . 218
11.2.3 Stability results* . . . 219
11.3 Mechanical systems* . . . 221
11.4 Rates of convergence for linear systems* . . . 224
11.5 Linearization and exponential stability . . . 225

**III Input-Output Properties** **227**

**12 Input output response** **229**
12.1 Response of continuous-time systems . . . 229
12.1.1 State response and the convolution integral . . . 229
12.1.2 Output response . . . 231
12.2 Discretization and zero-order hold . . . 235
12.2.1 Input-output discretization . . . 238
12.3 Response of discrete-time systems* . . . 239
12.3.1 State response and the convolution sum . . . 239
12.3.2 Input-output response and the pulse response matrix . . . 239

**13 Observability** **243**
13.1 Observability . . . 243
13.2 Main result . . . 244
13.3 Development of main result . . . 250
13.4 The unobservable subspace . . . 253
13.5 Unobservable modes . . . 255
13.5.1 PBH Test . . . 256
13.5.2 Existence of unobservable modes* . . . 259
13.5.3 A nice transformation* . . . 260
13.6 Observability grammians* . . . 265
13.6.1 Infinite time observability grammian . . . 268
13.7 Discrete time . . . 270
13.7.1 Main observability result . . . 270

13.8 Exercises . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 271

**14 Controllability 275**
14.1 Controllability . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 275
14.2 Main controllability result . . . . . . . . . . . . . . . . . . . . . . . . . . . . 276
14.3 The controllable subspace . . . . . . . . . . . . . . . . . . . . . . . . . . . . 280
14.4 Proof of the controllable subspace lemma* . . . . . . . . . . . . . . . . . . . . 284
14.4.1 Finite time controllability grammian . . . . . . . . . . . . . . . . . . . . 284
14.4.2 Proof of the controllable subspace lemma . . . . . . . . . . . . . . . . . . 286
14.5 Uncontrollable modes . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 287
14.5.1 Eigenvectors revisited: left eigenvectors . . . . . . . . . . . . . . . . . . . 287
14.5.2 Uncontrollable eigenvalues and modes . . . . . . . . . . . . . . . . . . . . 289
14.5.3 Existence of uncontrollable modes* . . . . . . . . . . . . . . . . . . . . . 289
14.5.4 A nice transformation* . . . . . . . . . . . . . . . . . . . . . . . . . . . 291
14.6 PBH test . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 295
14.6.1 PBH test for uncontrollable eigenvalues . . . . . . . . . . . . . . . . . . . 295
14.6.2 PBH controllability test . . . . . . . . . . . . . . . . . . . . . . . . . . . 297
14.7 Other characterizations of controllability* . . . . . . . . . . . . . . . . . . . . 298
14.7.1 Infinite time controllability grammian . . . . . . . . . . . . . . . . . . . . 298
14.8 Controllability, observability, and duality . . . . . . . . . . . . . . . . . . . . . 299
14.9 Discrete-time . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 300
14.9.1 Main controllability theorem . . . . . . . . . . . . . . . . . . . . . . . . . 300

# IV Control Design 301

**15 Stabilizability and state feedback 303**
15.1 Stabilizability and state feedback . . . . . . . . . . . . . . . . . . . . . . . . . 303
15.2 Eigenvalue placement (pole placement) by state feedback . . . . . . . . . . . . 304
15.2.1 Single input systems . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 305
15.2.2 Multi-input systems . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 310
15.3 Uncontrollable eigenvalues (you can't touch them) . . . . . . . . . . . . . . . . 312
15.4 Controllable canonical form . . . . . . . . . . . . . . . . . . . . . . . . . . . 315
15.4.1 Transformation to controllable canonical form* . . . . . . . . . . . . . . . 316
15.4.2 Eigenvalue placement by state feedback* . . . . . . . . . . . . . . . . . . . 319
15.5 Discrete-time systems . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 321
15.5.1 Dead beat controllers . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 321
15.6 Stabilization of nonlinear systems . . . . . . . . . . . . . . . . . . . . . . . . 322
15.6.1 Continuous-time controllers . . . . . . . . . . . . . . . . . . . . . . . . . 322
15.6.2 Discrete-time controllers . . . . . . . . . . . . . . . . . . . . . . . . . . . 322
15.7 Appendix . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 326

**16 Detectability and observers 329**
16.1 Observers, state estimators and detectability . . . . . . . . . . . . . . . . . . . 329
16.2 Eigenvalue placement for estimation error dynamics . . . . . . . . . . . . . . . 331

16.3 Unobservable modes (you can't see them) . . . . . . . . . . . . . . . . . . . . 332
16.4 Observable canonical form* . . . . . . . . . . . . . . . . . . . . . . . . . . . 333
16.5 Discrete-time systems . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 334

**17 Climax: output feedback controllers 337**
17.1 Memoryless (static) output feedback . . . . . . . . . . . . . . . . . . . . . . 337
17.2 Dynamic output feedback . . . . . . . . . . . . . . . . . . . . . . . . . . . . 338
17.3 Observer based controllers . . . . . . . . . . . . . . . . . . . . . . . . . . . 340
17.4 Discrete-time systems . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 343
17.4.1 Memoryless output feedback . . . . . . . . . . . . . . . . . . . . . . 343
17.4.2 Dynamic output feedback . . . . . . . . . . . . . . . . . . . . . . . 343
17.4.3 Observer based controllers . . . . . . . . . . . . . . . . . . . . . . . 344

**18 Constant output tracking in the presence of constant disturbances 347**
18.1 Zeros . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 347
18.1.1 A state space characterization of zeros . . . . . . . . . . . . . . . . 347
18.2 Tracking and disturbance rejection with state feedback . . . . . . . . . . . . 350
18.2.1 Nonrobust control . . . . . . . . . . . . . . . . . . . . . . . . . . . . 352
18.2.2 Robust controller . . . . . . . . . . . . . . . . . . . . . . . . . . . . 352
18.3 Measured Output Feedback Control . . . . . . . . . . . . . . . . . . . . . . . 355
18.3.1 Observer based controllers . . . . . . . . . . . . . . . . . . . . . . . 357

**19 Lyapunov revisited* 361**
19.1 Stability, observability, and controllability . . . . . . . . . . . . . . . . . . 361
19.2 A simple stabilizing controller . . . . . . . . . . . . . . . . . . . . . . . . . . 365

**20 Performance* 367**
20.1 The $L_2$ norm . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 367
20.1.1 The rms value of a signal . . . . . . . . . . . . . . . . . . . . . . . . 367
20.1.2 The $L_2$ norm of a time-varying matrix . . . . . . . . . . . . . . . . 368
20.2 The $H_2$ norm of an LTI system . . . . . . . . . . . . . . . . . . . . . . . . 368
20.3 Computation of the $H_2$ norm . . . . . . . . . . . . . . . . . . . . . . . . . 369

**21 Linear quadratic regulators (LQR) 373**
21.1 The linear quadratic optimal control problem . . . . . . . . . . . . . . . . . 373
21.2 The algebraic Riccati equation (ARE) . . . . . . . . . . . . . . . . . . . . . 375
21.3 Linear state feedback controllers . . . . . . . . . . . . . . . . . . . . . . . . 376
21.4 Infimal cost* . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 377
21.4.1 Reduced cost stabilizing controllers. . . . . . . . . . . . . . . . . . . 377
21.4.2 Cost reducing algorithm . . . . . . . . . . . . . . . . . . . . . . . . 377
21.4.3 Infimal cost . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 380
21.5 Optimal stabilizing controllers . . . . . . . . . . . . . . . . . . . . . . . . . . 383
21.5.1 Existence of stabilizing solutions to the ARE . . . . . . . . . . . . . 385
21.6 Summary . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 386
21.7 MATLAB . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 389

21.8 Minimum energy controllers . . . 390
21.9 $H_2$ optimal controllers* . . . 392
21.10 Exercises . . . 394
21.11 Appendix . . . 395
21.11.1 Proof of fact 1 . . . 395

**22 LQG controllers** **397**
22.1 LQ observer . . . 397
22.2 LQG controllers . . . 398
22.3 $H_2$ Optimal controllers* . . . 399
22.3.1 LQ observers . . . 399
22.3.2 LQG controllers . . . 400

**23 Systems with inputs and outputs: II*** **401**
23.1 Minimal realizations . . . 401
23.2 Exercises . . . 402

**24 Time-varying systems*** **403**
24.1 Linear time-varying systems (LTVs) . . . 403
24.1.1 Linearization about time-varying trajectories . . . 403
24.2 The state transition matrix . . . 404
24.3 Stability . . . 405
24.4 Systems with inputs and outputs . . . 406
24.5 DT . . . 406

**25 Appendix A: Complex stuff** **409**
25.1 Complex numbers . . . 409
25.1.1 Complex functions . . . 411
25.2 Complex vectors and $\mathbb{C}^n$ . . . 412
25.3 Complex matrices and $\mathbb{C}^{m\times n}$ . . . 412

**26 Appendix B: Norms** **415**
26.1 The Euclidean norm . . . 415

**27 Appendix C: Some linear algebra** **417**
27.1 Linear equations . . . 417
27.2 Subspaces . . . 424
27.3 Basis . . . 425
27.3.1 Span (not Spam) . . . 425
27.3.2 Linear independence . . . 426
27.3.3 Basis . . . 431
27.4 Range and null space of a matrix . . . 432
27.4.1 Null space . . . 432
27.4.2 Range . . . 434
27.4.3 Solving linear equations revisited . . . 437
27.5 Coordinate transformations . . . 440

27.5.1 Coordinate transformations and linear maps . . . . . . . . . . . . . . . 440
27.5.2 The structure of a linear map . . . . . . . . . . . . . . . . . . . . . . 442

**28 Appendix D: Inner products** **445**
28.1 The usual scalar product on $\mathbb{R}^n$ and $\mathbb{C}^n$ . . . . . . . . . . . . . . . . . . . . . 445
28.2 Orthogonality . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 446
28.2.1 Orthonormal basis . . . . . . . . . . . . . . . . . . . . . . . . . 447
28.3 Hermitian matrices . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 449
28.4 Positive and negative (semi)definite matrices . . . . . . . . . . . . . . . . 450
28.4.1 Quadratic forms . . . . . . . . . . . . . . . . . . . . . . . . . . 450
28.4.2 Definite matrices . . . . . . . . . . . . . . . . . . . . . . . . . . 451
28.4.3 Semi-definite matrices . . . . . . . . . . . . . . . . . . . . . . . 452
28.5 Singular value decomposition . . . . . . . . . . . . . . . . . . . . . . . . 453

# Preface

These notes were developed to accompany the corresponding course in the Network Mathematics Graduate Programme at the Hamilton Institute. The purpose of the course is to introduce some basic concepts and results in the analysis and control of linear and nonlinear dynamical systems. There is a lot more material in these notes than can be covered in twenty hours. However, the extra material should prove useful to the student who wants to pursue a specific topic in more depth. Sections marked with an asterisk will not be covered. Although proofs are given for most of the results presented here, they will not be covered in the course.

# Part I

# Representation of Dynamical Systems

# Chapter 1

# Introduction

---

The purpose of this course is to introduce some basic concepts and tools which are useful in the analysis and control of dynamical systems. The concept of a dynamical system is very general; it refers to anything which evolves with time. A communication network is a dynamical system. Vehicles (aircraft, spacecraft, motorcycles, cars) are dynamical systems. Other engineering examples of dynamical systems include, metal cutting machines such as lathes and milling machines, robots, chemical plants, and electrical circuits. Even civil engineering structures such as bridges and skyscrapers are examples; think of a structure subject to strong winds or an earthquake. The concept of a dynamical system is not restricted to engineering; non-engineering examples include plants, animals, yourself, and the economy.

A system interacts with its environment via inputs and outputs. Inputs can be considered to be exerted on the system by the environment whereas outputs are exerted by the system on the environment. Inputs are usually divided into control inputs and disturbance inputs. In an aircraft, the deflection of a control surface such as an elevator would be considered a control input; a wind gust would be considered a disturbance input. An actuator is a physical device used for the implementation of a control input. Some examples of actuators are: the elevator on an aircraft; the throttle twistgrip on a motorcycle; a valve in a chemical plant.

Outputs are usually divided into performance outputs and measured outputs. Performance outputs are those outputs whose behavior or performance you are interested in, for example, heading and speed of an aircraft. The measured outputs are the outputs you actually measure, for example, speed of an aircraft. Usually, all the performance outputs are measured. Sensors are the physical devices used to obtain the measured outputs. Some examples of sensors are: altimeter and airspeed sensor on an aircraft; a pressure gauge in a chemical plant.

A fundamental concept is describing the behavior of a dynamical system is the state of the system. A precise definition will be given later. However, a main desirable property of a system state is that the current state uniquely determines all future states, that is, if one knows the current state of the system and all future inputs then, one can predict the future state of the system.

Feedback is a fundamental concept in system control. In applying a control input to a system the controller usually takes into account the behavior of the system; the controller bases its control inputs on the measured outputs of the plant (the system under control). Control based on feedback is called closed loop control. The term open loop control is usually used when one applies a control input which is pre-specified function of time; hence no feedback is involved.

The state space description of an input-output system usually involves a time variable $t$ and three sets of variables:

- State variables: $x_1, x_2, \cdots, x_n$
- Input variables: $u_1, u_2, \cdots, u_m$
- Output variables: $y_1, y_2, \cdots, y_p$

In a continuous-time system, the time-variable $t$ can be any real number. For discrete-time systems, the time-variable only takes integer values, that is, $\ldots, -2, -1, 0, 1, 2, \ldots$.

For continuous-time systems, the description usually takes the following form

$$\boxed{\begin{array}{rcl} \dot{x}_1 &=& F_1(x_1, x_2, \cdots, x_n, u_1, u_2, \cdots, u_m) \\ \dot{x}_2 &=& F_2(x_1, x_2, \cdots, x_n, u_1, u_2, \cdots, u_m) \\ &\vdots& \\ \dot{x}_n &=& F_n(x_1, x_2, \cdots, x_n, u_1, u_2, \cdots, u_m) \end{array}} \tag{1.1}$$

and

$$\boxed{\begin{array}{rcl} y_1 &=& H_1(x_1, x_2, \cdots, x_n, u_1, u_2, \cdots, u_m) \\ y_2 &=& H_2(x_1, x_2, \cdots, x_n, u_1, u_2, \cdots, u_m) \\ &\vdots& \\ y_p &=& H_p(x_1, x_2, \cdots, x_n, u_1, u_2, \cdots, u_m) \end{array}} \tag{1.2}$$

where the overhead dot indicates differentiation with respect to time. The first set of equations are called the state equations and the second set are called the output equations.

For discrete-time systems, the description usually takes the following form

$$\boxed{\begin{array}{rcl} x_1(k{+}1) &=& F_1(x_1(k), x_2(k), \cdots, x_n(k), u_1(k), u_2(k), \cdots, u_m(k)) \\ x_2(k{+}1) &=& F_2(x_1(k), x_2(k), \cdots, x_n(k), u_1(k), u_2(k), \cdots, u_m(k)) \\ &\vdots& \\ x_n(k{+}1) &=& F_n(x_1(k), x_2(k), \cdots, x_n(k), u_1(k), u_2(k), \cdots, u_m(k)) \end{array}} \tag{1.3}$$

and

$$\boxed{\begin{array}{rcl} y_1 &=& H_1(x_1, x_2, \cdots, x_n, u_1, u_2, \cdots, u_m) \\ y_2 &=& H_2(x_1, x_2, \cdots, x_n, u_1, u_2, \cdots, u_m) \\ &\vdots& \\ y_p &=& H_p(x_1, x_2, \cdots, x_n, u_1, u_2, \cdots, u_m) \end{array}} \tag{1.4}$$

The first set of equations are called the state equations and the second set are called the output equations.

## 1.1 Ingredients

*Dynamical systems*
*Linear algebra*
*Applications*
MATLAB (including Simulink)

## 1.2 Some notation

*Sets :* $s \in \mathcal{S}$
$\mathbb{Z}$ represents the set of integers
$\mathbb{R}$ represents the set of real numbers
$\mathbb{C}$ represents the set of complex numbers
*Functions:* $f : \mathcal{S} \rightarrow \mathcal{T}$

## 1.3 MATLAB

Introduce yourself to MATLAB

```
%matlab
>> lookfor
>> help
>> quit
```

Learn the representation, addition and multiplication of real and complex numbers

# Chapter 2

# State space representation of dynamical systems

In this section, we consider a bunch of simple physical systems and demonstrate how one can obtain a state space mathematical models of these systems. We begin with some linear examples.

## 2.1 Linear examples

### 2.1.1 A first example

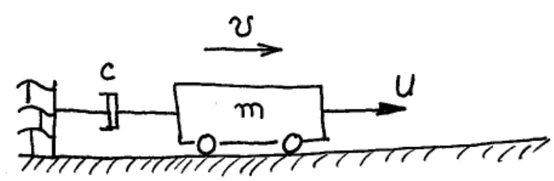



Figure 2.1: First example

Consider a small cart of mass $m$ which is constrained to move along a horizontal line. The cart is subject to viscous friction with damping coefficient $c$; it is also subject to an input force which can be represented by a real scalar variable $u(t)$ where the real scalar variable $t$ represents time. Let the real scalar variable $v(t)$ represent the velocity of the cart at time $t$; we will regard this as the output of the cart. Then, the motion of the cart can be described by the following first order ordinary differential equation (ODE):

$$m\dot{v}(t) = -cv(t) + u$$

Introducing the state $x := v$ results in

$$\begin{aligned} \dot{x} &= ax + bu \\ y &= x \end{aligned}$$

where $a := -c/m < 0$ and $b := 1/m$.

### 2.1.2 The unattached mass

Consider a small cart of mass $m$ which is constrained to move without friction along a horizontal line. It is also subject to an input force which can be represented by a real scalar variable $u(t)$. Let $q(t)$ be the horizontal displacement of the cart from a fixed point on its line of motion; we will regard $y = q$ as the output of the cart.

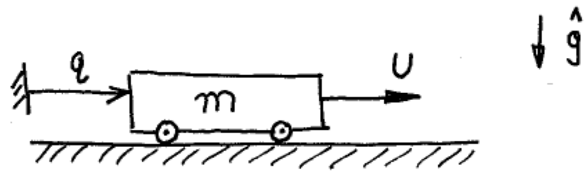



Figure 2.2: The unattached mass

Application of Newton's second law to the unattached mass illustrated in Figure 2.2 results in

$$m\ddot{q} = u$$

Introducing the state variables,

$$x_1 := q \qquad \text{and} \qquad x_2 := \dot{q},$$

yields the following state space description:

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= u/m \\ y &= x_1 \end{aligned}$$

### 2.1.3 Spring-mass-damper

Consider a system which can be modeled as a simple mechanical system consisting of a body of mass $m$ attached to a base via a linear spring of spring constant $k$ and linear dashpot with damping coefficient $c$. The base is subject to an acceleration $u$ which we will regard as the input to the system. As output $y$, we will consider the force transmitted to the mass from the spring-damper combination. Letting $q$ be the deflection of the spring, the motion of the system can be described by

$$m(\ddot{q} + u) = -c\dot{q} - kq - mg$$

and $y = -kq - c\dot{q}$ where $g$ is the gravitational acceleration constant. Introducing $x_1 := q$ and $x_2 := \dot{q}$ results in the following state space description:

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -(k/m)x_1 - (c/m)x_2 - u - g \\ y &= -kx_1 - cx_2\,. \end{aligned}$$

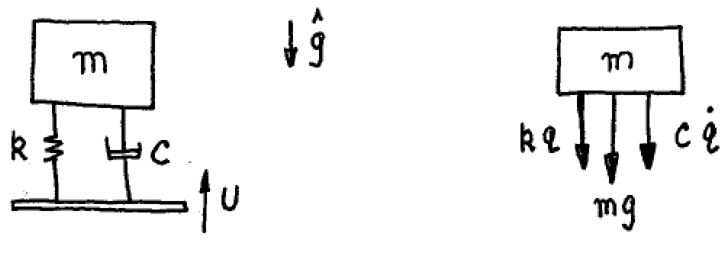



Figure 2.3: Spring-mass-damper with exciting base

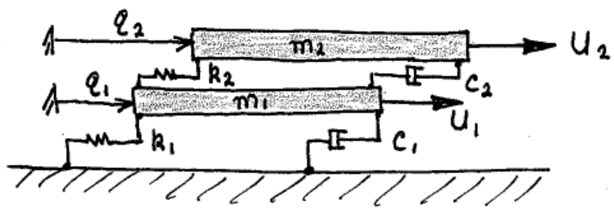



Figure 2.4: A simple structure

### 2.1.4 A simple structure

Consider a structure consisting of two floors. The scalar variables $q_1$ and $q_2$ represent the lateral displacement of the floors from their nominal positions. Application of Newton's second law to each floor results in

$$
\begin{array}{ccccccccccc}
m_1\ddot{q}_1 & + & (c_1+c_2)\dot{q}_1 & + & (k_1+k_2)q_1 & - & c_2\dot{q}_2 & - & k_2q_2 & = & u_1 \\
m_2\ddot{q}_2 & - & c_2\dot{q}_1 & - & k_2q_1 & + & c_2\dot{q}_2 & + & k_2q_2 & = & u_2
\end{array}
$$

Here $u_2$ is a control input resulting from a force applied to the second floor and $u_1$ is a disturbance input resulting from a force applied to the first floor. We have not considered any outputs here.

## 2.2 Nonlinear examples

### 2.2.1 A first nonlinear system

Recall the first example, and suppose now that the friction force on the cart is due to Coulomb (or dry) friction with coefficient of friction $\mu > 0$. Then we have

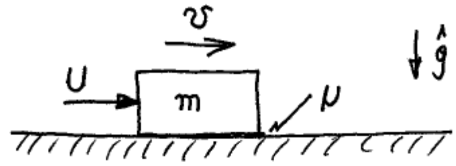



Figure 2.5: A first nonlinear system

$$m\dot{v} \;=\; -\mu m g \,\mathrm{sgm}\,(v) + u$$

where

$$\mathrm{sgm}\,(v) := \begin{cases} -1 & \text{if} \quad v < 0 \\ 0 & \text{if} \quad v = 0 \\ 1 & \text{if} \quad v > 0 \end{cases}$$

and $g$ is the gravitational acceleration constant of the planet on which the block resides. With $x := v$ we obtain

$$\begin{aligned} \dot{x} &= -\alpha \,\mathrm{sgm}\,(x) + bu \\ y &= x \end{aligned}$$

and $\alpha := \mu g$ and $b = 1/m$.

### 2.2.2 Planar pendulum

Consider a planar rigid body which is constrained to rotate about a horizontal axis which is perpendicular to the body. Here the input $u$ is a torque applied to the pendulum and the output is the angle $\theta$ that the pendulum makes with a vertical line.

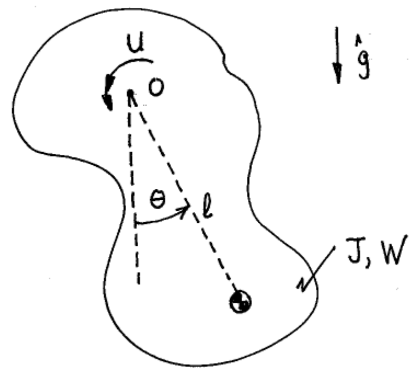



Figure 2.6: Simple pendulum

The motion of this system is governed by

$$J\ddot{\theta} + Wl\sin\theta = u$$

where $J > 0$ is the moment of inertia of the body about its axis of rotation, $W > 0$ is the weight of the body and $l$ is the distance between the mass center of the body and the axis of rotation. Introducing state variables

$$x_1 := \theta \qquad \text{and} \qquad x_2 := \dot{\theta}$$

results in the following nonlinear state space description:

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -a\sin x_1 + b_2 u \\ y &= x_1 \end{aligned}$$

where $a := Wl/J > 0$ and $b_2 = 1/J$.

### 2.2.3 Attitude dynamics of a rigid body

The equations describing the rotational motion of a rigid body are given by Euler's equations of motion. If we choose the axes of a body-fixed reference frame along the principal axes of inertia of the rigid body with origin at the center of mass, Euler's equations of motion take the simplified form

$$\begin{aligned} I_1\dot{\omega}_1 &= (I_2 - I_3)\omega_2\omega_3 \\ I_2\dot{\omega}_2 &= (I_3 - I_1)\omega_3\omega_1 \\ I_3\dot{\omega}_3 &= (I_1 - I_2)\omega_1\omega_2 \end{aligned}$$

Figure 2.7: Attitude dynamics of a rigid body

where $\omega_1, \omega_2, \omega_3$ denote the components of the body angular velocity vector with respect to the body principal axes, and the positive scalars $I_1, I_2, I_3$ are the principal moments of inertia of the body with respect to its mass center.

### 2.2.4 Body in central force motion

Figure 2.8: Body in central force motion

$$\begin{array}{ccccccl} \ddot{r} & - & r\omega^2 & + & g(r) & = 0 \\ r\dot{\omega} & + & 2\dot{r}\omega & & & = 0 \end{array}$$

For the simplest situation in orbit mechanics ( a “satellite” orbiting YFHB)

$$g(r) = \mu / r^2 \qquad \mu = GM$$

where $G$ is the universal constant of gravitation and $M$ is the mass of YFHB.

### 2.2.5 Double pendulum on cart

Consider the the double pendulum on a cart illustrated in Figure 2.9. The motion of this system can be described by the cart displacement $y$ and the two pendulum angles $\theta_1$, $\theta_2$. The input $u$ is a force applied to the cart. Application of your favorite laws of mechanics results in the following equations of motion:

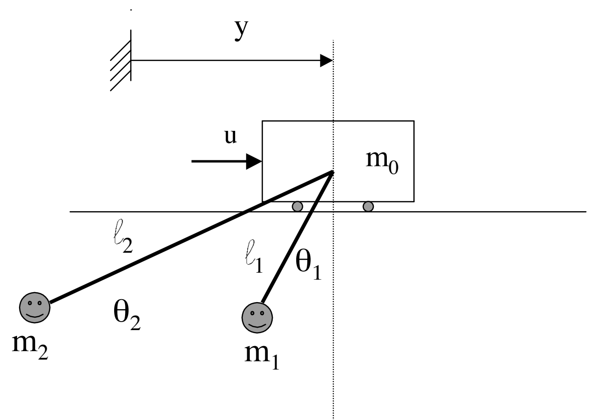



Figure 2.9: Double pendulum on cart

$$
\boxed{\begin{array}{rcrcl}
(m_0+m_1+m_2)\ddot{y} - m_1 l_1 \cos\theta_1\,\ddot{\theta}_1 - m_2 l_2 \cos\theta_2\,\ddot{\theta}_2 & + & m_1 l_1 \sin\theta_1 \dot{\theta}_1^2 + m_2 l_2 \sin\theta_2 \dot{\theta}_2^2 & = & u \\
-m_1 l_1 \cos\theta_1\,\ddot{y} + m_1 l_1^2 \ddot{\theta}_1 & + & m_1 l_1 g \sin\theta_1 & = & 0 \\
-m_2 l_2 \cos\theta_2\,\ddot{y} + m_2 l_2^2 \ddot{\theta}_2 & + & m_2 l_2 g \sin\theta_2 & = & 0
\end{array}}
$$

### 2.2.6 Two-link robotic manipulator

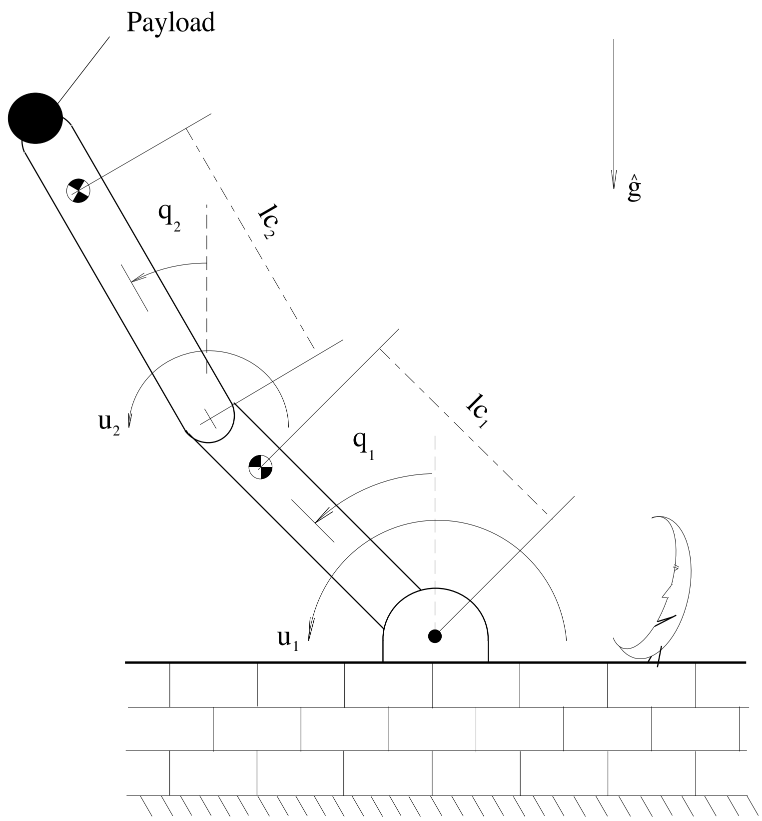



Figure 2.10: A simplified model of a two link manipulator

The coordinates $q_1$ and $q_2$ denote the angular location of the first and second links relative to the local vertical, respectively. The second link includes a payload located at its end. The masses of the first and the second links are $m_1$ and $m_2$, respectively. The moments of inertia of the first and the second links about their centers of mass are $I_1$ and $I_2$, respectively. The locations of the center of mass of links one and two are determined by $lc_1$ and $lc_2$, respectively; $l_1$ is the length of link 1. The equations of motion for the two arms are described by:

$$\begin{array}{ccccccccc} m_{11}\ddot{q}_1 & + & m_{12}\cos(q_1-q_2)\ddot{q}_2 & + & c_1\sin(q_1-q_2)\dot{q}_2^2 & + & g_1\sin(q_1) & = & u_1 \\ m_{21}\cos(q_1-q_2)\ddot{q}_1 & + & m_{22}\ddot{q}_2 & + & c_2\sin(q_1-q_2)\dot{q}_1^2 & + & g_2\sin(q_2) & = & u_2 \end{array}$$

where

$$\begin{array}{lll} m_{11} = I_1 + m_1 lc_1^2 + m_2 l_1^2\,, & m_{12} = m_{21} = m_2 l_1 lc_2\,, & m_{22} = I_2 + m_2 lc_2^2 \\ g_1 = -(m_1 lc_1 + m_2 l_1)g\,, & g_2 = -m_2 lc_2 g & \\ c_1 = m_2 l_1 lc_2\,, & c_2 = -m_2 l_1 lc_2 & \end{array}$$

## 2.3 Discrete-time examples

All the examples considered so far are continuous-time examples. Now we consider discrete-time examples; here the time variable $k$ is an integer, that is, $\cdots, -2, -1, 0, 1, 2, \cdots$. Sometimes, a discrete-time system results from the discretization of a continuous-time system (see first example) and sometimes it arises naturally (see second example).

### 2.3.1 The discrete unattached mass

Here we consider a discretization of the unattached mass of Section 2.1.2 described by

$$m\ddot{q} = u\,.$$

Suppose that the input to this system is given by a zero order hold as follows:

$$u(t) = u_d(k) \qquad \text{for} \quad kT \leq t < (k{+}1)T$$

where $k$ is an integer, that is, $k = \ldots, -2, -1, 0, 1, 2\ldots$ and $T > 0$ is some specified time interval; we call it the sampling time. Thus the input to the system is constant over each sampling interval $[kT,\,(k{+}1)T)$.

Suppose we sample the state of the system at integer multiples of the sampling time and let

$$x_1(k) = q(kT) \qquad x_2(k) := \dot{q}(kT)$$

for every integer $k$. Note that, for any $k$ and any $t$ with $kT \leq t \leq (k{+}1)T$ we have

$$\dot{q}(t) = \dot{q}(kT) + \int_{kT}^{t} \ddot{q}(\tau)\,d\tau = x_2(k) + \int_{kT}^{t} \frac{1}{m}u(k)\,dt = x_2(k) + \frac{t-kT}{m}u(k)\,.$$

Hence,

$$x_2(k{+}1) = \dot{q}((k{+}1)T) = x_2(k) + \frac{T}{m}u(k)$$

and

$$\begin{aligned} x_1(k{+}1) &= q((k{+}1)T) = q(kT) + \int_{kT}^{(k+1)T} \dot{q}(t)\,dt = x_2(k) + \int_{kT}^{(k+1)T} x_2(k) + \frac{t-kT}{m}u(k)\,dt \\ &= x_1(k) + Tx_2(k) + \frac{T^2}{2m}u(k)\,. \end{aligned}$$

Thus, the sampled system is described by the two first order linear difference equations:

$$\boxed{\begin{aligned} x_1(k{+}1) &= x_1(k) + Tx_2(k) + \tfrac{T^2}{2m}u(k) \\ x_2(k{+}1) &= x_2(k) + \tfrac{T}{m}u(k) \end{aligned}}$$

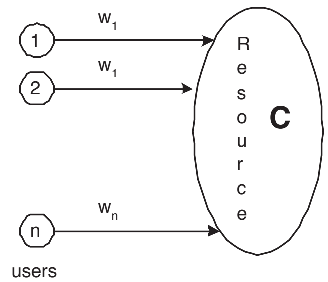
and the output equation:

$$\boxed{y(k) = x_1(k)}$$

This is a discrete-time state space description.

### 2.3.2 Additive increase multiplicative decrease (AIMD) algorithm for resource allocation

Consider a resource such as a communications router which is serving $n$ users and has a




maximum capacity of $c$. Ideally, we would like the resource to be fully utilized at all times. Each user determines how much capacity it will request without knowledge of the usage of the other users. How does each user determine its use of the resource? One approach is to use an additive increase multiplicative decrease (AIMD) algorithm. In this approach, each user increases its use of the resource linearly with time (additive increase phase) until it is notified that the resource has reached maximum capacity. It then (instantaneously) reduces its usage to a fraction of its usage at notification (multiplicative decrease phase). It again starts to increase its usage linearly with time.

To describe AIMD, let $w_i(t) \geq 0$ represent the $i$-th users share of the resource capacity at time $t$. Each user increases its share linearly with time until the resource reaches maximum capacity $c$. When maximum capacity is reached, we refer to this as a congestion event. Suppose we number the congestion events consecutively with index $k = 1, 2, \cdots$ and let $t_k$ denote the time at which the $k$-th congestion event occurs. Then.

$$w_1(t_k) + w_2(t_k) + \cdots + w_n(t_k) = c$$

Suppose that immediately after a congestion event, each user $i$ decreases its share of the resource to $\beta_i$ times its share at congestion where

$$0 < \beta_i < 1\,.$$

This is the multiplicative decrease phase of the algorithm. If $t_k^+$ represents a time immediately after the $k$-th congestion event then,

$$w_i(t_k^+) = \beta_i w_i(t_k)$$

Following the multiplicative decrease of its share after a congestion event, each user $i$ increases its share linearly in time at a rate $\alpha_i > 0$ until congestion occurs again, that is,

$$w_i(t) = \beta_i w_i(t_k) + \alpha_i(t - t_k)$$

for $t_k < t \leq t_{k+1}$. This is the additive increase phase of the algorithm.

Let $T_k$ be the time between the $k$-th congestion event and congestion event $k+1$. Then $T_k = t_{k+1} - t_k$ and

$$w_i(t_{k+1}) = \beta_i w_i(t_k) + \alpha_i T_k$$

Abusing notation and letting $w_i(k) = w_i(t_k)$, the AIMD algorithm is described by

$$w_i(k+1) = \beta_i w_i(k) + \alpha_i T_k \qquad \text{for} \quad i = 1, \cdots, n \tag{2.1}$$

where the inter-congestion time $T_k$ is determined by the congestion condition

$$w_1(k+1) + \cdots + w_n(k+1) = c$$

Using this, (2.1) and

$$w_1(k) + \cdots + w_n(k) = c$$

results in

$$(1-\beta_1)w_1(k) + \cdots + (1-\beta_n)w_n(k) = (\alpha_1 + \cdots + \alpha_n)T_k$$

Solving the above equation for $T_k$ and substitution into (2.1) results in

$$\boxed{w_i(k+1) = \beta_i w_i(k) + \gamma_i(1-\beta_1)w_1(k) + \cdots + \gamma_i(1-\beta_n)w_n(k)} \tag{2.2}$$

for $i = 1, \cdots, n.$ where

$$\gamma_i := \frac{\alpha_i}{\alpha_1 + \cdots + \alpha_n}\,.$$

Note that

$$\gamma_1 + \cdots + \gamma_n = 1\,. \tag{2.3}$$

Not also that the above system is described by $n$ first order difference equations.

## 2.4 General representation

### 2.4.1 Continuous-time

With the exception of the discrete-time examples, all of the preceding systems can be described by a bunch of first order ordinary differential equations of the form

$$\boxed{\begin{array}{rcl} \dot{x}_1 &=& F_1(x_1, x_2, \ldots, x_n, u_1, u_2, \ldots, u_m) \\ \dot{x}_2 &=& F_2(x_1, x_2, \ldots, x_n, u_1, u_2, \ldots, u_m) \\ &\vdots& \\ \dot{x}_n &=& F_n(x_1, x_2, \ldots, x_n, u_1, u_2, \ldots, u_m) \end{array}}$$

Such a description of a dynamical system is called a state space description; the real scalar variables, $x_i(t)$ are called the state variables; the real scalar variables, $u_i(t)$ are called the input variables and the real scalar variable $t$ is called the time variable. If the system has outputs, they are described by

$$\boxed{\begin{array}{rcl} y_1 &=& H_1(x_1, \ldots, x_n,\, u_1, \ldots, u_m) \\ y_2 &=& H_2(x_1, \ldots, x_n,\, u_1, \ldots, u_m) \\ &\vdots& \\ y_p &=& H_p(x_1, \ldots, x_n,\, u_1, \ldots, u_m) \end{array}}$$

where the real scalar variables, $y_i(t)$ are called the output variables

When a system has no inputs or the inputs are fixed at some constant values, the system is described by

$$\boxed{\begin{array}{rcl} \dot{x}_1 &=& f_1(x_1, x_2, \ldots, x_n) \\ \dot{x}_2 &=& f_2(x_1, x_2, \ldots, x_n) \\ &\vdots& \\ \dot{x}_n &=& f_n(x_1, x_2, \ldots, x_n) \end{array}}$$

**Limitations of the above description.** The above description cannot handle

Systems with delays.
Systems described by partial differential equations.

### Higher order ODE descriptions

**Single equation.** (Recall the spring-mass-damper system.) Consider a dynamical system described by a single $n^{th}$- order differential equation of the form

$$\boxed{F(q, \dot{q}, \ldots, q^{(n)}, u) = 0}$$

where $q(t)$ is a real scalar and $q^{(n)} := \frac{d^n q}{dt^n}$. To obtain an equivalent state space description, we proceed as follows.

(a) First solve for the highest order derivative $q^{(n)}$ of $q$ as a function of $q, \dot{q}, \ldots, q^{(n-1)}$ and $u$ to obtain something like:

$$q^{(n)} = a(q, \dot{q}, \ldots, q^{(n-1)}, u)$$

(b) Now introduce state variables,

$$\begin{aligned} x_1 &:= q \\ x_2 &:= \dot{q} \\ &\vdots \\ x_n &:= q^{(n-1)} \end{aligned}$$

to obtain the following state space description:

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= x_3 \\ &\vdots \\ \dot{x}_{n-1} &= x_n \\ \dot{x}_n &= a(x_1, x_2, \ldots, x_n, u) \end{aligned}$$

**Multiple equations.** (Recall the simple structure and the pendulum cart system.) Consider a dynamical system described by $N$ scalar differential equations in $N$ scalar variables:

$$\begin{aligned}
F_1(q_1, \dot{q}_1, \dots, q_1^{(n_1)},\ q_2, \dot{q}_2, \dots, q_2^{(n_2)},\ \dots,\ q_N, \dot{q}_N, \dots,\ q_N^{(n_N)},\ u_1, u_2, \dots, u_m\ ) &= 0 \\
F_2(q_1, \dot{q}_1, \dots, q_1^{(n_1)},\ q_2, \dot{q}_2, \dots, q_2^{(n_2)},\ \dots,\ q_N, \dot{q}_N, \dots,\ q_N^{(n_N)},\ u_1, u_2, \dots, u_m\ ) &= 0 \\
&\vdots \\
F_N(q_1, \dot{q}_1, \dots, q_1^{(n_1)},\ q_2, \dot{q}_2, \dots, q_2^{(n_2)},\ \dots,\ q_N, \dot{q}_N, \dots,\ q_N^{(n_N)},\ u_1, u_2, \dots, u_m\ ) &= 0
\end{aligned}$$

where $t, q_1(t), q_2(t), \dots, q_N(t)$ are real scalars. Note that $q_i^{(n_i)}$ is the highest order derivative of $q_i$ which appears in the above equations. To obtain an equivalent state space description, we proceed as follows.

(a) Solve for the highest order derivatives, $q_1^{(n_1)},\ q_2^{(n_2)},\ \dots,\ q_N^{(n_N)}$ of $q_1, \cdots, q_N$ which appear in the differential equations to obtain something like:

$$\begin{aligned}
q_1^{(n_1)} &= a_1(\, q_1, \dot{q}_1, \dots q_1^{(n_1-1)},\ q_2, \dot{q}_2, \dots, q_2^{(n_2-1)},\ \dots,\ q_N, \dot{q}_N, \dots, q_N^{(n_N-1)},\ u_1, u_2, \dots, u_m\,) \\
q_2^{(n_2)} &= a_2(\, q_1, \dot{q}_1, \dots q_1^{(n_1-1)},\ q_2, \dot{q}_2, \dots, q_2^{(n_2-1)},\ \dots,\ q_N, \dot{q}_N, \dots, q_N^{(n_N-1)},\ u_1, u_2, \dots, u_m\,) \\
&\vdots \\
q_N^{(n_N)} &= a_N(\, q_1, \dot{q}_1, \dots q_1^{(n_1-1)},\ q_2, \dot{q}_2, \dots, q_2^{(n_2-1)},\ \dots,\ q_N, \dot{q}_N, \dots, q_N^{(n_N-1)},\ u_1, u_2, \dots, u_m\,)
\end{aligned}$$

If one cannot uniquely solve for the above derivatives, then one cannot obtain a unique state space description

(b) As state variables, consider each $q_i$ variable and its derivatives up to but not including its highest order derivative. One way to do this is as follows. Let

$$\begin{array}{llll}
x_1 := q_1 & x_2 := \dot{q}_1 & \dots & x_{n_1} := q_1^{(n_1-1)} \\
x_{n_1+1} := q_2, & x_{n_1+2} := \dot{q}_2 & \dots & x_{n_1+n_2} := q_2^{(n_2-1)} \\
 & & \vdots & \\
x_{n_1+\ldots+n_{N-1}+1} := q_N & x_{n_1+\ldots+n_{N-1}+2} := \dot{q}_N & \dots & x_n := q_N^{(n_N-1)}
\end{array}$$

where

$$n := n_1 + n_2 + \ldots + n_N$$

to obtain

$$\begin{aligned}
\dot{x}_1 &= x_2 \\
\dot{x}_2 &= x_3 \\
&\vdots \\
\dot{x}_{n_1-1} &= x_{n_1} \\
\dot{x}_{n_1} &= a_1(x_1, x_2, \ldots, x_n,\, u_1, u_2, \ldots, u_m) \\
\dot{x}_{n_1+1} &= x_{n_1+2} \\
&\vdots \\
\dot{x}_{n_1+n_2} &= a_2(x_1, x_2, \ldots, x_n,\, u_1, u_2, \ldots, u_m) \\
&\vdots \\
\dot{x}_n &= a_N(x_1, x_2, \ldots, x_n,\, u_1, u_2, \ldots, u_m)
\end{aligned}$$

**Example 1**

$$\begin{aligned}
\ddot{q}_1 + \dot{q}_2 + 2q_1 &= 0 \\
-\ddot{q}_1 + \dot{q}_1 + \dot{q}_2 + 4q_2 &= 0
\end{aligned}$$

The highest order derivatives of $q_1$ and $q_2$ appearing in these equations are $\ddot{q}_1$ and $\dot{q}_2$, respectively.

Solving for $\ddot{q}_1$ and $\dot{q}_2$, we obtain

$$\begin{aligned}
\ddot{q}_1 &= -q_1 + \frac{1}{2}\dot{q}_1 + 2q_2 \\
\dot{q}_2 &= -q_1 - \frac{1}{2}\dot{q}_1 - 2q_2\,.
\end{aligned}$$

Introducing state variables $x_1 := q_1$, $x_2 := \dot{q}_1$, and $x_3 = q_2$ we obtain the following state space description:

$$\begin{aligned}
\dot{x}_1 &= x_2 \\
\dot{x}_2 &= -x_1 + \frac{1}{2}x_2 + 2x_3 \\
\dot{x}_3 &= -x_1 - \frac{1}{2}x_2 - 2x_3
\end{aligned}$$

### 2.4.2 Discrete-time

Recall the AIMD algorithm of Section 2.3.2. The general discrete-time system is described by a bunch of first order difference equations of the form:

$$\begin{aligned}
x_1(k+1) &= F_1(x_1(k), x_2(k), \ldots, x_n(k),\, u_1(k), u_2(k), \ldots, u_m(k)) \\
x_2(k+1) &= F_2(x_1(k), x_2(k), \ldots, x_n(k),\, u_1(k), u_2(k), \ldots, u_m(k)) \\
&\vdots \\
x_n(k+1) &= F_n(x_1(k), x_2(k), \ldots, x_n(k),\, u_1(k), u_2(k), \ldots, u_m(k))
\end{aligned}$$

The real scalar variables, $x_i(k)$, $i = 1, 2, \ldots, n$ are called the state variables; the real scalar variables, $u_i(k)$, $i = 1, 2, \ldots, m$ are called the input variables and the integer variable $k$ is called the time variable. If the system has outputs, they are described by

$$\boxed{\begin{array}{rcl} y_1 &=& H_1(x_1, \ldots, x_n,\, u_1, \ldots, u_m) \\ y_2 &=& H_2(x_1, \ldots, x_n,\, u_1, \ldots, u_m) \\ &\vdots& \\ y_p &=& H_p(x_1, \ldots, x_n,\, u_1, \ldots, u_m) \end{array}}$$

where the real scalar variables, $y_i(k)$, $i = 1, 2, \ldots, p$ are called the output variables

When a system has no inputs or the inputs are fixed at some constant values, the system is described by

$$\boxed{\begin{array}{rcl} x_1(k+1) &=& f_1(x_1(k), x_2(k), \ldots, x_n(k)) \\ x_2(k+1) &=& f_2(x_1(k), x_2(k), \ldots, x_n(k)) \\ &\vdots& \\ x_n(k+1) &=& f_n(x_1(k), x_2(k), \ldots, x_n(k)) \end{array}}$$

### Higher order difference equation descriptions

One converts higher order difference equation descriptions into state space descriptions in a similar manner to that of converting higher order ODE descriptions into state space descriptions; see Example 2.3.1.

### 2.4.3 Exercises

**Exercise 1** Consider a system described by a single $n^{th}$-order linear differential equation of the form

$$q^{(n)} + a_{n-1}q^{(n-1)} + \dots a_1\dot{q} + a_0 q = 0$$

where $q(t) \in \mathbb{R}$ and $q^{(n)} := \frac{d^n q}{dt^n}$. By appropriate definition of state variables, obtain a first order state space description of this system.

**Exercise 2** By appropriate definition of state variables, obtain a first order state space description of the following systems.

(i) The 'simple structure'

(ii) The 'body in central force motion'

(iii)

$$\begin{aligned} \ddot{q}_1 + q_1 + 2\dot{q}_2 = 0 \\ \ddot{q}_1 + \dot{q}_2 + q_2 = 0 \end{aligned}$$

where $q_1, q_2$ are real scalars.

**Exercise 3** By appropriate definition of state variables, obtain a first order state space description of the following systems where $q_1$ and $q_2$ are real scalars.

(i)

$$\begin{aligned} 2\ddot{q}_1 + \ddot{q}_2 + \sin q_1 &= 0 \\ \ddot{q}_1 + 2\ddot{q}_2 + \sin q_2 &= 0 \end{aligned}$$

(ii)

$$\begin{aligned} \ddot{q}_1 + \dot{q}_2 + q_1^3 = 0 \\ \dot{q}_1 + \dot{q}_2 + q_2^3 = 0 \end{aligned}$$

**Exercise 4** By appropriate definition of state variables, obtain a first order state space description of the following systems where $q_1, q_2$ are real scalars.

(i)

$$\begin{aligned} 2\ddot{q}_1 + 3\ddot{q}_2 + \dot{q}_1 - q_2 &= 0 \\ \ddot{q}_1 + 2\ddot{q}_2 - \dot{q}_2 + q_1 &= 0 \end{aligned}$$

(ii)

$$\dddot{q}_2 + 2\dot{q}_1 + q_2 - q_1 = 0$$
$$\ddot{q}_2 + \dot{q}_1 - q_2 + q_1 = 0$$

**Exercise 5** By appropriate definition of state variables, obtain a first order state space description of the following systems where $q_1, q_2$ are real scalars.

(i)

$$\begin{aligned} 3\ddot{q}_1 - \ddot{q}_2 + 2\dot{q}_1 + 4q_2 &= 0 \\ -\ddot{q}_1 + 3\ddot{q}_2 + 2\dot{q}_1 - 4q_2 &= 0 \end{aligned}$$

(ii)

$$3\ddot{q}_1 - \dot{q}_2 + 4q_2 - 4q_1 = 0$$
$$-\ddot{q}_1 + 3\dot{q}_2 - 4q_2 + 4q_1 = 0$$

**Exercise 6** Obtain a state-space description of the following single-input single-output system with input $u$ and output $y$.

$$\begin{aligned} \ddot{q}_1 + \dot{q}_2 + q_1 &= 0 \\ \ddot{q}_1 - \dot{q}_2 + q_2 &= u \\ y &= \ddot{q}_1 \end{aligned}$$

**Exercise 7** Consider the second order difference equation

$$q(k+2) = q(k+1) + q(k)$$

Obtain the solution corresponding to

$$q(0) = 1, \quad q(1) = 1$$

and $k = 0, 1, \ldots, 10$. Obtain a state space description of this system.

**Exercise 8** Obtain a state space representation of the following system:

$$q(k+n) + a_{n-1}q(k+n-1) + \ldots + a_1 q(k+1) + a_0 q(k) = 0$$

where $q(k) \in \mathbb{R}$.

## 2.5 Vectors

When dealing with systems containing many variables, the introduction of vectors for system description can considerably simplify system analysis and control design.

### 2.5.1 Vector spaces and $\mathbb{R}^n$

A scalar is a real or a complex number. The symbols $\mathbb{R}$ and $\mathbb{C}$ represent the set of real and complex numbers, respectively. In this section, all the definitions and results are given for real scalars. However, they also hold for complex scalars; to get the results for complex scalars, simply replace 'real' with 'complex' and $\mathbb{R}$ with $\mathbb{C}$.

Consider any positive integer $n$. A real $n$-vector $x$ is an ordered $n$-tuple of real numbers, $x_1, x_2, \ldots, x_n$. This is usually written as $x = (x_1\,, x_2\,, \cdots\,, x_n)$ or

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \qquad \text{or} \qquad x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

The real numbers $x_1, x_2, \ldots, x_n$ are called the scalar components of $x$ and $x_i$ is called the i-th component. The symbol $\mathbb{R}^n$ represents the set of ordered $n$-tuples of real numbers.

#### Addition

The addition of any two real $n$-vectors $x$ and $y$ yields another real $n$-vector $x + y$ which is defined by:

$$x + y \;=\; \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} := \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix}$$

Zero element of $\mathbb{R}^n$:

$$0 := \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Note that we are using the same symbol, 0, for a zero scalar and a zero vector.

The negative of a vector:

$$-x := \begin{pmatrix} -x_1 \\ -x_2 \\ \vdots \\ -x_n \end{pmatrix}$$

**Properties of addition**

(a) (Commutative). For each pair $x, y$ in $\mathbb{R}^n$,

$$x + y = y + x$$

(b) (Associative). For each $x$, $y$, $z$ in $\mathbb{R}^n$,

$$(x + y) + z = x + (y + z)$$

(c) There is an element 0 in $\mathbb{R}^n$ such that for every $x$ in $\mathbb{R}^n$,

$$x + 0 = x$$

(d) For each $x$ in $\mathbb{R}^n$, there is an element $-x$ in $\mathbb{R}^n$ such that

$$x + (-x) = 0$$

**Scalar multiplication.**

The multiplication of an element $x$ of $\mathbb{R}^n$ by a real scalar $\alpha$ yields an element of $\mathbb{R}^n$ and is defined by:

$$\alpha x \;=\; \alpha \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \;:=\; \begin{pmatrix} \alpha x_1 \\ \alpha x_2 \\ \vdots \\ \alpha x_n \end{pmatrix}$$

**Properties of scalar multiplication**

(a) For each scalar $\alpha$ and pair $x, y$ in $\mathbb{R}^n$

$$\alpha(x + y) = \alpha x + \alpha y$$

(b) For each pair of scalars $\alpha, \beta$ and $x$ in $\mathbb{R}^n$,

$$(\alpha + \beta)x = \alpha x + \beta x$$

(c) For each pair of scalars $\alpha, \beta$, and $x$ in $\mathbb{R}^n$,

$$\alpha(\beta x) = (\alpha\beta)x$$

(d) For each $x$ in $\mathbb{R}^n$,

$$1x = x$$

**Vector space**

Consider *any* set $\mathcal{V}$ equipped with an addition operation and a scalar multiplication operation. Suppose the addition operation assigns to each pair of elements $x, y$ in $\mathcal{V}$ a unique element $x + y$ in $\mathcal{V}$ and it satisfies the above four properties of addition (with $\mathbb{R}^n$ replaced by $\mathcal{V}$). Suppose the scalar multiplication operation assigns to each scalar $\alpha$ and element $x$ in $\mathcal{V}$ a unique element $\alpha x$ in $\mathcal{V}$ and it satisfies the above four properties of scalar multiplication (with $\mathbb{R}^n$ replaced by $\mathcal{V}$). Then this set (along with its addition and scalar multiplication) is called a vector space. Thus $\mathbb{R}^n$ equipped with its definitions of addition and scalar multiplication is a specific example of a vector space. We shall meet other examples of vectors spaces later. An element $x$ of a vector space is called a vector. A vector space with real (complex) scalars is called a real (complex) vector space.

As a more abstract example of a vector space, let $\mathcal{V}$ be the set of continuous real-valued functions which are defined on the interval $[0, \infty)$; thus, an element $x$ of $\mathcal{V}$ is a function which maps $[0, \infty)$ into $\mathbb{R}$, that is, $x : [0, \infty) \longrightarrow \mathbb{R}$. Addition and scalar multiplication are defined as follows. Suppose $x$ and $y$ are any two elements of $\mathcal{V}$ and $\alpha$ is any scalar, then

$$\begin{aligned} (x+y)(t) &= x(t) + y(t) \\ (\alpha x)(t) &= \alpha x(t) \end{aligned}$$

*Subtraction* in a vector space is defined by:

$$x - y := x + (-y)$$

Hence in $\mathbb{R}^n$,

$$x - y = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} - \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 - y_1 \\ x_2 - y_2 \\ \vdots \\ x_n - y_n \end{pmatrix}$$

### 2.5.2 $\mathbb{R}^2$ and pictures

An element of $\mathbb{R}^2$ can be represented in a plane by a point or a directed line segment.

### 2.5.3 Derivatives

Suppose $x(\cdot)$ is a function of a real variable $t$ where $x(t)$ is an $n$-vector. Then

$$\dot{x} := \frac{dx}{dt} := \begin{pmatrix} \frac{dx_1}{dt} \\ \\ \frac{dx_2}{dt} \\ \\ \vdots \\ \\ \frac{dx_n}{dt} \end{pmatrix} = \begin{pmatrix} \dot{x}_1 \\ \\ \dot{x}_2 \\ \\ \vdots \\ \\ \dot{x}_n \end{pmatrix}$$

### 2.5.4 MATLAB

Representation of real and complex vectors
Addition and substraction of vectors
Multiplication of a vector by a scalar

## 2.6 Vector representation of dynamical systems

Recall the general descriptions (continuous and discrete) of dynamical systems given in Section 2.4. We define the state (vector) $x$ as the vector with components, $x_1, x_2, \cdots, x_n$, that is,

$$x := \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

If the system has inputs, we define the input (vector) $u$ as the vector with components, $u_1, u_2, \cdots, u_m$, that is,

$$u := \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_m \end{pmatrix}.$$

If the system has outputs, we define the output (vector) $y$ as the vector with components, $y_1, y_2, \cdots, y_p$, that is,

$$y := \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{pmatrix}.$$

We introduce the vector valued functions $F$ and $H$ defined by

$$F(x,u) := \begin{pmatrix} F_1(x_1, x_2, \ldots, x_n,\, u_1, u_2, \cdots, u_m) \\ F_2(x_1, x_2, \ldots, x_n,\, u_1, u_2, \cdots, u_m) \\ \vdots \\ F_n(x_1, x_2, \ldots, x_n,\, u_1, u_2, \cdots, u_m) \end{pmatrix}$$

and

$$H(x,u) := \begin{pmatrix} H_1(x_1, x_2, \ldots, x_n,\, u_1, u_2, \cdots, u_m) \\ H_2(x_1, x_2, \ldots, x_n,\, u_1, u_2, \cdots, u_m) \\ \vdots \\ H_p(x_1, x_2, \ldots, x_n,\, u_1, u_2, \cdots, u_m) \end{pmatrix}$$

respectively.

**Continuous-time systems.** The general representation of a continuous-time dynamical system can be compactly described by the following equations:

$$\boxed{\begin{array}{rcl} \dot{x} &=& F(x,u) \\ y &=& H(x,u) \end{array}} \tag{2.4}$$

where $x(t)$ is an $n$-vector, $u(t)$ is an $m$-vector, $y(t)$ is a $p$-vector and the real variable $t$ is the time variable. The first equation above is a first order vector differential equation and is called the state equation. The second equation is called the output equation.

When a system has no inputs of the input vector vector is constant, it can be described by

$$\dot{x} = f(x)\,.$$

A system described by the above equations is called autonomous or time-invariant because the right-hand sides of the equations do not depend explicitly on time $t$. For the first part of the course, we will only concern ourselves with these systems.

However, one can have a system containing time-varying parameters. In this case the system might be described by

$$\begin{aligned} \dot{x} &= F(t,x,u) \\ y &= H(t,x,u) \end{aligned}$$

that is, the right-hand sides of the differential equation and/or the output equation depend explicitly on time. Such a system is called non-autonomous or time-varying. We will look at them later.

**Other systems described by higher order differential equations** Consider a system described by

$$A_N y^{(N)} + A_{N-1} y^{(N-1)} + \cdots + A_0 y = B_M u^{(M)} + B_{M-1} u^{(M-1)} + \cdots + B_0 u$$

where $y(t)$ is a $p$-vector, $u(t)$ is an $m$-vector, $N \geq M$, and the matrix $A_N$ is invertible. When $M = 0$, we can obtain a state space description with

$$x = \begin{pmatrix} y \\ \vdots \\ y^{(N-1)} \end{pmatrix}$$

When $M \geq 1$, we will see later how to put such a system into the standard state space form (2.4).

**Discrete-time systems.** The general representation of a continuous-time dynamical system can be compactly described by the following equations:

$$\boxed{\begin{aligned} x(k+1) &= F(x(k), u(k)) \\ y(k) &= H(x(k), u(k)) \end{aligned}} \tag{2.5}$$

where $x(k)$ is an $n$-vector, $u(k)$ is an $m$-vector, $y(k)$ is a $p$-vector and the integer variable $k$ is the time variable. The first equation above is a first order vector difference equation and is called the state equation. The second equation is called the output equation. A system described by the above equations is called autonomous or time-invariant because the right-hand sides of the equations do not depend explicitly on time $k$.

However, one can have a system containing time-varying parameters. In this case the system might be described by

$$
\begin{aligned}
x(k+1) &= F(k, x(k), u(k)) \\
y(k) &= H(k, x(k), u(k))
\end{aligned}
$$

that is, the right-hand sides of the differential equation and/or the output equation depend explicitly on time. Such a system is called non-autonomous or time-varying.

## 2.7 Solutions and equilibrium states : continuous-time

### 2.7.1 Equilibrium states

Consider a system described by

$$\dot{x} = f(x) \tag{2.6}$$

A solution or motion of this system is any continuous function $x(\cdot)$ satisfying $\dot{x}(t) = f(x(t))$ for all $t$.

An equilibrium solution is the simplest type of solution; it is constant for all time, that is, it satisfies

$$x(t) \equiv x^e$$

for some fixed state vector $x^e$. The state $x^e$ is called an equilibrium state. Since an equilibrium solution must satisfy the above differential equation, all equilibrium states must satisfy the equilibrium condition:

$$\boxed{f(x^e) = 0}$$

or, in scalar terms,

$$\begin{aligned}
f_1(x_1^e, x_2^e, \ldots, x_n^e) &= 0 \\
f_2(x_1^e, x_2^e, \ldots, x_n^e) &= 0 \\
&\vdots \\
f_n(x_1^e, x_2^e, \ldots, x_n^e) &= 0
\end{aligned}$$

Conversely, if a state $x^e$ satisfies the above equilibrium condition, then there is a solution satisfying $x(t) \equiv x^e$; hence $x^e$ is an equilibrium state.

**Example 2** Spring mass damper. With $u = 0$ this system is described by

$$\begin{aligned}
\dot{x}_1 &= x_2 \\
\dot{x}_2 &= -(k/m)x_1 - (c/m)x_2 - g\,.
\end{aligned}$$

Hence equilibrium states are given by:

$$\begin{aligned}
x_2^e &= 0 \\
-(k/m)x_1^e - (c/m)x_2^e - g &= 0
\end{aligned}$$

Hence

$$x^e = 0 = \left( \;-mg/k//0\; \right)$$

*This system has a single equilibrium state.*

**Example 3** The unattached mass. With $u = 0$ this system is described by

$$\begin{aligned}
\dot{x}_1 &= x_2 \\
\dot{x}_2 &= 0\,.
\end{aligned}$$

Hence

$$x^e = \begin{pmatrix} x_1^e \\ 0 \end{pmatrix}$$

where $x_1^e$ is arbitrary. *This system has an infinite number of equilibrium states.*

**Example 4** Pendulum. With $u = 0$, this system is described by

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

The equilibrium condition yields:

$$x_2^e = 0 \qquad \text{and} \qquad \sin(x_1^e) = 0\,.$$

Hence, all equilibrium states are of the form

$$x^e = \begin{pmatrix} m\pi \\ 0 \end{pmatrix}$$

where $m$ is an arbitrary integer. Physically, there are *only two distinct equilibrium states*

$$x^e = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \qquad \text{and} \qquad x^e = \begin{pmatrix} \pi \\ 0 \end{pmatrix}$$

**Higher order ODEs**

$$F(q, \dot{q}, \ldots, q^{(n)}) = 0$$

An equilibrium solution is the simplest type of solution; it is constant for all time, that is, it satisfies

$$q(t) \equiv q^e$$

for some fixed scalar $q^e$. Clearly, $q^e$ must satisfy

$$F(q^e, 0, \ldots, 0) = 0 \tag{2.7}$$

For the state space description of this system introduced earlier, all equilibrium states are given by

$$x^e = \begin{pmatrix} q^e \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

where $q^e$ solves (2.7).

**Multiple higher order ODEs**

Equilibrium solutions

$$q_i(t) \equiv q_i^e\,, \qquad i = 1, 2, \ldots, N$$

Hence

$$\begin{aligned} F_1(q_1^e, 0, \ldots,\ q_2^e, 0, \ldots, \ldots,\ q_N^e, \ldots, 0\ ) &= 0 \\ F_2(q_1^e, 0, \ldots,\ q_2^e, 0, \ldots, \ldots,\ q_N^e, \ldots, 0\ ) &= 0 \\ &\vdots \\ F_N(q_1^e, 0, \ldots,\ q_2^e, 0, \ldots, \ldots,\ q_N^e, \ldots, 0\ ) &= 0 \end{aligned} \tag{2.8}$$

For the state space description of this system introduced earlier, all equilibrium states are given by

$$x^e = \begin{pmatrix} q_1^e \\ 0 \\ \vdots \\ q_2^e \\ 0 \\ \vdots \\ \vdots \\ q_N^e \\ \vdots \\ 0 \end{pmatrix}$$

where $q_1^e, q_2^e, \ldots, q_N^e$ solve (2.8).

**Example 5** Central force motion in inverse square gravitational field

$$\begin{array}{lllll} \ddot{r} & - & r\omega^2 & + & \mu/r^2 & = 0 \\ r\dot{\omega} & + & 2\dot{r}\omega & & & = 0 \end{array}$$

Equilibrium solutions

$$r(t) \equiv r^e\,, \qquad \omega(t) \equiv \omega^e$$

Hence,

$$\dot{r}, \ddot{r}, \dot{\omega} = 0$$

This yields

$$\begin{array}{llll} -r^e(\omega^e)^2 & + & \mu/(r^e)^2 & = 0 \\ 0 & & & = 0 \end{array}$$

Thus there are infinite number of equilibrium solutions given by:

$$\omega^e = \pm\sqrt{\mu/(r^e)^3}$$

where $r^e$ is arbitrary. Note that, for this state space description, an equilibrium state corresponds to a circular orbit.

### 2.7.2 Exercises

**Exercise 9** Find all equilibrium states of the following systems

(i) The first nonlinear system.

(ii) The attitude dynamics system.

(iii) The two link manipulator.

### 2.7.3 Controlled equilibrium states

Consider now a system with inputs described by

$$\dot{x} = F(x, u) \tag{2.9}$$

Suppose the input is constant and equal to $u^e$, that is, $u(t) \equiv u^e$. Then the resulting system is described by

$$\dot{x}(t) = F(x(t), u^e)$$

The equilibrium states $x^e$ of this system are given by $F(x^e, u^e) = 0$. This leads to the following definition.

*A state $x^e$ is a* controlled equilibrium state *of the system, $\dot{x} = F(x, u)$, if there is a constant input $u^e$ such that*

$$\boxed{F(x^e, u^e) = 0}$$

**Example 6** (*Controlled pendulum*)

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

Any state of the form

$$x^e = \begin{pmatrix} x_1^e \\ 0 \end{pmatrix}$$

is a controlled equilibrium state. The corresponding constant input is $u^e = \sin(x_1^e)$.

## 2.8 Solutions and equilibrium states: discrete-time

### 2.8.1 Equilibrium states

Consider a discrete-time system described by

$$x(k+1) = f(x(k))\,. \tag{2.10}$$

A solution of (2.10) is a sequence

$$\big(\ x(0),\quad x(1),\quad x(2),\quad \ldots\ \big)$$

which satisfies (2.5).

**Example 7** Simple scalar nonlinear system

$$x(k+1) = x(k)^2$$

Sample solutions are:

$$\begin{aligned} &(0, 0, 0, \ldots) \\ &(-1, 1, 1, \ldots) \\ &(2, 4, 16, \ldots) \end{aligned}$$

Equilibrium solutions and states

$$x(k) \equiv x^e$$

$$\boxed{f(x^e) = x^e}$$

In scalar terms,

$$\begin{aligned} f_1(x_1^e, x_2^e, \ldots, x_n^e) &= x_1^e \\ f_2(x_1^e, x_2^e, \ldots, x_n^e) &= x_2^e \\ &\vdots \\ f_n(x_1^e, x_2^e, \ldots, x_n^e) &= x_m^e \end{aligned}$$

**Example 8** Simple scalar nonlinear system

$$x(k+1) = x(k)^2$$

All equilibrium states are given by:

$$(x^e)^2 = x^e$$

Solving yields

$$x^e = 0, 1$$

**Example 9** The discrete unattached mass

$$\begin{aligned} x_1^e + Tx_2^e &= x_1^e \\ x_2^e &= x_2^e \end{aligned}$$

$$x^e = \begin{pmatrix} x_1^e \\ 0 \end{pmatrix}$$

where $x_1^e$ is arbitrary.

**Example 10 (AIMD algoritm)** Recalling (2.2) we see that an equilibrium share vector $w$ must satisfy

$$w_i = \beta_i w_i + \gamma_i(1-\beta_1)w_1 + \cdots + \gamma_i(1-\beta_n)w_n \qquad \text{for} \quad i = 1, \cdots, n\,. \tag{2.11}$$

Letting

$$\eta_i = \frac{(1-\beta_i)w_1}{\gamma_i}$$

above equations simplify to

$$\eta_i = \gamma_1\eta_1 + \cdots + \gamma_n\eta_n \qquad \text{for} \quad i = 1, \cdots, n\,.$$

Thus every solution must satisfy

$$\eta_i = K \qquad \text{for} \quad i = 1, \cdots, n\,.$$

for some $K$. Since $\gamma_1 + \cdots + \gamma_n = 1$, it should be $K$ can be any real number. Hence

$$w_i = \frac{\gamma_i K}{1-\beta_i} \qquad \text{for} \quad i = 1, \cdots, n\,.$$

Considering $w_1 + \cdots + w_n = c$, the constant $K$ is uniquely given by

$$K = \frac{c}{\frac{\gamma_1}{1-\beta_1} + \cdots + \frac{\gamma_n}{1-\beta_n}}\,.$$

### 2.8.2 Controlled equilibrium states

Consider now a system with inputs described by

$$x(k+1) = F(x(k), u(k)) \tag{2.12}$$

Suppose the input is constant and equal to $u^e$, that is, $u(k) \equiv u^e$. Then the resulting system is described by

$$x(k+1) = F(x(k), u^e)$$

The equilibrium states $x^e$ of this system are given by $x^e = F(x^e, u^e)$. This leads to the following definition.

*A state $x^e$ is a* controlled equilibrium state *of the system, $x(k+1) = F(x(k), u(k))$, if there is a constant input $u^e$ such that*

$$\boxed{F(x^e, u^e) = x^e}$$

## 2.9 Numerical simulation

### 2.9.1 MATLAB

```
>> help ode23

 ODE23  Solve differential equations, low order method.
        ODE23 integrates a system of ordinary differential equations using
        2nd and 3rd order Runge-Kutta formulas.
        [T,Y] = ODE23('yprime', T0, Tfinal, Y0) integrates the system of
        ordinary differential equations described by the M-file YPRIME.M,
        over the interval T0 to Tfinal, with initial conditions Y0.
        [T, Y] = ODE23(F, T0, Tfinal, Y0, TOL, 1) uses tolerance TOL
        and displays status while the integration proceeds.

        INPUT:
        F     - String containing name of user-supplied problem description.
                Call: yprime = fun(t,y) where F = 'fun'.
                t      - Time (scalar).
                y      - Solution column-vector.
                yprime - Returned derivative column-vector; yprime(i) = dy(i)/dt.
        t0    - Initial value of t.
        tfinal- Final value of t.
        y0    - Initial value column-vector.
        tol   - The desired accuracy. (Default: tol = 1.e-3).
        trace - If nonzero, each step is printed. (Default: trace = 0).

        OUTPUT:
        T  - Returned integration time points (column-vector).
        Y  - Returned solution, one solution column-vector per tout-value.

        The result can be displayed by: plot(tout, yout).

        See also ODE45, ODEDEMO.
```

```
>> help ode45

 ODE45  Solve differential equations, higher order method.
        ODE45 integrates a system of ordinary differential equations using
        4th and 5th order Runge-Kutta formulas.
        [T,Y] = ODE45('yprime', T0, Tfinal, Y0) integrates the system of
        ordinary differential equations described by the M-file YPRIME.M,
        over the interval T0 to Tfinal, with initial conditions Y0.
        [T, Y] = ODE45(F, T0, Tfinal, Y0, TOL, 1) uses tolerance TOL
        and displays status while the integration proceeds.

        INPUT:
        F      - String containing name of user-supplied problem description.
                 Call: yprime = fun(t,y) where F = 'fun'.
                 t      - Time (scalar).
                 y      - Solution column-vector.
                 yprime - Returned derivative column-vector; yprime(i) = dy(i)/dt.
        t0     - Initial value of t.
        tfinal- Final value of t.
        y0     - Initial value column-vector.
        tol    - The desired accuracy. (Default: tol = 1.e-6).
        trace  - If nonzero, each step is printed. (Default: trace = 0).

        OUTPUT:
        T  - Returned integration time points (column-vector).
        Y  - Returned solution, one solution column-vector per tout-value.

        The result can be displayed by: plot(tout, yout).

        See also ODE23, ODEDEMO.
```

### 2.9.2 Simulink

Simulink is part of Matlab. In Simulink, one describes a system graphically. If one has a Simulink model of a system there are several useful Matlab operations which can be applied to the system, for example, one may readily find the equilibrium state of the system.

## 2.10 Exercises

**Exercise 10** Obtain a state-space description of the following system.

$$\begin{aligned}\dddot{q}_1 + \dot{q}_2 + q_1 &= 0 \\ \ddot{q}_1 - \dot{q}_2 + q_2 &= 0\end{aligned}$$

**Exercise 11** Recall the 'two link manipulator' example in the notes.

(i) By appropriate definition of state variables, obtain a first order state space description of this system.

(ii) Find all equilibrium states.

(iii) Numerically simulate this system using MATLAB. Use the following data and initial conditions:

| $m_1$ | $l_1$ | $lc_1$ | $I_1$ | $m_2$ | $l_2$ | $lc_2$ | $I_2$ | $m_{payload}$ |
|---|---|---|---|---|---|---|---|---|
| $kg$ | $m$ | $m$ | $kg.m^2$ | $kg$ | $m$ | $m$ | $kg.m^2$ | $kg$ |
| 10 | 1 | 0.5 | 10/12 | 5 | 1 | 0.5 | 5/12 | 0 |

$$\begin{aligned}q_1(0) &= \frac{\pi}{2}; \quad \dot{q}_1(0) = 0 \\ q_2(0) &= \frac{\pi}{4}; \quad \dot{q}_2(0) = 0\end{aligned}$$

**Exercise 12** Recall the two pendulum cart example in the notes.

(a) Obtain all equilibrium configurations corresponding to $u = 0$.

(b) Consider the following parameter sets

| | $m_0$ | $m_1$ | $m_2$ | $l_1$ | $l_2$ | $g$ |
|---|---|---|---|---|---|---|
| P1 | 2 | 1 | 1 | 1 | 1 | 1 |
| P2 | 2 | 1 | 1 | 1 | 0.99 | 1 |
| P3 | 2 | 1 | 0.5 | 1 | 1 | 1 |
| P4 | 2 | 1 | 1 | 1 | 0.5 | 1 |

and initial conditions,

| | $y$ | $\theta_1$ | $\theta_2$ | $\dot{y}$ | $\dot{\theta}_1$ | $\dot{\theta}_2$ |
|---|---|---|---|---|---|---|
| IC1 | 0 | $-10°$ | $10°$ | 0 | 0 | 0 |
| IC2 | 0 | $10°$ | $10°$ | 0 | 0 | 0 |
| IC3 | 0 | $-90°$ | $90°$ | 0 | 0 | 0 |
| IC4 | 0 | $-90.01°$ | $90°$ | 0 | 0 | 0 |
| IC5 | 0 | $100°$ | $100°$ | 0 | 0 | 0 |
| IC6 | 0 | $100.01°$ | $100°$ | 0 | 0 | 0 |
| IC7 | 0 | $179.99°$ | $0°$ | 0 | 0 | 0 |

Simulate the system using the following combinations:

$$\begin{aligned}
P1: &\quad IC1, IC2, IC3, IC4, IC5, IC6, IC7 \\
P2: &\quad IC1, IC2, IC3, IC5.IC7 \\
P3: &\quad IC1, IC2, IC5, IC7 \\
P4: &\quad IC1, IC2, IC3, IC4, IC5, IC6, IC7
\end{aligned}$$

# Chapter 3

# Linear time-invariant (LTI) systems

To discuss linear systems, we need matrices. In this section we briefly review some properties of matrices. All the definitions and results of this section are stated for real scalars and matrices. However, they also hold for complex scalars and matrices; to obtain the results for the complex case, simply replace 'real' with 'complex' and $\mathbb{R}$ with $\mathbb{C}$.

## 3.1 Matrices

An $m \times n$ matrix $A$ is an array of scalars consisting of $m$ rows and $n$ columns.

$$A = \underbrace{\left.\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}\right\}}_{n \text{ columns}} m \text{ rows}$$

The scalars $a_{ij}$ are called the elements of $A$. If the scalars are real numbers, $A$ is called a real matrix. The set of real $m \times n$ matrices is denoted by $\mathbb{R}^{m \times n}$

Recall that we can do the following things with matrices.

Matrix addition: $A + B$

Multiplication of a matrix by a scalar: $\alpha A$

Zero matrices: $0$

Negative of the matrix $A$: $-A$

One can readily show that $\mathbb{R}^{m \times n}$ with the usual definitions of addition and scalar multiplication is a *vector space*.

Matrix multiplication: $AB$

Some properties:

$$
\begin{array}{rcll}
(AB)C & = & A(BC) & \qquad\text{(associative)} \\
\\
\text{In general} \quad AB & \neq & BA & \qquad\text{(non-commutative)} \\
\\
A(B+C) & = & AB+AC & \\
(B+C)A & + & BA+CA & \\
\\
A(\alpha B) & = & \alpha AB & \\
(\alpha A)B & = & \alpha AB &
\end{array}
$$

Identity matrix in $\mathbb{R}^{n\times n}$:

$$
I \quad := \quad \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix} = \text{diag}\,(1,1,\dots,1)
$$

$$
AI = IA = A
$$

Inverse of a square $(n=m)$ matrix $A$: $A^{-1}$

$$
AA^{-1} = A^{-1}A = I
$$

Some properties:

$$
\begin{array}{rcl}
(\alpha A)^{-1} & = & \dfrac{1}{\alpha}A^{-1} \\
(AB)^{-1} & = & B^{-1}A^{-1}
\end{array}
$$

Transpose of $A$: $A^T$

Some properties:

$$
\begin{array}{rcl}
(A+B)^T & = & A^T+B^T \\
(\alpha A)^T & = & \alpha A^T \\
\\
(AB)^T & = & B^T A^T
\end{array}
$$

Determinant of a square matrix $A$ : Associated with any square matrix $A$ is a scalar called the determinant of $A$ and is denoted by $\det A$.

Some properties:

$$\begin{aligned}\det(AB) &= \det(A)\det(B)\\ \det(A^T) &= \det(A)\end{aligned}$$

Note that, in general,

$$\begin{aligned}\det(A+B) &\neq \det(A)+\det(B)\\ \det(\alpha A) &\neq \alpha\det(A)\end{aligned}$$

The following fact provides a very useful property of determinants.

**Fact:** *A square matrix is invertible iff its determinant is non-zero.*

**Exercise 13** Compute the determinant and inverse of the following matrix.

$$\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$

**Exercise 14** Prove that if $A$ is invertible then,

$$\det(A^{-1}) = \frac{1}{\det(A)}$$

### MATLAB

*Representation of real and complex matrices*
*Addition and multiplication of matrices*
*Multiplication of a matrix by a scalar*
*Matrix powers*
*Transpose of a matrix*
*Inverse of a matrix*

```
>> help zeros

 ZEROS  All zeros.
        ZEROS(N) is an N-by-N matrix of zeros.
        ZEROS(M,N) or ZEROS([M,N]) is an M-by-N matrix of zeros.
        ZEROS(SIZE(A)) is the same size as A and all zeros.

>> help eye

 EYE    Identity matrix.
        EYE(N) is the N-by-N identity matrix.
        EYE(M,N) or EYE([M,N]) is an M-by-N matrix with 1's on
        the diagonal and zeros elsewhere.
```

```
        EYE(SIZE(A)) is the same size as A.

>> help det

 DET    Determinant.
        DET(X) is the determinant of the square matrix X.

>> help inv

 INV    Matrix inverse.
        INV(X) is the inverse of the square matrix X.
        A warning message is printed if X is badly scaled or
        nearly singular.
```

**Exercise 15** Determine which of the following matrices are invertible. Obtain the inverses of those which are. Check your answers using MATLAB.

$$\text{(a)} \quad \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \qquad \text{(b)} \quad \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \qquad \text{(c)} \quad \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad \text{(d)} \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

**Exercise 16** Determine which of the following matrices are invertible. Using MATLAB, check your answers and determine the inverse of those that are invertible.

$$\text{(a)} \quad \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \qquad \text{(b)} \quad \begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix} \qquad \text{(c)} \quad \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

### Powers and polynomials of a matrix

Suppose $A$ is a square matrix. We can define powers of a matrix as follows:

$$\begin{aligned} A^0 &:= I \\ A^1 &:= A \\ A^2 &:= AA \\ &\vdots \\ A^{k+1} &:= AA^k \end{aligned}$$

We can also define polynomials of a matrix as follows. If

$$p(s) = a_0 + a_1 s + a_2 s^2 + \ldots + a_m s^m$$

where $a_0, \ldots, a_m$ are scalars, then,

$$\boxed{p(A) := a_0 I + a_1 A + a_2 A^2 + \ldots + a_m A^m}$$

**Exercise 17** Compute $p(A)$ for the following polynomial-matrix pairs. Check your answers using MATLAB.

(a)

$$p(s) = s^3 - 3s^2 + 3s - 1 \qquad A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

(b)

$$p(s) = s^2 - 3s \qquad A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

(c)

$$p(s) = s^3 + s^2 + s + 1 \qquad A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}$$

(d)

$$p(s) = s^3 - 2s \qquad A = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

## 3.2 Linear time-invariant systems

### 3.2.1 Continuous-time

A continuous-time, linear time-invariant (LTI) input-output system can be described by

$$\boxed{\begin{array}{rcl} \dot{x} &=& Ax + Bu \\ y &=& Cx + Du \end{array}}$$

where the $n$-vector $x(t)$ is the state vector at time $t$, the $m$-vector $u(t)$ is the input vector at time $t$, and the $p$-vector $y(t)$ is the output vector at time $t$. The $n \times n$ matrix $A$ is sometimes called the system matrix. The matrices $B$, $C$, and $D$ have dimensions $n \times m$, $p \times n$, and $p \times m$, respectively. In scalar terms, this system is described by:

$$\begin{array}{rcl} \dot{x}_1 &=& a_{11}x_1 + \ldots + a_{1n}x_n + b_{11}u_1 + \ldots + b_{1m}u_m \\ &\vdots& \\ \dot{x}_n &=& a_{n1}x_1 + \ldots + a_{nn}x_n + b_{n1}u_1 + \ldots + b_{nm}u_m \end{array}$$

and

$$\begin{array}{rcl} y_1 &=& c_{11}x_1 + \ldots + c_{1n}x_n + d_{11}u_1 + \ldots + d_{1m}u_m \\ &\vdots& \\ y_p &=& c_{p1}x_1 + \ldots + c_{pn}x_n + d_{p1}u_1 + \ldots + d_{pm}u_m \,. \end{array}$$

A system with no inputs is described by

$$\dot{x} = Ax \,. \tag{3.1}$$

**Example 11** The unattached mass

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ 1/m \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 0 \end{pmatrix}, \quad D = 0 \,.$$

**Example 12** Spring-mass-damper

$$A = \begin{pmatrix} 0 & 1 \\ -k/m & -c/m \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \quad C = \begin{pmatrix} -k & -c \end{pmatrix}, \quad D = 0 \,.$$

**Other systems.** It should be mentioned that the results of this section and most of these notes only apply to finite-dimensional systems, that is, systems whose state space is finite-dimensional. As an example of a system not included here, consider a system with a time delay $T$ described by

$$\begin{array}{rcl} \dot{q}(t) &=& A_0 q(t) + A_1 q(t-T) + Bu(T) \\ y(t) &=& Cq(t) + Du(t) \end{array}$$

As another example, consider any system described by a partial differtial equation, for example

$$\frac{\partial^2 q}{\partial t^2} + k^2 \frac{\partial^4 q}{\partial \eta^4} = 0 \,.$$

### 3.2.2 Discrete-time

A discrete-time linear time-invariant (LTI) input-output system can be described by

$$\boxed{\begin{array}{rcl} x(k+1) &=& Ax(k) + Bu(k) \\ y(k) &=& Cx(k) + Du(k) \end{array}}$$

where the $n$-vector $x(k)$ is the state vector at time $k$, the $m$-vector $u(k)$ is the input vector at time $k$, and the $p$-vector $y(k)$ is the output vector at time $k$. The $n \times n$ matrix $A$ is sometimes called the system matrix. The matrices $B$, $C$, and $D$ have dimensions $n \times m$, $p \times n$, and $p \times m$, respectively.
In scalar terms:

$$\begin{array}{rcl} x_1(k+1) &=& a_{11}x_1(k) + \ldots + a_{1n}x_n(k) + b_{11}u_1(k) + \ldots + b_{1m}u_m(k) \\ &\vdots& \\ x_n(k+1) &=& a_{n1}x_1(k) + \ldots + a_{nn}x_n(k) + b_{n1}u_1(k) + \ldots + b_{nm}u_m(k) \end{array}$$

and

$$\begin{array}{rcl} y_1 &=& c_{11}x_1 + \ldots + c_{1n}x_n + d_{11}u_1 + \ldots + d_{1m}u_m \\ &\vdots& \\ y_p &=& c_{p1}x_1 + \ldots + c_{pn}x_n + d_{p1}u_1 + \ldots + d_{pm}u_m \end{array}$$

A system with no inputs is described by

$$x(k+1) = Ax(k)\,. \tag{3.2}$$

**Example 13** The discrete unattached mass

$$A = \begin{pmatrix} 1 & T \\ 0 & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} T^2/2m \\ T/m \end{pmatrix}, \qquad C = \begin{pmatrix} 1 & 0 \end{pmatrix}, \qquad D = 0\,.$$

**Example 14** AIMD This is a system with no inputs and

$$A = \begin{pmatrix} \beta_1 + \gamma_1(1-\beta_1) & \gamma_1(1-\beta_2) & \cdots & \gamma_1(1-\beta_n) \\ \gamma_2(1-\beta_1) & \beta_2 + \gamma_2(1-\beta_2) & \cdots & \gamma_2(1-\beta_n) \\ \vdots & & \ddots & \vdots \\ \gamma_n(1-\beta_1) & \gamma_n(1-\beta_2) & \cdots & \beta_n + \gamma_n(1-\beta_n) \end{pmatrix}$$

## 3.3 Linearization about an equilibrium solution

In this section we present systematic procedures for approximating a nonlinear system by a linear system. These procedures rely heavily on derivatives.

### 3.3.1 Derivative as a matrix

Consider an $m$-vector-valued function $f$ of an $n$-vector variable $x$. Suppose that each of the partial derivatives, $\frac{\partial f_i}{\partial x_j}(x^*)$ exist and are continuous about some $x^*$.

The derivative of $f$ (Jacobian of $f$) at $x^*$ is the following $m \times n$ matrix:

$$Df(x^*) = \frac{\partial f}{\partial x}(x^*) := \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x^*) & \frac{\partial f_1}{\partial x_2}(x^*) & \dots & \frac{\partial f_1}{\partial x_n}(x^*) \\ \frac{\partial f_2}{\partial x_1}(x^*) & \frac{\partial f_2}{\partial x_2}(x^*) & \dots & \frac{\partial f_2}{\partial x_n}(x^*) \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(x^*) & \frac{\partial f_m}{\partial x_2}(x^*) & \dots & \frac{\partial f_m}{\partial x_n}(x^*) \end{pmatrix}.$$

So,

$$Df(x^*)_{ij} = \frac{\partial f_i}{\partial x_j}(x^*)\,,$$

that is, *the $ij$-th element of $Df$ is the partial derivative of $f_i$ with respect to $x_j$*. Sometimes $Df$ is written as $\frac{\partial f}{\partial x}$.

**Example 15** If

$$f(x) = \begin{pmatrix} x_1 x_2 \\ \cos(x_1) \\ e^{x_2} \end{pmatrix}.$$

then

$$Df(x) = \begin{pmatrix} x_2 & x_1 \\ -\sin(x_1) & 0 \\ 0 & e^{x_2} \end{pmatrix}.$$

Some properties:

$$\begin{array}{rcll} D(f+g) & = & Df + Dg & \\ D(\alpha f) & = & \alpha Df & (\alpha \text{ is a constant scalar}) \end{array}$$

$$\begin{array}{lrclrcl} \text{(Constant function:)} & f(x) & \equiv & c & Df(x) & = & 0 \\ \text{(Linear function:)} & f(x) & = & Ax & Df(x) & = & A \end{array}$$

All linearization is based on the following fundamental property of the derivative. If $x^*$ and $x$ are any two vectors and $x$ is 'close' to $x^*$, then

$$\boxed{f(x) \approx f(x^*) + \frac{\partial f}{\partial x}(x^*)(x - x^*)}$$

This says that, near $x^*$, the function can be approximated by a linear function plus a constant term.

### 3.3.2 Linearization of continuous-time systems

Consider a continuous-time system described by

$$\begin{aligned} \dot{x} &= F(x,u) \\ y &= H(x,u) \end{aligned} \tag{3.3}$$

and suppose $x^e$ is an equilibrium state corresponding to a constant input $u^e$; hence,

$$F(x^e,u^e) = 0\,.$$

Let $y^e$ be the corresponding constant output, that is,

$$y^e = H(x^e,u^e)\,.$$

Sometimes we call the triple $(x^e,u^e,y^e)$ a trim condition of the system. Introducing the perturbed state $\delta x$, the perturbed input $\delta u$ and the perturbed output $\delta y$ defined by

$$\delta x := x - x^e, \qquad \delta u := u - u^e, \qquad \delta x := u - u^e,$$

the system can also be described by

$$\begin{aligned} \delta\dot{x} &= F(x^e{+}\delta x, u^e{+}\delta u) \\ \delta y &= H(x^e{+}\delta x, u^e{+}\delta u) - H(x^e,u^e) \end{aligned}$$

Define now the following derivative matrices:

$$\frac{\partial F}{\partial x}(x^e, u^e) \quad := \quad \left(\frac{\partial F_i}{\partial x_j}(*)\right) \quad = \quad \begin{pmatrix} \frac{\partial F_1}{\partial x_1}(*) & \frac{\partial F_1}{\partial x_2}(*) & \dots & \frac{\partial F_1}{\partial x_n}(*) \\ \frac{\partial F_2}{\partial x_1}(*) & \frac{\partial F_2}{\partial x_2}(*) & \dots & \frac{\partial F_2}{\partial x_n}(*) \\ \vdots & \vdots & & \vdots \\ \frac{\partial F_n}{\partial x_1}(*) & \frac{\partial F_n}{\partial x_2}(*) & \dots & \frac{\partial F_n}{\partial x_n}(*) \end{pmatrix}$$

$$\frac{\partial F}{\partial u}(x^e, u^e) \quad := \quad \left(\frac{\partial F_i}{\partial u_j}(*)\right) \quad = \quad \begin{pmatrix} \frac{\partial F_1}{\partial u_1}(*) & \frac{\partial F_1}{\partial u_2}(*) & \dots & \frac{\partial F_1}{\partial u_m}(*) \\ \frac{\partial F_2}{\partial u_1}(*) & \frac{\partial F_2}{\partial u_2}(*) & \dots & \frac{\partial F_2}{\partial u_m}(*) \\ \vdots & \vdots & & \vdots \\ \frac{\partial F_n}{\partial u_1}(*) & \frac{\partial F_n}{\partial u_2}(*) & \dots & \frac{\partial F_n}{\partial u_m}(*) \end{pmatrix}$$

$$\frac{\partial H}{\partial x}(x^e, u^e) \quad := \quad \left(\frac{\partial H_i}{\partial x_j}(*)\right) \quad = \quad \begin{pmatrix} \frac{\partial H_1}{\partial x_1}(*) & \frac{\partial H_1}{\partial x_2}(*) & \dots & \frac{\partial H_1}{\partial x_n}(*) \\ \frac{\partial H_2}{\partial x_1}(*) & \frac{\partial H_2}{\partial x_2}(*) & \dots & \frac{\partial H_2}{\partial x_n}(*) \\ \vdots & \vdots & & \vdots \\ \frac{\partial H_p}{\partial x_1}(*) & \frac{\partial H_p}{\partial x_2}(*) & \dots & \frac{\partial H_p}{\partial x_n}(*) \end{pmatrix}$$

$$\frac{\partial H}{\partial u}(x^e, u^e) \quad := \quad \left(\frac{\partial H_i}{\partial u_j}(*)\right) \quad = \quad \begin{pmatrix} \frac{\partial H_1}{\partial u_1}(*) & \frac{\partial H_1}{\partial u_2}(*) & \dots & \frac{\partial H_1}{\partial u_m}(*) \\ \frac{\partial H_2}{\partial u_1}(*) & \frac{\partial H_2}{\partial u_2}(*) & \dots & \frac{\partial H_2}{\partial u_m}(*) \\ \vdots & \vdots & & \vdots \\ \frac{\partial H_p}{\partial u_1}(*) & \frac{\partial H_p}{\partial u_2}(*) & \dots & \frac{\partial H_p}{\partial u_m}(*) \end{pmatrix}$$

where $* = (x^e, u^e)$. When $x$ is 'close' to $x^e$ and $u$ is 'close' to $u^e$, that is, when $\delta x$ and $\delta u$ are 'small',

$$\begin{aligned} F(x^e{+}\delta x,\ u^e{+}\delta u) \quad &\approx \quad F(x^e, u^e) + \frac{\partial F}{\partial x}(x^e, u^e)\delta x + \frac{\partial F}{\partial u}(x^e, u^e)\delta u \\ &= \quad \frac{\partial F}{\partial x}(x^e, u^e)\delta x + \frac{\partial F}{\partial u}(x^e, u^e)\delta u \end{aligned}$$

and

$$H(x^e+\delta x,\ u^e+\delta u) \approx H(x^e,u^e) + \frac{\partial H}{\partial x}(x^e,u^e)\delta x + \frac{\partial H}{\partial u}(x^e,u^e)\delta u$$

Hence

$$\begin{aligned}
\delta\dot{x} &\approx \frac{\partial F}{\partial x}(x^e,u^e)\delta x + \frac{\partial F}{\partial u}(x^e,u^e)\delta u \\
\delta y &\approx \frac{\partial H}{\partial x}(x^e,u^e)\delta x + \frac{\partial H}{\partial u}(x^e,u^e)\delta u
\end{aligned}$$

This leads to the following definition:

**The linearization of system (3.3) about** $(x^e, u^e)$**:**

$$\boxed{\begin{aligned}
\delta\dot{x} &= A\,\delta x + B\,\delta u \\
\delta y &= C\,\delta x + D\,\delta u
\end{aligned}} \tag{3.4}$$

where

$$A = \frac{\partial F}{\partial x}(x^e,u^e)\,, \qquad B = \frac{\partial F}{\partial u}(x^e,u^e)\,, \qquad C = \frac{\partial H}{\partial x}(x^e,u^e)\,, \qquad D = \frac{\partial H}{\partial u}(x^e,u^e)\,.$$

**Example 16** (Simple pendulum.) The linearization of this system about any $(x^e, u^e)$ is given by (3.4) with

$$A = \begin{pmatrix} 0 & 1 \\ -\cos(x_1^e) & 0 \end{pmatrix} \qquad B = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \qquad C = \begin{pmatrix} 1 & 0 \end{pmatrix} \qquad D = 0$$

For eqm. state

$$x^e = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

we obtain

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

For eqm. state

$$x^e = \begin{pmatrix} \pi \\ 0 \end{pmatrix}$$

we obtain

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

### 3.3.3 Implicit linearization

This is the way to go when starting from a higher order ODE description.

*It is usually easier to first linearize and then obtain a state space description than vice-versa.*

For a single equation,

$$F(q, \dot{q}, \ldots, q^{(n)}, u) = 0$$

linearization about an equilibrium solution $q(t) \equiv q^e$ corresponding to constant input $u(t) \equiv u^e$ is defined by:

$$\frac{\partial F}{\partial q}(*)\delta q \quad + \quad \frac{\partial F}{\partial \dot{q}}(*)\delta \dot{q} \quad + \quad \ldots \quad + \quad \frac{\partial F}{\partial q^{(n)}}(*)\delta q^{(n)} \quad + \quad \frac{\partial F}{\partial u}(*)\delta u \quad = \quad 0$$

where $(*) = (q^e, 0, 0, \ldots, 0)$, $\delta q = q - q^e$ and $\delta u = u - u^e$.

For multiple equations, see next example.

**Example 17** Orbit mechanics

$$\begin{aligned} \ddot{r} - r\omega^2 + \mu/r^2 &= 0 \\ r\dot{\omega} + 2\dot{r}\omega &= 0 \end{aligned}$$

Linearization about equilibrium solutions corresponding to constant values $r^e, \omega^e$ of $r$ and $\omega$, respectively, results in:

$$\begin{aligned} \delta\ddot{r} - (\omega^e)^2\delta r - (2r^e\omega^e)\delta\omega - (2\mu/(r^e)^3)\delta r &= 0 \\ \dot{\omega}^e\delta r + r^e\delta\dot{\omega} + +2\omega^e\delta\dot{r} + 2\dot{r}^e\delta\omega &= 0 \end{aligned}$$

Using the relationship

$$(\omega^e)^2 = \mu/(r^e)^3$$

we obtain

$$\begin{aligned} \delta\ddot{r} - 3(\omega^e)^2\delta r - (2r^e\omega^e)\delta\omega &= 0 \\ r^e\delta\dot{\omega} + 2\omega^e\delta\dot{r} &= 0 \end{aligned}$$

Introducing state variables:

$$x_1 := \delta r, \qquad x_2 := \delta\dot{r} \qquad x_3 := \delta\omega$$

we obtain a LTI system with system matrix

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 3(\omega^e)^2 & 0 & 2r^e\omega^e \\ 0 & -2\omega^e/r^e & 0 \end{pmatrix}$$

**Example: vehicle lateral dynamics**

Consider a vehicle moving in a horizonal plane. We can describe the motion of this vehicle with the velocity $\bar{v}$ of its center of gravity (CG) and its yaw rate $\dot{\psi}$. The velocity of the mass center can be described by its magnitude $v$ (the speed of the CG) and the vehicle sideslip angle $\beta$ which is the angle between $\bar{v}$ and the vehicle. We will use these three variables to describe the state of the vehicle. As inputs we consider the steering angle $\delta$ of the front wheels and the braking and acceleration forces $F_{fx}$ and $F_{rx}$ on the front and rear wheels of the vehicle.

Application of Newton's second law to the vehicle mass center results in

$$\begin{aligned} m\dot{v} &= F_{rx}\cos\beta + F_{ry}\sin\beta + F_{fx}\cos(\beta-\delta) + F_{fx}\sin(\beta-\delta) && (3.5a) \\ mv(\dot{\beta}+\dot{\psi}) &= -F_{rx}\sin\beta + F_{ry}\cos\beta - F_{fx}\sin(\beta-\delta) + F_{ry}\cos(\beta-\delta) && (3.5b) \end{aligned}$$

where $F_{fx}, F_{rx}$ are the longitudinal forces on the front and rear wheels due to braking and acceleration; $F_{fy}, F_{ry}$ are the lateral forces on the front and rear wheels. Considerations of angular momentum yield

$$J_z\ddot{\psi} = a(F_{fy}\cos\delta + F_{fx}\sin\delta) - bF_{ry} \tag{3.6}$$

where $J_z$ is the moment of inertia of the vehicle about the vertical line through its CG, and $a$, $b$ are the distances between the mass center and the front and rear of the vehicle.

The lateral force $F_{ry}$ depends on the rear sideslip angle $\alpha_r$ which is the angle between the rear tires and the velocity of the rear of the vehicle. The lateral force $F_{fy}$ depends on the front sideslip angle $\alpha_f$ which is the angle between the front tires and the velocity of the front of the vehicle. Kinematical considerations yield the following relationships:

$$\begin{aligned} \tan\alpha_r &= \tan\beta - \frac{b\dot{\psi}}{v\cos\beta} && (3.7a) \\ \tan(\alpha_f+\delta) &= \tan\beta + \frac{a\dot{\psi}}{v\cos\beta} && (3.7b) \end{aligned}$$

Assuming the lateral forces are zero for zero sideslip angles, and considering $F_{rx} = F_{fx} = \delta = 0$, the vehicle has an equilibrium with $\beta = \dot{\psi} = \alpha_r = \alpha_f = 0$ and $v$ an arbitrary constant.

Linearization of the nonlinear equations about these equilibrium conditions results in

$$\begin{aligned} \delta\dot{v} &= a \\ mv(\dot{\beta}+\dot{\psi}) &= -C_{\alpha f}\alpha_f - C_{\alpha r}\alpha_r \\ J_z\ddot{\psi} &= -aC_{\alpha f}\alpha_f + bC_{\alpha r}\alpha_r \\ \alpha_r &= \beta - \frac{b\dot{\psi}}{v} \\ \alpha_f &= \beta + \frac{a\dot{\psi}}{v} - \delta \end{aligned}$$

where $a = (F_{rx} + F_{fx})/m$ and

$$C_{\alpha f} = -\frac{dF_{yf}}{d\alpha_f} \qquad \text{and} \qquad C_{\alpha r} = -\frac{dF_{yr}}{d\alpha_r}$$

are the front and rear "axle stiffnesses".

Notice that the dynamics of $\delta v$ can be decoupled from that of $(\beta, \dot{\psi})$. With $x = (\beta, \dot{\psi})$ and $u = \delta$, the lateral vehicle dynamics can be described by $\dot{x} = Ax + Bu$ where

$$A = \begin{pmatrix} \frac{-(C_{\alpha f}+C_{\alpha r})}{mv} & \frac{bC_{\alpha r}-aC_{\alpha f}}{mv^2} - 1 \\ \\ \frac{bC_{\alpha r}-aC_{\alpha f}}{J_z} & -\frac{a^2C_{\alpha f}+b^2C_{\alpha r}}{J_z v} \end{pmatrix}, \qquad B = \begin{pmatrix} \frac{C_{\alpha f}}{mv} \\ \\ \frac{aC_{\alpha f}}{J_z} \end{pmatrix} \tag{3.8}$$

### 3.3.4 Linearization of discrete-time systems

Consider a discrete-time system described by

$$\begin{aligned} x(k+1) &= F(x(k), u(k)) \\ y(k) &= H(x(k), u(k)) \end{aligned} \tag{3.9}$$

and suppose $x^e, u^e, y^e$ is a trim condition, that is,

$$\begin{aligned} F(x^e, u^e) &= x^e \\ H(x^e, u^e) &= y^e \end{aligned}$$

**The linearization of system (3.9) about** $(x^e, u^e)$**:**

$$\boxed{\begin{aligned} \delta x(k+1) &= A\,\delta x(k) + B\,\delta u(k) \\ \delta y(k) &= C\,\delta x(k) + D\,\delta u(k) \end{aligned}} \tag{3.10}$$

where

$$A = \frac{\partial F}{\partial x}(x^e, u^e)\,, \qquad B = \frac{\partial F}{\partial u}(x^e, u^e)\,, \qquad C = \frac{\partial H}{\partial x}(x^e, u^e)\,, \qquad D = \frac{\partial H}{\partial u}(x^e, u^e)\,.$$

**Example 18** Simple scalar nonlinear system

$$x(k+1) = x(k)^2$$

We have $x^e = 0, 1$. For $x^e = 0$:

$$\delta x(k+1) = 0$$

For $x^e = 1$:

$$\delta x(k+1) = 2\delta x(k)$$

## Exercises

**Exercise 18** Consider the two-degree-of-freedom spring-mass-damper system described by

$$\begin{aligned} m\ddot{q}_1 + C(\dot{q}_1 - \dot{q}_2) + K(q_1 - q_2) + c\dot{q}_1 + kq_1 &= 0 \\ m\ddot{q}_2 + C(\dot{q}_2 - \dot{q}_1) + K(q_2 - q_1) + c\dot{q}_2 + kq_2 &= 0 \end{aligned}$$

Obtain an $A$-matrix for a state space description of this system.

**Exercise 19** Linearize each of the following systems about the zero equilibrium state.

(i)

$$\begin{aligned} \dot{x}_1 &= (1 + x_1^2)x_2 \\ \dot{x}_2 &= -x_1^3 \end{aligned}$$

(ii)

$$\begin{aligned} \dot{x}_1 &= \sin x_2 \\ \dot{x}_2 &= (\cos x_1)x_3 \\ \dot{x}_3 &= e^{x_1}x_2 \end{aligned}$$

**Exercise 20**

(a) Obtain all equilibrium states of the following system:

$$\begin{aligned} \dot{x}_1 &= 2x_2(1 - x_1) - x_1 \\ \dot{x}_2 &= 3x_1(1 - x_2) - x_2 \end{aligned}$$

(b) Linearize the above system about the zero equilibrium state.

**Exercise 21** For each of the following systems, linearize about each equilibrium solution and obtain the system $A$-matrix for a state space representation of these linearized systems.

(a)

$$\ddot{q} + (q^2 - 1)\dot{q} + q = 0\,.$$

where $q(t)$ is a scalar.

(b)

$$\ddot{q} + \dot{q} + q - q^3 = 0$$

where $q(t)$ is a scalar.

(c)

$$\begin{aligned} (M + m)\ddot{q} + ml\,\ddot{\theta}\cos\theta - ml\,\dot{\theta}^2\sin\theta + kq &= 0 \\ ml\ddot{q}\cos\theta + ml^2\,\ddot{\theta} + mgl\,\sin\theta &= 0 \end{aligned}$$

where $q(t)$ and $\theta(t)$ are scalars.

(d)

$$\ddot{q} + 0.5\dot{q}|\dot{q}| + q = 0\,.$$

where $q(t)$ is a scalar.

**Exercise 22**

(a) Obtain all equilibrium solutions of the following system:

$$\begin{aligned}(\cos q_1)\ddot{q}_1 + (\sin q_1)\ddot{q}_2 \quad &+ \quad (\sin q_2)\dot{q}_1^2 + \sin q_2 = 0 \\ -(\sin q_1)\ddot{q}_1 + (\cos q_1)\ddot{q}_2 \quad &+ \quad (\cos q_2)\dot{q}_2^2 + \sin q_1 = 0\end{aligned}$$

(b) Linearize the above system about its zero solution.

**Exercise 23 (Simple pendulum in drag)**

Figure 3.1: Pendulum in drag

Recall the simple pendulum in drag whose motion is described by

$$ml\ddot{\theta} + \kappa V(l\dot{\theta} - w\sin\theta) + mg\sin\theta = 0$$

where

$$V = \sqrt{l^2\dot{\theta}^2 + w^2 - 2lw\sin(\theta)\dot{\theta}} \qquad \text{with} \qquad \kappa = \frac{\rho S C_D}{2}$$

and
$g$ is the gravitational acceleration constant of the earth,
$S$ is the reference area associated with the ball,
$C_D$ is the drag coefficient of the ball,
$\rho$ is the air density.

(a) Obtain the equilibrium values $\theta^e$ of $\theta$.

(b) Linearize the system about $\theta^e$.

(c) Obtain an expression for the $A$ matrix for a state space representation of the linearized system.

(d) Compare the behavior of the nonlinear system with that of the linearized system for the following parameters.

$$\begin{array}{lll} l = 0.1 \text{ m} & m = 10 \text{ grams} & g = 9.81 \text{ m/s}^2 \\ C_D = 0.2 & S = .01 \text{ m}^2 & \rho = 0.3809 \text{ kg/m}^3 \end{array}$$

Consider the following cases:

| $\theta^e$ | $w$ (m/s) |
|---|---|
| 0 | 0 |
| 0 | 5 |
| 0 | 15 |
| 0 | 20 |
| 180° | 20 |

Illustrate your results with time histories of $\delta\theta$ in degrees. Comment on your results.

# Chapter 4

# Transfer functions

## 4.1 $\longleftarrow$ Transfer functions $\longleftarrow$

### 4.1.1 The Laplace transform

A state space description of an input-output system can be regarded as a time-domain description. For linear time-invariant systems, transfer functions provide another way of describing input-output systems; this description is sometimes called a frequency domain description.

To discuss transfer functions of continuous-time systems, we need the Laplace transform. Suppose $f$ is a function of time $t$. Then the Laplace transform of $f$, which we denote by $\hat{f}$ or $\mathcal{L}(f)$, is a function of a complex variable $s$ and is defined by

$$\hat{f}(s) = \mathcal{L}(f)(s) = \int_0^\infty e^{-st} f(t) dt\,. \tag{4.1}$$

Note that $f(t)$ could be a scalar, a vector or a matrix.

We will need the following very useful relationship:

$$\boxed{\mathcal{L}(\dot{f})(s) = s\hat{f}(s) - f(0)} \tag{4.2}$$

The following example illustrates the use of the above relationship.

**Example 19** Consider $f(t) = e^{at}$ where $a$ is any scalar. Since $f$ satisfies

$$\dot{f} = af$$

and $f(0) = 0$, we obtain that

$$s\hat{f}(s) - 1 = a\hat{f}(s)\,.$$

Hence, as expected, $\hat{f}(s) = 1/(s-a)$.

### 4.1.2 Transfer functions

Consider now the LTI system

$$\begin{aligned}\dot{x}(t) &= Ax(t) + Bu(t)\\ y(t) &= Cx(t) + Du(t)\end{aligned}$$

along with the initial condition

$$x(0) = x_0 \,.$$

Taking the Laplace transform of the state equation results in

$$s\hat{x}(s) - x_0 = A\hat{x}(s) + B\hat{u}(s)$$

which can be rewritten as

$$(sI - A)\hat{x}(s) = x_0 + B\hat{u}(s)$$

where $\hat{x}$ and $\hat{u}$ are the Laplace transforms of $x$ and $u$, respectively, Whenever the matrix $sI - A$ is invertible, the above equation may be uniquely solved for $\hat{x}$ yielding

$$\hat{x}(s) = (sI - A)^{-1}x_0 + (sI - A)^{-1}B\hat{u}(s)\,.$$

Taking the Laplace transform of the above output equation yields

$$\hat{y}(s) = C\hat{x}(s) + D\hat{u}(s)$$

where $\hat{y}$ is the Laplace transform of $y$. Using the expression for $\hat{x}$ now results in

$$\hat{y}(s) = C(sI - A)^{-1}x_0 \quad + \quad \hat{G}(s)\hat{u}(s)$$

where the transfer function (matrix) $\hat{G}(s)$ is defined by

$$\boxed{\hat{G}(s) = C(sI - A)^{-1}B + D}$$

**Zero initial conditions response.** Suppose that $x(0) = 0$. Then, the relationship between the Laplace transform of the input $\hat{u}$ and the Laplace transform of the output $\hat{y}$ is given by the simple relationship

$$\boxed{\hat{y} = \hat{G}\hat{u}}$$

Sometimes this is represented by

$$\hat{y} \longleftarrow \boxed{\hat{G}} \longleftarrow \hat{u}$$

For a SISO system, we obtain that, for $\hat{u} \neq 0$,

$$\hat{G} = \frac{\hat{y}}{\hat{u}}\,.$$

*Thus, for a SISO system, the transfer function is the ratio of the Laplace transform of the output to the Laplace transform of the input all initial conditions are zero.*

**Example 20**

$$\begin{aligned} \dot{x}(t) &= -x(t) + u(t) \\ y(t) &= x(t) \end{aligned}$$

Here

$$\hat{G}(s) = C(sI - A)^{-1} + D = \frac{1}{s+1}$$

**Example 21** Unattached mass

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= u \\ y &= x_1 \end{aligned}$$

Here

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad B = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \qquad C = \begin{pmatrix} 1 & 0 \end{pmatrix} \qquad D = 0$$

Hence

$$\begin{aligned} \hat{G}(s) &= C(sI - A)^{-1}B \\ &= \frac{1}{s^2} \end{aligned}$$

**Exercise 24** Show that the transfer function of

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -\alpha_0 x_1 - \alpha_1 x_2 + u \\ y &= \beta_0 x_1 + \beta_1 x_2 \end{aligned}$$

is given by

$$\hat{G}(s) = \frac{\beta_1 s + \beta_0}{s^2 + \alpha_1 s + \alpha_0}$$

**Exercise 25** Show that the transfer function of

$$\begin{aligned} \dot{x}_1 &= -\alpha_0 x_2 + \beta_0 u \\ \dot{x}_2 &= x_1 - \alpha_1 x_2 + \beta_1 u \\ y &= x_2 \end{aligned}$$

is given by

$$\hat{G}(s) = \frac{\beta_1 s + \beta_0}{s^2 + \alpha_1 s + \alpha_0}$$

**MATLAB**

```
>> help ss2tf

 SS2TF  State-space to transfer function conversion.
        [NUM,DEN] = SS2TF(A,B,C,D,iu)  calculates the transfer function:

                        NUM(s)          -1
                H(s) = -------- = C(sI-A) B + D
                        DEN(s)
        of the system:
                .
                x = Ax + Bu
                y = Cx + Du

        from the iu'th input.  Vector DEN contains the coefficients of the
        denominator in descending powers of s.  The numerator coefficients
        are returned in matrix NUM with as many rows as there are
        outputs y.
```

**Example 22** MIMO (Multi-input multi-output) B&B Consider the system illustrated in Figure 4.1 which has two inputs and two outputs. Its motion can be described by

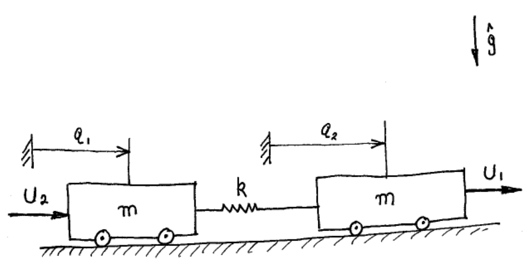



Figure 4.1: B&B

$$
\begin{aligned}
m\ddot{q}_1 &= \quad k(q_2 - q_1) &+ \quad u_1 \\
m\ddot{q}_2 &= -k(q_2 - q_1) &+ \quad u_2 \\
y_1 &= q_1 \\
y_2 &= q_2
\end{aligned}
$$

With $x_1 = q_1, x_2 = q_2, x_3 = \dot{q}_1, x_4 = \dot{q}_2$, we have

$$
A = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -k/m & k/m & 0 & 0 \\ k/m & -k/m & 0 & 0 \end{pmatrix} \qquad B = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1/m & 0 \\ 0 & 1/m \end{pmatrix}
$$

$$
C = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \qquad D = 0
$$

Hence

$$
\begin{aligned}
\hat{G}(s) &= C(sI-A)^{-1}B+D \\
&= \begin{pmatrix} \frac{ms^2+k}{ms^2(ms^2+2k)} & \frac{k}{ms^2(ms^2+2k)} \\ \\ \frac{k}{ms^2(ms^2+2k)} & \frac{ms^2+k}{ms^2(ms^2+2k)} \end{pmatrix}
\end{aligned}
$$

Note that the above calculation requires one to invert the $4 \times 4$ matrix $sI - A$

• We can also compute $\hat{G}(s)$ as follows. Taking the Laplace transform of the original second order differential equations with zero initial conditions, we get

$$
\begin{aligned}
ms^2\hat{q}_1 &= -k(\hat{q}_1 - \hat{q}_2) + \hat{u}_1 \\
ms^2\hat{q}_2 &= k(\hat{q}_1 - \hat{q}_2) + \hat{u}_2 \\
\hat{y}_1 &= \hat{q}_1 \\
\hat{y}_2 &= \hat{q}_2
\end{aligned}
$$

hence,

$$
\begin{pmatrix} ms^2+k & -k \\ -k & ms^2+k \end{pmatrix} \begin{pmatrix} \hat{q}_1 \\ \hat{q}_2 \end{pmatrix} = \begin{pmatrix} \hat{u}_1 \\ \hat{u}_2 \end{pmatrix}
$$

Solving, yields

$$
\begin{aligned}
\hat{y} = \begin{pmatrix} \hat{q}_1 \\ \hat{q}_2 \end{pmatrix} &= \frac{1}{\Delta(s)} \begin{pmatrix} ms^2+k & k \\ k & ms^2+k \end{pmatrix} \begin{pmatrix} \hat{u}_1 \\ \hat{u}_2 \end{pmatrix} \\
\Delta(s) &= ms^2(ms^2+2k)
\end{aligned}
$$

So,

$$
\hat{G}(s) = \begin{pmatrix} \frac{ms^2+k}{ms^2(ms^2+2k)} & \frac{k}{ms^2(ms^2+2k)} \\ \\ \frac{k}{ms^2(ms^2+2k)} & \frac{ms^2+k}{ms^2(ms^2+2k)} \end{pmatrix}
$$

Using this second method, the matrix to be inverted was only $2 \times 2$.

## MATLAB

```
>> help ss2tf

 SS2TF  State-space to transfer function conversion.
        [NUM,DEN] = SS2TF(A,B,C,D,iu)  calculates the transfer function:

                        NUM(s)          -1
                H(s) = -------- = C(sI-A) B + D
                        DEN(s)
```

```
of the system:
        .
        x = Ax + Bu
        y = Cx + Du

from the iu'th input.  Vector DEN contains the coefficients of the
denominator in descending powers of s.  The numerator coefficients
are returned in matrix NUM with as many rows as there are
outputs y.
```

See also ltimodels, ss and tf.

**Rational transfer functions.**

A scalar valued function of a complex variable is rational if it can be expressed as the ratio of two polynomials. It is proper if the order of the numerator polynomial is less than or equal to the order of the denominator polynomial. It is strictly proper if the order of the numerator polynomial is strictly less than that of the denominator polynomial. A matrix valued function of a complex variable is rational, proper or strictly proper if each element of the matrix function is respectively rational, proper or strictly proper.

**Example 23**

$$\begin{array}{rclll} \hat{G}(s) & = & \frac{s}{s^2+1} & \text{rational,} & \text{strictly proper} \\ \hat{G}(s) & = & \frac{s^2}{s^2+1} & \text{rational,} & \text{proper} \\ \hat{G}(s) & = & \frac{s^3}{s^2+1} & \text{rational,} & \text{not proper} \\ \hat{G}(s) & = & \frac{e^{-s}}{s^2+1} & \text{not rational} & \end{array}$$

We now demonstrate the following result:

*The transfer function of a finite-dimensional LTI system is always rational and proper.*

The above result follows from the fact that every element $\hat{G}_{ij}$ of the transfer function of a finite dimensional LTI system can be expressed as

$$\hat{G}_{ij} = \frac{N_{ij}}{\Delta} \tag{4.3}$$

where $\Delta(s) = \det(sI - A)$ is called the characteristic polynomial of $A$. It is a monic polynomial of order $n$. A monic polynomial is one whose coefficient of its highest order term is one. Thus

$$\det(sI - A) = s^n + \alpha_{n-1}s^{n-1} + \cdots + \alpha_1 s + \alpha_0$$

for some scalars $\alpha_0, \cdots, \alpha_{n-1}$. The order of the polynomial $\hat{N}_{ij}$ is at most $n$. It is less than $n$ if $D$ is zero.

To demonstrate the above fact, recall that the inverse of an invertible $n \times n$ matrix $M$ can be expressed as

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

where each element of the matrix $\operatorname{adj}(M)$ is equal to $\pm$ the determinant of an $(n-1) \times (n-1)$ submatrix of $M$. Hence $(sI - A)^{-1}$ can be expressed as

$$(sI - A)^{-1} = \frac{1}{\det(sI - A)} \operatorname{adj}(sI - A)$$

where each element of the matrix $\operatorname{adj}(sI-A)$ is equal to $\pm$ the determinant of an $(n-1)\times(n-1)$ submatrix of $sI - A$. Hence, if $A$ is $n \times n$, then each element of $\operatorname{adj}(sI - A)$ is a polynomial in $s$ whose degree is at most $n - 1$. Also, $\det(sI - A)$ is a polynomial of order $n$.

From the above form of $(sI - A)^{-1}$ it follows that if

$$\hat{G}(s) = C(sI - A)^{-1}B$$

where $A$ is $n \times n$, then

$$\hat{G}(s) = \frac{1}{\Delta(s)} N(s)$$

where $\Delta$ is the $n$-th order characteristic polynomial of $A$ , that is, $\Delta(s) = \det(sI - A)$ and each element $N_{ij}$ of $N$ is a polynomial of order less than or equal to $n-1$. So,

$$\hat{G}_{ij} = \frac{N_{ij}}{\Delta}$$

that is, each element $\hat{G}_{ij}$ of $\hat{G}$ is a rational function (ratio of two polynomials) and is strictly proper (order of numerator strictly less than order of denominator). When each element of $\hat{G}$ is rational and strictly proper, we say that $\hat{G}$ is rational and strictly proper.

If

$$\hat{G}(s) = C(sI - A)^{-1}B + D$$

then

$$\hat{G}(s) = \frac{1}{\Delta(s)} N(s)$$

where each element $N_{ij}$ of $N$ is a polynomial of order less than or equal to $n$. So, each element of $\hat{G}$ is a rational function and is proper (order of numerator less than or equal to order of denominator). When each element of $\hat{G}$ is rational and proper, we say that $\hat{G}$ is rational and proper.

So, we see that if $\hat{G}$ is the transfer function of a finite-dimensional LTI system then, $\hat{G}$ is rational and proper. Is the converse true? That is, is every proper rational $\hat{G}$ the transfer function of some finite-dimensional system and if so, what is(are) the system(s). This is answered in the next section.

**Invariance of transfer function under state transformations.** The transfer function of a linear time-invariant system describes the input output behavior of the system when all initial conditions are zero. Hence, it should be independent of the state variables used in a state space description of the system. We now demonstrate this explicitly.

Suppose $\hat{G}$ is the transfer function of a system described by

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y &= Cx + Du \end{aligned}$$

Then $\hat{G}(s) = C(sI - A)^{-1}B + D$. Consider any state transformation

$$x = T\xi$$

where $T$ is square and invertible. Then, the system can also be described by

$$\begin{aligned} \dot{\xi} &= \bar{A}\xi + \bar{B}u \\ y &= \bar{C}\xi + \bar{D}u \end{aligned}$$

with

$$\begin{array}{ll} \bar{A} = T^{-1}AT & \qquad \bar{B} = T^{-1}B \\ \bar{C} = CT & \qquad \bar{D} = D \end{array}$$

It can be readily verified that

$$\bar{C}(sI - \bar{A})^{-1}\bar{B} + \bar{D} = C(sI - A)^{-1}B + D$$

that is, the transfer function matrix is unaffected by a state transformation.

**Systems described by higher order differential equations.** Consider a system described by

$$A_l y^{(l)} + A_{l-1} y^{(l-1)} + \cdots + A_0 y = B_l u^{(l)} + B_{l-1} u^{(l-1)} + \cdots + B_0 u$$

where the input $u(t)$ is an $m$-vector, the output $y(t)$ is a $p$-vector, $l \geq 1$, and the matrix $A_l$ is invertible. How do we put such a system into state space form? By taking the Laplace transform of the above equation and considering zero initial values for $y, \dot{y}, \cdots, y^{(l-1)}$ and $u, \dot{u}, \cdots, u^{(l-1)}$, we obtain

$$[s^l A_l + s^{l-1} A_{l-1} + \cdots + A_0]\hat{y}(s) = [s^l B_l + s^{l-1} B_{l-1} + \cdots + B_0]\hat{u}(s)$$

Hence $\hat{y}(s) = \hat{G}(s)\hat{u}(s)$ where

$$\hat{G}(s) = [s^l A_l + s^{l-1} A_{l-1} + \cdots + A_0]^{-1}[s^l B_l + s^{l-1} B_{l-1} + \cdots + B_0]$$

To obtain a state space description of the system under consideration, we could look for matrices $A$, $B$, $C$, $D$ such that $\hat{G}(s) = C(sI - A)^{-1}B + D$. We consider this problem in a later section.

**General input-output systems.** Consider *any linear* input-output system described by

$$y(t) = \int_0^t G(t-\tau)u(\tau)\,d\tau$$

where $G$ is Laplace transformable. Then,

$$\hat{y} = \hat{G}\hat{u}$$

where $\hat{G}$ is the Laplace transform of $G$.

**Example 24**

$$\begin{aligned} \dot{x}(t) &= -x(t) + u(t-1) \\ y(t) &= x(t) \end{aligned}$$

Taking Laplace transform with $x(0) = 0$ yields

$$\begin{aligned} s\hat{x}(s) &= -\hat{x}(s) + e^{-s}\hat{u}(s) \\ \hat{y}(s) &= \hat{x}(s) \end{aligned}$$

Hence $\hat{y}(s) = \hat{G}(s)\hat{u}(s)$ with

$$\hat{G}(s) = \frac{e^{-s}}{s+1}$$

**Exercise 26** What is the transfer function of the following system?

$$\begin{aligned} \dot{x}_1(t) &= x_2(t) \\ \dot{x}_2(t) &= u(t-2) \\ y(t) &= x_1(t) \end{aligned}$$

**Exercise 27** Obtain the transfer function of the following system

$$\begin{aligned} \dot{x}(t) &= -x(t) + 2x(t-h) + u(t) \\ y(t) &= x(t) \end{aligned}$$

### 4.1.3 Poles and zeros

**Poles.** Suppose $\hat{G}$ is a matrix valued function of a complex variable. A complex number $\lambda$ is defined to be a pole of $\hat{G}$ if for some element $\hat{G}_{ij}$ of $\hat{G}$,

$$\lim_{s\to\lambda} \hat{G}_{ij}(s) = \infty\,.$$

**Example 25**

| $\hat{G}(s)$ | poles |
|---|---|
| $\frac{1}{s^2+3s+2}$ | -1, -2 |
| $\frac{s+1}{s^2+3s+2}$ | -2 |
| $\begin{pmatrix} \frac{1}{s} & \frac{e^{-s}}{s+1} \\ \frac{s^2}{s+1} & \frac{s+1}{s+2} \end{pmatrix}$ | 0, -1, -2 |

When $\hat{G}(s) = C(sI - A)^{-1}B + D$ it should be clear that if $\lambda$ is a pole of $\hat{G}$ then $\lambda$ must be an eigenvalue of $A$. However, as the following example illustrates, the converse is not always true.

**Example 26** If,

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad B = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \qquad C = \begin{pmatrix} -1 & 1 \end{pmatrix} \qquad D = 0\,,$$

then

$$\det(sI - A) = s^2 - 1$$

and

$$\hat{G}(s) = C(sI - A)^{-1}B = \frac{s-1}{s^2-1} = \frac{1}{s+1}\,.$$

Here $A$ has eigenvalues $-1$ and $1$ whereas $\hat{G}$ only has a single pole at $-1$.

We say that *a complex number $\lambda$ is a* pole of order $m$ *of a transfer function $\hat{G}$ if $\lambda$ is a pole of $\hat{G}$ and $m$ is the smallest integer such that $\lambda$ is not a pole of*

$$(s - \lambda)^m \hat{G}\,.$$

**Example 27** The transfer function

$$\hat{G}(s) = \frac{s+1}{(s-3)(s+4)^2}$$

has two poles: 3 and $-4$. The order of the pole at 3 is 1 whereas the order of the pole at $-4$ is 2.

Suppose $\lambda_1, \lambda_2, \cdots, \lambda_l$ are the poles of a rational transfer function $\hat{G}$. Then, it should be clear that we can express $\hat{G}$ as

$$\hat{G}(s) = \frac{1}{d(s)} N(s) \tag{4.4}$$

where

$$d(s) = (s - \lambda_1)^{m_1}(s - \lambda_2)^{m_2} \cdots (s - \lambda_l)^{m_l}\,,$$

with $m_i$ being the order of the pole $\lambda_i$, and $N$ is a matrix with polynomial elements.

**Zeroes.** Here, we only define zeroes for scalar-input scalar-output transfer functions. The definition of zeroes for MIMO systems is much more complicated. We say that a complex number $\eta$ is a zero of a scalar-input scalar-output transfer function $\hat{G}$ if

$$\hat{G}(\eta) = 0 \,. \tag{4.5}$$

**The significance of zeroes.** Consider a general LTI SISO system described by

$$y(t) = \int_0^t G(t-\tau)u(\tau)\, d\tau$$

Then, the transfer function $\hat{G}$ of this system is simply the Laplace transform of the impulse response $G$, that is,

$$\hat{G}(s) = \int_0^\infty e^{-st}G(t)\, dt \,.$$

If we assume that the impulse response $G$ converges to zero exponentially and $s$ is not a pole of $\hat{G}$, then it can be shown that the steady-state response of this system to an input

$$u(t) = e^{st}$$

is given by the output

$$\boxed{y(t) = \hat{G}(s)e^{st}}$$

So, we have the following conclusion:

*If $\eta$ is a zero of $\hat{G}$ then, the steady-state response of the system to the input $u(t) = e^{\eta t}$ is zero.*

Suppose that $\hat{G}$ is the transfer function of a system with a *real* impulse response that is,

$$\hat{G}(s) = \int_0^\infty e^{-st}G(t)\, dt$$

where $G(t)$ is real. Using the property that $e^{-\bar{s}t} = \overline{e^{-st}}$, the realness of $G(t)$ implies that $e^{-\bar{s}t}G(t) = \overline{e^{-st}}G(t) = \overline{e^{-st}G(t)}$; hence

$$\hat{G}(\bar{s}) = \int_0^\infty e^{-\bar{s}t}G(t)\, dt = \int_0^\infty \overline{e^{-st}G(t)}\, dt = \overline{\int_0^\infty e^{-st}G(t)\, dt} = \overline{\hat{G}(s)} \,,$$

that is,

$$\hat{G}(\bar{s}) = \overline{\hat{G}(s)} \,. \tag{4.6}$$

Let $\phi(s)$ be the argument of the complex number $\hat{G}(s)$, that is

$$\hat{G}(s) = |\hat{G}(s)|e^{\jmath\phi(s)} \tag{4.7}$$

where $\phi(s)$ is real. We now show that the steady state response of this system to

$$u(t) = e^{\alpha t}\sin(\omega t) \,,$$

where $\alpha$ and $\omega$ are real numbers, is given by

$$\boxed{y(t) = |\hat{G}(s)|e^{\alpha t}\sin(\omega t+\phi(s))} \tag{4.8}$$

where $s = \alpha + \jmath\omega$. Thus the output has the same form as the input; the non-negative real scalar $|\hat{G}(s)|$ can be regarded as the **system gain** at $s$ and $\phi(s)$ is the **system phase shift** at $s$.

We also obtain the following conclusion:

*If $\eta = \alpha + \jmath\omega$ is a zero of $\hat{G}$ then, the steady-state response of the system to the input $u(t) = e^{\alpha t}\sin(\omega t)$ is zero.*

We also see that the steady state response to a purely sinusoidal input

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

is given by

$$\boxed{y(t) = |\hat{G}(\jmath\omega)|\sin(\omega t+\phi(\jmath\omega))} \tag{4.9}$$

To demonstrate the above result, we first note that $u(t) = e^{\alpha t}\sin(\omega t)$ can be expressed as $u(t) = (e^{st} - e^{\bar{s}t})/(2\jmath)$ where $s = \alpha + \jmath\omega$. So, the corresponding steady state response is given by

$$y(t) = [\hat{G}(s)e^{st} + \hat{G}(\bar{s})e^{\bar{s}t}]/2\jmath\,.$$

Since $\overline{\hat{G}(s)} = \hat{G}(\bar{s})$, we obtain that

$$y(t) = [\hat{G}(s)e^{st} - \overline{\hat{G}(s)e^{st}}]/(2\jmath) = \Im\left(\hat{G}(s)e^{st}\right)\,.$$

Noting that

$$\hat{G}(s)e^{st} = |\hat{G}(s)|e^{\jmath\phi(s)}e^{\alpha t}e^{\jmath\omega t} = |\hat{G}(s)|e^{\alpha t}e^{\jmath(\omega t+\phi(s))}\,,$$

it now follows that

$$y(t) = |\hat{G}(s)|e^{\alpha t}\sin(\omega t + \phi(s))\,.$$

**Example 28 (Tuned vibration absorber.)** Consider a mechanical system subject to a sinusoidal disturbance input:

$$M\dot{q} + C\dot{q} + Kq = w(t)$$

where $w(t) = A\sin(\omega t)$.

Attaching a vibration absorber to the system results in

$$\begin{aligned} M\ddot{q} + C\dot{q} + Kq - k(\eta - q) &= w & (4.10)\\ m\ddot{\eta} + k(\eta - q) &= 0\,. & (4.11) \end{aligned}$$

Taking the Laplace transform of the above equations yields:

$$\begin{aligned} (Ms^2 + Cs + K + k)\hat{q} - k\hat{\eta} &= \hat{w} & (4.12)\\ (ms^2 + k)\hat{\eta} - k\hat{q} &= 0\,. & (4.13) \end{aligned}$$

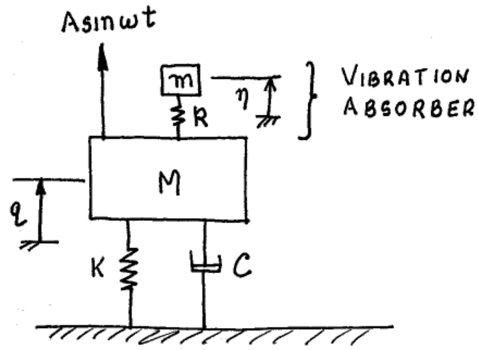



Figure 4.2: Tuned vibration absorber

We see that $\hat{q} = \hat{G}\hat{w}$ where

$$\hat{G} = \frac{ms^2 + k}{\Delta(s)}$$

and

$$\Delta(s) = (Ms^2 + Cs + K + k)(ms^2 + k) - k^2 \,.$$

This transfer function has zeros at $\pm \jmath\sqrt{k/m}$. if we choose

$$k/m = \omega^2$$

then the steady state response of $q$ to $w$ will be zero.

### 4.1.4 Discrete-time

The transfer function matrix for

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) \\ y(k) &= Cx(k) + Du(k) \end{aligned}$$

is defined by

$$\hat{G}(z) = C(zI - A)^{-1}B + D$$

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

### 4.1.5 Exercises

**Exercise 28** (BB in Laundromat.) Consider

$$\begin{aligned} m\ddot{\phi}_1 - m\Omega^2\phi_1 + \tfrac{k}{2}(\phi_1 - \phi_2) &= b_1 u \\ m\ddot{\phi}_2 - m\Omega^2\phi_2 - \tfrac{k}{2}(\phi_1 - \phi_2) &= b_2 u \\ y &= c_1\phi_1 + c_2\phi_2 \end{aligned}$$

For each of the following cases, obtain the system transfer function and its poles; consider $\omega := \sqrt{k/m} > \Omega$.

(a)

$$b_1 = 1 \qquad b_2 = 0$$
$$c_1 = 1 \qquad c_2 = 0$$

(b) (Self excited.)

$$b_1 = 1 \qquad b_2 = -1$$
$$c_1 = 1 \qquad c_2 = 0$$

(c) (Mass center measurement.)

$$b_1 = 1 \qquad b_2 = 0$$
$$c_1 = 1 \qquad c_2 = 1$$

(d) (Self excited and mass center measurement.)

$$b_1 = 1 \qquad b_2 = -1$$
$$c_1 = 1 \qquad c_2 = 1$$

Comment on your results.

## 4.2 Some transfer function properties

### 4.2.1 Power series expansion and Markov parameters

Suppose $\hat{G}$ is a transfer function for $(A, B, C, D)$, that is $\hat{G}(s) = C(sI - A)^{-1}B + D$. Then we will show that, for sufficiently large $s$, we have the following power series expansion:

$$\hat{G}(s) = D + \sum_{k=1}^{\infty} \frac{1}{s^k} CA^{k-1}B = D + \frac{1}{s}CB + \frac{1}{s^2}CAB + \frac{1}{s^3}CA^2B + \cdots .$$

To see this recall that, for sufficiently large $s$,

$$(sI - A)^{-1} = \frac{1}{s}I + \frac{1}{s^2}A + \frac{1}{s^3}A^2 + \cdots .$$

The result follows from this. The coefficient matrices

$$D, \quad CB, \quad CAB, \quad CA^2B, \quad \cdots$$

are called the Markov parameters of $\hat{G}$.

### 4.2.2 More properties*

Recall that a transfer function $\hat{G}$ for a finite dimensional system must be rational and proper, that is, it can be expressed as

$$\hat{G} = \frac{1}{d}N$$

where $d$ is a scalar valued polynomial and $N$ is a matrix of polynomials whose degrees are not greater than that of $d$. Without loss of generality, we consider $d$ to be monic. Then we can express $d$ and $N$ as

$$\begin{aligned} d(s) &= s^n + d_{n-1}s^{n-1} + \cdots + d_1 s + d_0 \\ N(s) &= s^n N_n + s^{n-1}N_{n-1} + \cdots + sN_1 + N_0 \end{aligned}$$

where $d_0, \ldots, d_{n-1}$ are scalars and $N_0, \ldots, N_n$ are constant matrices whose dimensions are the same as those of $\hat{G}(s)$. We now show that $\hat{G} = \frac{1}{d}N$ is a transfer function for a system $(A, B, C, D)$ if and only if $d$ satisfies

$$CA^k d(A)B = 0 \qquad \text{for} \quad k = 0, 1, 2, \cdots . \tag{4.14}$$

and

$$\begin{aligned} N_n &= D \\ N_{n-1} &= CB + d_{n-1}D \\ N_{n-2} &= CAB + d_{n-1}CB + d_{n-2}D \\ &\vdots \\ N_0 &= CA^{n-1}B + d_{n-1}CA^{n-2}B + \cdots + d_1CB + d_0D \end{aligned} \tag{4.15}$$

To obtain the above result, we first use the power series expansion for $\hat{G}$ to obtain

$$\frac{1}{d(s)}N(s) = D + \frac{1}{s}CB + \frac{1}{s^2}CAB + \frac{1}{s^3}CA^2B + \cdots$$

Multiplying both sides by $d(s)$ yields

$$\begin{aligned} & s^n N_n + s^{n-1}N_{n-1} + \cdots + sN_1 + N_0 \\ = \; & [s^n + d_{n-1}s^{n-1} + \cdots + d_1 s + d_0][D + \frac{1}{s}CB + \frac{1}{s^2}CAB + \frac{1}{s^3}CA^2B + \cdots] \end{aligned}$$

Since the coefficients of like powers of $s$ on both sides of the equation must be equal, this equation holds if and only if

$$\begin{aligned} N_n &= D \\ N_{n-1} &= CB + d_{n-1}D \\ N_{n-2} &= CAB + d_{n-1}CB + d_{n-2}D \\ &\vdots \\ N_0 &= CA^{n-1}B + d_{n-1}CA^{n-2}B + \cdots + d_1CB + d_0D \end{aligned}$$

and

$$\begin{aligned} 0 &= CA^nB + d_{n-1}CA^{n-1}B + \cdots + d_0CB \\ 0 &= CA^{n+1}B + d_{n-1}CA^nB + \cdots + d_0CAB \\ &\vdots \\ 0 &= CA^{n+k}B + d_{n-1}CA^{n+k-1}B + \cdots + d_0CA^kB \\ &\vdots \end{aligned}$$

Now note that the first set of equations above is the same as (4.15) while the second set is equivalent to (4.14) ∎

Suppose $\hat{G}$ is a transfer function for $(A, B, C, D)$ and $d$ is *any* polynomial for which $d(A) = 0$. It follows immediately that (4.15) holds. Hence we can express $\hat{G}$ as $\hat{G} = \frac{1}{d}N$ where the coefficients of $N$ are computed from (4.14).

**A proof of the Cayley Hamilton Theorem.** Consider any square matrix $A$ and let $\hat{G}(s) = (sI - A)^{-1}$. The function $\hat{G}$ is the transfer function for the system with $B = I$, $C = I$ and $D = 0$. Hence $\hat{G}$ can be expressed as $\hat{G} = \frac{1}{d}N$ where $d$ is the characteristic polynomial of $A$. Considering $k = 1$ in (4.14), it now follows that $d(A) = 0$.

## 4.3 State space realization of transfer functions

Here we consider the following type of problem: Given a matrix valued function $\hat{G}$ of a complex variable $s$, find constant matrices $A, B, C, D$ (or determine that none exist) so that

$$\hat{G}(s) = C(sI - A)^{-1}B + D\,. \tag{4.16}$$

When the above holds, we say that $(A, B, C, D)$ is a finite-dimensional state space realization of $\hat{G}$.

**Theorem 1** *Suppose $\hat{G}$ is a matrix valued function of a complex variable. Then $\hat{G}$ has a finite-dimensional state space realization if and only if $\hat{G}$ is rational and proper.*

PROOF: We have already seen the necessity of $\hat{G}$ being proper rational. We prove sufficiency in the next two sections by demonstrating specific realizations. ■

**Minimal realizations.** It should be clear from Section 4.1 that a given transfer function has an infinite number of state space realizations. One might expect that the dimension of every state space realization is the same. This is not even true as the next example illustrates. We say that a realization is a minimal realization if its dimension is less than or equal to that of any other realization. Clearly, all minimal realizations have the same dimension. Later on we will characterize minimal state space realizations.

**Example 29** Consider a SISO system of dimension two described by

$$A = \begin{pmatrix} 1 & -3 \\ -3 & 1 \end{pmatrix} \qquad B = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad C = \begin{pmatrix} 1 & 0 \end{pmatrix} \qquad D = 0$$

Here

$$\hat{G}(s) = C(sI - A)^{-1}B + D = \frac{s-4}{s^2 - 2s - 8} = \frac{1}{s+2}$$

This also has the following one dimensional realization

$$A = -2 \qquad B = 1 \qquad C = 1 \qquad D = 0$$

Note that this system is stable, whereas the original system is unstable.

## 4.4 Realization of SISO transfer functions

### 4.4.1 Controllable canonical form realization

Suppose that $\hat{G}$ is a proper rational scalar function of a complex variable $s$. If $\hat{G} \neq 0$, then for some positive integer $n$, the function $\hat{G}$ can be expressed as

$$\boxed{\hat{G}(s) = \frac{\beta_{n-1}s^{n-1} + \cdots + \beta_1 s + \beta_0}{s^n + \alpha_{n-1}s^{n-1} + \cdots + \alpha_1 s + \alpha_0} \quad + \quad d} \tag{4.17}$$

Note that the denominator polynomial is monic. This is illustrated in the following example.

**Example 30** Consider

$$G(s) = \frac{4s+8}{-2s+10}$$

This is a proper rational function. Also $4s + 8 = -2(-2s + 10) + 28$ which results in

$$G(s) = \frac{-2(-2s+10)+28}{-2s+10} = -2 + \frac{28}{-2s+10} = \frac{-14}{s-5} - 2\,.$$

When $\hat{G}$ is expressed in the form shown in (4.17) we now show that

$$\hat{G}(s) = C(sI - A)^{-1}B + D$$

where

$$\hat{G}(s) = C(sI - A)^{-1}B + D$$

where

$$\boxed{\begin{array}{lll} A & = & \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & & \ddots & & \vdots \\ \vdots & \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 1 \\ -\alpha_0 & -\alpha_1 & -\alpha_2 & \dots & -\alpha_{n-2} & -\alpha_{n-1} \end{pmatrix} \qquad B = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ \vdots \\ 0 \\ 1 \end{pmatrix} \\ \\ C & = & \begin{pmatrix} \beta_0 & \beta_1 & \beta_2 & \dots & \beta_{n-2} & \beta_{n-1} \end{pmatrix} \qquad D = d \end{array}} \tag{4.18}$$

It then follows that one can obtain a state space realization of any proper rational scalar SISO transfer function.

To demonstrate the above claim, let $\Delta$ be the denominator polynomial in the above expression for $\hat{G}$, that is,

$$\Delta(s) := \alpha_0 + \alpha_1 s + \ldots + s^n$$

and define

$$v(s) := \begin{pmatrix} 1 \\ s \\ \vdots \\ s^{n-1} \end{pmatrix}.$$

Then,

$$(sI - A)v(s) \;=\; \begin{pmatrix} s & -1 & 0 & \dots & 0 & 0 \\ 0 & s & -1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & s & -1 \\ \alpha_0 & \alpha_1 & \alpha_2 & \dots & \alpha_{n-2} & s+\alpha_{n-1} \end{pmatrix} \begin{pmatrix} 1 \\ s \\ \vdots \\ \vdots \\ s^{s-2} \\ s^{n-1} \end{pmatrix} \;=\; \begin{pmatrix} 0 \\ 0 \\ \vdots \\ \vdots \\ 0 \\ \Delta(s) \end{pmatrix} \;=\; \Delta(s)B$$

Hence, whenever $s$ is not an eigenvalue of $A$,

$$(sI - A)^{-1}B = \frac{1}{\Delta(s)}v(s)$$

and

$$\begin{aligned} C(sI - A)^{-1}B \;&=\; \frac{1}{\Delta(s)} \begin{pmatrix} \beta_0 & \beta_1 & \dots & \beta_{n-1} \end{pmatrix} \begin{pmatrix} 1 \\ s \\ \vdots \\ s^{n-1} \end{pmatrix} \\ &=\; \frac{\beta_0 + \beta_1 s + \cdots + \beta_{n-1}s^{n-1}}{\alpha_0 + \alpha_1 s + \cdots + \alpha_{n-1}s^{n-1} + s^n} \end{aligned}$$

So,

$$\hat{G}(s) = C(sI - A)^{-1}B + D$$

■

We will see later why this realization is called controllable.

### 4.4.2 Observable canonical form realization

Suppose $\hat{G}(s)$ is a proper rational scalar function of a complex variable $s$. If $\hat{G} \neq 0$, then, for some positive integer $n$,

$$\hat{G}(s) = \frac{\beta_{n-1}s^{n-1} + \cdots + \beta_1 s + \beta_0}{s^n + \alpha_{n-1}s^{n-1} + \cdots + \alpha_1 s + \alpha_0} \quad + \quad d$$

We will show that

$$\hat{G}(s) = C(sI - A)^{-1}B + D$$

where

$$A = \begin{pmatrix} 0 & 0 & \dots & 0 & -\alpha_0 \\ 1 & 0 & \dots & 0 & -\alpha_1 \\ 0 & 1 & \ddots & 0 & -\alpha_2 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & 1 & -\alpha_{n-1} \end{pmatrix} \qquad B = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \vdots \\ \beta_{n-1} \end{pmatrix}$$

$$C = \begin{pmatrix} 0 & 0 & \dots & 0 & 1 \end{pmatrix} \qquad D = d$$

• Noting that $C(sI - A)^{-1}B + D$ is a scalar and using the results of the previous section, we have

$$C(sI - A)^{-1}B + D = B^T(sI - A^T)^{-1}C^T + D = \hat{G}(s)$$

■

We will see later why this realization is called observable.

**MATLAB**

```
>> help tf2ss

 TF2SS  Transfer function to state-space conversion.
        [A,B,C,D] = TF2SS(NUM,DEN)  calculates the state-space
        representation:
                .
                x = Ax + Bu
                y = Cx + Du

        of the system:
                        NUM(s)
                H(s) = --------
                        DEN(s)

        from a single input.  Vector DEN must contain the coefficients of
        the denominator in descending powers of s.  Matrix NUM must
        contain the numerator coefficients with as many rows as there are
        outputs y.  The A,B,C,D matrices are returned in controller
        canonical form.  This calculation also works for discrete systems.
        To avoid confusion when using this function with discrete systems,
        always use a numerator polynomial that has been padded with zeros
        to make it the same length as the denominator.  See the User's
        guide for more details.
```

**Example 31** Consider

$$\hat{G}(s) = \frac{s^2}{s^2+1}$$

This can be written as

$$\hat{G}(s) = \frac{-1}{s^2+1} + 1$$

hence a state space representation is given by

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \qquad B = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \qquad C = \begin{pmatrix} -1 & 0 \end{pmatrix} \qquad D = 1$$

or,

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -x_1 + u \\ y &= -x_1 + u \end{aligned}$$

Matlab time:

```
>>  num =  [1 0 0]
>>  den =  [1 0 1]
>> [a,b,c,d]=tf2ss(num,den)

a =
     0    -1
     1     0

b =
     1
     0

c =
     0    -1

d =
     1
```

## 4.5 Realization of MIMO systems

Suppose $\hat{G}$ is an $p \times m$ proper rational transfer function. Because $\hat{G}$ is a proper rational function it has a representation of the form

$$\hat{G}(s) = \frac{1}{d(s)}N(s) + D \tag{4.19}$$

where $D$ is a constant $p \times m$ matrix, $N$ is a $p \times m$ matrix of polynomials and $d$ is a scalar valued monic polynomial.

### 4.5.1 Controllable realizations

Sometimes it is more convenient to express $\hat{G}$ as

$$\hat{G}(s) = N(s)\Delta(s)^{-1} + D \tag{4.20}$$

where $\Delta$ is a $p \times m$ matrix of polynomials. Let us consider this case. Note that if $\hat{G}$ is expressed as in (4.19) then, $\hat{G}$ can be expressed as in (4.20) with $\Delta(s) = d(s)I$ where $I$ is the $m \times m$ identity matrix.

Without loss of generality, we can express $N$ and $\Delta$ as

$$\begin{array}{rcr} N(s) & = & s^{l-1}N_{l-1} + s^{l-2}N_{l-2} + \cdots + sN_1 + N_0 \\ \Delta(s) & = & s^lI + s^{l-1}D_{l-1} + s^{l-2}D_{l-2} + \cdots + sD_1 + D_0 \end{array} \tag{4.21}$$

Here, $N_0$, …, $N_{l-1}$ are constant $p \times m$ matrices while $D_0$, …, $D_{l-1}$ are constant $m \times m$ matrices. With $n = ml$, let $A$, $B$ and $C$ be the following matrices of dimensions $n \times n$, $n \times m$

and $p \times n$, respectively,

$$A = \begin{pmatrix} 0 & I & 0 & \dots & 0 & 0 \\ 0 & 0 & I & \dots & 0 & 0 \\ \vdots & \vdots & & \ddots & & \vdots \\ \vdots & \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 & I \\ -D_0 & -D_1 & -D_2 & \dots & -D_{l-2} & -D_{l-1} \end{pmatrix}, \qquad B = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ \vdots \\ 0 \\ I \end{pmatrix}, \tag{4.22a}$$

$$C = \begin{pmatrix} N_0 & N_1 & N_2 & \dots & N_{l-2} & N_{l-1} \end{pmatrix}. \tag{4.22b}$$

Note that in the above matrices, $I$ and $0$ represent the $m \times m$ identity and zero matrices, respectively. We claim that $(A, B, C, D)$ is a realization of $\hat{G}$. This is proven at the end of this section.

Note that, when $G(s) = \frac{1}{d(s)} N(s)$ and $d(s) = \alpha_0 + \alpha_1 s + \cdots + \alpha_{n-1} s^{n-1} + s^n$, we have

$$A = \begin{pmatrix} 0 & I & 0 & \dots & 0 & 0 \\ 0 & 0 & I & \dots & 0 & 0 \\ \vdots & \vdots & & \ddots & & \vdots \\ \vdots & \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 & I \\ -\alpha_0 I & -\alpha_1 I & -\alpha_2 I & \dots & -\alpha_{l-2} I & -\alpha_{l-1} I \end{pmatrix}, \qquad B = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ \vdots \\ 0 \\ I \end{pmatrix}, \tag{4.23a}$$

$$C = \begin{pmatrix} N_0 & N_1 & N_2 & \dots & N_{l-2} & N_{l-1} \end{pmatrix} \tag{4.23b}$$

where $I$ and $0$ represent the $m \times m$ identity and zero matrices, respectively.

**Example 32** Consider

$$\hat{G}(s) = \begin{pmatrix} \frac{3}{s-2} \\ \\ \frac{s+4}{s^2+2s+3} \end{pmatrix}$$

The lowest common denominator (LCD) of the two elements of $\hat{G}(s)$ is

$$(s-2)(s^2+2s+3) = s^3 - s - 6 =: d(s)$$

and $\hat{G}(s) = N(s)d(s)^{-1}$ where

$$N(s) = \begin{pmatrix} 3s^2 + 6s + 9 \\ s^2 + 2s - 8 \end{pmatrix} = s^2 \begin{pmatrix} 3 \\ 1 \end{pmatrix} + s \begin{pmatrix} 6 \\ 2 \end{pmatrix} + \begin{pmatrix} 9 \\ -8 \end{pmatrix}$$

Hence, a controllable realization of $\hat{G}$ is given by

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 6 & 1 & 0 \end{pmatrix} \qquad B = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \qquad C = \begin{pmatrix} 9 & 6 & 3 \\ -8 & 2 & 1 \end{pmatrix} \qquad D = 0$$

**Example 33** Consider the strictly proper rational function,

$$\hat{G}(s) = \begin{pmatrix} \frac{1}{s+1} & \frac{1}{s^2+3s+2} \\ \\ 0 & \frac{1}{s+2} \end{pmatrix}.$$

Here $\hat{G} = \frac{1}{d}N$ where

$$d(s) = s^2 + 3s + 2$$

and

$$N(s) = \begin{pmatrix} s+2 & 1 \\ \\ 0 & s+1 \end{pmatrix} = s \begin{pmatrix} 1 & 0 \\ \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 2 & 1 \\ \\ 0 & 1 \end{pmatrix}$$

Hence a realization of $\hat{G}$ is given by

$$A = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -2 & 0 & -3 & 0 \\ 0 & -2 & 0 & -3 \end{pmatrix} \qquad B = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad C = \begin{pmatrix} 2 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix} \qquad D = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Here we show that $(A, B, C, D)$ as defined above is a realization of $\hat{G}$. Let $V$ be the $n \times m$ matrix of polynomials defined by

$$V(s) = \begin{pmatrix} I \\ sI \\ \vdots \\ s^{l-1}I \end{pmatrix}.$$

where $I$ represents the $m \times m$ identity matrix. Then we have

$$\begin{aligned} (sI - A)V(s) &= \begin{pmatrix} sI & -I & 0 & \dots & 0 & 0 \\ 0 & sI & -I & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & sI & -I \\ D_0 & D_1 & D_2 & \dots & D_{l-2} & sI + D_{l-1} \end{pmatrix} \begin{pmatrix} I \\ sI \\ \vdots \\ \vdots \\ s^{l-2}I \\ s^{l-1}I \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ \vdots \\ 0 \\ \Delta(s) \end{pmatrix} \\ &= B\Delta(s) \end{aligned}$$

Hence, whenever $s$ is not an eigenvalue of $A$,

$$(sI - A)^{-1}B = V(s)\Delta(s)^{-1}$$

Now note that

$$CV(s) = \begin{pmatrix} N_0 & N_1 & \cdots & N_{l-1} \end{pmatrix} \begin{pmatrix} I \\ sI \\ \vdots \\ \vdots \\ s^{l-2}I \\ s^{l-1}I \end{pmatrix} = N(s).$$

Hence

$$C(sI-A)^{-1}B = CV(s)\Delta(s)^{-1} = N(s)\Delta(s)^{-1}$$

and

$$\hat{G}(s) = C(sI-A)^{-1}B + D\,.$$

■

### 4.5.2 Observable realizations

Sometimes it is more convenient to express $\hat{G}$ as

$$\hat{G}(s) = \Delta(s)^{-1}N(s) + D \tag{4.24}$$

where $\Delta$ is a $p\times p$ matrix of polynomials. Note that if $\hat{G}$ is expressed as in (4.19) then, $\hat{G}$ can be expressed as in (4.24) with $\Delta(s) = d(s)I$ where $I$ is the $p\times p$ identity matrix.

Without loss of generality, we can express $N$ and $\Delta$ as

$$\begin{array}{rcr} N(s) & = & s^{l-1}N_{l-1} + s^{l-2}N_{l-2} + \cdots + sN_1 + N_0 \\ \Delta(s) & = & s^lI + s^{l-1}D_{l-1} + s^{l-2}D_{l-2} + \cdots + sD_1 + D_0 \end{array} \tag{4.25}$$

Here, $N_0,\ \ldots,\ N_{l-1}$ are constant $p\times m$ matrices while $D_0,\ \ldots,\ D_{l-1}$ are constant $p\times p$ matrices With $n = pl$, let $A$, $B$ and $C$ be the following matrices of dimensions $n\times n$, $n\times m$ and $p\times n$, respectively,

$$A \;=\; \begin{pmatrix} -D_{l-1} & I & 0 & \ldots & 0 & 0 \\ -D_{l-2} & 0 & I & \ldots & 0 & 0 \\ \vdots & \vdots & & \ddots & & \vdots \\ \vdots & \vdots & & & \ddots & \vdots \\ -D_1 & 0 & 0 & \ldots & 0 & I \\ -D_0 & 0 & 0 & \ldots & 0 & 0 \end{pmatrix}, \qquad B = \begin{pmatrix} N_{n-1} \\ N_{n-2} \\ \vdots \\ \vdots \\ N_1 \\ N_0 \end{pmatrix}, \tag{4.26a}$$

$$C \;=\; \begin{pmatrix} I & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}. \tag{4.26b}$$

Note that in the above matrices, $I$ and $0$ represent the $p\times p$ identity and zero matrices, respectively. We claim that $(A,B,C,D)$ is a realization of $\hat{G}$. This can be shown as follows.

Let $V$ be the $p\times n$ matrix of polynomials defined by

$$V(s) = \begin{pmatrix} s^{l-1}I & s^{l-2}I & \cdots & sI & I \end{pmatrix}.$$

where $I$ represents the $p \times p$ identity matrix. Then we have

$$
\begin{aligned}
V(s)(sI - A) &= \begin{pmatrix} s^{l-1}I & s^{l-2}I & \cdots & \cdots & sI & I \end{pmatrix} \begin{pmatrix} sI + D_{l-1} & -I & 0 & \ldots & 0 & 0 \\ D_{l-2} & sI & -I & \ldots & 0 & 0 \\ \vdots & \vdots & & \ddots & & \vdots \\ \vdots & \vdots & & & \ddots & \vdots \\ D_1 & 0 & 0 & \ldots & 0 & -I \\ D_0 & 0 & 0 & \ldots & 0 & sI \end{pmatrix} \\
&= \begin{pmatrix} \Delta(s) & 0 & \cdots & \cdots & 0 & 0 \end{pmatrix} \\
&= \Delta(s)C
\end{aligned}
$$

Hence, whenever $s$ is not an eigenvalue of $A$,

$$C(sI - A)^{-1} = \Delta(s)^{-1}V(s)$$

Now note that

$$V(s)B = \begin{pmatrix} s^{l-1}I & s^{l-2}I & \cdots & sI & I \end{pmatrix} \begin{pmatrix} N_{l-1} \\ N_{l-2} \\ \vdots \\ N_1 \\ N_0 \end{pmatrix} = N(s)\,.$$

Hence

$$C(sI - A)^{-1}B = \Delta(s)^{-1}V(s)B = \Delta(s)^{-1}N(s)$$

and

$$\hat{G}(s) = C(sI - A)^{-1}B + D\,.$$

■

### 4.5.3 Alternative realizations*

**SIMO systems**

Suppose $\hat{G}$ is an $p \times 1$ transfer matrix function and

$$\hat{G} = \begin{pmatrix} \hat{g}_1 \\ \vdots \\ \hat{g}_p \end{pmatrix}$$

that is $\hat{g}_i$ is the $i$-th element of $\hat{G}$. Choosing a monic polynomial $\Delta$ as the least common denominator polynomial of $\hat{g}_1, \ldots, \hat{g}_p$, we can write each $\hat{g}_i$ as

$$\hat{g}_i(s) = \frac{n_i(s)}{\Delta(s)} + d_i$$

where

$$\Delta(s) = s^n + \alpha_{n-1}s^{n-1} + \cdots + \alpha_1 s + \alpha_0$$

and

$$n_i(s) = \beta^i_{n-1}s^{n-1} + \cdots + \beta^i_1 s + \beta^i_0$$

Letting

$$A = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & & \ddots & & \vdots \\ \vdots & \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 1 \\ -\alpha_0 & -\alpha_1 & -\alpha_2 & \dots & -\alpha_{n-2} & -\alpha_{n-1} \end{pmatrix} \qquad B = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

$$C = \begin{pmatrix} \beta^1_0 & \beta^1_1 & \beta^1_2 & \dots & \beta^1_{n-2} & \beta^1_{n-1} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ \beta^p_0 & \beta^p_1 & \beta^p_2 & \dots & \beta^p_{n-2} & \beta^p_{n-1} \end{pmatrix} \qquad D = \begin{pmatrix} d_1 \\ \vdots \\ d_p \end{pmatrix}$$

we obtain

$$C(sI - A)^{-1}B + D = \hat{G}(s)$$

**Example 34** Consider

$$\hat{G}(s) = \begin{pmatrix} \frac{3}{s-2} \\ \\ \frac{s+4}{s^2+2s+3} \end{pmatrix}$$

The lowest common denominator (LCD) of the two elements of $\hat{G}(s)$ is

$$(s-2)(s^2+2s+3) = s^3 - s - 6 =: \Delta(s)$$

and $\hat{G}(s)$ can be expressed as

$$\hat{G}(s) = \begin{pmatrix} \frac{3s^2+6s+9}{\Delta(s)} \\ \\ \frac{s^2+2s-8}{\Delta(s)} \end{pmatrix}$$

Hence, a controllable realization of $\hat{G}$ is given by

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 6 & 1 & 0 \end{pmatrix} \qquad B = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \qquad C = \begin{pmatrix} 9 & 6 & 3 \\ -8 & 2 & 1 \end{pmatrix}$$

**MIMO systems**

Suppose $\hat{G}$ is an $p \times m$ transfer matrix function and

$$\hat{G} = \left( \begin{array}{ccc} \hat{G}_1 & \dots & \hat{G}_m \end{array} \right)$$

that is $\hat{G}_i$ is the $i$-th column of $\hat{G}$. Using the results of the previous section, for $i = 1, \dots, m$, we can obtain matrices $A_i, B_i, C_i, D_i$ such that

$$\hat{G}_i(s) = C_i(sI - A_i)^{-1}B_i + D_i$$

Let

$$A = \begin{pmatrix} A_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & A_m \end{pmatrix} \qquad B = \begin{pmatrix} B_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & B_m \end{pmatrix}$$

$$C = \left( \begin{array}{ccc} C_1 & \dots & C_m \end{array} \right) \qquad D = \left( \begin{array}{ccc} D_1 & \dots & D_m \end{array} \right)$$

Then

$$C(sI - A)^{-1}B + D = \hat{G}(s)$$

• Nonminimal realization, in general.

**Example 35**

$$\hat{G}(s) = \begin{pmatrix} \frac{1}{s} & \frac{1}{s^2} \\ & \\ 0 & \frac{1}{s} \end{pmatrix}$$

The first and second columns of $\hat{G}$ have realizations

$$A_1 = 0 \qquad B_1 = 1 \qquad C_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad D_1 = 0$$

and

$$A_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad B_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \qquad C_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad D_2 = 0$$

respectively. Hence a realization of $\hat{G}$ is given by

$$A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \qquad B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix} \qquad C = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad D = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Note that this transfer function also has the following lower dimensional realization:

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad C = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad D = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

## Exercises

**Exercise 29** Obtain a state space realization of the transfer function,

$$\hat{G}(s) = \frac{s^2 + 3s + 2}{s^2 + 5s + 6}\,.$$

Is your realization minimal?

**Exercise 30** Obtain a state space realization of the transfer function,

$$\hat{G}(s) = \begin{pmatrix} 1 & \dfrac{1}{s-1} \\ \dfrac{1}{s+1} & \dfrac{1}{s^2-1} \end{pmatrix}.$$

**Exercise 31** Obtain a state space realization of the following transfer function.

$$\hat{G}(s) = \begin{pmatrix} \frac{s+2}{s+1} & \frac{1}{s+3} \\ \frac{5}{s+1} & \frac{5s+1}{s+2} \end{pmatrix}$$

**Exercise 32** Obtain a state space realization of the transfer function,

$$\hat{G}(s) = \begin{pmatrix} \dfrac{1}{s-1} & \dfrac{1}{s-1} \\ \dfrac{1}{s+1} & -\dfrac{1}{s+1} \end{pmatrix}.$$

**Exercise 33** Obtain a state space realization of the transfer function,

$$\hat{G}(s) = \begin{pmatrix} \dfrac{s}{s-1} & \dfrac{1}{s+1} \end{pmatrix}.$$

Is your realization minimal?

**Exercise 34** Obtain a state space representation of the following transfer function.

$$\hat{G}(s) = \begin{pmatrix} \dfrac{s^2+1}{s^2-1} & \dfrac{2}{s-1} \end{pmatrix}$$

**Exercise 35** Obtain a state space representation of the following input-output system.

$$2\frac{d^3y}{dt^3} + 6\frac{d^2y}{dt^2} + 4\frac{dy}{dt} + 2y = 6\frac{d^2u}{dt^2} + 10\frac{du}{dt} + 8u$$

**Exercise 36** (a) Obtain a state space realization of the following single-input single-output system.

$$\ddot{y} - 3\dot{y} - 4y = \ddot{u} - 2\dot{u} - 8u$$

(b) Is your realization minimal?

**Exercise 37** Obtain a state space realization of the following input-output system.

$$\begin{aligned} \ddot{y}_1 - y_1 + y_2 &= \dot{u}_1 + u_1 \\ \ddot{y}_2 + y_2 - y_1 &= u_2 \end{aligned}$$

**Exercise 38** Obtain a state space realization of the following input-output system.

$$\begin{aligned} \dot{y}_1 &= \dot{u}_1 + u_2 \\ \dot{y}_2 &= u_1 + \dot{u}_2 \end{aligned}$$

**Exercise 39** Obtain a state space realization of the following input-output system.

$$\begin{aligned} \dot{y}_1 + y_2 &= \dot{u}_2 + u_1 \\ \dot{y}_2 + y_1 &= \dot{u}_1 + u_2 \end{aligned}$$

**Exercise 40** Consider the single-input single-output system described by

$$\begin{aligned} \dot{x}_1 &= -2x_2 + u \\ \dot{x}_2 &= -2x_1 + u \\ y &= x_1 - x_2 \end{aligned}$$

(a) Compute $e^{At}$ for the matrix,

$$A = \begin{pmatrix} 0 & -2 \\ -2 & 0 \end{pmatrix}.$$

(b) Obtain the impulse response of the system.

(c) What is the transfer function of this system?

**Exercise 41** Consider the single-input single-output system described by

$$\begin{aligned} \dot{x}_1 &= -3x_1 + x_2 + w \\ \dot{x}_2 &= x_1 - 3x_2 + w \\ y &= x_1 + x_2 \end{aligned}$$

(a) Compute $e^{At}$ for the matrix,

$$A = \begin{pmatrix} -3 & 1 \\ 1 & -3 \end{pmatrix}$$

(b) Obtain the impulse response of the system.

(c) What is the transfer function of this system?

(d) What are the poles and zeros of the transfer function?

# Part II

# System Behavior and Stability

# Chapter 5

# Behavior of (LTI) systems : I

## 5.1 Initial stuff

### 5.1.1 Continuous-time

Here we look at the behavior of linear time-invariant systems described by

$$\boxed{\dot{x} = Ax} \tag{5.1}$$

where the state $x(t)$ is a complex $n$-vector and $A$ is a complex $n \times n$ matrix. (The reason for complexifying things will be seen shortly.) Sometimes $A$ is called the system matrix.

**Solutions.** By a solution of system (5.1) we mean any continuous function $x(\cdot)$ which is defined on some time interval $[t_0, \infty)$ and which satisfies (5.1). Sometimes we are interested in solutions which satisfy an initial condition of the form

$$\boxed{x(t_0) = x_0} \tag{5.2}$$

where $t_0$ is called the initial time and $x_0$ is called the initial state. As a consequence of the linearity of the system, one can prove the following result:

*For each initial condition, system (5.1) has a unique solution.*

Shortly, we will study the nature of these solutions.

Since the system is time-invariant, knowledge of the solutions for zero initial time, that is $t_0 = 0$, yields knowledge of the solutions for any other initial time. For suppose $\tilde{x}(\cdot)$ is a solution corresponding to $\tilde{x}(0) = x_0$. Then, the solution corresponding to $x(t_0) = x_0$ is given by

$$x(t) = \tilde{x}(t-t_0)\,.$$

Since the system is linear, one can readily show that *solutions depends linearly on the initial state;* more specifically, there is a matrix valued function $\Phi$ (called the state transition matrix) such that at each time $t$ the solution due to $x(0) = x_0$ is uniquely given by

$$x(t) = \Phi(t)x_0$$

Later, we develop explicit expressions for $\Phi$.

### 5.1.2 Discrete-time

Here we look at the behavior of linear time-invariant systems described by

$$\boxed{x(k+1) = Ax(k)}$$

The state $x(k)$ is a complex $n$-vector and $A$ is a complex $n \times n$ matrix. Sometimes $A$ is called the system matrix.

**Solutions:** Consider *initial condition*:

$$x(k_0) = x_0 \tag{5.3}$$

Since the system is time-invariant we can, wlog (without loss of generality), consider only zero initial time, that is, $k_0 = 0$. For suppose $x(\cdot)$ is a solution corresponding

$$\tilde{x}(k_0) = x_0$$

Then

$$\tilde{x}(k) = x(k-k_0)$$

where $x(\cdot)$ is the solution corresponding to (5.3).

For $k \geq k_0 = 0$, we have existence and uniqueness of solutions; specifically,

$$\boxed{x(k) = A^k x_0}$$

*Proof:* (Example of a proof by induction) Suppose there is a $k^* \geq 0$ for which the above holds, that is,

$$x(k^*) = A^{k^*} x_0$$

Then

$$\begin{aligned} x(k^*+1) &= Ax(k^*) \\ &= AA^{k^*} x_0 \\ &= A^{(k^*+1)} x_0 \end{aligned}$$

So, if it holds for $k^*$ it holds for $k^* + 1$. Since, $x(0) = A^0 x_0$, it holds for $k^* = 0$. It now follows by induction that it holds for all $k^* \geq 0$. ■

*Solution depends linearly on initial state*; specifically,

$$x(k) = \Phi(k)x(0)\,, \qquad \text{for} \quad k \geq 0$$

where $\Phi(k) = A^k$.

## 5.2 Scalar systems

### 5.2.1 Continuous-time

Consider a scalar system described by

$$\boxed{\dot{x} = ax}$$

where $x(t)$ and $a$ are scalars. All solutions are given by:

$$\boxed{x(t) = e^{at}x_0 \; ; \qquad x_0 := x(0)}$$

**Real $a$.** So,

(a) If $a < 0$ you decay with increasing age.

(b) If $a > 0$ you grow with increasing age.

(c) If $a = 0$ you stay the same; every state is an equilibrium state.

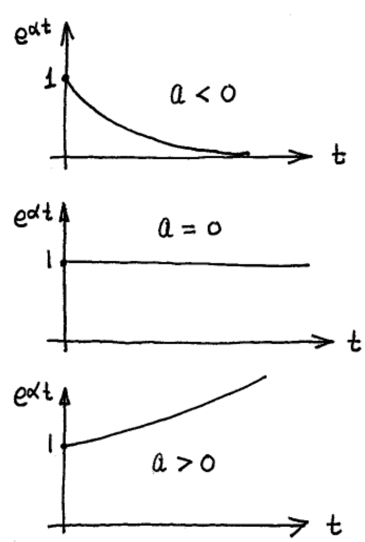



Figure 5.1: The three types of modes for real scalar systems: $\alpha = a$

**Example 36** (A first example.) Here

$$a = -c/m < 0$$

Thus, all solutions are of the form $e^{-(c/m)t}x_0$; hence they decay exponentially.

**Complex $a$.** Suppose $a$ is a genuine complex number (unreal); that is,

$$a = \alpha + \jmath\omega$$

where $\alpha$ and $\omega$ are real numbers. We call $\alpha$ the real part of $a$ and denote this by $\alpha = \Re(a)$. Similarly, we call $\omega$ the imaginary part of $a$ and denote this by $\omega = \Im(a)$.

We now show that the growth behavior of $e^{at}$ depends only on the real part of $a$. To this end, we first note that

$$e^{at} \;=\; e^{(\alpha+\jmath\omega)t} = e^{(\alpha t+\jmath\omega t)} = e^{\alpha t}e^{\jmath\omega t}\,,$$

and, hence, $|e^{at}| = |e^{\alpha t}|\,|e^{\jmath\omega t}|$. Since $e^{\alpha t}$ is positive, $|e^{\alpha t}| = e^{\alpha t}$. Since,

$$e^{\jmath\omega t} = \cos\omega t + \jmath\sin\omega t\,,$$

we obtain that $|e^{\jmath\omega t}| = 1$. Hence,

$$\boxed{|e^{at}| = e^{\alpha t} \qquad \text{where} \qquad \alpha = \Re(a)}$$

Thu, $|x(t)| = |e^{at}x_0| = |e^{at}||x_0| = e^{\alpha t}|x_0|$. So,

(a) If $\Re(a) < 0$ you decay with increasing age.

(b) If $\Re(a) > 0$ you grow with increasing age.

(c) If $\Re(a) = 0$ your magnitude remains constant.

### 5.2.2 First order complex and second order real.

Suppose $a$ is a genuine complex number (unreal); that is,

$$a = \alpha + \jmath\omega \qquad \text{where} \qquad \alpha = \Re(a),\;\; \omega = \Im(a)$$

Since

$$x = \xi_1 + \jmath\xi_2 \qquad \text{where} \qquad \xi_1 = \Re(x),\;\; \xi_2 = \Im(x)$$

Then

$$\begin{array}{rcl} \dot{\xi}_1 &=& \alpha\xi_1 - \omega\xi_2 \\ \dot{\xi}_2 &=& \omega\xi_1 + \alpha\xi_2 \end{array} \qquad = \qquad \underbrace{\begin{pmatrix} \alpha & -\omega \\ \omega & \alpha \end{pmatrix}\xi}_{\text{axisymmetric spacecraft}}$$

**Exercise 42** Verify the last set of equations.

Also,

$$\begin{array}{rcl} \xi_1(t) & = & e^{\alpha t}\cos(\omega t)\xi_{10} - e^{\alpha t}\sin(\omega t)\xi_{20} \\ \xi_2(t) & = & e^{\alpha t}\sin(\omega t)\xi_{10} + e^{\alpha t}\cos(\omega t)\xi_{20} \end{array} = \begin{pmatrix} e^{\alpha t}\cos(\omega t) & -e^{\alpha t}\sin(\omega t) \\ e^{\alpha t}\sin(\omega t) & e^{\alpha t}\cos(\omega t) \end{pmatrix} \xi_0$$

where $\xi_{10} = \xi_1(0), \xi_{20} = \xi_2(0)$

**Exercise 43** Verify the last set of equations. Hint: $e^{\jmath\omega t} = \cos\omega t + \jmath \sin\omega t$

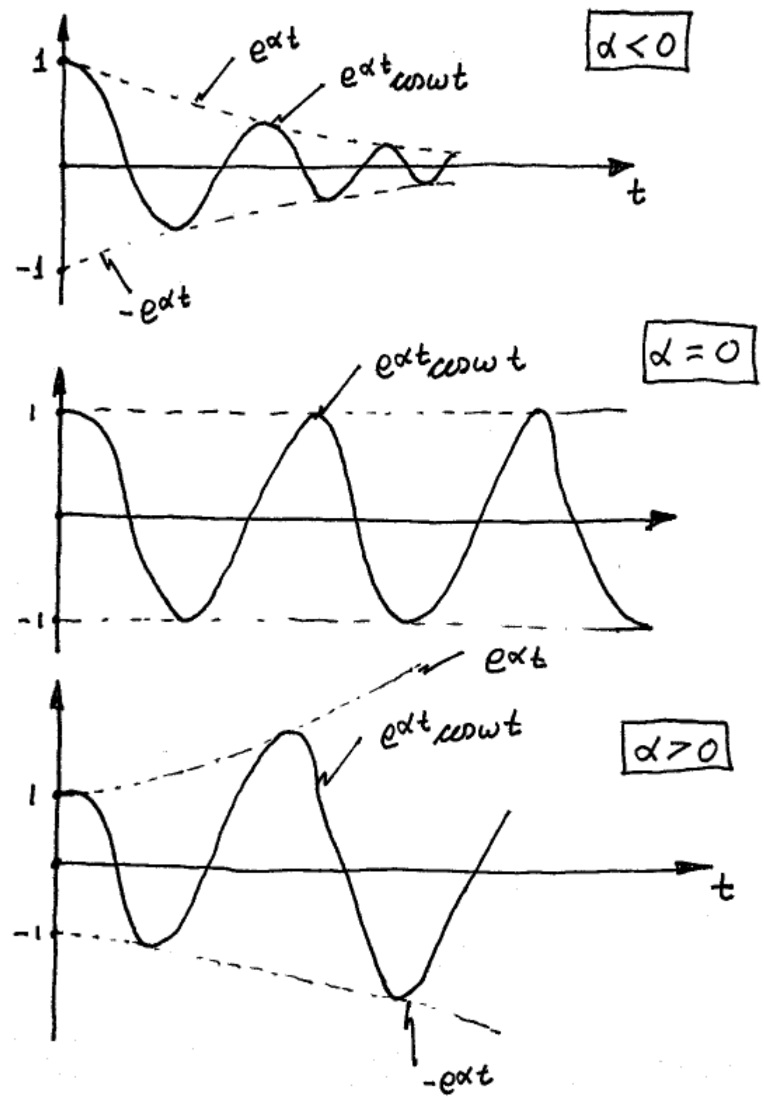



Figure 5.2: The three complex modes

### 5.2.3 DT

$$\boxed{x(k+1) = ax(k)}$$

where $x(k)$ and $a$ are scalars. All solutions are given by:

$$\boxed{x(k) = a^k x_0 \ ; \qquad x_0 := x(0)}$$

Hence, $|x(k)| = |a^k x_0| = |a|^k |x_0|$. So,

(a) If $|a| < 1$ you decay with increasing age.

(b) If $|a| > 1$ you grow with increasing age.

(c) If $|a| = 1$ your magnitude remains the same.

(d) If $a = 1$ you stay the same; every state is an equilibrium state.

## 5.3 Eigenvalues and eigenvectors

Shortly, we shall see that the behavior of a linear time invariant system is completely determined by the eigenvalues and eigenvectors of its system matrix.

Suppose $A$ is a complex $n \times n$ matrix. (The use of the word complex also includes real. To indicate that something is complex but not real we say genuine complex.)

**Defn.** *A complex number $\lambda$ is an* eigenvalue *of $A$ if there is a nonzero vector $v$ such that*

$$\boxed{Av = \lambda v}$$

*The nonzero vector $v$ is called an* eigenvector *of $A$ corresponding to $\lambda$.*

So, an eigenvector of $A$ is a nonzero complex $n$-vector with the property that there is a complex number $\lambda$ such that the above relationship holds or, equivalently,

$$\boxed{(\lambda I - A)v = 0}$$

Since $v$ is nonzero, $\lambda I - A$ must be singular; so,

$$\boxed{\det(\lambda I - A) = 0}$$

• The characteristic polynomial of $A$:

$$\text{charpoly}(A) := \det(sI - A)$$

Note that $\det(sI - A)$ is an $n$-th order monic polynomial, that is, it has the form

$$\det(sI - A) = a_0 + a_1 s + \ldots + s^n$$

• We conclude that a complex number $\lambda$ is an eigenvalue of $A$ if and only if it is a root of the $n$th order characteristic polynomial of $A$. Hence, $A$ has at least one eigenvalue and at most $n$ distinct eigenvalues. Suppose $A$ has $l$ distinct eigenvalues, $\lambda_1, \lambda_2, \ldots, \lambda_l$; since these are the distinct roots of the characteristic polynomial of $A$, we must have

$$\det(sI - A) = \prod_{i=1}^{l} (s - \lambda_i)^{m_i}$$

• The integer $m_i$ is called the algebraic multiplicity of $\lambda_i$

• It should be clear that if $v^1$ and $v^2$ are any two eigenvectors corresponding to the same eigenvalue $\lambda$ and $\xi_1$ and $\xi_2$ are any two numbers then (provided it is nonzero) $\xi_1 v^1 + \xi_2 v^2$ is also an eigenvector for $\lambda$. The set of eigenvectors corresponding to $\lambda$ along with the zero vector is called the eigenspace of $A$ associated with $\lambda$. This is simply the null space of $\lambda I - A$

• The geometric multiplicity of $\lambda$ is the nullity of $\lambda I - A$, that is, it is the maximum number of linearly independent eigenvectors associated with $\lambda$.

• It follows that $A$ is invertible if and only if all its eigenvalues are nonzero.

**Example 37** (An eigenvalue-eigenvector calculation)

$$A = \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix}$$

The characteristic polynomial of $A$ is given by

$$\begin{aligned} \det(sI - A) &= \det \begin{pmatrix} s-3 & 1 \\ 1 & s-3 \end{pmatrix} \\ &= (s-3)(s-3) - 1 = s^2 - 6s + 8 \\ &= (s-2)(s-4) \end{aligned}$$

The roots of this polynomial yield *two distinct real eigenvalues:*

$$\lambda_1 = 2, \qquad \lambda_2 = 4$$

To compute eigenvectors for $\lambda_1$, we use $(\lambda_1 I - A)v = 0$ to obtain

$$\begin{aligned} -v_1 + v_2 &= 0 \\ v_1 - v_2 &= 0 \end{aligned}$$

which is equivalent to

$$v_1 - v_2 = 0$$

So, only one linearly independent eigenvector for $\lambda_1$; let's take

$$v = v^1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

In a similar fashion, $\lambda_2$ has one linearly independent eigenvector; we take

$$v^2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

To check the above calculations, note that:

$$Av^1 = \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \lambda_1 v^1$$

Similarly for $\lambda_2$.

**Example 38** (Another eigenvalue-eigenvector calculation)

$$A = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

The characteristic polynomial of $A$ is given by

$$\begin{aligned} \det(sI - A) &= \det \begin{pmatrix} s-1 & -1 \\ 1 & s-1 \end{pmatrix} &= (s-1)(s-1)+1 \\ &= s^2 - 2s + 2 \end{aligned}$$

Computing the roots of this polynomial yields *two distinct complex eigenvalues*:

$$\lambda_1 = 1 + \jmath, \qquad \lambda_2 = 1 - \jmath$$

To compute eigenvectors for $\lambda_1$, we use $(\lambda_1 I - A)v = 0$ to obtain

$$\begin{aligned} \jmath v_1 - v_2 &= 0 \\ v_1 + \jmath v_2 &= 0 \end{aligned}$$

which is equivalent to

$$-\jmath v_1 + v_2 = 0$$

So, only one linearly independent eigenvector; let's take

$$v = v^1 = \begin{pmatrix} 1 \\ \jmath \end{pmatrix}$$

(We could also take $\begin{pmatrix} \jmath & 1 \end{pmatrix}^T$.) In a similar fashion, $\lambda_2$ has one linearly independent eigenvector; we take

$$v^2 = \begin{pmatrix} 1 \\ -\jmath \end{pmatrix}$$

To check the above calculations, note that

$$Av^1 = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ \jmath \end{pmatrix} = \begin{pmatrix} 1+\jmath \\ -1+\jmath \end{pmatrix} = (1+\jmath) \begin{pmatrix} 1 \\ \jmath \end{pmatrix} = \lambda_1 v^1$$

Similarly for $\lambda_2$.

**Example 39**

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

**Exercise 44** Compute expressions for eigenvalues and linearly independent eigenvectors of

$$A = \begin{pmatrix} \alpha & \omega \\ -\omega & \alpha \end{pmatrix} \qquad\qquad \alpha, \omega \neq 0$$

### 5.3.1 Real $A$

Recall Examples 37 and 38. They illustrate the following facts.

• If $A$ is real, its eigenvalues and eigenvectors occur in complex conjugate pairs.

To see this, suppose $\lambda$ is an eigenvalue of $A$ with eigenvector $v$. Then

$$Av = \lambda v$$

Taking complex conjugate of both sides of this equation yields

$$\bar{A}\bar{v} = \bar{\lambda}\bar{v}$$

Since $A$ is real, $\bar{A} = A$; hence

$$A\bar{v} = \bar{\lambda}\bar{v}$$

that is, $\bar{\lambda}$ is an eigenvalue of $A$ with eigenvector $\bar{v}$ ■

• Real eigenvectors for real eigenvalues.

• Genuine complex eigenvectors for (genuine) complex eigenvalues.

• The real and imaginary parts of a complex eigenvector are linearly independent.

**Example 40** Consider any non-zero angular velocity matrix of the form:

$$A = \begin{pmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{pmatrix}$$

$$\text{charpoly}(A) = s(s^2 + \Omega^2), \qquad \Omega := \sqrt{\omega_1^2 + \omega_2^2 + \omega_3^2}$$

Hence, $A$ has one real eigenvalue 0 and two complex conjugate eigenvalues, $\jmath\Omega, -\jmath\Omega$.

### 5.3.2 MATLAB

```
>> help eig

 EIG    Eigenvalues and eigenvectors.
        EIG(X) is a vector containing the eigenvalues of a square
        matrix X.

        [V,D] = EIG(X) produces a diagonal matrix D of
        eigenvalues and a full matrix V whose columns are the
        corresponding eigenvectors so that X*V = V*D.

        [V,D] = EIG(X,'nobalance') performs the computation with
        balancing disabled, which sometimes gives more accurate results
        for certain problems with unusual scaling.

>> help poly

 POLY   Characteristic polynomial.
        If A is an N by N matrix, POLY(A) is a row vector with
        N+1 elements which are the coefficients of the
        characteristic polynomial, DET(lambda*EYE(A) - A) .
        If V is a vector, POLY(V) is a vector whose elements are
        the coefficients of the polynomial whose roots are the
        elements of V . For vectors, ROOTS and POLY are inverse
        functions of each other, up to ordering, scaling, and
        roundoff error.
        ROOTS(POLY(1:20)) generates Wilkinson's famous example.

>> help roots

 ROOTS  Find polynomial roots.
        ROOTS(C) computes the roots of the polynomial whose coefficients
        are the elements of the vector C. If C has N+1 components,
        the polynomial is C(1)*X^N + ... + C(N)*X + C(N+1).
```

### 5.3.3 Companion matrices

These matrices will be useful later in control design. We met these matrices when constructing state space descriptions for systems described by a single higher order linear differential equation.

**Exercise 45** Compute the characteristic polynomial of

$$\begin{pmatrix} 0 & 1 \\ -a_0 & -a_1 \end{pmatrix}$$

Hence, write down the characteristic polynomials for the $A$ matrices associated with the unattached mass and the damped linear oscillator.

**Exercise 46** Compute the characteristic polynomial of

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -a_0 & -a_1 & -a_2 \end{pmatrix}$$

**Fact 1** *The characteristic polynomial of*

$$\begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & & \ddots & & \vdots \\ \vdots & \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 1 \\ -a_0 & -a_1 & -a_2 & \dots & -a_{n-2} & -a_{n-1} \end{pmatrix}$$

*is given by*

$$p(s) = a_0 + a_1 s + \ldots + a_{n-1}s^{n-1} + s^n$$

*Hence, given any bunch of $n$ complex numbers, there is a companion matrix whose eigenvalues are exactly these numbers.*

**Exercise 47** What is the real $2 \times 2$ companion matrix with eigenvalues $1 + \jmath$ , $1 - \jmath$?

**Exercise 48** What is the companion matrix whose eigenvalues are $-1$, $-2$, and $-3$?

## 5.4 Behavior of continuous-time systems

### 5.4.1 System significance of eigenvectors and eigenvalues

In this section, we demonstrate that the eigenvalues and eigenvectors of a square matrix $A$ play a fundamental role in the behavior of the linear system described by

$$\dot{x} = Ax\,. \tag{5.4}$$

We first demonstrate the following result.

*The system $\dot{x} = Ax$ has a non-zero solution of the form*

$$\boxed{x(t) = e^{\lambda t} v} \tag{5.5}$$

*if and only if $\lambda$ is an eigenvector of $A$ and $v$ is a corresponding eigenvector.*

To demonstrate the above claim, note that if $x(t) = e^{\lambda t}v$, then

$$\dot{x} = e^{\lambda t}\lambda v \qquad \text{and} \qquad Ax(t) = Ae^{\lambda t}v = e^{\lambda t}Av\,.$$

Hence, $x(t) = e^{\lambda t}v$ is a non-zero solution of $\dot{x} = Ax$ if and only if $v$ is nonzero and

$$e^{\lambda t}\lambda v = e^{\lambda t}Av\,.$$

Since $e^{\lambda t}$ is nonzero, the above condition is equivalent to

$$Av = \lambda v\,.$$

Since $v$ is non-zero, this means $\lambda$ is an eigenvector of $A$ and $v$ is a corresponding eigenvector. ■

A solution of the form $e^{\lambda t}v$ is a very special type of solution and is sometimes called a mode of the system. Note that if $x(t) = e^{\lambda t}v$ is a solution, then $x(0) = v$, that is, $v$ is the initial value of $x$. Hence, we can make the following statement.

*If $v$ is an eigenvector of $A$, then the solution to $\dot{x} = Ax$ with initial condition $x(0) = v$ is given by $x(t) = e^{\lambda t}v$ where $\lambda$ is the eigenvalue corresponding to $v$.*

If $v$ is an eigenvector then, $e^{\lambda t}v$ is also an eigenvector for each $t$. Also, considering the magnitude of $e^{\lambda t}v$, we obtain

$$||e^{\lambda t}v|| = |e^{\lambda t}|\,\|v\| = e^{\alpha t}\,\|v\|$$

where $\alpha$ is the real part of $\lambda$. So,

(a) Once an eigenvector, always an eigenvector!

(b) If $\Re(\lambda) < 0$ you decay with increasing age.

(c) If $\Re(\lambda) > 0$ you grow with increasing age.

(d) If $\Re(\lambda) = 0$ your magnitude remains constant.

(e) If $\lambda = 0$ you stay put; every eigenvector corresponding to $\lambda = 0$ is an equilibrium state.

**Example 41** The matrix $A$ of example 37 had 2 as an eigenvalue with corresponding eigenvector $[\ 1 \quad 1\ ]^T$. Hence with initial state,

$$x_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

the system $\dot{x} = Ax$ has the following solution:

$$x(t) = \begin{pmatrix} e^{2t} \\ e^{2t} \end{pmatrix}$$

What is the solution for

$$x_0 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}?$$

**Exercise 49** Prove the following statement. *Suppose that $v^1, \ldots, v^m$ are eigenvectors of $A$ with corresponding eigenvalues, $\lambda_1, \ldots, \lambda_m$. Then the solution of (5.4) with initial state*

$$x_0 = \xi_{10}v^1 + \ldots + \xi_{m0}v^m\,,$$

*where $\xi_{10}, \ldots, \xi_{m0}$ are scalars, is given by*

$$x(t) = \xi_{10}e^{\lambda_1 t}v^1 + \ldots + \xi_{m0}e^{\lambda_m t}v^m\,.$$

**Behaviour in the $v$ subspace.** Suppose $v$ is an eigenvector of $A$ with corresponding eigenvalue $\lambda$ and consider any initial state of the form $x_0 = \xi_0 v$. Thus, the initial state $x_0$ is in the 1-dimensional subspace spanned by $v$. We have shown that the solution corresponding to $x(0) = x_0$ has the form $x(t) = \xi(t)v$; hence $x(t)$ remains in this subspace. Forgetting about the actual solution, we note that $\dot{\xi}v = \dot{x} = Ax = A(\xi v) = \lambda\xi v$. Hence

$$\dot{\xi} = \lambda\xi \qquad\qquad \xi(0) = \xi_0$$

In other words, the behavior of the system in the 1-dimensional subspace spanned by $v$ is that of a scalar system.

*Some state space pictures*

**Real systems and complex eigenvalues.** Suppose we are interested in the behavior of a 'real' system with real state $x$ and real system matrix $A$ and suppose $\lambda$ is a (genuine) complex eigenvalue of $A$, that is,

$$Av = \lambda v$$

for a nonzero $v$. Since $A$ is real and $\lambda$ is complex, the eigenvector $v$ is complex. The complex system can start with a complex initial state but the 'real' system cannot. So is the above result useful? It is, just wait.

Since $A$ is real, $\bar{\lambda}$ is also an eigenvalue of $A$ with corresponding eigenvector $\bar{v}$; hence the (complex) solution corresponding to $x(0) = \bar{v}$ is given by

$$e^{\bar{\lambda}t}\bar{v} = \overline{e^{\lambda t}v}$$

Consider now the real vector

$$u := \Re(v) = \frac{1}{2}(v + \bar{v})$$

as an initial state. Since the solution depends linearly on initial state, the solution for this real initial state is

$$\frac{1}{2}(e^{\lambda t}v + \overline{e^{\lambda t}v}\,) \quad = \quad \Re(\,e^{\lambda t}v\,)$$

Suppose we let

$$\alpha = \Re(\lambda) \qquad \omega = \Im(\lambda) \qquad w = \Im(v)\,.$$

Then

$$\lambda = \alpha + \jmath\omega \qquad \text{and} \qquad v = u + \jmath w\,.$$

Hence,

$$\begin{aligned} e^{\lambda t}v &= e^{\alpha t}(\cos\omega t + \jmath \sin\omega t)(u + \jmath w) \\ &= [e^{\alpha t}\cos(\omega t)\,u - e^{\alpha t}\sin(\omega t)\,w\,] \; + \quad \jmath\left(e^{\alpha t}\sin(\omega t)\,u + e^{\alpha t}\cos(\omega t)\,w\right) \end{aligned}$$

Hence, the solution due to $x(0) = u$ is

$$e^{\alpha t}\cos(\omega t)\,u - e^{\alpha t}\sin(\omega t)\,w$$

Similarly, the solution due to initial state $w$ is $\Im(e^{\lambda t}v)$, that is,

$$e^{\alpha t}\sin(\omega t)\,u + e^{\alpha t}\cos(\omega t)\,w$$

• So, if the initial state is of the form

$$x(0) = \xi_{10}u + \xi_{20}w$$

where $\xi_{10}, \xi_{20}$ are any two real scalars (that is, the system starts in the real subspace spanned by the real span of $u, w$), the resulting solution is given by

$$x(t) = \xi_1(t)u + \xi_2(t)w$$

where

$$\begin{array}{rcl}\xi_1(t) & = & e^{\alpha t}\cos(\omega t)\xi_{10} + e^{\alpha t}\sin(\omega t)\xi_{20} \\ \xi_2(t) & = & -e^{\alpha t}\sin(\omega t)\xi_{10} + e^{\alpha t}\cos(\omega t)\xi_{20}\end{array} \quad = \quad \begin{pmatrix} e^{\alpha t}\cos(\omega t) & e^{\alpha t}\sin(\omega t) \\ -e^{\alpha t}\sin(\omega t) & e^{\alpha t}\cos(\omega t)\end{pmatrix}\xi_0$$

Note that, if we let $A_0 = (\xi_{10}^2 + \xi_{20}^2)^{\frac{1}{2}}$ and let $\phi_0$ be the unique number in the interval $[0, 2\pi)$ defined by

$$\sin(\phi_0) = \xi_{10}/A_0 \qquad \text{and} \quad \cos(\phi_0) = \xi_{20}/A_0$$

then

$$\begin{array}{rcl}\xi_1(t) & = & A_0 e^{\alpha t}\sin(\omega t + \phi_0) \\ \xi_2(t) & = & A_0 e^{\alpha t}\cos(\omega t + \phi_0)\end{array}$$

and

$$x(t) = A_0 e^{\alpha t}\left(\sin(\omega t + \phi_0)u + \cos(\omega t + \phi_0)w\right)$$

So, we conclude that for real systems:

(a) If the initial state is in the real 2D subspace spanned by the real and imaginary parts of an complex eigenvector, then so is the resulting solution.

(b) If $\Re(\lambda) < 0$ you decay with increasing age.

(c) If $\Re(\lambda) > 0$ you grow with increasing age.

(d) If $\Re(\lambda) = 0$ your magnitude remains constant and if $\lambda \neq 0$ you oscillate forever.

(e) If $\lambda = 0$ you stay put; every eigenvector corresponding to $\lambda = 0$ is an equilibrium state.

**Example 42** Recall example 38. Considering

$$v = v^1 = \begin{pmatrix} 1 \\ \jmath \end{pmatrix}$$

we have

$$u = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad w = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

So,

$$x = x_1 u + x_2 w$$

Hence, all solutions are given by

$$\begin{array}{rcl}x_1(t) & = & e^{\alpha t}\cos(\omega t)x_{10} + e^{\alpha t}\sin(\omega t)x_{20} \\ x_2(t) & = & -e^{\alpha t}\sin(\omega t)x_{10} + e^{\alpha t}\cos(\omega t)x_{20}\end{array} \quad = \quad \begin{pmatrix} e^{\alpha t}\cos(\omega t) & e^{\alpha t}\sin(\omega t) \\ -e^{\alpha t}\sin(\omega t) & e^{\alpha t}\cos(\omega t)\end{pmatrix}x_0$$

**Behavior in the $u-w$ subspace.** We have shown that if the state starts in the 2-dimensional subspace spanned by $u, w$, ( let's call it the $u-w$ subspace) it remains there, that is, if $x_0 = \xi_{10}u + \xi_{20}w$ then $x(t) = \xi_1(t)u + \xi_2(t)w$. Forgetting about the actual solution, we now show that motion in this subspace is described by:

$$\boxed{\begin{array}{rcrcr} \dot{\xi}_1 & = & \alpha\xi_1 & + & \omega\xi_2 \\ \dot{\xi}_2 & = & -\omega\xi_1 & + & \alpha\xi_2 \end{array} \quad = \quad \begin{pmatrix} \alpha & \omega \\ -\omega & \alpha \end{pmatrix} \xi}$$

To see this, we first equate real and imaginary parts of

$$A(u + \jmath w) = (\alpha + \jmath\omega)(u + \jmath w)$$

to obtain

$$\begin{array}{rcrcr} Au & = & \alpha u & - & \omega w \\ Aw & = & \omega u & + & \alpha w \end{array}$$

Now

$$\begin{aligned} \dot{\xi}_1 u + \dot{\xi}_2 w &= \dot{x} \\ &= Ax \\ &= A(\xi_1 u + \xi_2 w) \\ &= \xi_1 Au + \xi_2 Aw \\ &= (\alpha\xi_1 + \omega\xi_2)u + (-\omega\xi_1 + \alpha\xi_2)w \end{aligned}$$

The desired result now follows by equating coefficients of $u$ and $w$ on both sides of the equation.

**Example 43 (Beavis and Butthead)** Consider $u_1 = u_2 = 0$ in the system in Figure 5.3.

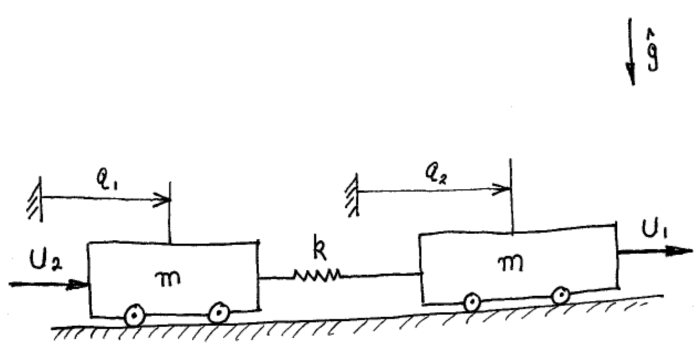



Figure 5.3: Beavis & Butthead

$$\begin{aligned} m\ddot{q}_1 &= k(q_2 - q_1) \\ m\ddot{q}_2 &= -k(q_2 - q_1) \end{aligned}$$

Introducing state variables

$$x_1 := q_1 \,, \quad x_2 := q_2 \,, \quad x_3 := \dot{q}_1 \,, \quad x_4 := \dot{q}_2 \,,$$

this system is described by $\dot{x} = Ax$ where $A$ is the block partitioned matrix given by

$$A = \begin{pmatrix} 0 & I \\ A_{21} & 0 \end{pmatrix} \,, \qquad A_{21} = k/m \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \,.$$

The characteristic polynomial of $A$ is given by

$$p(s) = s^2(s^2 + \omega^2) \,, \qquad \omega = \sqrt{2k/m}$$

Hence, $A$ has three eigenvalues

$$\lambda_1 = \jmath\omega \,, \qquad \lambda_2 = -\jmath\omega \,, \qquad \lambda_3 = 0 \,.$$

Corresponding to each eigenvalue $\lambda_i$, there is at most one linearly independent eigenvector $v^i$, for example,

$$v^1 = \begin{pmatrix} 1 \\ -1 \\ \jmath\omega \\ -\jmath\omega \end{pmatrix} \,, \qquad v^2 = \overline{v^1} \,, \qquad v^3 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \,.$$

Consider now the subspace $\mathcal{S}^1$ spanned by the real and imaginary parts of $v^1$:

$$\mathcal{S}^1 = \text{span} \left\{ \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix} , \begin{pmatrix} 0 \\ 0 \\ \omega \\ -\omega \end{pmatrix} \right\} = \text{span} \left\{ \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix} , \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix} \right\} .$$

A little reflection reveals that states in this subspace correspond to those states for which $q_2 = -q_1$ and $\dot{q}_2 = -\dot{q}_1$. A state trajectory which remains in this subspace is characterized by the condition that

$$q_2(t) \equiv -q_1(t) \,.$$

We have seen that whenever a motion originates in this subspace it remains there; also all motions in this subspace are purely oscillatory and consist of terms of the form $\cos(\omega t)$ and $\sin(\omega t)$.

Consider now the subspace $\mathcal{S}^3$ spanned by the eigenvector $v^3$. Every state in this subspace has the form

$$\begin{pmatrix} q_1 \\ q_1 \\ 0 \\ 0 \end{pmatrix} ,$$

that is, $q_1$ is arbitrary, $q_2 = q_1$, and $\dot{q}_1 = \dot{q}_2 = 0$. Since the eigenvalue corresponding to this eigenspace is zero, every state in this subspace is an equilibrium state.

### 5.4.2 Solutions for nondefective matrices

Recall that the algebraic multiplicity of an eigenvalue of a matrix is the number of times that the eigenvalue is repeated as a root of the characteristic polynomial of the matrix. Also the geometric multiplicity of an eigenvalue is the dimension of its eigenspace, or equivalently, the maximum number of linearly independent eigenvectors associated with that eigenvalue. It can be proven that the geometric multiplicity of an eigenvalue is always less than or equal to its algebraic multiplicity. An eigenvalue whose geometric multiplicity equals its algebraic multiplicity is call a nondefective eigenvalue. Otherwise, it is called defective. If an eigenvalue is not repeated, that is, it is not a repeated root of the matrix characteristic polynomial, then it has algebraic multiplicity one. Since its geometric multiplicity must be greater than zero and less than or equal to its algebraic multiplicity, it follows that its geometric multiplicity equals its algebraic multiplicity; hence the eigenvalue is nondefective.

A matrix with the property that all its eigenvalues are nondefective is called a nondefective matrix. Otherwise, it is called defective. When is a matrix $A$ nondefective? Some examples are:

(a) $A$ has $n$ *distinct* eigenvalues where $A$ is $n \times n$.

(b) $A$ is hermitian that is, $A^* = A$.

(c) $A$ is unitary, that is, $A^*A = I$.

Here $A^*$ is the complex conjugate transpose of $A$.

We have the following result.

**Fact 2** *An $n \times n$ matrix has $n$ linearly independent eigenvectors if and only if $A$ is nondefective.*

**Example 44** The following matrix is a simple example of a defective matrix.

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

**Example 45** The following matrix is another example example of a defective matrix.

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -\omega^2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -\omega^2 & 0 \end{pmatrix}$$

where $\omega$ is any real number.

Suppose $A$ is a nondefective matrix and consider any initial condition: $x(0) = x_0$. Since $A$ is nondefective, it has a basis of eigenvectors $v^1, \cdots, v^n$. Hence, there are unique scalars $\xi_{10}, \xi_{10}, \ldots, \xi_{n0}$ such that $x_0$ can be expressed as

$$x_0 = \xi_{10}v^1 + \xi_{20}v^2 + \ldots + \xi_{n0}v^n \,.$$

From this it follows that the solution due to $x(0) = x_0$ is given by

$$\boxed{x(t) = e^{\lambda_1 t}\xi_{10}v^1 + e^{\lambda_2 t}\xi_{20}v^2 + \ldots + e^{\lambda_n t}\xi_{n0}v^n} \tag{5.6}$$

Note that each term $e^{\lambda_i t}\xi_{i0}v^i$ on the right-hand-side of the above equation is a mode. This results in the following fundamental observation.

*Every motion (solution) of a linear time-invariant system with nondefective $A$ matrix is simply a sum of modes.*

### 5.4.3 Defective matrices and generalized eigenvectors

Consider a square matrix $A$. Then we have the following fact.

**Fact 3** *An eigenvalue $\lambda$ for $A$ is defective if and only if there is a nonzero vector $g$ such that*

$$(A - \lambda I)^2 g = 0 \qquad \textit{but} \qquad (A - \lambda I)g \neq 0\,.$$

The vector $g$ is called a generalized eigenvector for $A$.

**Example 46**

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

Consider $\lambda = 0$. Then

$$g = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

is a generalized eigenvector of $A$.

We now demonstrate the consequence of a system having a defective eigenvalue. So, suppose $\lambda$ is a defective eigenvalue for a square matrix $A$. Let $g$ be any corresponding generalized eigenvector for $\lambda$. We now show that

$$\boxed{x(t) = e^{\lambda t}g + te^{\lambda t}(A - \lambda I)g}$$

is a solution of $\dot{x} = Ax$. To see this, differentiate $x$ to obtain

$$\begin{aligned}
\dot{x}(t) &= e^{\lambda t}\lambda g + e^{\lambda t}(A - \lambda I)g + te^{\lambda t}\lambda(A - \lambda I)g \\
&= e^{\lambda t}Ag + te^{\lambda t}\lambda(A - \lambda I)g
\end{aligned}$$

Since $(A - \lambda I)^2 g = 0$, we have

$$A(A - \lambda I)g = \lambda(A - \lambda I)g\,.$$

Hence

$$\begin{aligned}
\dot{x}(t) &= e^{\lambda t}Ag + te^{\lambda t}A(A - \lambda I)g = A(e^{\lambda t}g + te^{\lambda t}(A - \lambda I)g) \\
&= Ax(t)\,.
\end{aligned}$$

■

Consider the function given by $te^{\lambda t}$ and note that

$$|te^{\lambda t}| = |t|e^{\alpha t}$$

where $\alpha$ is the real part of $\lambda$. So, if $\alpha$ is negative, the magnitude of the above function decays exponentially. If $\alpha$ is positive, the magnitude of the above function grows exponentially. If $\alpha$ is zero, the magnitude of the above function is given by $|t|$ which is also unbounded. From this discussion, we see that *the defectiveness of an eigenvalue is only significant when the eigenvalue lies on the imaginary axis of the complex plane, that is when it has zero real part.*

**Example 47 (Resonance)** Consider a simple spring-mass system, of mass $m$ and spring constant $k$, subject to a sinusoidal force $w(t) = A\sin(\omega t + \phi)$ where $\omega = \sqrt{k/m}$ is the unforced natural frequency of the spring mass system and $A$ and $\phi$ are arbitrary constants. Letting $q$ be the displacement of the mass from its unforced equilibrium position, the motion of this system can be described by

$$m\ddot{q} + kq = w\,.$$

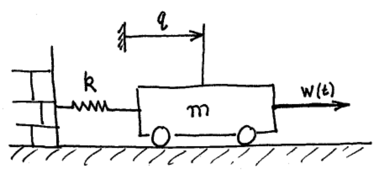



Figure 5.4: Resonance

Since $w(t) = A\sin(\omega t + \phi)$ it satisfies the following differential equation:

$$\ddot{w} + \omega^2 w = 0\,.$$

If we introduce state variables,

$$x_1 = q\,, \qquad x_2 = \dot{q}\,, \quad x_3 = w\,, \qquad \ddot{x}_4\,, = \dot{w}$$

the sinusoidally forced spring-mass system can be described by $\dot{x} = Ax$ where

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -\omega^2 & 0 & 1/m & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -\omega^2 & 0 \end{pmatrix}.$$

This matrix has characteristic polynomial

$$p(s) = (s^2 + \omega^2)^2\,.$$

Hence $A$ has two distinct imaginary eigenvalues $\lambda_1 = \jmath\omega$ and $\lambda_2 = -\jmath\omega$, both of which are repeated twice as roots of the characteristic polynomial. All eigenvalues for $\lambda_1$ and $\lambda_2$ are of the form $cv^1$ and $cv^2$, respectively, where

$$v^1 = \begin{pmatrix} 1 \\ \jmath\omega \\ 0 \\ 0 \end{pmatrix} \qquad \text{and} \qquad v^2 = \begin{pmatrix} 1 \\ -\jmath\omega \\ 0 \\ 0 \end{pmatrix}.$$

Since these eigenvalues have algebraic multiplicity two and geometric multiplicity one, they are defective. A generalized eigenvector corresponding to $\lambda_1$ and $\lambda_2$ is given by

$$g^1 = \begin{pmatrix} 1/2k \\ 0 \\ 1 \\ \jmath\omega \end{pmatrix} \qquad \text{and} \qquad g^2 = \begin{pmatrix} 1/2k \\ 0 \\ 1 \\ -\jmath\omega \end{pmatrix},$$

respectively. Hence this system has solutions with terms of the form $t\sin(\omega t)$.

**Exercises**

**Exercise 50 (Beavis and Butthead at the laundromat)**

$$\begin{aligned} m\ddot{q}_1 - m\Omega^2 q_1 + \tfrac{k}{2}(q_1 - q_2 - 2l_0) &= 0 \\ m\ddot{q}_2 - m\Omega^2 q_2 - \tfrac{k}{2}(q_1 - q_2 - 2l_0) &= 0 \end{aligned}$$

with

$$\omega := \sqrt{k/m} > \Omega$$

Figure 5.5: At the laundromat

(a) Obtain the equilibrium values $q_1^e, q_2^e$ of $q_1, q_2$.

(b) With

$$\begin{aligned} x_1 &= \delta q_1 := q_1 - q_1^e \\ x_2 &= \delta q_2 := q_2 - q_2^e \\ x_3 &= \dot{q}_1 \\ x_4 &= \dot{q}_2 \end{aligned}$$

Obtain a state space description of the form $\dot{x} = Ax$.

(c) Obtain expressions for the eigenvalues of $A$. Write down expressions for the complex Jordan form and the real Jordan form of $A$?

(d) For the rest of the exercise, consider $\omega^2 = 2\Omega^2$. Obtain expressions for linearly independent eigenvectors of $A$.

(e) For the rest of the exercise, let $\Omega = 1$. Numerically simulate with initial conditions

$$\delta q_1(0) = 1, \quad \delta q_2(0) = -1, \quad \dot{q}_1(0) = 0, \quad \dot{q}_2(0) = 0$$

On the same graph, plot $\delta q_1(t)$ and $-\delta q_2(t)$ vs $t$. Why is $\delta q_2(t) \equiv -\delta q_1(t)$? Plot $\delta q_2(t)$ vs $\delta q_1(t)$.

(f) Numerically simulate with initial conditions

$$\delta q_1(0) = 1, \quad \delta q_2(0) = 1, \quad \dot{q}_1(0) = -1, \quad \dot{q}_2(0) = -1$$

On the same graph, plot $\delta q_1(t)$ and $\delta q_2(t)$ vs $t$. Why is $\delta q_2(t) \equiv \delta q_1(t)$? Why does the solution decay to zero? Plot $\delta q_2(t)$ vs $\delta q_1(t)$.

(g) Numerically simulate with initial conditions

$$\delta q_1(0) = 1, \quad \delta q_2(0) = 1, \quad \dot{q}_1(0) = 1, \quad \dot{q}_2(0) = 1$$

On the same graph, plot $\delta q_1(t)$ and $\delta q_2(t)$ vs $t$. Why is $\delta q_2(t) \equiv \delta q_1(t)$? Is this solution bounded? Plot $\delta q_2(t)$ vs $\delta q_1(t)$.

(h) Numerically simulate with initial conditions

$$\delta q_1(0) = 1, \quad \delta q_2(0) = -1, \quad \dot{q}_1(0) = \jmath, \quad \dot{q}_2(0) = -\jmath$$

Compute $||x(t)||$ and show (by plotting $||x||$ vs $t$) that it is constant. Why is it constant?

**Exercise 51** Consider the delay differential equation described by

$$\dot{x}(t) = A_1 x(t) + A_2 x(t-1)$$

where $t \in \mathbb{R}$ and $x(t) \in \mathbb{C}^n$. Show that if a complex number $\lambda$ satisfies

$$\det(\lambda I - A_1 - e^{-\lambda} A_2) = 0$$

then the delay differential equation has a solution of the form

$$e^{\lambda t} v$$

where $v \in \mathbb{C}^n$ is nonzero.

**Exercise 52** Consider a system described by

$$M\ddot{q} + C\dot{q} + Kq = 0$$

where $q(t)$ is an $N$-vector and $M$, $C$ and $K$ are square matrices. Suppose $\lambda$ is a complex number which satisfies

$$\det(\lambda^2 M + \lambda C + K) = 0\,.$$

Show that the above system has a solution of the form

$$q(t) = e^{\lambda t} v$$

where $v$ is a constant $N$-vector.

## 5.5 Behavior of discrete-time systems

### 5.5.1 System significance of eigenvectors and eigenvalues

Consider now a discrete time system described by

$$x(k+1) = Ax(k) \tag{5.7}$$

We now demonstrate the following result:

*The discrete-time system (5.7) has a solution of the form*

$$\boxed{x(k) = \lambda^k v} \tag{5.8}$$

*if and only if $\lambda$ is an eigenvalue of $A$ and $v$ is a corresponding eigenvector.*

To prove this, we first suppose that $\lambda$ is an eigenvalue of $A$ with $v$ as a corresponding eigenvector. We need to show that $x(k) = \lambda^k v$ is a solution of (5.7). Since $Av = \lambda v$, it follows that $x(k+1) = \lambda^{(k+1)} v = \lambda^k \lambda v = \lambda^k A v = A(\lambda^k v) = Ax(k)$, that is $x(k+1) = Ax(k)$; hence $x(\cdot)$ is a solution of (5.7).

Now suppose that $x(k) = \lambda^k v$ is a solution of (5.7). Then $\lambda v = x(1) = Ax(0) = Av$, that is, $Av = \lambda v$; hence $\lambda$ is an eigenvalue of $A$ and $v$ is a corresponding eigenvector. ■

A solution of the form $\lambda^k v$ is a very special type of solution and is sometimes called a mode of the system. Note that if $x(k) = \lambda^k v$ is a solution, then $x(0) = v$, that is, $v$ is the initial value of $x$. Hence, we can make the following statement.

*If $v$ is an eigenvector of $A$, then the solution to $x(k+1) = Ax(k)$ with initial condition $x(0) = v$ is given by $x(k) = \lambda^k v$ where $\lambda$ is the eigenvalue corresponding to $v$.*

If $v$ is an eigenvector then, $\lambda^k v$ is also an eigenvector for each $k$. Also, considering the magnitude of $\lambda^k v$, we obtain

$$||\lambda^k v|| = |\lambda^k| \, \|v\| = |\lambda|^k \, \|v\|$$

So,

(a) Once an eigenvector, always an eigenvector!

(b) If $|\lambda| < 1$ you decay with increasing age.

(c) If $|\lambda| > 1$ you grow with increasing age.

(d) If $|\lambda| = 1$ and $\lambda \neq 1$ you oscillate forever.

• If $\lambda = 1$ and $x(0) = v$, then $x(k) = v$ for all $k$, that is $v$ is an equilibrium state for the system.
• If $\lambda = 0$ and $x(0) = v$, then $x(k) = 0$ for all $k > 0$, that is *the state of the system goes to zero in one step.*

### 5.5.2 All solutions for nondefective matrices

Recall that an eigenvalue whose geometric multiplicity equals its algebraic multiplicity is called a nondefective eigenvalue. Otherwise, it is considered defective. A matrix with the property that all its eigenvalues are nondefective is called a nondefective matrix. Otherwise, it is considered defective. Recall also that an $n \times n$ matrix has $n$ linearly independent eigenvectors if and only if $A$ is nondefective.

**Example 48** The following matrix is a simple example of a defective matrix.

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

Suppose $A$ is a nondefective matrix and consider any initial condition: $x(0) = x_0$. Since $A$ is nondefective, it has a basis of eigenvectors $v^1, \cdots, v^n$. Hence,

$$x_0 = \xi_{10} v^1 + \xi_{20} v^2 + \ldots + \xi_{n0} v^n$$

From this it follows that

$$\boxed{x(k) = \lambda_1^k \xi_{10} v^1 + \lambda_2^k \xi_{20} v^2 + \ldots + \lambda_n^k \xi_{n0} v^n} \tag{5.9}$$

Note that each term $\lambda_i^k \xi_{i0} v^i$ on the right-hand-side of the above equation is a mode. This results in the following fundamental observation.

*Every motion (solution) of a discrete-time linear time-invariant system with nondefective $A$ matrix is simply a sum of modes.*

### 5.5.3 Some solutions for defective matrices

Consider a square matrix $A$ and recall that an eigenvalue $\lambda$ for $A$ is defective if and only if there is a nonzero vector $g$ such that

$$(A - \lambda I)^2 g = 0 \qquad \text{but} \qquad (A - \lambda I) g \neq 0\,.$$

The vector $g$ is called a generalized eigenvector for $A$.

**Example 49**

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

Consider $\lambda = 0$. Then

$$g = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

is a generalized eigenvector of $A$.

We now demonstrate the consequence of a system having a defective eigenvalue. So, suppose $\lambda$ is a defective eigenvalue for a square matrix $A$. Let $g$ be any corresponding generalized eigenvector for $\lambda$. We now show that

$$\boxed{x(k) = \lambda^k g + k\lambda^{(k-1)}(A - \lambda I)g}$$

is a solution of $x(k+1) = Ax(k)$. To see this, we first note that, since $(A - \lambda I)^2 g = 0$, we have

$$A(A - \lambda I)g = \lambda(A - \lambda I)g\,.$$

Hence

$$\begin{aligned} Ax(k) &= \lambda^k Ag + k\lambda^{(k-1)}A(A - \lambda I)g \\ &= \lambda^k Ag + k\lambda^k(A - \lambda I)g\,. \end{aligned}$$

Now note that

$$\begin{aligned} x(k+1) &= \lambda^{(k+1)}g + (k+1)\lambda^k(A - \lambda I)g \\ &= \lambda^{(k+1)}g + k\lambda^k(A - \lambda I)g + \lambda^k Ag - \lambda^{(k+1)}g \\ &= \lambda^k Ag + k\lambda^k(A - \lambda I)g\,. \end{aligned}$$

Hence $x(k+1) = Ax(k)$ ■

Consider the function given by $k\lambda^{k-1}$ and note that

$$|k\lambda^{k-1}| = |k||\lambda|^{k-1}\,.$$

So, if $|\lambda| < 1$, the magnitude of the above function decays exponentially. If $|\lambda| > 1$, the magnitude of the above function grows exponentially. If $|\lambda| = 1$, the magnitude of the above function is given by $|k|$ which is also unbounded. From this discussion, we see that *the defectiveness of an eigenvalue is only significant when the eigenvalue lies on the unit circle in the complex plane, that is when its magnitude is one.*

## 5.6 Similarity transformations

**Motivating example.** Consider

$$\begin{aligned}\dot{x}_1 &= 3x_1 - x_2 \\ \dot{x}_2 &= -x_1 + 3x_2\end{aligned}$$

Suppose we define new states:

$$\begin{aligned}\xi_1 &= (x_1 + x_2)/2 \\ \xi_2 &= (x_1 - x_2)/2\end{aligned}$$

Then,

$$\begin{aligned}\dot{\xi}_1 &= 2\xi_1 \\ \dot{\xi}_2 &= 4\xi_2\end{aligned}$$

That looks better!
Where did new states come from? First note that

$$\begin{aligned}x_1 &= \xi_1 + \xi_2 \\ x_2 &= \xi_1 - \xi_2\end{aligned}$$

and we have the following geometric interpretation. Since,

$$\begin{aligned}x &= x_1\begin{pmatrix}1\\0\end{pmatrix} + x_2\begin{pmatrix}0\\1\end{pmatrix} \\ &= \xi_1\begin{pmatrix}1\\1\end{pmatrix} + \xi_2\begin{pmatrix}1\\-1\end{pmatrix}\end{aligned}$$

the scalars $x_1$, $x_2$ are the coordinates of the vector $x$ wrt the usual basis for $\mathbb{R}^2$; the scalars $\xi_1$, $\xi_2$ are the coordinates of $x$ wrt the basis

$$\begin{pmatrix}1\\1\end{pmatrix} \qquad \begin{pmatrix}1\\-1\end{pmatrix}$$

Where did these new basis vectors come from? These new basis vectors are eigenvectors of the matrix

$$A = \begin{pmatrix}3 & -1\\-1 & 3\end{pmatrix}$$

and the system is described by $\dot{x} = Ax$.

**Coordinate transformations.** Suppose we have $n$ scalar variables $x_1, x_2, \dots, x_n$ and we implicitly define new scalar variables $\xi_1, \xi_2, \dots, \xi_n$ by

$$\boxed{x = T\xi} \qquad \text{where} \qquad x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \text{ and } \xi = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{pmatrix}$$

and $T$ is an $n \times n$ invertible matrix. Then, $\xi$ is explicitly given by

$$\xi = T^{-1}x$$

We can obtain a geometric interpretation of this change of variables as follows. First, observe that

$$\underbrace{\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}}_{x} = x_1 \underbrace{\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}}_{e^1} + x_2 \underbrace{\begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}}_{e^2} + \dots + x_n \underbrace{\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}}_{e^n},$$

that is the scalars $x_1, x_2, \dots, x_n$ are the coordinates of the vector $x$ wrt the standard basis $e^1, e^2, \dots, e^n$, or, $x$ is the coordinate vector of itself wrt the standard basis. Suppose

$$T = \begin{pmatrix} t^1 & \dots & t^j & \dots & t^n \end{pmatrix}$$

that is, $t^j$ is the $j$-th column of $T$. Since $T$ is invertible, its columns, $t^1, t^2, \dots, t^n$, form a basis for $\mathbb{C}^n$. Also, $x = T\xi$ can be written as

$$x = \xi_1 t^1 + \xi_2 t^2 + \dots + \xi_n t^n$$

From this we see that $\xi_1, \xi_2, \dots, \xi_n$ are the coordinates of $x$ wrt to the new basis and the vector $\xi$ is the coordinate vector of the vector $x$ wrt this new basis. So $x = T\xi$ defines a coordinate transformation.

**Similarity transformations.** Consider now a CT (DT) system described by

$$\dot{x} = Ax \qquad\qquad (\, x(k{+}1) = Ax(k)\,)$$

and suppose we define new state variables by $x = T\xi$ where $T$ is nonsingular; then the behavior of $\xi$ is governed by

$$\dot{\xi} = \Lambda\xi \qquad\qquad (\, \xi(k{+}1) = \Lambda\xi(k)\,)$$

where

$$\boxed{\Lambda = T^{-1}AT}$$

*A square matrix $A$ is said to be* similar *to another square matrix $\Lambda$ if there exists a nonsingular matrix $T$ such that $\Lambda = T^{-1}AT$.*

Recalling that the columns $t^1, t^2, \dots, t^n$ of $T$ form a basis, we have now the following very useful result.

**Useful result.** Suppose

$$At^j = \alpha_{1j}t^1 + \alpha_{2j}t^2 + \ldots + \alpha_{nj}t^n$$

Then the matrix $\Lambda = T^{-1}AT$ is uniquely given by

$$\Lambda_{ij} = \alpha_{ij}$$

that is $\Lambda_{ij}$ is the $i$-th coordinate of the vector $At^j$ wrt the basis $t^1, t^2, \cdots, t^n$. Thus, the $j$-th column of the matrix $\Lambda$ is the coordinate vector of $At^j$ wrt the basis $t^1, t^2, \cdots, t^n$.

Proof. Premultiply $\Lambda = T^{-1}AT$ by $T$ to get:

$$T\Lambda = AT$$

On the right we have

$$\begin{aligned} AT &= A\begin{pmatrix} t^1 & \ldots & t^j & \ldots & t^n \end{pmatrix} \\ &= \begin{pmatrix} At^1 & \ldots & At^j & \ldots & At^n \end{pmatrix} \end{aligned}$$

And on the left,

$$T\Lambda = \begin{pmatrix} t^1 & t^2 & \ldots & \ldots & t^n \end{pmatrix} \begin{pmatrix} \Lambda_{11} & \ldots & \Lambda_{1j} & \ldots & \Lambda_{1n} \\ \Lambda_{21} & \ldots & \Lambda_{2j} & \ldots & \Lambda_{2n} \\ \vdots & & \vdots & & \vdots \\ \vdots & & \vdots & & \vdots \\ \Lambda_{n1} & \ldots & \Lambda_{nj} & \ldots & \Lambda_{nn} \end{pmatrix}$$

$$= \left( \Lambda_{11}t^1 + \cdots + \Lambda_{n1}t^n \vdots \quad \ldots\ldots \quad \vdots \underbrace{\Lambda_{1j}t^1 + \ldots + \Lambda_{nj}t^n}_{j\text{-th column}} \vdots \quad \ldots\ldots \quad \vdots \Lambda_{1n}t^1 + \ldots + \Lambda_{nn}t^n \right)$$

Comparing the expressions for the $j$-th columns of $AT$ and $T\Lambda$ yields

$$\begin{aligned} \Lambda_{1j}t^1 + \ldots + \Lambda_{nj}t^n &= At^j \\ &= \alpha_{1j}t^1 + \ldots + \alpha_{nj}t^n \end{aligned}$$

Since the $b_i$'s form a basis we must have $\Lambda_{ij} = \alpha_{ij}$. ∎

- Note that

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

**Example 50** Recall example 37. Note that the eigenvectors $v^1$ and $v^2$ are linearly independent; hence the matrix

$$T := \begin{pmatrix} v^1 \, v^2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

is invertible. Since

$$
\begin{aligned}
Av^1 &= 2v^1 \\
Av^2 &= 4v^2
\end{aligned}
$$

we can use the above result and immediately write down (without computing $T^{-1}$)

$$
\Lambda = T^{-1}AT = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}
$$

Explicitly compute $T^{-1}$ and then $T^{-1}AT$ to check.

**Example 51** Recalling example 38 we can follow the procedure of the last example and define

$$
T := \begin{pmatrix} 1 & 1 \\ \jmath & -\jmath \end{pmatrix}
$$

to obtain (without computing $T^{-1}$)

$$
\Lambda = T^{-1}AT = \begin{pmatrix} 1+\jmath & 0 \\ 0 & 1-\jmath \end{pmatrix}
$$

You may now compute $T^{-1}$ and check this.

**Example 52** Suppose $A$ is a $4 \times 4$ matrix and $v^1, v^2, v^3, v^4$ are four linearly independent vectors which satisfy:

$$
Av^1 = v^2\,, \qquad Av^2 = v^3\,, \qquad Av^3 = v^4\,, \qquad Av^4 = 3v^1 - v^2 + 6v^3
$$

If we let $V$ be the matrix whose columns are $v^1, v^2, v^3, v^4$, then $V$ is invertible and we can immediately write down that

$$
V^{-1}AV = \begin{pmatrix} 0 & 0 & 0 & 3 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 6 \\ 0 & 0 & 1 & 0 \end{pmatrix}
$$

**Example 53** B&B

**Some properties:**

**Exercise 53** Prove the following statements:

(a) If $A$ is similar to $B$, then $B$ is similar to $A$

(b) If $A$ is similar to $B$ and $B$ is similar to $C$, then $A$ is similar to $C$.

**Exercise 54** Show that if $A$ is similar to $\Lambda$ then the following hold:

(a) $\det A = \det \Lambda$

(b) charpoly $A$ = charpoly $\Lambda$ and, hence, $A$ and $\Lambda$ have the same eigenvalues with the same algebraic multiplicities.

**Exercise 55** Suppose $A$ is a $3 \times 3$ matrix, $b$ is a 3-vector,

$$A^3 + 2A^2 + A + 1 = 0$$

and the matrix

$$T = \begin{pmatrix} b & Ab & A^2b \end{pmatrix}$$

is invertible. What is the matrix $\Lambda := T^{-1}AT$? Specify each element of $\Lambda$ explicitly.

## 5.7 Hotel California

### 5.7.1 Invariant subspaces

Suppose $A$ is a square matrix. We have already met several subspaces associated with $A$, namely, its range, null space, and eigenspaces. These subspaces are all examples of invariant subspaces for $A$

*We say that a subspace $\mathcal{S}$ is an* invariant subspace *for $A$ if for each vector $x$ in $\mathcal{S}$, the vector $Ax$ is also in $\mathcal{S}$.*

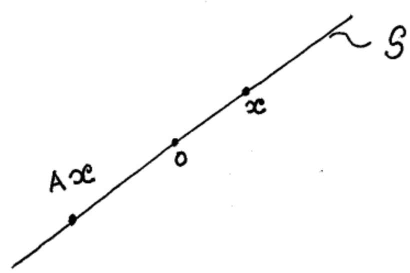



Figure 5.6: Invariant subspace

The significance of invariant subspaces for LTI systems lies in the following result which is proven later.

*If the state of a linear-time-invariant system starts in an invariant subspace of the system matrix, it never leaves the subspace.*

**A coordinate transformation.** Suppose $\mathcal{S}$ is an invariant subspace for a square matrix $A$. If $\mathcal{S}$ has dimension $m$, let

$$t^1, t^2, \cdots, t^m$$

be a basis for $\mathcal{S}$. If $A$ is $n \times n$, extend the above basis to a basis

$$t^1, \dots, t^m, t^{m+1}, \cdots, t^n$$

for $\mathbb{C}^n$ (This can always be done.) Now introduce the coordinate transformation

$$x = T\xi$$

where

$$T := \begin{pmatrix} t^1 & t^2 & \dots & t^n \end{pmatrix}$$

Letting

$$\xi = \begin{pmatrix} \zeta \\ \eta \end{pmatrix},$$

where $\zeta$ is an $m$-vector and $\eta$ is an $(n-m)$-vector, we obtain

$$x = \zeta_1 t^1 + \dots + \zeta_m t^m + \eta_1 t^{m+1} + \dots + \eta_{n-m} t^n$$

So

$$x \text{ is in } \mathcal{S} \qquad \text{if and only if} \qquad \eta = 0$$

Let

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

Using our 'useful result on similarity transformations' and the fact that $\mathcal{S}$ is invariant for $A$, it follows that $\Lambda$ must have the following structure:

$$\Lambda = \begin{pmatrix} \Lambda_{11} & \Lambda_{21} \\ 0 & \Lambda_{22} \end{pmatrix}$$

**System significance of invariant subspaces.** Consider a system described by

$$\dot{x} = Ax$$

and suppose $\mathcal{S}$ is an invariant subspace for $A$. Introduce the above coordinate transformation to obtain

$$\dot{\xi} = \Lambda\xi$$

or

$$\begin{aligned} \dot{\zeta} &= \Lambda_{11}\zeta + \Lambda_{12}\eta \\ \dot{\eta} &= \Lambda_{22}\eta \end{aligned}$$

Looking at the differential equation $\dot{\eta} = \Lambda_{22}\eta$, we see that if $\eta(0) = 0$, then $\eta(t) \equiv 0$; hence, if $x(0)$ is in $\mathcal{S}$ then $x(t)$ is always in $\mathcal{S}$. In other words, if you are in $\mathcal{S}$, you can check out any time, but you can never leave.....

### 5.7.2 More on invariant subspaces*

Suppose $\mathcal{S}_1, \mathcal{S}_2, \ldots, \mathcal{S}_q$ are $q$ invariant subspaces of a square matrix $A$ with the property that there is basis

$$t^1, \ldots, t^n$$

where

$$t^1, \ldots, t^{n_1}$$

is a basis for $\mathcal{S}_1$,

$$t^{n_1+1}, \ldots, t^{n_1+n_2}$$

is a basis for $\mathcal{S}_2$,

$$t^{n_1+\cdots+n_{q-1}+1}, \ldots, t^n$$

is a basis for $\mathcal{S}_q$,

$$\dot{x} = Ax$$

Coordinate transformation

$$x = T\xi$$

where

$$T = \begin{pmatrix} t^1 & t^2 & \dots & t^n \end{pmatrix}$$

that is, $t^i$ is the $i$-th column of $T$. Consider

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

Claim:

$$\Lambda = \text{blockdiag}\,(\Lambda_1,\, \dots,\, \Lambda_q)$$

## 5.8 Diagonalizable systems

### 5.8.1 Diagonalizable matrices

*We say that a square matrix $A$ is* diagonalizable *if it is similar to a diagonal matrix, that is, there is an invertible matrix $T$ and a diagonal matrix $\Lambda$ such that*

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

In this section, we show that $A$ is diagonalizable if and only if it has the following property: For each distinct eigenvalue of $A$,

$$\boxed{\text{geometric multiplicity} = \text{algebraic multiplicity}}$$

Recall that the geometric multiplicity of an eigenvalue is the dimension of its eigenspace, or equivalently, the maximum number of linearly independent eigenvectors associated with that eigenvalue. An eigenvalue whose geometric multiplicity equals its algebraic multiplicity is call a nondefective eigenvalue. Otherwise, it is considered defective. A matrix with the property that all its eigenvalues are nondefective is called a nondefective matrix. Otherwise, it is considered defective. When is a matrix $A$ nondefective? Some examples are:

(a) $A$ has $n$ *distinct* eigenvalues where $A$ is $n \times n$.

(b) $A$ is *hermitian* that is, $A^* = A$.

We first need the following result.

**Fact 4** *Suppose that geometric multiplicity equals algebraic multiplicity for each distinct eigenvalue of a square matrix $A$. Then $A$ has a* basis of eigenvectors*:*

$$v^1, v^2, \ldots, v^n$$

**Example 54**

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

**Theorem 2** *Suppose that for each distinct eigenvalue of $A$, its geometric multiplicity is the same as its algebraic multiplicity . Then $A$ is similar to a diagonal matrix $\Lambda$ whose diagonal elements are the eigenvalues of $A$. (If an eigenvalue has multiplicity $m$, it appears $m$ times on the diagonal of $\Lambda$)*

PROOF. Let

$$T := \begin{pmatrix} v^1 & v^2 & \ldots & v^n \end{pmatrix}$$

where $v^1, \ldots, v^n$ is a basis of eigenvectors for $A$. Let $\lambda_i$ be the eigenvalue corresponding to $v^i$. Then,

$$Av^i = \lambda_i v^i$$

Defining

$$\Lambda := T^{-1}AT$$

it follows from the 'useful result on similarity transformations' that

$$\begin{aligned}\Lambda &= \text{diag}\,(\lambda_1, \lambda_2, \ldots, \lambda_n) \\ &:= \underbrace{\begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix}}_{\text{Complex Jordan form of } A}\end{aligned}$$

So $A$ is similar to the diagonal matrix $\Lambda$. ∎

**Example 55**

$$A = \begin{pmatrix} 4 & -2 & -2 \\ 0 & 2 & -2 \\ 0 & -2 & 2 \end{pmatrix}$$

The characteristic polynomial of $A$ is

$$p(s) = \det(sI - A) = (s-4)[(s-2)(s-2) - 4] = s(s-4)^2$$

So $A$ has eigenvalues,

$$\lambda_1 = 0, \qquad \lambda_2 = 4$$

with algebraic multiplicities 1 and 2.

The equation $(A - \lambda I)v = 0$ yields:

For $\lambda_1 = 0$,

$$\text{rref}(A - \lambda_1 I) = \text{rref}\begin{pmatrix} 4 & -2 & -2 \\ 0 & 2 & -2 \\ 0 & -2 & 2 \end{pmatrix} = \text{rref}\begin{pmatrix} 4 & 0 & -4 \\ 0 & 2 & -2 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}$$

Hence, $\lambda_1$ has geometric multiplicity 1 and its eigenspace is spanned by

$$v^1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

For $\lambda_2 = 4$,

$$\text{rref}(A - \lambda_2 I) = \text{rref}\begin{pmatrix} 0 & -2 & -2 \\ 0 & -2 & -2 \\ 0 & -2 & -2 \end{pmatrix} = \text{rref}\begin{pmatrix} 0 & -2 & -2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Hence, $\lambda_2$ has geometric multiplicity 2 and its eigenspace is spanned by

$$v^2 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \qquad \text{and} \qquad v^3 = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}$$

Since the geometric multiplicity of each eigenvalue is equal to its algebraic multiplicity, $A$ is diagonalizable and its Jordan form is

$$\Lambda = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{pmatrix}$$

A quick check with MATLAB yields $\Lambda = T^{-1}AT$ where $T = (v^1 v^2 v^3)$.

### 5.8.2 Diagonalizable continuous-time systems

Consider a continuous-time system

$$\dot{x} = Ax$$

and suppose $A$ is an $n \times n$ diagonalizable matrix, that is, there is a nonsingular matrix $T$ such that

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

is diagonal. Hence, the diagonal elements of $\Lambda$ are the eigenvalues (including possible multiplicities) $\lambda_1, \lambda_2, \ldots, \lambda_n$ of $A$. Introducing the state transformation,

$$x = T\xi$$

we have $\xi = T^{-1}x$ and the system behavior is now equivalently described by

$$\dot{\xi} = \Lambda\xi\,.$$

This yields

$$\boxed{\begin{array}{rcl} \dot{\xi}_1 & = & \lambda_1\xi_1 \\ \dot{\xi}_2 & = & \lambda_2\xi_2 \\ & \vdots & \\ \dot{\xi}_n & = & \lambda_n\xi_n \end{array}}$$

that is, $n$ decoupled first order systems. Hence, for $i = 1, 2, \ldots, n$,

$$\xi_i(t) = e^{\lambda_i t}\xi_{i0} \tag{5.10}$$

where $\xi_{i0} = \xi(0)$. In vector form, we obtain

$$\xi(t) = \begin{pmatrix} e^{\lambda_1 t} & 0 & \ldots & 0 \\ 0 & e^{\lambda_2 t} & \ldots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \ldots & e^{\lambda_n t} \end{pmatrix} \xi_0 \tag{5.11}$$

where $\xi_0 = \begin{pmatrix} \xi_{10} & \xi_{20} & \cdots & \xi_{n0} \end{pmatrix}^T$. Since $x(t) = T\xi(t)$, we have

$$x(t) = \xi_1(t)v^1 + \xi_2(t)v^2 + \ldots + \xi_n(t)v^n$$

where $v^1, v^2, \ldots, v^n$ are the columns of $T$. It now follows from (5.10) that the solution for $x(t)$ is given by

$$\boxed{x(t) = e^{\lambda_1 t}\xi_{10}v^1 + e^{\lambda_2 t}\xi_{20}v^2 + \ldots + e^{\lambda_n t}\xi_{n0}v^n} \tag{5.12}$$

Note that each term $e^{\lambda_i t}\xi_{i0}v^i$ on the right-hand-side of the above equation is a mode. This results in the following fundamental observation.

*Every motion (solution) of a diagonalizable linear time-invariant system is simply a sum of modes.*

Suppose $x_0$ is the initial value of $x$ at $t = 0$, that is, $x(0) = x_0$. Then, the corresponding initial value $\xi_0 = \xi(0)$ of $\xi$ is given by

$$\xi_0 = T^{-1}x_0 \, .$$

**State transition matrix.** Since $x(t) = T\xi(t)$ and $\xi_0 = T^{-1}x_0$, it follows from (5.11) that

$$x(t) = \Phi(t)x_0 \tag{5.13}$$

where the state transition matrix $\Phi(\cdot)$ is given by

$$\boxed{\Phi(t) := T \begin{pmatrix} e^{\lambda_1 t} & 0 & \ldots & 0 \\ 0 & e^{\lambda_2 t} & \ldots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \ldots & e^{\lambda_n t} \end{pmatrix} T^{-1}} \tag{5.14}$$

**Example 56** Consider a CT system with $A$ from example 37, that is,

$$A = \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix}$$

Here,

$$T = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

and $\lambda_1 = 2, \lambda_2 = 4$.

Hence,

$$\begin{aligned} \Phi(t) &= T \begin{pmatrix} e^{2t} & 0 \\ 0 & e^{4t} \end{pmatrix} T^{-1} \\ &= \frac{1}{2} \begin{pmatrix} e^{2t} + e^{4t} & e^{2t} - e^{4t} \\ e^{2t} - e^{4t} & e^{2t} + e^{4t} \end{pmatrix} \end{aligned}$$

and all solutions are given by

$$\begin{aligned} x_1(t) &= \frac{1}{2}(\,e^{2t} + e^{4t}\,)x_{10} + \frac{1}{2}(\,e^{2t} - e^{4t}\,)x_{20} \\ x_2(t) &= \frac{1}{2}(\,e^{2t} - e^{4t}\,)x_{10} + \frac{1}{2}(\,e^{2t} + e^{4t}\,)x_{20} \end{aligned}$$

**Exercise 56** Compute the state transition matrix for the CT system corresponding to the $A$ matrix in example 55.

### 5.8.3 Diagonalizable DT systems

The story here is basically the same as that for CT. Consider a DT system

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

and suppose $A$ is diagonalizable, that is, there is a nonsingular matrix $T$ such that

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

is diagonal and the diagonal elements of $\Lambda$ are the eigenvalues $\lambda_1, \dots, \lambda_n$ of $A$.

Introducing the state transformation,

$$x = T\xi$$

we have $\xi = T^{-1}x$ and its behavior is described by:

$$\xi(k+1) = \Lambda\xi(k)$$

Hence,

$$\begin{aligned}
\xi_1(k+1) &= \lambda_1\xi_1(k) \\
\xi_2(k+1) &= \lambda_2\xi_2(k) \\
&\vdots \\
\xi_n(k+1) &= \lambda_n\xi_n(k)
\end{aligned}$$

If

$$x(0) = x_0$$

then

$$\xi(0) = \xi_0 = T^{-1}x_0$$

So,

$$\xi_i(k) = \lambda_i^k \xi_{i0}, \qquad i = 1, 2, \dots, n$$

or,

$$\xi(k) = \begin{pmatrix} \lambda_1^k & 0 & \dots & 0 \\ 0 & \lambda_2^k & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n^k \end{pmatrix} \xi_0$$

Since

$$x = \xi_1 v^1 + \xi_2 v^2 + \dots + \xi_n v^n$$

where $v^1, v^2, \ldots, v^n$ are the columns of $T$, the solution for $x(k)$ is given by

$$\boxed{x(k) = \lambda_1^k \xi_{10} v^1 + \lambda_2^k \xi_{20} v^2 + \ldots + \lambda_n^k \xi_{n0} v^n}$$

Or,

$$x(k) = \Phi(k) x_0$$

where the *DT state transition matrix* $\Phi(\cdot)$ is given by

$$\Phi(k) := T \begin{pmatrix} \lambda_1^k & 0 & \ldots & 0 \\ 0 & \lambda_2^k & \ldots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n^k \end{pmatrix} T^{-1}$$

**Example 57** Consider a DT system with $A$ from example 37, that is,

$$A = \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix}$$

Here,

$$T = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

and $\lambda_1 = 2, \lambda_2 = 4$.

Hence,

$$\begin{aligned} \Phi(k) &= T \begin{pmatrix} 2^k & 0 \\ 0 & 4^k \end{pmatrix} T^{-1} \\ &= \frac{1}{2} \begin{pmatrix} 2^k + 4^k & 2^k - 4^k \\ 2^k - 4^k & 2^k + 4^k \end{pmatrix} \end{aligned}$$

and all solutions are given by

$$\begin{aligned} x_1(t) &= (\,2^k + 4^k\,)x_{10} + (\,2^k - 4^k\,)x_{20} \\ x_2(t) &= (\,2^k - 4^k\,)x_{10} + (\,2^k + 4^k\,)x_{20} \end{aligned}$$

## 5.9 Real Jordan form for real diagonalizable systems*

Suppose $A$ is real and has a complex eigenvalue

$$\lambda = \alpha + \jmath\omega$$

where $\alpha, \omega$ are real. Then any corresponding eigenvector $v$ is complex and can be expressed as

$$v = u + \jmath w$$

where $u, w$ are real and are linearly independent over the real scalars.

Since $A$ is real, $\bar{\lambda}$ is also an eigenvector of $A$ with eigenvector $\bar{v}$. So we can replace each $v$ and $\bar{v}$ in the basis of eigenvectors with the corresponding $u$ and $w$ to obtain a new basis; then we define a coordinate transformation where the columns of $T$ are these new basis vectors.

The relationship,

$$Av = \lambda v$$

and $A$ real implies

$$\begin{aligned} Au &= \alpha u - \omega w \\ Aw &= \omega u + \alpha w \end{aligned}$$

Hence, using the 'useful result on similarity transformations', we can choose $T$ so that $\Lambda = T^{-1}AT$ is given by

$$\begin{aligned} \Lambda &= \text{block-diag}\left(\lambda_1, \ldots, \lambda_l, \begin{pmatrix} \alpha_1 & \omega_1 \\ -\omega_1 & \alpha_1 \end{pmatrix}, \ldots, \begin{pmatrix} \alpha_m & \omega_m \\ -\omega_m & \alpha_m \end{pmatrix}\right) \\ &= \underbrace{\begin{pmatrix} \lambda_1 & & & & & \\ & \ddots & & & & \\ & & \lambda_l & & & \\ & & & \begin{pmatrix} \alpha_1 & \omega_1 \\ -\omega_1 & \alpha_1 \end{pmatrix} & & \\ & & & & \ddots & \\ & & & & & \begin{pmatrix} \alpha_m & \omega_m \\ -\omega_m & \alpha_m \end{pmatrix} \end{pmatrix}}_{\text{Real Jordan form of } A} \end{aligned}$$

where $\lambda_1, \ldots, \lambda_l$ are real eigenvalues of $A$ and $\alpha_1 + \jmath\omega_1, \ldots, \alpha_m + \jmath\omega_m$ are genuine complex eigenvalues.

**Example 58** Recall example 38.

**Example 59**

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

By inspection (recall companion matrices),

$$\det(sI - A) = s^4 - 1$$

So eigenvalues are:

$$\lambda_1 = 1 \quad \lambda_2 = -1 \quad \lambda_2 = \jmath \quad \lambda_3 = -\jmath$$

Since there are $n = 4$ distinct eigenvalues, the complex Jordan form is:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & \jmath & 0 \\ 0 & 0 & 0 & -\jmath \end{pmatrix}$$

The real Jordan form is:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}$$

### 5.9.1 CT systems

The solutions of

$$\dot{x} = Ax, \qquad x(0) = x_0$$

are given by

$$x(t) = \Phi(t)x_0$$

where the state transition matrix $\Phi(\cdot)$ is given by

$$\Phi(t) = T\hat{\Phi}(t)T^{-1}$$

with $\hat{\Phi}(t)$ being the block diagonal matrix:

$$\text{block-diag}\left(e^{\lambda_1 t}, \ldots, e^{\lambda_l t}, \begin{pmatrix} e^{\alpha_1 t}\cos(\omega_1 t) & e^{\alpha_1 t}\sin(\omega_1 t) \\ -e^{\alpha_1 t}\sin(\omega_1 t) & e^{\alpha_1 t}\cos(\omega_1 t) \end{pmatrix}, \ldots, \begin{pmatrix} e^{\alpha_m t}\cos(\omega_m t) & e^{\alpha_m t}\sin(\omega_m t) \\ -e^{\alpha_m t}\sin(\omega_m t) & e^{\alpha_m t}\cos(\omega_m t) \end{pmatrix}\right)$$

## 5.10 Diagonalizable second order systems

Only two cases for real Jordan form

$$\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \qquad\qquad \begin{pmatrix} \alpha & \omega \\ -\omega & \alpha \end{pmatrix}$$

where all the above scalars are real.

### 5.10.1 State plane portraits for CT systems

#### Real eigenvalues

#### Genuine complex eigenvalues

## 5.11 Exercises

**Exercise 57** Compute the eigenvalues and eigenvectors of the matrix,

$$A = \begin{pmatrix} 2 & -3 & 0 \\ 2 & -3 & 0 \\ 3 & -5 & 1 \end{pmatrix}$$

.

**Exercise 58** Compute the eigenvalues and eigenvectors of the matrix,

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 1 & 2 & 1 \\ -1 & 2 & 3 \end{pmatrix}$$

.

**Exercise 59** Compute the eigenvalues and eigenvectors of the following matrix.

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

**Exercise 60** Compute expressions for eigenvalues and linearly independent eigenvectors of

$$A = \begin{pmatrix} \alpha & \omega \\ -\omega & \alpha \end{pmatrix} \qquad\qquad \alpha, \omega \neq 0$$

**Exercise 61** Find expressions for the eigenvalues and linearly independent eigenvectors of matrices of the form:

$$\begin{pmatrix} \beta & \gamma \\ \gamma & \beta \end{pmatrix}$$

where $\gamma$ and $\beta$ are real with $\gamma \geq 0$. Obtain state plane portraits for

$$\begin{aligned} \beta &= 0 & \gamma &= 1 \\ \beta &= -3 & \gamma &= 2 \\ \beta &= 3 & \gamma &= 2 \end{aligned}$$

**Exercise 62** Consider an LTI system described by $\dot{x} = Ax$.
(a) Suppose that the vectors

$$\begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

are eigenvectors of $A$ corresponding to eigenvalues $-2$ and 3, respectively. What is the response $x(t)$ of the system to the initial condition

$$x(0) = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}?$$

(b) Suppose $A$ is a real matrix and the vector

$$\begin{pmatrix} 1 \\ \jmath \\ -1 \\ -\jmath \end{pmatrix}$$

is an eigenvector of $A$ corresponding to the eigenvalue $2 + 3\imath$. What is the response $x(t)$ of the system to the initial condition

$$x(0) = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}?$$

**Exercise 63** Suppose $A$ is a square matrix and the vectors

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

are eigenvectors of $A$ corresponding to eigenvalues $-1$ and 1, respectively. What is the response $x(t)$ of the system $\dot{x} = Ax$ to the following initial conditions.

$$(a) \quad x(0) = \begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix} \qquad (b) \quad x(0) = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix} \qquad (c) \quad x(0) = \begin{pmatrix} 0 \\ 2 \\ 4 \end{pmatrix}?$$

(d) Suppose $A$ is a real matrix and the vector

$$\begin{pmatrix} 1 + 2\jmath \\ 1 - \jmath \\ 3 \end{pmatrix}$$

is an eigenvector of $A$ corresponding to the eigenvalue $2 - 3\jmath$. What is the response $x(t)$ of the system $\dot{x} = Ax$ to the initial condition

$$x(0) = \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix}?$$

**Exercise 64** Suppose $A$ is a real square matrix and the vectors

$$\begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix} \quad \begin{pmatrix} 1 \\ 2 \\ 4 \\ 8 \end{pmatrix} \quad \begin{pmatrix} 1 \\ \jmath \\ -1 \\ -\jmath \end{pmatrix}$$

are eigenvectors of $A$ corresponding to eigenvalues $-1$ and $2$ and $\jmath$, respectively. What is the response $x(t)$ of the system $\dot{x} = Ax$ to the following initial conditions.

$$(a) \quad x(0) = \begin{pmatrix} 2 \\ -2 \\ 2 \\ -2 \end{pmatrix} \qquad (b) \quad x(0) = \begin{pmatrix} -1 \\ -2 \\ -4 \\ -8 \end{pmatrix} \qquad (c) \quad x(0) = \begin{pmatrix} 0 \\ 3 \\ 3 \\ 9 \end{pmatrix} \qquad (d) \quad x(0) = \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}?$$

**Exercise 65** Suppose $A$ is a real square matrix with characteristic polynomial

$$p(s) = (s+2)(s^2+4)^2$$

Suppose that all eigenvectors of $A$ are nonzero multiples of

$$\begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1+\jmath \\ -\jmath \\ 0 \\ 1 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} 0 \\ 1-\jmath \\ \jmath \\ 0 \\ 1 \end{pmatrix}$$

What is the solution to $\dot{x} = Ax$ with the following initial conditions.

$$(a) \quad x(0) = \begin{pmatrix} 2 \\ 4 \\ 6 \\ 8 \\ 0 \end{pmatrix}, \qquad (b) \quad x(0) = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \qquad (c) \quad x(0) = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}$$

(d) All all solutions of the system $\dot{x} = Ax$ bounded? Justify your answer.

**Exercise 66** Consider a discrete-time LTI system described by $x(k+1) = Ax(k)$.
(a) Suppose that the vectors

$$\begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

are eigenvectors of $A$ corresponding to eigenvalues $-2$ and $3$, respectively. What is the response $x(k)$ of the system to the initial condition

$$x(0) = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}?$$

(b) Suppose $A$ is a real matrix and the vector

$$\begin{pmatrix} 1 \\ \jmath \\ -1 \\ -\jmath \end{pmatrix}$$

is an eigenvector of $A$ corresponding to the eigenvalue $2+3\imath$. What is the response $x(k)$ of the system to the initial condition

$$x(0) = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}?$$

**Exercise 67** Determine whether or not the following matrix is nondefective.

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & 1 & 1 \end{pmatrix}$$

**Exercise 68** In vibration analysis, one usually encounters systems described by

$$M\ddot{q} + Kq = 0$$

where $q(t)$ is an $N$-vector and $M$ and $K$ are square matrices. Suppose $\omega$ is a real number which satisfies

$$\det(\omega^2 M - K) = 0\,.$$

Show that the above system has a solution of the form

$$q(t) = \sin(\omega t)v$$

where $v$ is a constant $N$-vector.

**Exercise 69** Determine whether or not the matrix,

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & -2 & 3 \end{pmatrix},$$

is diagonalizable, that is, determine whether or not there exists an invertible matrix $T$ such that $T^{-1}AT$ is diagonal. If $A$ is diagonalizable, determine $T$.

**Problem 1** Recall that a square matrix is said to be diagonalizable if there exists an invertible matrix $T$ so that $T^{-1}AT$ is diagonal. Determine whether or not each of the following matrices are diagonalizable. If not diagonalizable, give a reason. If diagonalizable, obtain $T$.

(a)

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

(b)

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

(c)

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

# Chapter 6

# $e^{At}$

Recall that all solutions to the scalar system $\dot{x} = ax$ satisfy $x(t) = e^{a(t-t_0)}x(t_0)$ for all $t$ and $t_0$. Recall also the power series expansion for $e^{at}$:

$$e^{at} = 1 + at + \frac{1}{2!}(at)^2 + \cdots + \frac{1}{k!}(at)^k + \cdots$$

The main result of this chapter is that all solutions of $\dot{x} = Ax$ satisfy

$$\boxed{x(t) = e^{A(t-t_0)}x(t_0)} \tag{6.1}$$

for all $t$ and $t_0$ where

$$\boxed{e^{At} = 1 + At + \frac{1}{2!}(At)^2 + \cdots + \frac{1}{k!}(At)^k + \cdots} \tag{6.2}$$

## 6.1 State transition matrix

### 6.1.1 Continuous time

Consider the initial value problem specified by

$$\dot{x} = Ax \qquad \text{and} \qquad x(t_0) = x_0\,. \tag{6.3}$$

It can be shown that for each initial time $t_0$ and each initial state $x_0$ there is an unique solution $x(\cdot)$ to the above initial value problem and this solution is defined for all time $t$. Let $\tilde{\phi}(t; t_0, x_0)$ be the value of this solution at $t$, that is, for any $t_0$ and $x_0$, the solution to (6.3) is given by $x(t) = \tilde{\phi}(t; t_0, x)$. Since the system $\dot{x} = Ax$ is time-invariant, we must have $\tilde{\phi}(t; t_0, x_0) = \phi(t-t_0; x_0)$ for some function $\phi$. Since the system under consideration is linear, one can show that, for any $t$ and $t_0$, the solution at time $t$ depends linearly on the initial state $x_0$ at time $t_0$; hence there exists a matrix $\Phi(t-t_0)$ such that $\phi(t-t_0, x_0) = \Phi(t-t_0)x_0$. We call the matrix valued function $\Phi(\cdot)$ the state transition matrix for the system $\dot{x} = Ax$. So, we have concluded that there is a matrix valued function $\Phi(\cdot)$ such that for any $t_0$ and $x_0$, the solution $x(\cdot)$ to the above initial value problem is given by

$$\boxed{x(t) = \Phi(t-t_0)x_0}$$

We now show that $\Phi$ must satisfy the following matrix differential equation and initial condition:

$$\boxed{\begin{array}{rcl} \dot{\Phi} & = & A\Phi \\ \Phi(0) & = & I \end{array}}$$

To demonstrate this, consider any initial state $x_0$ and let $x(\cdot)$ be the solution to (6.3) with $t_0 = 0$. Then $x(t) = \Phi(t)x_0$. Considering $t = 0$, we note that

$$x_0 = x(0) = \Phi(0)x_0\,;$$

that is $\Phi(0)x_0 = x_0$. Since the above holds for any $x_0$, we must have $\Phi(0) = I$.

We now note that for all $t$,

$$\dot{\Phi}(t)x_0 = \dot{x}(t) = Ax(t) = A\Phi(t)x_0\,;$$

that is $\dot{\Phi}(t)x_0 = A\Phi(t)x_0$. Since the above must hold for any $x_0$, we obtain that $\dot{\Phi}(t) = A\Phi(t)$.

Note that the above two conditions on $\Phi$ uniquely specify $\Phi$ (Why?) The purpose of this section is to examine $\Phi$ and look at ways of evaluating it.

### 6.1.2 Discrete time

$$x(k+1) = Ax(k) \qquad \text{and} \qquad x(k_0) = x_0$$

We have

$$\boxed{x(k) = \Phi(k-k_0)x_0}$$

The matrix valued function $\Phi(\cdot)$ is called the state transition matrix and is given by

$$\boxed{\Phi(k) = A^k}$$

Note that, if $A$ is singular, $\Phi(k)$ is not defined for $k < 0$. That was easy; now for CT.

## 6.2 Polynomials of a square matrix

### 6.2.1 Polynomials of a matrix

We define polynomials of a matrix as follows. Consider any polynomial $p$ of a scalar variable $s$ given by:

$$p(s) = a_0 + a_1 s + a_2 s^2 + \ldots + a_m s^m$$

where $a_0, \ldots, a_m$ are scalars. If $A$ is a square matrix, we define $p(A)$ as follows:

$$p(A) := a_0 I + a_1 A + a_2 A^2 + \ldots + a_m A^m$$

that is, we replace $s^k$ with $A^k$. (Recall that $s^0 = 1$ and $A^0 = I$)

The following properties can be readily deduced from the above definition.

(a) The matrices $A$ and $p(A)$ commute, that is,

$$Ap(A) = p(A)A$$

(b) Consider any nonsingular matrix $T$, and suppose $A = T\Lambda T^{-1}$. Then

$$p(A) = Tp(\Lambda)T^{-1}$$

(c) Suppose $A$ is diagonal, that is,

$$A = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

then

$$p(A) = \begin{pmatrix} p(\lambda_1) & 0 & \cdots & 0 \\ 0 & p(\lambda_2) & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & p(\lambda_n) \end{pmatrix}$$

(d) Suppose $A$ is diagonalizable, that is, for some nonsingular $T$,

$$A = T \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} T^{-1}$$

Then

$$p(A) = T \begin{pmatrix} p(\lambda_1) & 0 & \cdots & 0 \\ 0 & p(\lambda_2) & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & p(\lambda_n) \end{pmatrix} T^{-1}$$

(e) If $\lambda$ is an eigenvalue of $A$ with eigenvector $v$, then $p(\lambda)$ is an eigenvalue of $p(A)$ with eigenvector $v$.

Properties (b)-(d) are useful for computing $p(A)$.

**Example 60** For

$$A = \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix}$$

we found that

$$T = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

results in

$$\Lambda = T^{-1}AT = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}$$

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

$$\begin{aligned} p(A) &= Tp(\Lambda)T^{-1} \\ &= T\begin{pmatrix} p(2) & 0 \\ 0 & p(4) \end{pmatrix}T^{-1} \\ &= \frac{1}{2}\begin{pmatrix} p(2)+p(4) & p(2)-p(4) \\ p(2)-p(4) & p(2)+p(4) \end{pmatrix} \end{aligned}$$

Suppose $p$ is the characteristic polynomial of $A$. Since $2, 4$ are the eigenvalues of $A$, we have $p(2) = p(4) = 0$; hence

$$p(A) = 0$$

### 6.2.2 Cayley-Hamilton Theorem

The following result is a fundamental result in linear algebra and is very useful in systems and control.

**Theorem 3** *If $p$ is the characteristic polynomial of a square matrix $A$, then*

$$p(A) = 0\,.$$

PROOF. The proof is easy for a diagonalizable matrix $A$. Recall that such a matrix is similar to a diagonal matrix whose diagonal elements $\lambda_1, \dots, \lambda_n$ are the eigenvalues of $A$; hence $p(\lambda_1) = \dots = p(\lambda_n) = 0$. Now use property (d) above. We will leave the proof of the general nondiagonalizable case for another day. ■

Suppose the characteristic polynomial of $A$ is

$$\det(sI - A) = a_0 + a_1 s + \dots + a_{n-1}s^{n-1} + s^n$$

Then the Cayley Hamilton theorem states that

$$a_0 I + a_1 A + \dots + a_{n-1}A^{n-1} + A^n = 0$$

hence,

$$A^n = -a_0 I - a_1 A - \ldots - a_{n-1}A^{n-1}$$

From this one can readily show that for any $m \geq n$, $A^m$ can be expressed as a linear combination of $I, A, \ldots, A^{n-1}$. Hence, any polynomial of $A$ can be can be expressed as a linear combination of $I, A, \ldots, A^{n-1}$.

**Example 61** Consider a discrete-time system described by $x(k+1) = Ax(k)$ and suppose that all the eigenvalues of $A$ are zero. Then the characteristic polynomial of $A$ is given by $p(s) = s^n$ where $A$ is $n \times n$. It now follows from the Cayley Hamilton theorem that

$$A^n = 0$$

Since all solutions of the system satisfy $x(k) = A^k x(0)$ for $k = 0, 1, 2, \ldots$, it follows that

$$x(k) = 0 \qquad \text{for} \quad k \geq n.$$

Thus, *all solutions go to zero in at most $n$ steps.*

### Minimal polynomial of $A$

The minimal polynomial (minpoly) of a square matrix $A$ is the monic polynomial $q$ of lowest order for which $q(A) = 0$.

The order of minpoly $q$ is always less than or equal to the order of charpoly $p$; its roots are precisely the eigenvalues of $A$ and $p(s) = g(s)q(s)$.

**Example 62** The matrix

$$A = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

has charpoly $p(s) = s^2$ and minpoly $q(s) = s$. The matrix

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

has charpoly $p(s) = s^2$ and minimal poly $q(s) = s^2$.

• It follows from property (d) above that for a *diagonalizable matrix* $A$, its minimal polynomial is

$$\prod_{i=1}^{l}(s - \lambda_i)$$

where $\lambda_i$ are the *distinct* eigenvalues of $A$.

**Exercise 70** Without doing any matrix multiplications, compute $A^4$ for

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 564 & 1 & 0 & 0 \end{pmatrix}$$

Justify your answer.

## 6.3 Functions of a square matrix

Consider a complex valued function $f$ of a complex variable $\lambda$. Suppose $f$ is analytic in some open disk $\{\lambda \in \mathbb{C} : |\lambda| < \rho\}$ of radius $\rho > 0$. Then, provided $|\lambda| < \rho$, $f(\lambda)$ can be expressed as the sum of a convergent power series; specifically

$$f(\lambda) = \sum_{k=0}^{\infty} a_k \lambda^k = a_0 + a_1\lambda + a_2\lambda^2 + \dots$$

where

$$a_k = \frac{1}{k!} f^{(k)}(0) = \frac{1}{k!} \frac{d^k f}{d\lambda^k}(0)$$

Suppose $A$ is a square complex matrix and $|\lambda_i| < \rho$ for every eigenvalue $\lambda_i$ of $A$. Then, it can be shown that the power series $\{\sum_{k=0}^{n} a_k A^k\}_{n=0}^{\infty}$ is convergent. We define $f(A)$ to be the sum of this power series, that is,

$$f(A) := \sum_{k=0}^{\infty} a_k A^k = a_0 I + a_1 A + a_2 A^2 + \dots$$

**Example 63** Consider

$$f(\lambda) = \frac{1}{1-\lambda}$$

This is analytic in the open unit disk ($\rho = 1$) and

$$f(\lambda) = \sum_{k=0}^{\infty} \lambda^k = 1 + \lambda + \lambda^2 + \dots$$

Consider any matrix $A$ with eigenvalues in the open unit disk. Then the series $\{\sum_{k=0}^{n} A^k\}_{n=0}^{\infty}$ is convergent and we let

$$f(A) := \sum_{k=0}^{\infty} A^k$$

Multiplication of both sides of the above equality by $I - A$ yields

$$(I - A)f(A) = \sum_{k=0}^{\infty} (A^k - A^{k+1}) = I;$$

hence the matrix $I - A$ is invertible with inverse $f(A)$ and we obtain the following power series expansion for $(I - A)^{-1}$:

$$(I - A)^{-1} = \sum_{k=0}^{\infty} A^k = I + A + A^2 + \dots$$

**Power series expansion for $(sI - A)^{-1}$.** We present here an expansion which will be useful later on. Consider any square matrix $A$ and any complex number $s$ which satisfies:

$$|s| > |\lambda_i| \qquad \text{for every eigenvalue } \lambda_i \text{ of } A\,.$$

Then $sI - A$ is invertible and $(sI - A)^{-1} = \frac{1}{s}(I - \frac{1}{s}A)^{-1}$. Also, the eigenvalues of $\frac{1}{s}A$ are in the open unit disk; hence

$$(I - \frac{1}{s}A)^{-1} = \sum_{k=0}^{\infty} \frac{1}{s^k} A^k = I + \frac{1}{s}A + \frac{1}{s^2}A^2 + \ldots$$

and

$$\boxed{(sI - A)^{-1} = \sum_{k=0}^{\infty} \frac{1}{s^{k+1}} A^k = \frac{1}{s}I + \frac{1}{s^2}A + \frac{1}{s^3}A^2 + \ldots}$$

**Properties of $f(A)$.** The following properties can be readily deduced from the above definition of $f(A)$.

(a) The matrices $A$ and $f(A)$ commute, that is,

$$Af(A) = f(A)A\,.$$

(b) Consider any nonsingular matrix $T$ and suppose $A = T\Lambda T^{-1}$. Then

$$f(A) = Tf(\Lambda)T^{-1}\,.$$

(c) Suppose $A$ is diagonal, that is,

$$A = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

then

$$f(A) = \begin{pmatrix} f(\lambda_1) & 0 & \cdots & 0 \\ 0 & f(\lambda_2) & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & f(\lambda_n) \end{pmatrix}$$

(d) Suppose $A$ is diagonalizable, that is, for some nonsingular $T$,

$$A = T \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} T^{-1}$$

Then

$$f(A) = T \begin{pmatrix} f(\lambda_1) & 0 & \cdots & 0 \\ 0 & f(\lambda_2) & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & f(\lambda_n) \end{pmatrix} T^{-1}$$

(e) If $\lambda$ is an eigenvalue of $A$ with eigenvector $v$, then $f(\lambda)$ is an eigenvalue of $f(A)$ with eigenvector $v$.

- It follows from Cayley Hamilton, that $f(A)$ is a linear combination of $I, A, \ldots, A^{n-1}$

Properties (b)-(d) are useful for calculating $f(A)$. When $A$ is diagonalizable, one simply has to evaluate the scalars $f(\lambda_1), \ldots, f(\lambda_n)$ and then compute $T\,\mathrm{diag}\,(f(\lambda_1), \ldots, f(\lambda_n))\,T^{-1}$.

**Example 64** For

$$A = \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix}$$

we found that

$$T = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

results in

$$\Lambda = T^{-1}AT = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}$$

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

$$\begin{aligned} f(A) &= Tf(\Lambda)T^{-1} \\ &= T \begin{pmatrix} f(2) & 0 \\ 0 & f(4) \end{pmatrix} T^{-1} \\ &= \frac{1}{2} \begin{pmatrix} f(2)+f(4) & f(2)-f(4) \\ f(2)-f(4) & f(2)+f(4) \end{pmatrix} \end{aligned}$$

**Exercise 71** Compute an explicit expression (see last example) for $f(A)$ where

$$A = \begin{pmatrix} 9 & -2 \\ -2 & 6 \end{pmatrix}$$

### 6.3.1 The matrix exponential: $e^A$

Recall the exponential function:

$$f(\lambda) = e^{\lambda}$$

This function is analytic in the whole complex plane (that is $\rho = \infty$) and

$$\frac{d^k e^{\lambda}}{d\lambda^k} = e^{\lambda}$$

for $k = 1, 2, \ldots$; hence for any complex number $\lambda$,

$$e^{\lambda} \quad = \quad \sum_{k=0}^{\infty} \frac{1}{k!}\lambda^k \quad = \quad 1 + \lambda + \frac{1}{2}\lambda^2 + \frac{1}{3!}\lambda^3 + \ldots$$

Thus the exponential of any square matrix $A$ is defined by:

$$\boxed{e^A \quad := \quad \sum_{k=0}^{\infty} \frac{1}{k!}A^k \quad = \quad I + A + \frac{1}{2!}A^2 + \ldots}$$

and this power series converges for every square matrix $A$. Some properties:

(a) For any zero matrix 0,

$$e^0 = I$$

(b) Suppose $t$ is a scalar variable.

$$\frac{de^{At}}{dt} = Ae^{At} = e^{At}A$$

The following results also hold:

(i)

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

(ii) If $\lambda$ is an eigenvalue of $A$ with eigenvector $v$, then $e^{\lambda}$ is an eigenvalue of $e^A$ with eigenvector $v$. So, $e^A$ has no eigenvalue at 0 and hence:

(iii) $e^A$ is nonsingular.

(iv) The following result also holds

$$e^{A+B} = e^A e^B \qquad \text{if} \quad A \text{ and } B \text{ commute}$$

### 6.3.2 Other matrix functions

$$\cos(A)$$
$$\sin(A)$$

## 6.4 The state transition matrix: $e^{At}$

Using the above definitions and abusing notation ($e^{At}$ instead of $e^{tA}$),

$$\boxed{e^{At} := \sum_{k=0}^{\infty} \frac{1}{k!}(tA)^k \quad = \quad I + tA + \frac{1}{2!}(tA)^2 + \frac{1}{3!}(tA)^3 + \ldots}$$

Letting

$$\Phi(t) := e^{At}$$

we have

$$\begin{aligned} \Phi(0) &= I \\ \dot{\Phi} &= A\Phi \end{aligned}$$

Hence, the solution of

$$\dot{x} = Ax \qquad x(0) = x_0$$

is given by

$$\boxed{x(t) = e^{At}x_0}$$

## 6.5 Computation of $e^{At}$

### 6.5.1 MATLAB

```
>> help expm

 EXPM   Matrix exponential.
        EXPM(X) is the matrix exponential of X.  EXPM is computed using
        a scaling and squaring algorithm with a Pade approximation.
        Although it is not computed this way, if X has a full set
        of eigenvectors V with corresponding eigenvalues D, then
        [V,D] = EIG(X) and EXPM(X) = V*diag(exp(diag(D)))/V.
        See EXPM1, EXPM2 and EXPM3 for alternative methods.

        EXP(X) (that's without the M) does it element-by-element.
```

### 6.5.2 Numerical simulation

Let $\phi_i$ be the $i$-th column of the state transition matrix $\Phi(t) = e^{At}$; thus

$$\Phi(t) = \begin{pmatrix} \phi_1(t) & \phi_2(t) & \cdots & \phi_n(t) \end{pmatrix} .$$

Let $x$ be the unique solution to $\dot{x} = Ax$ and $x(0) = x_0$. Since $x(t) = \Phi(t)x_0$ we have

$$x(t) = x_{01}\phi_1(t) + \cdots + x_{0n}\phi_n$$

where $x_{0i}$ is the $i$-th component of $x_0$. Hence $\phi_i$, is the solution to

$$\dot{x} = Ax \qquad x(0) = e^i$$

where $e^i$ is all zeros except for its $i$-th component which is 1.

### 6.5.3 Jordan form

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

where $\Lambda = T^{-1}AT$ and $\Lambda$ is the Jordan form of $A$. We have already seen this for diagonalizable systems.

### 6.5.4 Laplace style

$$\dot{x} = Ax \qquad\qquad x(0) = x_0$$

Suppose

$$X(s) = \mathcal{L}(x)(s)$$

is the Laplace transform of $x(\cdot)$ evaluated at $s \in \mathbb{C}$. Taking the Laplace transform of $\dot{x} = Ax$ yields:

$$sX(s) - x_0 = AX(s)$$

Hence, except when $s$ is an eigenvalue of $A$, $sI - A$ is invertible and

$$\mathcal{L}(x)(s) = X(s) = (sI - A)^{-1}x_0$$

Since $x(t) = e^{At}x_0$ for all $x_0$, we must have

$$\boxed{\begin{aligned} \mathcal{L}(e^{At}) &= (sI - A)^{-1} \\ e^{At} &= \mathcal{L}^{-1}\left((sI - A)^{-1}\right) \end{aligned}}$$

**Example 65** Recall

$$A = \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix}$$

So,

$$(sI - A)^{-1} = \begin{pmatrix} \frac{s-3}{(s-2)(s-4)} & \frac{-1}{(s-2)(s-4)} \\ \frac{-1}{(s-2)(s-4)} & \frac{s-3}{(s-2)(s-4)} \end{pmatrix}$$

Since

$$\begin{aligned} \frac{s-3}{(s-2)(s-4)} &= \frac{1}{2}\left(\frac{1}{s-2} + \frac{1}{s-4}\right) \\ \frac{-1}{(s-2)(s-4)} &= \frac{1}{2}\left(\frac{1}{s-2} - \frac{1}{s-4}\right) \end{aligned}$$

and

$$\mathcal{L}^{-1}\left(\frac{1}{s-2}\right) = e^{2t} \qquad \mathcal{L}^{-1}\left(\frac{1}{s-4}\right) = e^{4t}$$

we have,

$$\begin{aligned} e^{At} &= \mathcal{L}^{-1}\left((sI - A)^{-1}\right) \\ &= \frac{1}{2}\begin{pmatrix} e^{2t} + e^{4t} & e^{2t} - e^{4t} \\ e^{2t} - e^{4t} & e^{2t} + e^{4t} \end{pmatrix} \end{aligned}$$

**Exercise 72** Compute $e^{At}$ at $t = \ln(2)$ for

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

using *all* methods of this section.

## 6.6 Sampled-data systems

Here we discuss the process of discretizing a continuous-time system to obtain a discrete-time system.

Consider a continuous-time LTI system described by

$$\dot{x}_c(t) = A_c x_c(t)\,.$$

Suppose we only look at the behavior of this system at discrete instants of time, namely

$$\cdots, -2T, T, 0, T, 2T, \cdots$$

where $T > 0$ is called the sampling time.

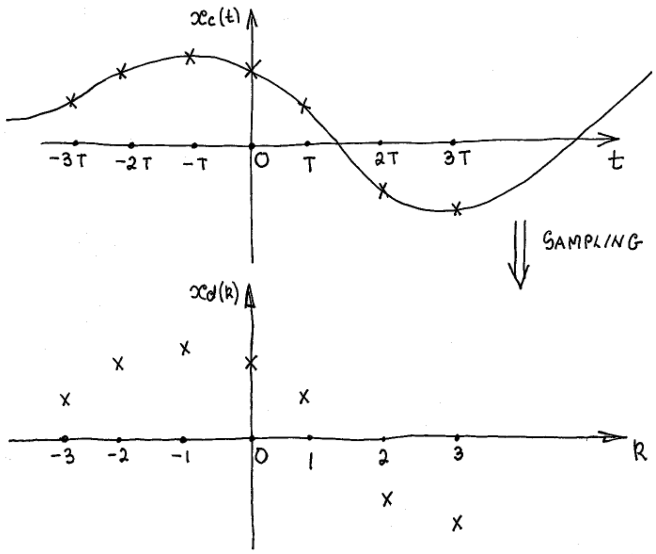



Figure 6.1: Sampling

If we introduce the discrete-time state defined

$$x_d(k) = x_c(kT) \qquad k = \ldots, -2, -1, 0, 1, 2, \ldots\,,$$

then,

$$\begin{aligned} x_d(k+1) &= x_c(\,(k+1)T\,) \\ &= e^{A_cT} x_c(kT) \\ &= e^{A_cT} x_d(k) \end{aligned}$$

Hence, this sampled data system is described by the discrete-time system

$$x_d(k+1) = A_d x_d(k)$$

where

$$\boxed{A_d = e^{A_cT}}$$

The eigenvalues of $A_d$ are

$$e^{\lambda_1 T}, \ldots e^{\lambda_l T}$$

where $\lambda_1, \ldots, \lambda_l$ are the eigenvalues of $A_c$. Note that,

$$\begin{array}{lll} |e^{\lambda_i T}| < 1 & \text{iff} & \Re(\lambda_i) < 0 \\ |e^{\lambda_i T}| = 1 & \text{iff} & \Re(\lambda_i) = 0 \\ |e^{\lambda_i T}| > 1 & \text{iff} & \Re(\lambda_i) > 0 \\ & & \\ e^{\lambda_i T} = 1 & \text{iff} & \lambda_i = 0 \end{array}$$

It should be clear that *all eigenvalues of $A_d$ are nonzero, that is $A_d$ in invertible.*

Note that if we approximate $e^{A_c T}$ by the first two terms in its power series, that is,

$$e^{A_c T} \approx I + A_c T$$

then,

$$A_d \approx I + A_c T\,.$$

## 6.7 Exercises

**Exercise 73** Suppose $A$ is a $3 \times 3$ matrix, $b$ is a 3-vector and

$$\det(sI - A) = s^3 + 2s^2 + s + 1$$

(a) Express $A^3 b$ in terms of $b, Ab, A^2 b$.

(b) Suppose that the matrix

$$T = \begin{pmatrix} b & Ab & A^2 b \end{pmatrix}$$

is invertible. What is the matrix $\Lambda := T^{-1} A T$? Specify each element of $\Lambda$ explicitly.

**Exercise 74** Suppose $A$ is a $3 \times 3$ matrix and

$$\det(sI - A) = s^3 + 2s^2 + s + 1$$

(a) Express $A^3$ in terms of $I, A, A^2$.

(b) Express $A^5$ in terms of $I, A, A^2$.

(c) Express $A^{-1}$ in terms of $I, A, A^2$

**Exercise 75** Suppose $A$ is a $3 \times 3$ matrix with eigenvalues 1, $2\pi\jmath$ and $-2\pi\jmath$. What are the eigenvalues of $e^A$.

**Exercise 76** Consider the differential equation

$$\dot{x} = Ax \tag{6.4}$$

where $A$ is a square matrix. Show that if $A$ has $\jmath 2\pi$ as an eigenvalue, then there is an nonzero initial state $x_0$ such that (6.4) has a solution $x$ which satisfies $x(1) = x(0) = x_0$.

**Exercise 77** Suppose $A$ is a square matrix.
(a) Obtain expressions for

$$\cos(At) \qquad \text{and} \qquad \sin(At)\,.$$

(b) Show that

$$\cos(At) = (e^{\jmath At} + e^{-\jmath At})/2 \qquad \text{and} \qquad \sin(At) = (e^{\jmath At} - e^{-\jmath At})/2\jmath$$

(c) Show that

$$\frac{d\cos(At)}{dt} = -A\sin(At) \qquad \text{and} \qquad \frac{d\sin(At)}{dt} = A\cos(At)\,.$$

**Exercise 78** Compute $e^{At}$ for the matrix

$$A = \begin{pmatrix} 3 & -5 \\ -5 & 3 \end{pmatrix}$$

**Exercise 79** Compute $e^{At}$ for the matrix,

$$A = \begin{pmatrix} -11 & 20 \\ -6 & 11 \end{pmatrix}.$$

**Exercise 80** Compute $e^{At}$ for the matrix

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

**Exercise 81** Suppose $A$ is a matrix with four distinct eigenvalues: $-1, +1-j, j$. Show that $A^4 = I$.

**Exercise 82** Suppose

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

What is $A^{20} + 4A^4$?

**Exercise 83** Compute $e^{At}$ for the matrix,

$$A = \begin{pmatrix} -3 & 2 \\ -4 & 3 \end{pmatrix}.$$

**Exercise 84** Compute $e^{At}$ for the matrix

$$A = \begin{pmatrix} 3 & -5 \\ -5 & 3 \end{pmatrix}$$

**Exercise 85** Compute $e^{At}$ at $t = \ln(2)$ for

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

using *all* four methods mentioned in the notes.

# Chapter 7

# Stability and boundedness

Consider a general nonlinear system described by

$$\dot{x} = f(x) \tag{7.1}$$

where $x(t)$ is a real $n$-vector and $t$ is a real scalar. By a solution of (7.1) we mean any continuous function $x(\cdot) : [0, t_1) \to \mathbb{R}^n$ with $t_1 > 0$, which satisfies $\dot{x}(t) = f(x(t))$ for $0 \le t < t_1$.

## 7.1 Boundedness of solutions

**Defn.** (Boundedness) *A solution $x(\cdot)$ is* bounded *if there exists $\beta \ge 0$ such that*

$$||x(t)|| \le \beta \qquad \textit{for all} \qquad t \ge 0$$

*A solution is* unbounded *if it is not bounded.*

It should be clear from the above definitions that a solution $x(\cdot)$ is bounded if and only if each component $x_i(\cdot)$ of $x(\cdot)$ is bounded. Also, a solution is unbounded if and only if at least one of its components is unbounded. So, $x(t) = [\ e^{-t} \quad e^{-2t}\ ]^T$ is bounded while $x(t) = [\ e^{-t} \quad e^{t}\ ]^T$ is unbounded.

**Example 66** All solutions of

$$\dot{x} = 0$$

are bounded.

**Example 67** All solutions of

$$\dot{x} = x - x^3$$

are bounded.

**Example 68** All solutions (except the zero solution) of

$$\dot{x} = x$$

are unbounded.

**Example 69** Consider

$$\dot{x} = x^2 \,, \qquad x(0) = x_0$$

If $x_0 > 0$, the corresponding solution has a finite escape time and is unbounded. If $x_0 < 0$, the corresponding solution is bounded.

**Example 70** Undamped oscillator.

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

All solutions are bounded.

**Boundedness and linear time-invariant systems.** Consider a general LTI (linear time-invariant) system

$$\dot{x} = Ax \tag{7.2}$$

Recall that every solution of this system has the form

$$x(t) = \sum_{i=1}^{l} \sum_{j=0}^{n_i - 1} t^j e^{\lambda_i t} v^{ij}$$

where $\lambda_1, \ldots, \lambda_l$ are the eigenvalues of $A$, $n_i$ is the index of $\lambda_i$, and the constant vectors $v^{ij}$ depend on initial state.

We say that an eigenvalue $\lambda$ of $A$ is non-defective if its index is one; this means that the algebraic multiplicity and the geometric multiplicity of $\lambda$ are the same. Otherwise we say $\lambda$ is defective.

Hence we conclude that all solutions of (7.2) are bounded if and only if for each eigenvalue $\lambda_i$ of $A$:

(b1) $\Re(\lambda_i) \leq 0$ and

(b2) if $\Re(\lambda_i) = 0$ then $\lambda_i$ is non-defective.

If there is an eigenvalue $\lambda_i$ of $A$ such that either

(u1) $\Re(\lambda_i) > 0$ or

(u2) $\Re(\lambda_i) = 0$ and $\lambda_i$ is defective

then, the system has some unbounded solutions.

**Example 71** Unattached mass

$$
\begin{aligned}
\dot{x}_1 &= x_2 \\
\dot{x}_2 &= 0
\end{aligned}
$$

Here

$$
A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}
$$

has a single eigenvalue 0. This eigenvalue has algebraic multiplicity 2 but geometric multiplicity 1; hence some of the solutions of the system are unbounded. One example is

$$
x(t) = \begin{pmatrix} t \\ 1 \end{pmatrix}
$$

**Example 72**

$$
\begin{aligned}
\dot{x}_1 &= 0 \\
\dot{x}_2 &= 0
\end{aligned}
$$

Here

$$
A = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}
$$

has a single eigenvalue 0. This eigenvalue has algebraic multiplicity 2 and geometric multiplicity 2. Hence all the solutions of the system are bounded. Actually every state is an equilibrium state and every solution is constant.

**Example 73 (Resonance)** Consider a simple linear oscillator subject to a sinusoidal input of amplitude $W$:

$$
\ddot{q} + q = W\sin(\omega t + \phi)
$$

Resonance occurs when $\omega = 1$. To see this, let

$$
x_1 := q\,, \quad x_2 := \dot{q}\,, \quad x_3 := W\sin(\omega t + \phi)\,, \quad x_4 := \omega W\cos(\omega t + \phi)
$$

to yield

$$
\dot{x} = Ax
$$

where

$$
A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -\omega^2 & 0 \end{pmatrix}
$$

If $\omega = 1$ then, $A$ has eigenvalues $\jmath$ and $-\jmath$. These eigenvalues have algebraic multiplicity two but geometric multiplicity one; hence the system has unbounded solutions.

## 7.2 Stability of equilibrium states

Suppose $x^e$ is an equilibrium state of the system $\dot{x} = f(x)$. Then, whenever $x(0) = x^e$, we have (assuming uniqueness of solutions) $x(t) = x^e$ for all $t \geq 0$. Roughly speaking, we say that $x^e$ is a stable equilibrium state for the system if the following holds. If the initial state of the system is close to $x^e$, then the resulting solution is close to $x^e$. The formal definition is as follows.

**Defn.** (Stability) *An equilibrium state $x^e$ is* stable *if for each $\epsilon > 0$ there exists $\delta > 0$ such that whenever $||x(0)-x^e|| < \delta$ one has $||x(t)-x^e|| < \epsilon$ for all $t \geq 0$.*

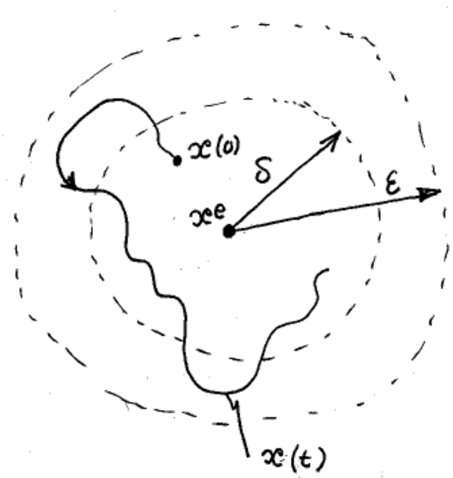



Figure 7.1: Stability of an equilibrium state

*An equilibrium state $x^e$ is said to be* unstable *if it is not stable.*

**Example 74**

$$\dot{x} = 0$$

Every equilibrium state is stable. (Choose $\epsilon = \delta$.)

**Example 75**

$$\dot{x} = x$$

The origin is unstable.

**Example 76**

$$\dot{x} = x - x^3$$

The origin is unstable; the remaining equilibrium states 1 and $-1$ are stable.

**Example 77** Undamped oscillator
The origin is stable.

**Example 78** Simple pendulum

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

$(0,0)$ is stable; $(\pi,0)$ is unstable.

**Example 79** Van der Pol oscillator

$$\begin{aligned}\dot{x}_1 &= x_2 \\ \dot{x}_2 &= (1-x_1^2)x_2 - x_1\end{aligned}$$

Figure 7.2: Van der Pol oscillator

The origin is unstable. However, all solutions are bounded.

**Stability of linear time-invariant systems.** It can be shown that every equilibrium state of a LTI system (7.2) is stable if and only if all eigenvalues $\lambda_i$ of $A$ satisfy conditions (b1) and (b2) above. Hence every equilibrium state of a LTI system is stable if and only if all solutions are bounded.
It can also be shown that every equilibrium state is unstable if and only if there is an eigenvalue $\lambda_i$ of $A$ which satisfies condition (u1) or (u2) above. Hence every equilibrium state of a LTI system is unstable if and only if there are unbounded solutions.

## 7.3 Asymptotic stability

### 7.3.1 Global asymptotic stability

**Defn.** (Global asymptotic stability) *An equilibrium state $x^e$ is* globally asymptotically stable (GAS) *if*

*(a) It is stable*

*(b) Every solution $x(\cdot)$ converges to $x^e$ with increasing time, that is,*

$$\lim_{t\to\infty} x(t) = x^e$$

If $x^e$ is a globally asymptotically stable equilibrium state, then there are no other equilibrium states and all solutions are bounded. In this case we say that the system $\dot{x} = f(x)$ is globally asymptotically stable.

**Example 80** The system

$$\dot{x} = -x$$

is GAS.

**Example 81** The system

$$\dot{x} = -x^3$$

is GAS.

### 7.3.2 Asymptotic stability

In asymptotic stability, we do not require that all solutions converge to the equilibrium state; we only require that all solutions which originate in some neighborhood of the equilibrium state converge to the equilibrium state.

**Defn.** (Asymptotic stability) *An equilibrium state $x^e$ is* asymptotically stable (AS) *if*

*(a) It is stable.*

*(b) There exists $R > 0$ such that whenever $||x(0)-x^e|| < R$ one has*

$$\lim_{t\to\infty} x(t) = x^e\,. \tag{7.3}$$

The region of attraction of an equilibrium state $x^e$ which is AS is the set of initial states which result in (7.3), that is it is the set of initial states which are attracted to $x^e$. Thus, the region of attraction of a globally asymptotically equilibrium state is the whole state space.

Figure 7.3: Asymptotic stability

**Example 82**

$$\dot{x} = -x$$

**Example 83**

$$\dot{x} = x - x^3$$

The equilibrium states $-1$ and $1$ are AS with regions of attraction $(-\infty, 0)$ and $(0, \infty)$, respectively.

**Example 84** Damped simple pendulum

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

The zero state is AS but not GAS.

**Example 85** Reverse Van der Pol oscillator

$$\begin{aligned} \dot{x}_1 &= -x_2 \\ \dot{x}_2 &= -(1 - x_1^2)x_2 + x_1 \end{aligned}$$

The zero state is AS but not GAS. Also, the system has unbounded solutions.

Figure 7.4: Reverse Van der Pol oscillator

**LTI systems.** For LTI systems, it should be clear from the general form of the solution that the zero state is AS if and only if all the eigenvalues $\lambda_i$ of $A$ have negative real parts, that is,

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

Also AS is equivalent to GAS.

**Example 86** The system

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

is GAS.

**Example 87** The system

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

is GAS.

## 7.4 Exponential stability

We now present the "strongest" form of stability considered in this section.

**Defn.** (Global exponential stability) *An equilibrium state $x^e$ is* globally exponentially stable (GES) *with* rate of convergence $\alpha > 0$ *if there exists $\beta > 0$ such that every solution satisfies*

$$||x(t)-x^e|| \quad \leq \quad \beta||x(0)-x^e|| \exp(-\alpha t) \qquad \textit{for all} \quad t \geq 0$$

**Example 88**

$$\dot{x} = -x$$

GES with rate $\alpha = 1$.

Note that global exponential stability implies global asymptotic stability, but, in general, the converse is not true. This is illustrated in the next example. For linear time-invariant systems, GAS and GES are equivalent.

**Example 89**

$$\dot{x} = -x^3$$

Solutions satisfy

$$x(t) = \frac{x_0}{\sqrt{1 + 2x_0^2 t}} \qquad \text{where} \qquad x_0 = x(0)\,.$$

GAS but not GES

**DEFN.** (Exponential stability) *An equilibrium state* $x^e$ *is* exponentially stable (ES) *with* rate of convergence $\alpha > 0$ *if there exists* $R > 0$ *and* $\beta > 0$ *such that whenever* $||x(0)-x^e|| < R$ *one has*

$$||x(t)-x^e|| \quad \leq \quad \beta||x(0)-x^e||\exp(-\alpha t) \qquad \textit{for all} \quad t \geq 0$$

Figure 7.5: Exponential stability

Note that exponential stability implies asymptotic stability, but, in general, the converse is not true.

**Example 90**

$$\dot{x} = -x^3$$

GAS but not even ES

**Example 91**

$$\dot{x} = -\frac{x}{1+x^2}$$

GAS, ES, but not GES

## 7.5 LTI systems

Consider a LTI system

$$\dot{x} = Ax \tag{7.4}$$

Recall that every solution of this system has the form

$$x(t) = \sum_{i=1}^{l}\sum_{j=0}^{n_i-1} t^j e^{\lambda_i t} v^{ij}$$

where $\lambda_1, \ldots, \lambda_l$ are the eigenvalues of $A$, the integer $n_i$ is the index of $\lambda_i$, and the constant vectors $v^{ij}$ depend on initial state. From this it follows that the stability properties of this system are completely determined by the location of its eigenvalues; this is summarized in the table below.

The following table summarizes the relationhip between the stability properties of a LTI system and the eigenproperties of its $A$-matrix. In the table, unless otherwise stated, a property involving $\lambda$ must hold for all eigenvalues $\lambda$ of $A$.

| Stability properties | Eigenproperties |
|---|---|
| Global exponential stability and boundedness | $\Re(\lambda) < 0$ |
| Stability and boundedness | $\Re(\lambda) \leq 0$<br>If $\Re(\lambda) = 0$ then $\lambda$ is non-defective. |
| Instability and some unbounded solutions | There is an eigenvalue $\lambda$ with $\Re(\lambda) > 0$ or $\Re(\lambda) = 0$ and $\lambda$ is defective. |

**Example 92** The system

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

is GES.

**Example 93** The system

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

is GES.

**Example 94** The system

$$\begin{aligned} \dot{x}_1 &= 0 \\ \dot{x}_2 &= 0 \end{aligned}$$

is stable about every equilibrium point.

**Example 95** The system

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= 0 \end{aligned}$$

is unstable about every equilibrium point.

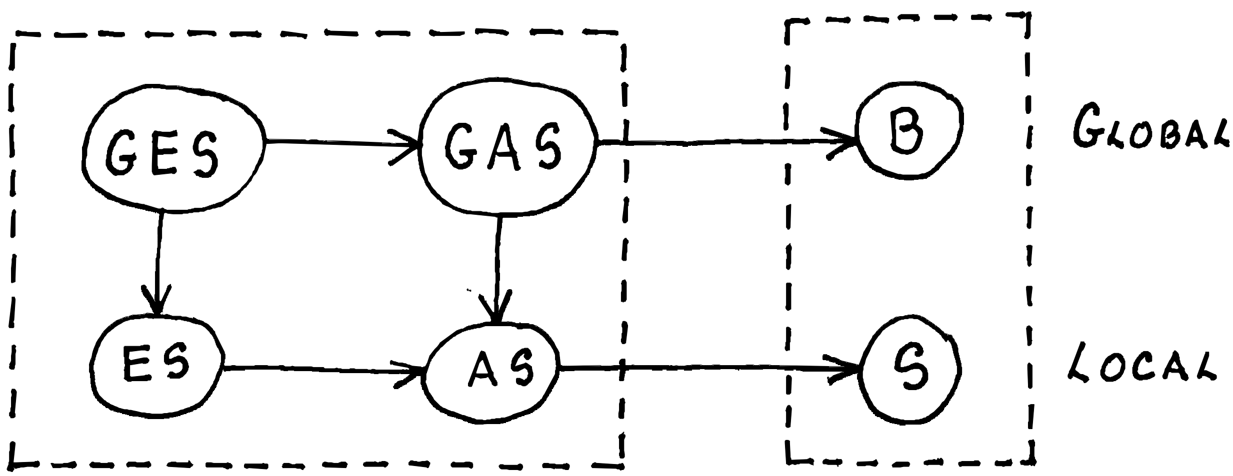



Figure 7.6: The big picture. For LTI systems, the concepts in each dashed box are equivalent

## 7.6 Linearization and stability

Consider a nonlinear time-invariant system described by

$$\dot{x} = f(x)$$

where $x(t)$ is an $n$-vector at each time $t$. Suppose $x^e$ is an equilibrium state for this system, that is, $f(x^e) = 0$, and consider the linearization of this system about $x^e$:

$$\delta\dot{x} = A\delta x \qquad \text{where} \quad A = \frac{\partial f}{\partial x}(x^e)\,.$$

The following results can be demonstrated using nonlinear Lyapunov stability theory.

**Stability.** *If all the eigenvalues of the $A$ matrix of the linearized system have negative real parts, then the nonlinear system is* exponentially stable *about $x^e$.*

**Instability.** *If at least one eigenvalue of the $A$ matrix of the linearized system has a positive real part, then the nonlinear system is* unstable *about $x^e$.*

**Undetermined.** *Suppose all the eigenvalues of the $A$ matrix of the linearized system have non-positive real parts and at least one eigenvalue of $A$ has zero real part. Then, based on the linearized system alone, one* cannot predict *the stability properties of the nonlinear system about $x^e$.*

Note that the first statement above is equivalent to the following statement. If the linearized system is exponentially stable, then the nonlinear system is exponentially stable about $x^e$.

**Example 96** (Damped simple pendulum.) Physically, the system

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

has two distinct equilibrium states: $(0, 0)$ and $(\pi, 0)$. The $A$ matrix for the linearization of this system about $(0, 0)$ is

$$A = \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}$$

Since all the eigenvalues of this matrix have negative real parts, the nonlinear system is exponentially stable about $(0, 0)$. The $A$ matrix corresponding to linearization about $(\pi, 0)$ is

$$A = \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix}$$

Since this matrix has an eigenvalue with a positive real part, the nonlinear system is unstable about $(\pi, 0)$

The following example illustrates the fact that if the eigenvalues of the $A$ matrix have non-positive real parts and there is at least one eigenvalue with zero real part, then, one cannot make any conclusions on the stability of the nonlinear system based on the linearization.

**Example 97** Consider the scalar nonlinear system:

$$\dot{x} = ax^3$$

This origin is GAS if $a < 0$, unstable if $a > 0$ and stable if $a = 0$. However, the linearization of this system about zero, given by

$$\delta\dot{x} = 0$$

is independent of $a$ and is stable.

The following example illustrates that instability of the linearized system does not imply instability of the nonlinear system.

**Example 98** Using nonlinear techniques, one can show that the following system is GAS about the origin.

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -x_1^3 - x_2^3 \end{aligned}$$

However, the linearization of this system, given by

$$\begin{aligned} \delta\dot{x}_1 &= \delta x_2 \\ \delta\dot{x}_2 &= 0\,, \end{aligned}$$

is unstable about the origin.

## Exercises

**Exercise 86** For each of the following systems, determine (from the state portrait) the stability properties of each equilibrium state. For AS equilibrium states, give the region of attraction.

(a)

$$\dot{x} = -x - x^3$$

(b)

$$\dot{x} = -x + x^3$$

(c)

$$\dot{x} = x - 2x^2 + x^3$$

**Exercise 87** If possible, use linearization to determine the stability properties of each of the following systems about the zero equilibrium state.

(i)

$$\begin{aligned} \dot{x}_1 &= (1 + x_1^2)x_2 \\ \dot{x}_2 &= -x_1^3 \end{aligned}$$

(ii)

$$\begin{aligned} \dot{x}_1 &= \sin x_2 \\ \dot{x}_2 &= (\cos x_1)x_3 \\ \dot{x}_3 &= e^{x_1}x_2 \end{aligned}$$

**Exercise 88** If possible, use linearization to determine the stability properties of each equilibrium state of the Lorenz system.

# Chapter 8

# Stability and boundedness: discrete time

## 8.1 Boundedness of solutions

Consider a general discrete-time nonlinear system described by

$$x(k+1) = f(x(k)) \tag{8.1}$$

where $x(k)$ is an $n$-vector and $k$ is an integer. By a solution of (8.1) we mean a sequence $x(\cdot) = (\, x(0), x(1), x(2), \, \ldots \,)$ which satisfies (8.1) for all $k \geq 0$.

**Defn.** (Boundedness of solutions) A solution $x(\cdot)$ is bounded if there exists $\beta \geq 0$ such that

$$||x(k)|| \leq \beta \qquad \text{for all} \qquad k \geq 0\,.$$

A solution is unbounded if it is not bounded.

It should be clear from the above definitions that a solution $x(\cdot)$ is bounded if and only if each component $x_i(\cdot)$ of $x(\cdot)$ is bounded. Also, a solution is unbounded if and only if at least one of its components is unbounded. So, $x(k) = [\ (0.5)^k \quad (-0.5)^k\ ]^T$ is bounded while $x(k) = [\ (0.5)^k \quad 2^k\ ]^T$ is unbounded.

**Example 99**

$$x(k+1) = 0$$

All solutions are bounded.

**Example 100**

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

All solutions are bounded.

**Example 101**

$$x(k+1) = -2x(k)$$

All solutions (except the zero solution) are unbounded.

**Linear time invariant (LTI) systems.** All solutions of the LTI system

$$x(k+1) = Ax(k) \tag{8.2}$$

are bounded if and only if for each eigenvalue $\lambda$ of $A$:

(b1) $|\lambda| \leq 1$ and

(b2) if $|\lambda| = 1$ then $\lambda$ is non-defective.

If there is an eigenvalue $\lambda$ of $A$ such that either

(u1) $|\lambda| > 1$ or

(u2) $|\lambda| = 1$ and $\lambda$ is defective.

then the system has some unbounded solutions.

**Example 102** Discrete unattached mass. Here

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

has a single eigenvalue 1. This eigenvalue has algebraic multiplicity 2 but geometric multiplicity 1; hence this eigenvalue is defective. So, some of the solutions of the system $x(k+1) = Ax(k)$ are unbounded. One example is

$$x(k) = \begin{pmatrix} k \\ 1 \end{pmatrix}$$

## 8.2 Stability of equilibrium states

Suppose $x^e$ is an equilibrium state of the system $x(k+1) = f(x(k))$. Then, whenever $x(0) = x^e$, we have $x(k) = x^e$ for all $k \geq 0$. Roughly speaking, we say that $x^e$ is a stable equilibrium state for the system if the following holds. If the initial state of the system is close to $x^e$, then the resulting solution is close to $x^e$. The formal definition is as follows.

**Defn.** (Stability) An equilibrium state $x^e$ is stable if for each $\epsilon > 0$ there exists $\delta > 0$ such that whenever $||x(0) - x^e|| < \delta$, one has $||x(k) - x^e|| < \epsilon$ for all $k \geq 0$.
$x^e$ is unstable if it is not stable.

**Example 103**

$$x(k+1) = -x(k)\,.$$

The origin is stable. (Consider $\delta = \epsilon$.)

**Example 104**

$$x(k+1) = -2x(k)\,.$$

The origin is unstable.

**Example 105**

$$x(k+1) = -x(k)^3\,.$$

The single equilibrium at the origin is stable, but, the system has unbounded solutions.

**LTI systems.** Every equilibrium state of a LTI system (8.2) is stable if and only if all eigenvalues $\lambda$ of $A$ satisfy conditions (b1) and (b2) above. Hence every equilibrium state of a LTI system is stable if and only if all solutions are bounded.

Every equilibrium state is unstable if and only if there is an eigenvalue $\lambda$ of $A$ which satisfies condition (u1) or (u2) above. Hence every equilibrium state of a LTI system is unstable if and only if there are unbounded solutions.

## 8.3 Asymptotic stability

### 8.3.1 Global asymptotic stability

**DEFN.** (Global asymptotic stability (GAS)) An equilibrium state $x^e$ is globally asymptotically stable (GAS) if

(a) It is stable

(b) Every solution $x(\cdot)$ converges to $x^e$, that is,

$$\lim_{k\to\infty} x(k) = x^e \,.$$

If $x^e$ is a globally asymptotically stable equilibrium state, then there are no other equilibrium states. In this case we say *the system (8.1) is globally asymptotically stable.*

**Example 106**

$$x(k+1) = \frac{1}{2}x(k)$$

GAS

**Example 107**

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

### 8.3.2 Asymptotic stability

**DEFN.** (Asymptotic stability (AS)) An equilibrium state $x^e$ is asymptotically stable (AS) if

(a) It is stable

(b) There exists $R > 0$ such that whenever $||x(0) - x^e|| < R$ one has

$$\lim_{k\to\infty} x(k) = x^e \tag{8.3}$$

The region of attraction of an equilibrium state $x^e$ which is AS is the set of initial states which result in (8.3), that is it is the set of initial states which are attracted to $x^e$. Thus, the region of attraction of a globally asymptotically equilibrium state is the whole state space.

**Example 108**

$$x(k+1) = x(k)^3$$

Origin is AS with region of attraction $[-1 \quad 1]$.

**LTI systems.** For LTI systems, it should be clear from the general form of the solution that the zero state is AS if and only if all the eigenvalues $\lambda_i$ of $A$ have magnitude less than one, that is,

$$|\lambda_i| < 1\,.$$

Also AS is equivalent to GAS.

## 8.4 Exponential stability

We now present the "strongest" form of stability considered in this section.

**DEFN.** (Global exponential stability) *An equilibrium state $x^e$ is* globally exponentially stable (GES) *with there exists $0 \leq \lambda < 1$ and $\beta > 0$ such that every solution satisfies*

$$||x(k)-x^e|| \quad \leq \quad \beta\lambda^k||x(0)-x^e|| \qquad \textit{for all} \quad k \geq 0$$

**Example 109**

$$x(k+1) = \frac{1}{2}x(k)$$

GES with $\gamma = \frac{1}{2}$.

Note that global exponential stability implies global asymptotic stability, but, in general, the converse is not true. This is illustrated in the next example. For linear time-invariant systems, GAS and GES are equivalent.

**Lemma 1** *Consider a system described by $x(k+1) = f(x(k))$ and suppose that for some scalar $\lambda \geq 0$,*

$$\|f(x)\| \leq \lambda\|x\|$$

*for all $x$. Then, every solution $x(\cdot)$ of the system satisfies*

$$\|x(k)\| \leq \lambda^k\|x(0)\|$$

*for all $k \geq 0$. In particular, if $\lambda < 1$ then, the system is globally exponentially stable.*

**Example 110**

$$x(k+1) = \frac{x(k)}{\sqrt{1+2x(k)^2}}$$

Solutions satisfy

$$x(k) = \frac{x_0}{\sqrt{1+2kx_0^2}} \qquad \text{where} \qquad x_0 = x(0)\,.$$

GAS but not GES

**Defn.** (Exponential stability) *An equilibrium state $x^e$ is* exponentially stable (ES) *if there exists $R > 0$, $0 \le \lambda < 1$ and $\beta > 0$ such that whenever $||x(0)-x^e|| < R$ one has*

$$||x(k)-x^e|| \quad \le \quad \beta\lambda^k ||x(0)-x^e|| \qquad \textit{for all} \quad k \ge 1$$

Note that exponential stability implies asymptotic stability, but, in general, the converse is not true.

**Example 111**

$$x(k+1) = \frac{x(k)}{\sqrt{1+2x(k)^2}}$$

Solutions satisfy

$$x(k) = \frac{x_0}{\sqrt{1+2kx_0^2}} \qquad \text{where} \qquad x_0 = x(0)\,.$$

GAS but not even ES

## 8.5 LTI systems

The following table summarizes the relationhip between the stability properties of a LTI system and the eigenproperties of its $A$-matrix. In the table, unless otherwise stated, a property involving $\lambda$ must hold for all eigenvalues $\lambda$ of $A$.

| Stability properties | eigenproperties |
|---|---|
| Asymptotic stability and boundedness | $\|\lambda\| < 1$ |
| Stability and boundedness | $\|\lambda\| \le 1$<br>If $\|\lambda\| = 1$ then $\lambda$ is non-defective |
| Instability and some unbounded solutions | There is an eigenvalue of $A$ with<br>$\|\lambda\| > 1$ or<br>$\|\lambda\| = 1$ and $\lambda$ is defective |

## 8.6 Linearization and stability

Consider a nonlinear time-invariant system described by

$$x(k+1) = f(x(k))$$

where $x(k)$ is an $n$-vector at each time $t$. Suppose $x^e$ is an equilibrium state for this system, that is, $f(x^e) = x^e$, and consider the linearization of this system about $x^e$:

$$\delta x(k+1) = A\delta x(k) \qquad \text{where} \quad A = \frac{\partial f}{\partial x}(x^e)\,.$$

The following results can be demonstrated using nonlinear Lyapunov stability theory.

**Stability.** *If all the eigenvalues of the $A$ matrix of the linearized system have magnitude less than one, then the nonlinear system is* exponentially stable *about $x^e$.*

**Instability.** *If at least one eigenvalue of the $A$ matrix of the linearized system has magnitude greater than one, then the nonlinear system is* unstable *about $x^e$.*

**Undetermined.** *Suppose all the eigenvalues of the $A$ matrix of the linearized system have magnitude less than or equal to one and at least one eigenvalue of $A$ has magnitude one. Then, based on the linearized system alone, one* cannot predict *the stability properties of the nonlinear system about $x^e$.*

Note that the first statement above is equivalent to the following statement. If the linearized system is exponentially stable, then the nonlinear system is exponentially stable about $x^e$.

**Example 112 (Newton's method)** Recall that Newton's method for a scalar function can be described by

$$x(k+1) = x(k) - \frac{g(x(k))}{g'(x(k))}$$

Here

$$f(x) = x - \frac{g(x)}{g'(x)}$$

So,

$$f'(x) = 1 - 1 + \frac{g(x)g''(x)}{g'(x)^2} = \frac{g(x)g''(x)}{g'(x)^2}\,.$$

At an equilibrium state $x^e$, we have $g(x^e) = 0$; hence $f'(x^e) = 0$ and the linearization about any equilibrium state is given by

$$\delta x(k+1) = 0\,.$$

Thus, every equilibrium state is exponentially stable.

# Chapter 9

# Basic Lyapunov theory

Suppose we are interested in the stability properties of the system

$$\dot{x} = f(x) \tag{9.1}$$

where $x(t)$ is a real $n$-vector at time $t$. If the system is linear, we can determine its stability properties from the properties of the eigenvalues of the system matrix. What do we do for a nonlinear system? We could linearize about each equilibrium state and determine the stability properties of the resulting linearizations. Under certain conditions this will tell us something about the local stability properties of the nonlinear system about its equilibrium states. However there are situations where linearization cannot be used to deduce even the local stability properties of the nonlinear system. Also, linearization tells us nothing about the global stability properties of the nonlinear system.

In general, we cannot explicitly obtain solutions for nonlinear systems. Lyapunov theory allows to say something about the stability properties of a system without knowing the form or structure of the solutions.

In this chapter, $V$ is a scalar-valued function of the state, that is $V : \mathbb{R}^n \to \mathbb{R}$. Suppose $V$ is continuously differentiable. Then, at any time $t$, the derivative of $V$ along a solution $x(\cdot)$ of system (9.1) is given by

$$\begin{aligned} \frac{dV}{dt}(x(t)) &= DV(x(t))\dot{x}(t) \\ &= DV(x(t))f(x(t)) \end{aligned}$$

where $DV(x)$ is the derivative of $V$ at $x$ and is given by

$$DV(x) = \begin{pmatrix} \frac{\partial V}{\partial x_1}(x) & \frac{\partial V}{\partial x_2}(x) & \ldots & \frac{\partial V}{\partial x_n}(x) \end{pmatrix}$$

Note that

$$DV(x)f(x) = \frac{\partial V}{\partial x_1}(x)f_1(x) + \frac{\partial V}{\partial x_2}(x)f_2(x) + \ldots + \frac{\partial V}{\partial x_n}(x)f_n(x)$$

In what follows, if a condition involves $DV$, then it is implicitly assumed that $V$ is continuously differentiable. Sometimes $DV$ is denoted by

$$\frac{\partial V}{\partial x}\,.$$

Also, when the system under consideration is fixed, $DVf$ is sometimes denoted by

$$\dot{V}\ .$$

This a slight but convenient abuse of notation.

## 9.1 Stability

### 9.1.1 Locally positive definite functions

**Defn.** (Locally positive definite function) A function $V$ is locally positive definite (lpd) about a point $x^e$ if

$$V(x^e) = 0$$

and there is a scalar $R > 0$ such that

$$V(x) > 0 \qquad \text{whenever} \quad x \neq x^e \quad \text{and} \quad ||x - x^e|| < R$$

Basically, a function is lpd about a point $x^e$ if it is zero at $x^e$ has a strict local minimum at $x^e$.

Figure 9.1: A locally positive definite function

**Example 113** (Scalar $x$) The following functions are lpd about zero.

$$\begin{aligned}
V(x) &= x^2 \\
V(x) &= 1 - e^{-x^2} \\
V(x) &= 1 - \cos x \\
V(x) &= x^2 - x^4
\end{aligned}$$

**Example 114**

$$\begin{aligned}
V(x) &= ||x||^2 \\
&= x_1^2 + x_2^2 + \ldots + x_n^2
\end{aligned}$$

Lpd about the origin.

**Quadratic forms.** Suppose $P$ is a real $n \times n$ symmetric matrix and is positive definite. Consider the quadratic form defined by

$$V(x) = x^T P x = \sum_{i=1}^{n} \sum_{j=1}^{n} P_{ij} x_i x_j \,.$$

Clearly $V(0) = 0$. Recalling the definition of a positive definite matrix, it follows that $V(x) = x^T Px > 0$ for all nonzero $x$. Hence $V$ is locally positive definite about the origin.

• The **second derivative of $V$ at $x$** is the square symmetric matrix given by:

$$D^2V(x) := \begin{pmatrix} \frac{\partial^2 V}{\partial^2 x_1}(x) & \frac{\partial^2 V}{\partial x_1 \partial x_2}(x) & \dots & \frac{\partial^2 V}{\partial x_1 \partial x_n}(x) \\ \frac{\partial^2 V}{\partial x_2 \partial x_1}(x) & \frac{\partial^2 V}{\partial^2 x_2}(x) & \dots & \frac{\partial^2 V}{\partial x_2 \partial x_n}(x) \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 V}{\partial x_n \partial x_1}(x) & \frac{\partial^2 V}{\partial x_n \partial x_2}(x) & \dots & \frac{\partial^2 V}{\partial^2 x_n}(x) \end{pmatrix}$$

that is,

$$D^2V(x)_{ij} = \frac{\partial^2 V}{\partial x_i \partial x_j}(x)$$

Sometimes $D^2V(x)$ is written as $\frac{\partial^2 V}{\partial x^2}(x)$ and is referred to as the **Hessian** of $V$.

The following lemma is sometimes useful in demonstrating that a function is lpd.

**Lemma 2** *Suppose $V$ is twice continuously differentiable and*

$$\begin{aligned} V(x^e) &= 0 \\ DV(x^e) &= 0 \\ D^2V(x^e) &> 0 \end{aligned}$$

*Then $V$ is a locally positive definite about $x^e$.*

**Example 115** Consider $V(x) = x^2 - x^4$ where $x$ is a scalar. Since $V(0) = DV(0) = 0$ and $D^2V(0) = 2 > 0$, $V$ is lpd about zero.

**Example 116** Consider

$$V(x) = 1 - \cos x_1 + \frac{1}{2}x_2^2$$

Clearly

$$V(0) = 0$$

Since

$$DV(x) = \begin{pmatrix} \sin x_1 & x_2 \end{pmatrix}$$

we have

$$DV(0) = 0$$

Also,

$$D^2V(x) = \begin{pmatrix} \cos x_1 & 0 \\ 0 & 1 \end{pmatrix}$$

Hence,

$$D^2V(0) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} > 0$$

Since $V$ satisfies the hypotheses of the previous lemma with $x^e = 0$, it is lpd about zero.

### 9.1.2 A stability result

If the equilibrium state of a nonlinear system is stable but not asymptotically stable, then one cannot deduce the stability properties of the equilibrium state of the nonlinear system from the linearization of the nonlinear system about that equilibrium state

**Theorem 4** (Stability) *Suppose there exists a function $V$ and a scalar $R > 0$ such that $V$ is locally positive definite about $x^e$ and*

$$DV(x)f(x) \leq 0 \qquad \textit{for} \quad ||x - x^e|| < R$$

*Then $x^e$ is a stable equilibrium state.*

If $V$ satisfies the hypotheses of the above theorem, then $V$ is said to be a Lyapunov function which guarantees the stability of $x^e$.

**Example 117**

$$\dot{x} = 0$$

Consider

$$V(x) = x^2$$

as a candidate Lyapunov function. Then $V$ is lpd about 0 and

$$DV(x)f(x) = 0$$

Hence (it follows from Theorem 4 that) the origin is stable.

**Example 118** (Undamped linear oscillator.)

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -kx_1 \qquad\qquad k > 0 \end{aligned}$$

Consider the total *energy*,

$$V(x) = \frac{1}{2}kx_1^2 + \frac{1}{2}x_2^2$$

as a candidate Lyapunov function. Then $V$ is lpd about the origin and

$$DV(x)f(x) = 0$$

Hence the origin is stable.

**Example 119** (Simple pendulum.)

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

Consider the total energy,

$$V(x) = 1 - \cos x_1 + \frac{1}{2}x_2^2$$

as a candidate Lyapunov function. Then $V$ is lpd about the origin and

$$DV(x)f(x) = 0$$

Hence the origin is stable

**Example 120** (Stability of origin for attitude dynamics system.) Recall

$$\begin{aligned}\dot{x}_1 &= \frac{(I_2 - I_3)}{I_1}x_2x_3 \\ \dot{x}_2 &= \frac{(I_3 - I_1)}{I_2}x_3x_1 \\ \dot{x}_3 &= \frac{(I_1 - I_2)}{I_3}x_1x_2\end{aligned}$$

where

$$I_1, I_2, I_3 > 0$$

Consider the kinetic energy

$$V(x) = \frac{1}{2}\left(I_1x_1^2 + I_2x_2^2 + I_3x_3^2\right)$$

as a candidate Lyapunov function. Then $V$ is lpd about the origin and

$$DV(x)f(x) = 0$$

Hence the origin is stable.

**Example 121** (Undamped Duffing system)

$$\begin{aligned}\dot{x}_1 &= x_2 \\ \dot{x}_2 &= x_1 - x_1^3\end{aligned}$$

As a candidate Lyapunov function for the equilibrium state $x^e = [\ 1 \quad 0\ ]^T$ consider the total energy

$$V(x) = \frac{1}{4}x_1^4 - \frac{1}{2}x_1^2 + \frac{1}{2}x_2^2 + \frac{1}{4}\,.$$

Since

$$DV(x) = \begin{pmatrix} x_1^3 - x_1 & x_2 \end{pmatrix} \quad \text{and} \quad D^2V(x) = \begin{pmatrix} 3x_1^2 - 1 & 0 \\ 0 & 1 \end{pmatrix}$$

we have $V(x^e) = 0$, $DV(x^e) = 0$ and $D^2V(x^e) > 0$, and it follows that $V$ is lpd about $x^e$. One can readily verify that $DV(x)f(x) = 0$. Hence, $x^e$ is stable.

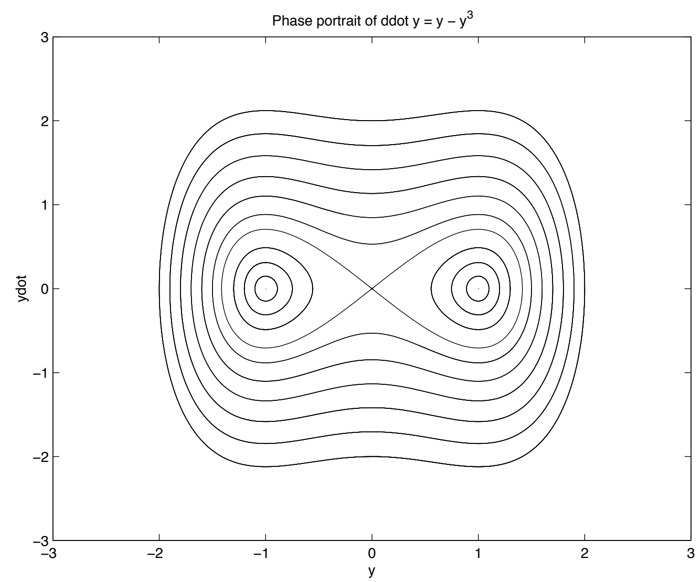



Figure 9.2: Duffing system

## Exercises

**Exercise 89** Determine whether or not the following functions are lpd. (a)

$$V(x) = x_1^2 - x_1^4 + x_2^2$$

(b)

$$V(x) = x_1 + x_2^2$$

(c)

$$V(x) = 2x_1^2 - x_1^3 + x_1x_2 + x_2^2$$

**Exercise 90** (Simple pendulum with Coulomb damping.) By appropriate choice of Lyapunov function, show that the origin is a stable equilibrium state for

$$\begin{aligned}\dot{x}_1 &= x_2 \\ \dot{x}_2 &= -\sin x_1 - c\,\mathrm{sgm}\,(x_2)\end{aligned}$$

where $c > 0$.

**Exercise 91** By appropriate choice of Lyapunov function, show that the origin is a stable equilibrium state for

$$\begin{aligned}\dot{x}_1 &= x_2 \\ \dot{x}_2 &= -x_1^3\end{aligned}$$

Note that the linearization of this system about the origin is unstable.

**Exercise 92** By appropriate choice of Lyapunov function, show that the origin is a stable equilibrium state for

$$\begin{aligned}\dot{x}_1 &= x_2 \\ \dot{x}_2 &= -x_1 + x_1^3\end{aligned}$$

## 9.2 Asymptotic stability

The following result presents conditions which guarantee that an equilibrium state is asymptotically stable.

**Theorem 5** (Asymptotic stability) *Suppose there exists a function $V$ and a scalar $R > 0$ such that $V$ is locally positive definite about $x^e$ and*

$$DV(x)f(x) < 0 \qquad \textit{for} \quad x \neq x^e \quad \textit{and} \quad ||x - x^e|| < R$$

*Then $x^e$ is an asymptotically stable equilibrium state for $\dot{x} = f(x)$.*

**Example 122**

$$\dot{x} = -x^3$$

Consider

$$V(x) = x^2$$

Then $V$ is lpd about zero and

$$DV(x)f(x) = -2x^4 < 0 \qquad \text{for} \quad x \neq 0$$

Hence the origin is AS.

**Example 123**

$$\dot{x} = -x + x^3$$

Consider

$$V(x) = x^2$$

Then $V$ is lpd about zero and

$$DV(x)f(x) = -2x^2(1 - x^2) < 0 \qquad \text{for} \quad |x| < 1,\ x \neq 0$$

Hence the origin is AS. Although the origin is AS, there are solutions which go unbounded in a finite time.

**Example 124**

$$\dot{x} = -\sin x$$

Consider

$$V(x) = x^2$$

Then $V$ is lpd about zero and

$$DV(x)f(x) = -2x\sin(x) < 0 \qquad \text{for} \quad |x| < \pi,\ x \neq 0$$

Hence the origin is AS.

**Example 125** *(Simple pendulum with viscous damping.)* Intuitively, we expect the origin to be an asymptotically stable equilibrium state for the damped simple pendulum:

$$\begin{aligned}\dot{x}_1 &= x_2 \\ \dot{x}_2 &= -\sin x_1 - cx_2\end{aligned}$$

where $c > 0$ is the damping coefficient. If we consider the total mechanical energy

$$V(x) = 1 - \cos x_1 + \frac{1}{2}x_2^2$$

as a candidate Lyapunov function, we obtain

$$DV(x)f(x) = -cx_2^2\,.$$

Since $DV(x)f(x) \leq 0$ for all $x$, we have stability of the origin. Since $DV(x)f(x) = 0$ whenever $x_2 = 0$, it follows that $DV(x)f(x) = 0$ for points arbitrarily close to the origin; hence $V$ does not satisfy the requirements of the above theorem for asymptotic stability.

Suppose we modify $V$ to

$$V(x) = \frac{1}{2}\lambda c^2 x_1^2 + \lambda c x_1 x_2 + \frac{1}{2}x_2^2 + 1 - \cos x_1$$

where $\lambda$ is any scalar with $0 < \lambda < 1$. Letting

$$P = \frac{1}{2}\begin{pmatrix} \lambda c^2 & \lambda c \\ \lambda c & 1 \end{pmatrix}$$

note that $P > 0$ and

$$\begin{aligned}V(x) &= x^T P x + 1 - \cos x_1 \\ &\geq x^T P x\end{aligned}$$

Hence $V$ is lpd about zero and we obtain

$$DV(x)f(x) = -\lambda c x_1 \sin x_1 - (1-\lambda)cx_2^2 < 0 \qquad \text{for} \quad ||x|| < \pi,\ x \neq 0$$

to satisfy the requirements of above theorem; hence the origin is AS.

## Exercises

**Exercise 93** By appropriate choice of Lyapunov function, show that the origin is an asymptotically stable equilibrium state for

$$\begin{aligned}\dot{x}_1 &= x_2 \\ \dot{x}_2 &= -x_1^5 - x_2\end{aligned}$$

**Exercise 94** By appropriate choice of Lyapunov function, show that the origin is a asymptotically stable equilibrium state for

$$\begin{aligned}\dot{x}_1 &= x_2 \\ \dot{x}_2 &= -x_1 + x_1^3 - x_2\end{aligned}$$

## 9.3 Boundedness

### 9.3.1 Radially unbounded functions

**Defn.** *A scalar valued function $V$ of the state $x$ is said to be* radially unbounded *if*

$$\boxed{\lim_{x\to\infty} V(x) = \infty}$$

**Example 126**

$$\begin{array}{rcll} V(x) &=& x^2 & \text{yes} \\ V(x) &=& 1-e^{-x^2} & \text{no} \\ V(x) &=& x^2-x & \text{yes} \\ V(x) &=& x^4-x^2 & \text{yes} \\ V(x) &=& x\sin x & \text{no} \\ V(x) &=& x^2(1-\cos x) & \text{no} \end{array}$$

**Example 127** Suppose $P$ is a real $n\times n$ matrix and is positive definite symmetric and consider the quadratic form defined by

$$V(x) = x^T P x$$

Since $P$ is real symmetric,

$$x^T P x \geq \lambda_{\min}(P)||x||^2$$

for all $x$, where $\lambda_{\min}(P)$ is the minimum eigenvalue of $P$. Since $P$ is positive definite, $\lambda_{\min}(P) > 0$. From this it should be clear that $V$ is radially unbounded.

The following lemma can be useful for guaranteeing radial unboundedness.

**Lemma 3** *Suppose $V$ is twice continuously differentiable, and there is a positive definite symmetric matrix $P$ and a scalar $R \geq 0$ such that*

$$D^2V(x) \geq P \qquad \textit{for} \quad ||x|| \geq R$$

*Then, $V$ is radially unbounded.*

### 9.3.2 A boundedness result

**Theorem 6** *Suppose there exists a radially unbounded function $V$ and a scalar $R \geq 0$ such that*

$$DV(x)f(x) \leq 0 \qquad \textit{for} \quad ||x|| \geq R$$

*Then all solutions of $\dot{x} = f(x)$ are bounded.*

Note that, in the above theorem, $V$ does not have to be positive away from the origin; it only has to be radially unbounded.

**Example 128** Recall

$$\dot{x} = x - x^3 \,.$$

Consider

$$V(x) = x^2$$

Since $V$ is radially unbounded and

$$DV(x)f(x) = -2x^2(x^2 - 1)$$

the hypotheses of the above theorem are satisfied with $R = 1$; hence all solutions are bounded. Note that the origin is unstable.

**Example 129** Duffing's equation

$$\begin{aligned}
\dot{x}_1 &= x_2 \\
\dot{x}_2 &= x_1 - x_1^3
\end{aligned}$$

Consider

$$V(x) = \frac{1}{2}x_2^2 - \frac{1}{2}x_1^2 + \frac{1}{4}x_1^4$$

It should be clear that $V$ is radially unbounded; also

$$DV(x)f(x) = 0 \leq 0 \qquad \text{for all} \quad x$$

So, the hypotheses of the above theorem are satisfied with any $R$; hence all solutions are bounded.

## Exercises

**Exercise 95** Determine whether or not the the following function is radially unbounded.

$$V(x) = x_1 - x_1^3 + x_1^4 - x_2^2 + x_2^4$$

**Exercise 96** (Forced Duffing's equation with damping.) Show that all solutions of the system

$$\begin{aligned}
\dot{x}_1 &= x_2 \\
\dot{x}_2 &= x_1 - x_1^3 - cx_2 + 1 \qquad c > 0
\end{aligned}$$

are bounded.
Hint: Consider

$$V(x) = \frac{1}{2}\lambda c^2 x_1^2 + \lambda c x_1 x_2 + \frac{1}{2}x_2^2 - \frac{1}{2}x_1^2 + \frac{1}{4}x_1^4$$

where $0 < \lambda < 1$. Letting

$$P = \frac{1}{2}\begin{pmatrix} \lambda c^2 & \lambda c \\ \lambda c & 1 \end{pmatrix}$$

note that $P > 0$ and

$$\begin{aligned}
V(x) &= x^T P x - \frac{1}{2}x_1^2 + \frac{1}{4}x_1^4 \\
&\geq x^T P x - \frac{1}{4}
\end{aligned}$$

**Exercise 97** Recall the Lorenz system

$$\begin{aligned}\dot{x}_1 &= \sigma(x_2 - x_1)\\ \dot{x}_2 &= rx_1 - x_2 - x_1x_3\\ \dot{x}_3 &= -bx_3 + x_1x_2\end{aligned}$$

with $b > 0$. Prove that all solutions of this system are bounded. (Hint: Consider $V(x) = rx_1^2 + \sigma x_2^2 + \sigma(x_3 - 2r)^2$.)

## 9.4 Global asymptotic stability

### 9.4.1 Positive definite functions

**Defn.** (Positive definite function.) *A function* $V$ *is* positive definite (pd) *if*

$$\begin{aligned} V(0) &= 0 \\ V(x) &> 0 \qquad \textit{for all} \quad x \neq 0 \\ \lim_{x\to\infty} V(x) &= \infty \end{aligned}$$

In the above definition, note the requirement that $V$ be *radially unbounded.*

**Example 130** Scalar $x$

$$\begin{aligned} V(x) &= x^2 && \text{pd} \\ V(x) &= 1 - e^{-x^2} && \text{lpd but not pd} \\ V(x) &= 1 - \cos x && \text{lpd but not pd} \end{aligned}$$

**Example 131**

$$V(x) = x_1^4 + x_2^2$$

**Example 132**

$$\begin{aligned} V(x) &= ||x||^2 \\ &= x_1^2 + x_2^2 + \ldots + x_n^2 \end{aligned}$$

**Example 133** Suppose $P$ is a real $n \times n$ matrix and is positive definite symmetric and consider the quadratic form defined by

$$V(x) = x^T P x$$

$V$ is a positive definite function.

**Lemma 4** *Suppose* $V$ *is twice continuously differentiable and*

$$\begin{aligned} V(0) &= 0 \\ DV(0) &= 0 \\ D^2V(x) &> 0 \qquad \textit{for all} \quad x \end{aligned}$$

*Then* $V(x) > 0$ *for all* $x$.

If $V$ satisfies the hypotheses of the above lemma, it is not guaranteed to be radially unbounded, hence it is not guaranteed to be positive definite. Lemma 3 can be useful for guaranteeing radial unboundedness. We also have the following lemma.

**Lemma 5** *Suppose $V$ is twice continuously differentiable,*

$$\begin{aligned} V(0) &= 0 \\ DV(0) &= 0 \end{aligned}$$

*and there is a positive definite symmetric matrix $P$ such that*

$$D^2V(x) \geq P \qquad \textit{for all} \quad x$$

*Then*

$$V(x) \geq \frac{1}{2}x^T P x$$

*for all $x$.*

### 9.4.2 A result on global asymptotic stability

**Theorem 7** (Global asymptotic stability) *Suppose there exists a positive definite function $V$ such that*

$$DV(x)f(x) < 0 \qquad \textit{for all} \quad x \neq 0$$

*Then the origin is a globally asymptotically stable equilibrium state for $\dot{x} = f(x)$.*

**Example 134**

$$\dot{x} = -x^3$$

$$V(x) = x^2$$

$$\begin{aligned} DV(x)f(x) &= -2x^4 \\ &< 0 \qquad \text{for all} \quad x \neq 0 \end{aligned}$$

We have GAS. Note that linearization of this system about the origin cannot be used to deduce the asymptotic stability of this system.

**Example 135** The first nonlinear system

$$\dot{x} = -\mathrm{sgm}\,(x)\,.$$

This system is not linearizable about its unique equilibrium state at the origin. Considering

$$V(x) = x^2$$

we obtain

$$\begin{aligned} DV(x)f(x) &= -2|x| \\ &< 0 \qquad \text{for all} \quad x \neq 0\,. \end{aligned}$$

Hence, we have GAS.

**Example 136** *(Linear globally stabilizing controller for inverted pendulum.)*

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

Consider

$$u = -k_1 x_1 - k_2 x_2 \qquad \text{with} \quad k_1 > 1,\ k_2 > 0$$

Closed loop system:

$$\begin{aligned}\dot{x}_1 &= x_2\\ \dot{x}_2 &= -k_1 x_1 + \sin x_1 - k_2 x_2\end{aligned}$$

Consider

$$V(x) = \frac{1}{2}\lambda k_2^2 x_1^2 + \lambda k_2 x_1 x_2 + \frac{1}{2}x_2^2 + \frac{1}{2}k_1 x_1^2 + \cos x_1 - 1$$

where $0 < \lambda < 1$. Then $V$ is pd (apply lemma 5) and

$$DV(x)f(x) = \lambda k_2(-k_1 x_1^2 + x_1 \sin x_1) - (1-\lambda)k_2 x_2^2$$

Since

$$|\sin x_1| \leq |x_1| \qquad \text{for all} \quad x_1$$

it follows that

$$x_1 \sin x_1 \leq x_1^2 \qquad \text{for all} \quad x_1$$

hence

$$DV(x)f(x) \leq -\lambda k_2(k_1-1)x_1^2 - (1-\lambda)k_2 x_2^2 \quad < 0 \qquad \text{for all} \quad x \neq 0$$

The closed loop system is GAS.

**Example 137** (Stabilization of origin for attitude dynamics system.)

$$\begin{aligned}\dot{x}_1 &= \frac{(I_2 - I_3)}{I_1}x_2 x_3 &+& \frac{u_1}{I_1}\\ \dot{x}_2 &= \frac{(I_3 - I_1)}{I_2}x_3 x_1 &+& \frac{u_2}{I_2}\\ \dot{x}_3 &= \frac{(I_1 - I_2)}{I_3}x_1 x_2 &+& \frac{u_3}{I_3}\end{aligned}$$

where

$$I_1, I_2, I_3 > 0$$

Consider any linear controller of the form

$$u_i = -k x_i \quad i = 1, 2, 3, \qquad k > 0$$

Closed loop system:

$$\begin{aligned}
\dot{x}_1 &= \left(\frac{I_2 - I_3}{I_1}\right) x_2 x_3 &-& \frac{kx_1}{I_1} \\
\dot{x}_2 &= \left(\frac{I_3 - I_1}{I_2}\right) x_3 x_1 &-& \frac{kx_2}{I_2} \\
\dot{x}_3 &= \left(\frac{I_1 - I_2}{I_3}\right) x_1 x_2 &-& \frac{kx_3}{I_3}
\end{aligned}$$

Consider the kinetic energy

$$V(x) = \frac{1}{2}\left(I_1 x_1^2 + I_2 x_2^2 + I_3 x_3^2\right)$$

as a candidate Lyapunov function. Then $V$ is pd and

$$\begin{aligned}
DV(x)f(x) &= -k(x_1^2 + x_2^2 + x_3^2) \\
&< 0 \qquad\qquad \text{for all} \quad x \neq 0
\end{aligned}$$

Hence the origin is GAS for the closed loop system.

## Exercises

**Exercise 98** Determine whether or not the following function is positive definite.

$$V(x) = x_1^4 - x_1^2 x_2 + x_2^2$$

**Exercise 99** Consider any scalar system described by

$$\dot{x} = -g(x)$$

where $g$ has the following properties:

$$\begin{aligned}
g(x) > 0 \quad &\text{for} \quad x > 0 \\
g(x) < 0 \quad &\text{for} \quad x < 0
\end{aligned}$$

Show that this system is GAS.

**Exercise 100** Recall the Lorenz system

$$\begin{aligned}
\dot{x}_1 &= \sigma(x_2 - x_1) \\
\dot{x}_2 &= rx_1 - x_2 - x_1 x_3 \\
\dot{x}_3 &= -bx_3 + x_1 x_2
\end{aligned}$$

Prove that if

$$b > 0 \qquad \text{and} \qquad 0 \leq r < 1\,,$$

then this system is GAS about the origin. (Hint: Consider $V(x) = x_1^2 + \sigma x_2^2 + \sigma x_3^2$.)

**Exercise 101** *(Stabilization of the Duffing system.)* Consider the Duffing system with a scalar control input $u(t)$:

$$\begin{aligned}\dot{x}_1 &= x_2 \\ \dot{x}_2 &= x_1 - x_1^3 + u\end{aligned}$$

Obtain a linear controller of the form

$$u = -k_1 x_1 - k_2 x_2$$

which results in a closed loop system which is GAS about the origin. Numerically simulate the open loop system ($u = 0$) and the closed loop system for several initial conditions.

## 9.5 Exponential stability

### 9.5.1 Global exponential stability

**Theorem 8** (Global exponential stability) *Suppose there exists a function $V$ and positive scalars $\alpha, \beta_1, \beta_2$ such that for all $x$,*

$$\beta_1 ||x||^2 \leq V(x) \leq \beta_2 ||x||^2$$

*and*

$$DV(x)f(x) \leq -2\alpha V(x)\,.$$

*Then, for the system $\dot{x} = f(x)$, the origin is a globally exponentially stable equilibrium state with rate of convergence $\alpha$. In particular, all solutions $x(\cdot)$ of the system satisfy*

$$\|x(t)\| \leq \sqrt{\beta_2/\beta_1}\,\|x(0)\|\,e^{-\alpha t} \qquad \textit{for all} \qquad t \geq 0\,.$$

Proof. See below.

**Example 138**

$$\dot{x} = -x$$

Considering

$$V(x) = x^2$$

we have

$$\begin{aligned}DV(x)f(x) &= -2x^2 \\ &= -2V(x)\end{aligned}$$

Hence, we have GES with rate of convergence 1.

**Example 139**

$$\dot{x} = -x - x^3$$

Considering

$$V(x) = x^2$$

we have

$$\begin{aligned} DV(x)f(x) &= -2x^2 - 2x^4 \\ &\leq -2V(x) \end{aligned}$$

Hence, we have GES with rate of convergence 1.

### 9.5.2 Proof of theorem 8

Consider any solution $x(\cdot)$ and let $v(t) = V(x(t))$. Then

$$\dot{v} \leq -2\alpha v\,.$$

### 9.5.3 Exponential stability

**Theorem 9** (Exponential stability) *Suppose there exists a function $V$ and positive scalars $R$, $\alpha, \beta_1, \beta_2$ such that, whenever $\|x\| \leq R$, one has*

$$\beta_1 ||x||^2 \leq V(x) \leq \beta_2 ||x||^2$$

*and*

$$DV(x)f(x) \leq -2\alpha V(x)$$

*Then, for system $\dot{x} = f(x)$, the state $x^e$ is an exponentially stable equilibrium state with rate of convergence $\alpha$.*

**Exercise 102** Consider the scalar system

$$\dot{x} = -x + x^3$$

As a candidate Lyapunov function for exponential stability, consider $V(x) = x^2$. Clearly, the condition on $V$ is satisfied with $\beta_1 = \beta_2 = 1$. Noting that

$$\dot{V} = -2x^2 + 2x^4 = -2(1 - x^2)x^2\,,$$

and considering $R = 1/2$, we obtain that whenever $|x| \leq R$, we have $\dot{V} \leq -2\alpha V$ where $\alpha = 3/4$. Hence, we have ES.

### 9.5.4 A special class of GES systems

Consider a system described by

$$\dot{x} = f(x) \tag{9.2}$$

and suppose that there exist two positive definite symmetric matrices $P$ and $Q$ such that

$$\boxed{x^T P f(x) \leq -x^T Q x\,.}$$

We will show that the origin is GES with rate

$$\boxed{\alpha := \lambda_{\min}(P^{-1}Q)}$$

where $\lambda_{\min}(P^{-1}Q) > 0$ is the smallest eigenvalue of $P^{-1}Q$.

As a candidate Lyapunov function, consider

$$V(x) = x^T P x\,.$$

Then

$$\lambda_{\min}(P)||x||^2 \leq V(x) \leq \lambda_{\max}(P)||x||^2$$

that is

$$\beta_1||x||^2 \leq V(x) \leq \beta_2||x||^2$$

with

$$\beta_1 = \lambda_{\min}(P) > 0 \qquad \text{and} \qquad \beta_2 = \lambda_{\max}(P) > 0\,.$$

Now note that

$$\begin{aligned} DV(x)f(x) &= 2x^T P f(x) \\ &\leq -2x^T Q x \end{aligned}$$

For any two positive-definite matrices $P$ and $Q$, one can show that all the eigenvalues of $P^{-1}Q$ are real positive and

$$x^T Q x \geq \lambda_{\min}(P^{-1}Q)\, x^T P x$$

where $\lambda_{\min}(P^{-1}Q) > 0$ is the smallest eigenvalue of $P^{-1}Q$. Thus,

$$\begin{aligned} DV(x)f(x) &\leq -2\lambda_{\min}(P^{-1}Q)x^T P x \\ &= -2\alpha V(x) \end{aligned}$$

Hence GES with rate $\alpha$.

**Example 140** Recall example 137.
Closed loop system:

$$\begin{aligned} \dot{x}_1 &= \left(\frac{I_2 - I_3}{I_1}\right) x_2 x_3 &- \frac{k x_1}{I_1} \\ \dot{x}_2 &= \left(\frac{I_3 - I_1}{I_2}\right) x_3 x_1 &- \frac{k x_2}{I_2} \\ \dot{x}_3 &= \left(\frac{I_1 - I_2}{I_3}\right) x_1 x_2 &- \frac{k x_3}{I_3} \end{aligned}$$

Considering

$$P = \begin{pmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_3 \end{pmatrix}$$

we have

$$x^T P f(x) = -k x^T x$$

that is, $Q = kI$. Hence, GES with rate

$$\begin{aligned} \alpha &= \lambda_{\min}(P^{-1}Q) \\ &= \lambda_{\min}(kP^{-1}) \\ &= k\lambda_{\min}(P^{-1}) \\ &= \frac{k}{\lambda_{\max}(P)} \\ &= \frac{k}{\max\{I_1, I_2, I_3\}} \end{aligned}$$

### 9.5.5 Summary

The following table summarizes the results of this chapter for stability about the origin.

| $V$ | $\dot{V}$ | $\Rightarrow$ stability properties |
|---|---|---|
| lpd | $\leq 0$ for $\|\|x\|\| \leq R$ | $\Rightarrow$ S |
| lpd | $< 0$ for $\|\|x\|\| \leq R$, $x \neq 0$ | $\Rightarrow$ AS |
| ru | $\leq 0$ for $\|\|x\|\| \geq R$ | $\Rightarrow$ B |
| pd | $< 0$ for $x \neq 0$ | $\Rightarrow$ GAS |
| $\beta_1\|\|x\|\|^2 \leq V(x) \leq \beta_2\|\|x\|\|^2$ | $\leq -2\alpha V(x)$ | $\Rightarrow$ GES |
| $\beta_1\|\|x\|\|^2 \leq V(x) \leq \beta_2\|\|x\|\|^2$ for $\|\|x\|\| \leq R$ | $\leq -2\alpha V(x)$ for $\|\|x\|\| \leq R$ | $\Rightarrow$ ES |

Figure 9.3: Lyapunov table

# Chapter 10

# Basic Lyapunov theory: discrete time*

Suppose we are interested in the stability properties of the system,

$$x(k+1) = f(x(k)) \tag{10.1}$$

where $x(k) \in \mathbb{R}^n$ and $k \in \mathbb{R}$. If the system is linear, we can determine its stability properties from the properties of the eigenvalues of the system matrix. What do we for a nonlinear system? We could linearize about each equilibrium state and determine the stability properties of the resulting linearizations. Under certain conditions (see later) this will tell us something about the local stability properties of the nonlinear system about its equilibrium states. However there are situations where linearization cannot be used to deduce even the local stability properties of the nonlinear system. Also, linearization tells us nothing about the global stability properties of the nonlinear system.

In general, we cannot explicitly obtain solutions for nonlinear systems. Lyapunov theory allows to say something about the stability properties of a system without knowing the form or structure of the solutions.

• In this chapter, $V$ is a scalar-valued function of the state, i.e. $V : \mathbb{R}^n \to \mathbb{R}$. At any time $k$, the one step change in $V$ along a solution $x(\cdot)$ of system (10.1) is given by

$$V(x(k+1)) - V(x(k)) = \Delta V(x(k))$$

where

$$\boxed{\Delta V(x) := V(f(x)) - V(x)}$$

## 10.1 Stability

**Theorem 10** (Stability) *Suppose there exists a locally positive definite function $V$ and a scalar $R > 0$ such that*

$$\Delta V(x) \leq 0 \qquad \textit{for} \quad ||x|| < R$$

*Then the origin is a stable equilibrium state.*

• If $V$ satisfies the hypotheses of the above theorem, then $V$ is said to be a *Lyapunov function* which guarantees the stability of origin.

**Example 141**

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

Consider

$$V(x) = x^2$$

as a *candidate Lyapunov function.* Then $V$ is a lpdf and

$$\Delta V(x) = 0$$

Hence (it follows from theorem 10 that) the origin is stable.

## 10.2 Asymptotic stability

**Theorem 11** (Asymptotic stability) *Suppose there exists a locally positive definite function $V$ and a scalar $R > 0$ such that*

$$\Delta V(x) < 0 \qquad \text{for} \quad x \neq 0 \quad \text{and} \quad ||x|| < R$$

*Then the origin is an asymptotically stable equilibrium state.*

**Example 142**

$$x(k+1) = \frac{1}{2}x(k)$$

Consider

$$V(x) = x^2$$

Then $V$ is a lpdf and

$$\begin{aligned} \Delta V(x) &= (\frac{1}{2}x)^2 - x^2 \\ &= -\frac{3}{4}x^2 \\ &< 0 \qquad \text{for} \quad x \neq 0 \end{aligned}$$

Hence the origin is AS.

**Example 143**

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

Consider

$$V(x) = x^2$$

Then $V$ is a lpdf and

$$\begin{aligned} \Delta V(x) &= (\frac{1}{2}x + x^2)^2 - x^2 \\ &= -x^2(\frac{3}{2} + x)(\frac{1}{2} - x) \\ &< 0 \qquad \text{for} \quad |x| < \frac{1}{2},\ x \neq 0 \end{aligned}$$

Hence the origin is AS.

## 10.3 Boundedness

**Theorem 12** *Suppose there exists a radially unbounded function $V$ and a scalar $R \geq 0$ such that*

$$\Delta V(x) \leq 0 \qquad \textit{for} \quad ||x|| \geq R$$

*Then all solutions of (10.1) are bounded.*

Note that, in the above theorem, $V$ does not have to positive away from the origin; it only has to be radially unbounded.

**Example 144**

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

Consider

$$V(x) = x^2$$

Since $V$ is radially unbounded and

$$\begin{aligned}
\Delta V(x) &= \left(\frac{2x}{1+x^2}\right)^2 - x^2 \\
&= -\frac{x^2(x^2+3)(x^2-1)}{(x^2+1)^2} \\
&\leq 0 \qquad \text{for} \quad |x| \geq 1
\end{aligned}$$

the hypotheses of the above theorem are satisfied with $R = 1$; hence all solutions are bounded. Note that the origin is unstable.

## 10.4 Global asymptotic stability

**Theorem 13** (Global asymptotic stability) *Suppose there exists a positive definite function $V$ such that*

$$\Delta V(x) < 0 \qquad \textit{for all} \quad x \neq 0$$

*Then the origin is a globally asymptotically stable equilibrium state.*

**Example 145**

$$x(k+1) = \frac{1}{2}x(k)$$

**Example 146**

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

Consider

$$V(x) = x^2$$

Then

$$\begin{aligned} \Delta V(x) &= \left(\frac{x}{1+x^2}\right)^2 - x^2 \\ &= -\frac{2x^4 + x^6}{(1+x^2)^2} \\ &< 0 \qquad \text{for all} \quad x \neq 0 \end{aligned}$$

Hence this system is GAS about zero.

**Exercise 103** Consider any scalar system described by

$$x(k+1) = g(x(k))$$

where $g$ has the following properties:

$$|g(x)| < |x| \quad \text{for} \quad x \neq 0$$

Show that this system is GAS.

## 10.5 Exponential stability

**Theorem 14** (Global exponential stability.) *Suppose there exists a function $V$ and scalars $\alpha, \beta_1, \beta_2$ such that for all $x$,*

$$\beta_1 ||x||^2 \leq V(x) \leq \beta_2 ||x||^2 \qquad\qquad \beta_1, \beta_2 > 0$$

*and*

$$V(f(x)) \;\; \leq \;\; \alpha^2 V(x) \qquad\qquad 0 \leq \alpha < 1$$

*Then, every solution satisfies*

$$||x(k)|| \leq \sqrt{\frac{\beta_2}{\beta_1}} \alpha^k ||x(0)|| \qquad \textit{for} \quad k \geq 0$$

*Hence, the origin is a globally exponentially stable equilibrium state with rate of convergence $\alpha$.*

PROOF.

**Example 147**

$$x(k+1) = -\frac{1}{2} x(k)$$

Considering

$$V(x) = x^2$$

we have

$$\begin{aligned} V(f(x)) &= \frac{1}{4}x^2 \\ &= (\frac{1}{2})^2 V(x) \end{aligned}$$

Hence, we have GES with rate of convergence $\alpha = \frac{1}{2}$.

**Example 148**

$$x(k+1) = \frac{1}{2}\sin(x(k))$$

Considering

$$V(x) = x^2$$

we have

$$\begin{aligned} V(f(x)) &= (\frac{1}{2}\sin x)^2 \\ &= (\frac{1}{2})^2 |\sin x|^2 \\ &\leq (\frac{1}{2})^2 |x|^2 \\ &= (\frac{1}{2})^2 V(x) \end{aligned}$$

Hence, we have GES with rate of convergence $\alpha = \frac{1}{2}$

# Chapter 11

# Lyapunov theory for linear time-invariant systems

The main result of this section is contained in Theorem 15.

## 11.1 Positive and negative (semi)definite matrices

### 11.1.1 Definite matrices

For any square matrix $n \times n$ matrix $P$ we can define an associated quadratic form by

$$x^*Px = \sum_{i=1}^{n}\sum_{j=1}^{n} p_{ij}x_i^*x_j$$

Recall that a square complex matrix $P$ is hermitian if $P^* = P$ where $P^*$ is the complex conjugate of $P$. If $P$ is real then hermitian is equivalent to symmetric, that is, $P^T = P$.

**Fact 5** *Every hermitian matrix $P$ has the following properties.*

*(a) All eigenvalues of $P$ are real.*

*(b) The scalar $x^*Px$ is real for all $x \in \mathbb{C}^n$. and satisfies*

$$\lambda_{min}(P)||x||^2 \leq x^*Px \leq \lambda_{max}(P)||x||^2$$

*where $\lambda_{min}(P)$ and $\lambda_{max}(P)$ denote the minimum and maximum eigenvalues of $P$.*

**Defn.** A hermitian matrix $P$ is positive definite (pd) if

$$x^*Px > 0$$

for all nonzero $x$. We denote this by $P > 0$. The matrix $P$ is negative definite (nd) if $-P$ is positive definite; we denote this by $P < 0$.

**Example 149** For

$$P = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$$

we have (note the completion of squares trick)

$$\begin{aligned} x^*Px &= x_1^*x_1 - x_1^*x_2 - x_2^*x_1 + 2x_2^*x_2 \\ &= (x_1 - x_2)^*(x_1 - x_2) + x_2^*x_2 \\ &= |x_1 - x_2|^2 + |x_2|^2 \end{aligned}$$

Clearly, $x^*Px \geq 0$ for all $x$. If $x^*Px = 0$, then $x_1 - x_2 = 0$ and $x_2 = 0$; hence $x = 0$. So, $P > 0$.

**Fact 6** *The following statements are equivalent for any hermitian matrix $P$.*

*(a) $P$ is positive definite.*

*(b) All the eigenvalues of $P$ are positive.*

*(c) All the leading principal minors of $P$ are positive.*

**Example 150** Consider

$$P = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$$

Since $p_{11} = 1 > 0$ and $\det(P) = 1 > 0$, we must have $P > 0$.

```
>> eig(P)
ans =
    0.3820
    2.6180
```

Note the positive eigenvalues.

**Exercise 104** Suppose $P$ is hermitian and $T$ is invertible. Show that $P > 0$ iff $T^*PT > 0$.

**Exercise 105** Suppose $P$ and $Q$ are two hermitian matrices with $P > 0$. Show that $P + \lambda Q > 0$ for all real $\lambda$ sufficiently small; i.e., there exists $\bar{\lambda} > 0$ such that whenever $|\lambda| < \bar{\lambda}$, one has $P + \lambda Q > 0$.

### 11.1.2 Semi-definite matrices*

**Defn.** A hermitian matrix $P$ is positive semi-definite (psd) if

$$x^*Px \geq 0$$

for all non-zero $x$. We denote this by $P \geq 0$

$P$ is *negative semi-definite (nsd)* if $-P$ is positive semi-definite; we denote this by $P \leq 0$

**Fact 7** *The following statements are equivalent for any hermitian matrix $P$.*

*(a) $P$ is positive semi-definite.*

*(b) All the eigenvalues of $P$ are non-negative.*

*(c) All the leading minors of $P$ are positive.*

**Example 151** This example illustrates that non-negativity of the leading principal minors of $P$ is not sufficient for $P \geq 0$.

$$P = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}$$

We have $p_{11} = 0$ and $\det(P) = 0$. However,

$$x^* P x = -|x_2|^2$$

hence, $P$ is not psd. Actually, $P$ is nsd.

**Fact 8** *Consider any $m \times n$ complex matrix $M$ and let $P = M^* M$. Then*

*(a) $P$ is hermitian and $P \geq 0$*

*(b) $P > 0$ iff rank $M = n$.*

**Example 152**

$$P = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

Since

$$P = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \end{pmatrix}$$

and

$$\text{rank} \begin{pmatrix} 1 & 1 \end{pmatrix} = 1$$

$P \geq 0$ but $P$ is not pd.

## 11.2 Lyapunov theory

### 11.2.1 Asymptotic stability results

The scalar system

$$\dot{x} = ax$$

is asymptotically stable if and only if $a+a^* < 0$. Consider now a general linear time-invariant system described by

$$\dot{x} = Ax \tag{11.1}$$

The generalization of $a + a^* < 0$ to this system is $A + A^* < 0$. We will see shortly that this condition is sufficient for asymptotic stability; however, as the following example illustrates, it is not necessary.

**Example 153** The matrix

$$A = \begin{pmatrix} -1 & 3 \\ 0 & -1 \end{pmatrix}$$

is asymptotically stable. However, the matrix

$$A + A^* = \begin{pmatrix} -2 & 3 \\ 3 & -2 \end{pmatrix}$$

is not negative definite.

Since stability is invariant under a similarity transformation $T$, i.e., the stability of $A$ and $T^{-1}AT$ are equivalent for any nonsingular $T$, we should consider the more general condition

$$T^{-1}AT + T^*A^*T^{-*} < 0$$

Introducing the hermitian matrix $P := T^{-*}T^{-1}$, and pre- and post-multiplying the above inequality by $T^{-*}$ and $T^{-1}$, respectively, yields

$$P > 0 \tag{11.2a}$$
$$PA + A^*P < 0 \tag{11.2b}$$

We now show that the existence of a hermitian matrix $P$ satisfying these conditions guarantees asymptotic stability.

**Lemma 6** *Suppose there is a hermitian matrix $P$ satisfying (11.2). Then system (11.1) is asymptotically stable*

PROOF. Suppose there exists a hermitian matrix $P$ which satisfies inequalities (11.2). Consider any eigenvalue $\lambda$ of $A$. Let $v \neq 0$ be an eigenvector corresponding to $\lambda$, i.e.,

$$Av = \lambda v$$

Then

$$v^*PAv = \lambda v^*Pv \; ; \qquad v^*A^*Pv = \bar{\lambda}v^*Pv$$

Pre- and post-multiplying inequality (11.2b) by $v^*$ and $v$, respectively, yields

$$\lambda v^* P v + \bar{\lambda} v^* P v < 0$$

i.e.,

$$2Re(\lambda) v^* P v < 0$$

Since $P > 0$, we must have $v^* P v > 0$; hence $Re(\lambda) < 0$. Since the above holds for every eigenvalue of $A$, system (11.1) is asymptotically stable. ■

**Remark 1** A hermitian matrix $P$ which satisfies inequalities (11.2) will be referred to as a *Lyapunov matrix* for (11.1) or $A$.

**Remark 2** Conditions (11.2) are referred to as *linear matrix inequalities (LMI's)*. In recent years, efficient numerical algorithms have been developed to solve LMI's or determine that a solution does not exist.

**Lyapunov functions.** To obtain an alternative interpretation of inequalities (11.2), consider the quadratic function $V$ of the state defined by

$$V(x) := x^* P x$$

Since $P > 0$, this function has a strict global minimum at zero. Using inequality (11.2b), it follows that along any non-zero solution $x(\cdot)$ of (11.1),

$$\begin{aligned} \frac{dV(x(t))}{dt} &= \dot{x}^* P x + x^* P \dot{x} \\ &= x^* \left( A^* P + P A \right) x \\ &< 0 \end{aligned}$$

i.e., $V(x(\cdot))$ is strictly decreasing along any non-zero solution. Intuitively, one then expects every solution to asymptotically approach the origin. The function $V$ is called a *Lyapunov function* for the system. The concept of a Lyapunov function readily generalizes to nonlinear systems and has proved a very useful tool in the analysis of the stability properties of nonlinear systems.

**Remark 3** Defining the hermitian matrix $S := P^{-1}$, inequalities (11.2) become

$$S > 0 \tag{11.3a}$$

$$AS + SA^* < 0 \tag{11.3b}$$

Hence the above lemma can be stated with $S$ replacing $P$ and the preceding inequalities replacing (11.2).

So far we have shown that if a LTI system has a Lyapunov matrix, then it is AS. Is the converse true? That is, does every AS LTI system have a Lyapunov matrix? And if this is true how does one find a Lyapunov matrix? To answer this question note that satisfaction of inequality (11.2b) is equivalent to

$$\boxed{PA + A^*P + Q = 0} \tag{11.4}$$

where $Q$ is a hermitian positive definite matrix. This linear matrix equation is known as the Lyapunov equation. So one approach to looking for Lyapunov matrices could be to choose a pd hermitian $Q$ and determine whether the Lyapunov equation has a pd hermitian solution for $P$.

We first show that if the system $\dot{x} = Ax$ is asymptotically stable and the Lyapunov equation (11.4) has a solution then, the solution is unique. Suppose $P_1$ and $P_2$ are two solutions to (11.4). Then,

$$(P_2 - P_1)A + A^*(P_2 - P_1) = 0\,.$$

Hence,

$$e^{A^*t}(P_2 - P_1)Ae^{At} + e^{A^*t}A^*(P_2 - P_1)e^{At} = 0$$

that is,

$$\frac{d\left(e^{A^*t}(P_2 - P_1)e^{At}\right)}{dt} = 0$$

This implies that, for all $t$,

$$e^{A^*t}(P_2 - P_1)e^{At} = e^{A^*0}(P_2 - P_1)e^{A0} = P_2 - P_1\,.$$

Since $\dot{x} = Ax$ is asymptotically stable, $\lim_{t\longrightarrow\infty} e^{At} = 0$. Hence

$$P_2 - P_1 = \lim_{t\longrightarrow\infty} e^{A^*t}(P_2 - P_1)e^{At} = 0\,.$$

From this it follows that $P_2 = P_1$.

The following lemma tells us that every AS LTI system has a Lyapunov matrix and a Lyapunov matrix can be obtained by solving the Lyapunov equation with any pd hermitian $Q$.

**Lemma 7** *Suppose system (11.1) is asymptotically stable. Then for every matrix $Q$, the Lyapunov equation (11.4) has a unique solution for $P$. If $Q$ is positive definite hermitian then this solution is positive-definite hermitian.*

PROOF. Suppose system (11.1) is asymptotically stable and consider any matrix $Q$. Let

$$\boxed{P := \int_0^\infty e^{A^*t}Qe^{At}dt}$$

This integral exists because each element of $e^{At}$ is exponentially decreasing.

To show that $P$ satisfies the Lyapunov equation (11.4), use the following properties of $e^{At}$,

$$e^{At}A = \frac{de^{At}}{dt} \qquad A^*e^{A^*t} = \frac{de^{A^*t}}{dt}$$

to obtain

$$
\begin{aligned}
PA + A^*P &= \int_0^\infty \left(e^{A^*t}Qe^{At}A + A^*e^{A^*t}Qe^{At}\right)\, dt \\
&= \int_0^\infty \left(e^{A^*t}Q\frac{de^{At}}{dt} + \frac{de^{A^*t}}{dt}Qe^{At}\right)\, dt \\
&= \int_0^\infty \frac{d\left(e^{A^*t}Qe^{At}\right)}{dt}\, dt \\
&= \lim_{t\to\infty}\int_0^t \frac{de^{A^*t}Qe^{At}}{dt}\, dt \\
&= \lim_{t\to\infty} e^{A^*t}Qe^{At} \;-\; Q \\
&= -Q
\end{aligned}
$$

We have already demonstrated uniqueness of solutions to (11.4),
Suppose $Q$ is pd hermitian. Then it should be clear that $P$ is pd hermitian. ■.

Using the above two lemmas, we can now state the main result of this section.

**Theorem 15** *The following statements are equivalent.*

*(a) The system $\dot{x} = Ax$ is asymptotically stable.*

*(b) There exist positive definite hermitian matrices $P$ and $Q$ satisfying the Lyapunov equation (11.4).*

*(c) For* every *positive definite hermitian matrix $Q$, the Lyapunov equation (11.4) has a unique solution for $P$ and this solution is hermitian positive-definite.*

Proof. The first lemma yields (b) $\Longrightarrow$ (a). The second lemma says that (a) $\Longrightarrow$ (c). Hence, (b) $\Longrightarrow$ (c).

To see that (c) $\Longrightarrow$ (b), pick any positive definite hermitian $Q$. So, (b) is equivalent to (c). Also, (c) $\Longrightarrow$ (a); hence (a) and (c) are equivalent.

■.

**Example 154**

$$
\begin{aligned}
\dot{x_1} &= x_2 \\
\dot{x_2} &= -x_1 + cx_2
\end{aligned}
$$

Here

$$A = \begin{pmatrix} 0 & 1 \\ -1 & c \end{pmatrix}$$

Choosing

$$Q = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

and letting

$$P = \begin{pmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{pmatrix}$$

(note we have taken $p_{21} = p_{12}$ because we are looking for symmetric solutions) the Lyapunov equation results in

$$\begin{aligned} -2p_{12} + 1 &= 0 \\ p_{11} + cp_{12} - p_{22} &= 0 \\ 2p_{12} + 2cp_{22} + 1 &= 0 \end{aligned}$$

We consider three cases:
*Case i)* $c = 0$. No solution; hence no GAS.
*Case ii)* $c = -1$. A single solution; this solution is pd; hence GAS.
*Case iii)* $c = 1$. A single solution; this solution is not pd; hence no GAS.

**Remark 4** One can state the above theorem replacing $P$ with $S$ and replacing (11.4) with

$$AS + SA^* + Q = 0 \tag{11.5}$$

### 11.2.2 MATLAB.

```
>> help lyap

 LYAP   Lyapunov equation.
        X = LYAP(A,C) solves the special form of the Lyapunov matrix
        equation:

                A*X + X*A' = -C

        X = LYAP(A,B,C) solves the general form of the Lyapunov matrix
        equation:

                A*X + X*B = -C

        See also DLYAP.
```

**Note.** In MATLAB, A' is the complex conjugate transpose of A (i.e. $A^*$). Hence, to solve (11.4) one must use P= lyap(A', Q).

```
>> q=eye(2);

>> a=[0 1; -1 -1];

>> p=lyap(a', q)
p =
    1.5000    0.5000
    0.5000    1.0000

>> p*a + a'*p
ans =
   -1.0000   -0.0000
         0   -1.0000

>> det(p)
ans =
    1.2500

>> eig(p)
ans =
    1.8090
    0.6910
```

**Exercise 106** Consider the Lyapunov equation

$$PA + A^*P + 2\alpha P + Q = 0$$

where $A$ is a square matrix, $Q$ is some positive-definite hermitian matrix and $\alpha$ is some positive scalar. *Show* that this equation has a unique solution for $P$ and this solution is positive-definite hermitian if and only if for every eigenvalue $\lambda$ of $A$,

$$Re(\lambda) < -\alpha$$

### 11.2.3 Stability results*

For stability, the following lemma is the analog of Lemma 11.2.

**Lemma 8** *Suppose there is a hermitian matrix $P$ satisfying*

$$\begin{aligned} P &> 0 && (11.6a) \\ PA + A^*P &\leq 0 && (11.6b) \end{aligned}$$

*Then system (11.1) is stable.*

We do not have the exact analog of lemma 7. However the following can be shown.

**Lemma 9** *Suppose system (11.1) is stable. Then there exist a positive definite hermitian matrix $P$ and a positive semi-definite hermitian matrix $Q$ satisfying the Lyapunov equation (11.4).*

The above lemma does *not* state that for every positive semi-definite hermitian $Q$ the Lyapunov equation has a solution for $P$. Also, when the Lyapunov equation has a solution, it is not unique. This is illustrated in the following example.

**Example 155**

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

With

$$P = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

we obtain

$$PA + A^*P = 0$$

In this example the Lyapunov equation has a pd solution with $Q = 0$; this solution is not unique; any matrix of the form

$$P = \begin{pmatrix} p & 0 \\ 0 & p \end{pmatrix}$$

(where $p$ is arbitrary) is also a solution.

If we consider the psd matrix

$$Q = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

the Lyapunov equation has no solution.

Combining the above two lemmas results in the following theorem.

**Theorem 16** *The following statements are equivalent.*

*(a) The system $\dot{x} = Ax$ is stable.*

*(b) There exist a positive definite hermitian matrix $P$ and a positive semi-definite matrix $Q$ satisfying the Lyapunov equation (11.4).*

## 11.3 Mechanical systems*

**Example 156** A simple structure

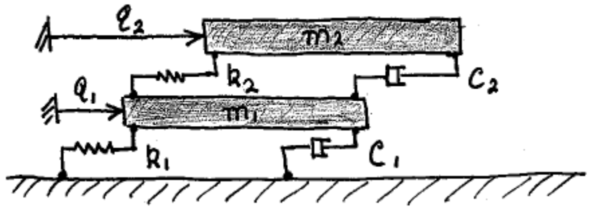



Figure 11.1: A simple structure

Recall

$$\begin{array}{ccccccccccc} m_1\ddot{q}_1 & + & (c_1+c_2)\dot{q}_1 & - & c_2\dot{q}_2 & + & (k_1+k_2)q_1 & - & k_2q_2 & = & 0 \\ m_2\ddot{q}_2 & - & c_2\dot{q}_1 & + & c_2\dot{q}_2 & - & k_2q_1 & + & k_2q_2 & = & 0 \end{array}$$

Letting

$$q = \begin{pmatrix} q_1 \\ q_2 \end{pmatrix}$$

this system can be described by the following second order vector differential equation:

$$M\ddot{q} + C\dot{q} + Kq = 0$$

where the symmetric matrices $M, C, K$ are given by

$$M = \begin{pmatrix} m_1 & 0 \\ 0 & m_2 \end{pmatrix} \qquad C = \begin{pmatrix} c_1+c_2 & -c_2 \\ -c_2 & c_2 \end{pmatrix} \qquad K = \begin{pmatrix} k_1+k_2 & -k_2 \\ -k_2 & k_2 \end{pmatrix}$$

Since $m_1, m_2 > 0$, we have $M > 0$.

If $k_1, k_2 > 0$ then $K > 0$ (why?) and if $c_1, c_2 > 0$ then $C > 0$.

Note also that

$$\begin{array}{lclcl} \text{kinetic energy} & = & \frac{1}{2}m_1\dot{q}_1^2 + \frac{1}{2}m_2\dot{q}_2^2 & = & \frac{1}{2}\dot{q}^T M\dot{q} \\[2ex] \text{potential energy} & = & \frac{1}{2}k_1q_1^2 + \frac{1}{2}k_2(q_2-q_1)^2 & = & \frac{1}{2}q^T Kq \end{array}$$

Consider now a general mechanical system described by

$$\boxed{M\ddot{q} + C\dot{q} + Kq = 0}$$

where $q(t)$ is a real $N$-vector of generalized coordinates which describes the configuraton of the system. The real matrix $M$ is the *inertia matrix* and satisfies

$$M^T = M > 0$$

We call $K$ and $C$ the *'stiffness' matrix* and *'damping' matrix* respectively. The *kinetic energy* of the system is given by

$$\frac{1}{2}\dot{q}^T M \dot{q}$$

With $x = \left(q^T, \dot{q}^T\right)^T$, this system has a state space description of the form $\dot{x} = Ax$ with

$$A = \begin{pmatrix} 0 & I \\ -M^{-1}K & -M^{-1}C \end{pmatrix}$$

**A first Lyapunov matrix.** Suppose

$$K^T = K > 0 \qquad \text{and} \qquad C^T = C \geq 0\,.$$

The *potential energy* of the system is given by

$$\frac{1}{2}q^T K q$$

Consider the following *candidate Lyapunov matrix*

$$P = \frac{1}{2}\begin{pmatrix} K & 0 \\ 0 & M \end{pmatrix}$$

Then $P^* = P > 0$ and

$$PA + A^*P + Q = 0$$

where

$$Q = \begin{pmatrix} 0 & 0 \\ 0 & C \end{pmatrix}$$

It should be clear that

$$PA + A^*P \leq 0$$

iff

$$C \geq 0$$

Hence,

$$\boxed{K^T = K > 0 \text{ and } C^T = C \geq 0 \text{ imply stability.}}$$

Note that

$$\begin{aligned} V(x) &= x^T P x = \frac{1}{2}\dot{q}^T M \dot{q} + \frac{1}{2}q^T K q = \text{total energy} \\ \dot{V}(x) &= -\dot{q}^T C \dot{q} \end{aligned}$$

**A second Lyapunov matrix.**

$$P = \frac{1}{2}\begin{pmatrix} K+\lambda C & \lambda M \\ \lambda M & M \end{pmatrix} \quad = \quad \frac{1}{2}\begin{pmatrix} K & 0 \\ 0 & M \end{pmatrix} + \frac{\lambda}{2}\begin{pmatrix} C & M \\ M & 0 \end{pmatrix}$$

For sufficiently small $\lambda$, the matrix $P$ is positive definite.

Now,

$$PA + A^*P + Q = 0$$

where

$$Q = \begin{pmatrix} \lambda K & 0 \\ 0 & C - \lambda M \end{pmatrix}$$

For sufficiently small $\lambda > 0$, the matrix $C - \lambda M$ is pd and, hence, $Q$ is pd. So,

$$\boxed{K^T = K > 0 \text{ and } C^T = C > 0 \text{ imply asymptotic stability.}}$$

## 11.4 Rates of convergence for linear systems*

**Theorem 17** *Consider an asymptotically stable linear system described by $\dot{x} = Ax$ and let $\alpha$ be any real number satisfying*

$$0 < \alpha < \bar{\alpha}$$

*where*

$$\bar{\alpha} := -\max\{\Re(\lambda) : \lambda \textit{ is an eigenvalue of } A\}\,.$$

*Then $\dot{x} = Ax$ is GES with rate $\alpha$.*

Proof: Consider any $\alpha$ satisfying $0 < \alpha < \bar{\alpha}$. As a consequence of the definition of $\bar{\alpha}$, all the eigenvalues of the matrix $A+\alpha I$ have negative parts. Hence, the Lyapunov equation

$$P(A+\alpha I) + (A+\alpha I)'P + I = 0\,.$$

has a unique solution for $P$ and $P = P^T > 0$. As a candidate Lyapunov for the system $\dot{x} = Ax$, consider $V(x) = x^T P x$. Then,

$$\begin{aligned}\dot{V} &= x^T P\dot{x} + \dot{x}^T P x\\ &= x^T PAx + (Ax)^T Px\\ &= x^T(PA + A^T P)x\,.\end{aligned}$$

From the above Lyapunov matrix equation, we obtain that

$$PA + A^T P = -2\alpha P - I\,;$$

hence,

$$\begin{aligned}\dot{V} &= -2\alpha x^T Px - x^T x\\ &\leq -2\alpha V(x)\,.\end{aligned}$$

Hence the system is globally exponentially stable with rate of convergence $\alpha$.

## 11.5 Linearization and exponential stability

Consider a nonlinear system

$$\dot{x} = f(x) \tag{11.7}$$

and suppose that $x^e$ is an equilibrium state for this system. We assume that $f$ is differential at $x^e$ and let

$$Df(x^e) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x^e) & \frac{\partial f_1}{\partial x_2}(x^e) & \cdots & \frac{\partial f_1}{\partial x_n}(x^e) \\ \frac{\partial f_2}{\partial x_1}(x^e) & \frac{\partial f_2}{\partial x_2}(x^e) & \cdots & \frac{\partial f_2}{\partial x_n}(x^e) \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1}(x^e) & \frac{\partial f_n}{\partial x_2}(x^e) & \cdots & \frac{\partial f_n}{\partial x_n}(x^e) \end{pmatrix}$$

where $x = (x_1, x_2, \cdots, x_n)$. Then the linearization of $\dot{x} = f(x)$ about $x^e$ is the linear system defined by

$$\delta\dot{x} = A\delta x \qquad \text{where} \quad A = Df(x^e)\,. \tag{11.8}$$

**Theorem 18** *Suppose $x^e$ is an equilibrium state of a nonlinear system of the form $\dot{x} = f(x)$ and the corresponding linearization is exponentially stable. Then the nonlinear system is exponentially stable about $x^e$.*

Proof. Suppose the linearization is exponentially stable and let

$$\bar{\alpha} := -\max\{\Re(\lambda) \,:\, \lambda \text{ is an eigenvalue of A }\}\,.$$

Since all the eigenvalues have negative real parts, we have $\bar{\alpha} > 0$. Consider now any $\alpha$ satisfying $0 < \alpha < \bar{\alpha}$. As a consequence of the definition of $\bar{\alpha}$, all the eigenvalues of the matrix $A+\alpha I$ have negative parts. Hence, the Lyapunov equation

$$P(A+\alpha I) + (A+\alpha I)'P + I = 0\,.$$

has a unique solution for $P$ and $P = P^T > 0$. As a candidate Lyapunov function for the nonlinear system, consider

$$V(x) = (x-x^e)'P(x-x^e)\,.$$

Recall that

$$f(x) = f(x^e) + Df(x^e)(x-x^e) + o(x)$$

where the "remainder term" has the following property:

$$\lim_{x\to 0,\, x\neq 0} \frac{o(x)}{\|x\|} = 0\,.$$

Hence

$$f(x) = A(x-x^e) + o(x)$$

and

$$\begin{aligned}
\dot{V} &= 2x^T P\dot{x} \\
&= x^T P f(x) \\
&= 2x^T PAx + 2x^T o(x) \\
&\leq x^T(PA + A^T P)x + 2\|x\| o(x)\,.
\end{aligned}$$

From the above Lyapunov matrix equation, we obtain that

$$PA + A^T P = -2\alpha P - I\,;$$

hence,

$$\begin{aligned}
\dot{V} &= -2x^T Px - x^T x + 2\|x\| o(x) \\
&= -2\alpha V(x) - \|x\|^2 + o(x)\|x\|\,.
\end{aligned}$$

As a consequence of the properties of $o$, there is a scalar $R > 0$ such that

$$o(x) \leq \|x\| \text{ when } \|x\| \leq R\,.$$

Hence, whenever $\|x\| \leq R$ we obtain that

$$\dot{V} \leq -2\alpha V(x)\,.$$

Hence the nonlinear system is exponentially stable about $x^e$ with rate of convergence $\alpha$.

# Part III

# Input-Output Properties

# Chapter 12

# Input output response

$$\begin{array}{ccccc} \text{OUTPUTS} & & & & \text{INPUTS} \\ \longleftarrow & & \boxed{\text{SYSTEM}} & & \longleftarrow \end{array}$$

## 12.1 Response of continuous-time systems

In general, a linear time-invariant continuous-time system with inputs and outputs is described by

$$\overset{y}{\longleftarrow} \boxed{\begin{array}{rcl} \dot{x} & = & Ax + Bu \\ y & = & Cx + Du \end{array}} \overset{u}{\longleftarrow}$$

where $A, B, C$ and $D$ are matrices of dimensions $n \times n$, $n \times m$, $p \times n$ and $p \times m$, respectively.

### 12.1.1 State response and the convolution integral

The solution to

$$\dot{x}(t) = Ax(t) + Bu(t) \qquad \text{with} \qquad x(t_0) = x_0 \tag{12.1}$$

is unique and is given by

$$\boxed{x(t) \quad = \quad e^{A(t-t_0)}x_0 + \int_{t_0}^{t} e^{A(t-\tau)}Bu(\tau)d\tau} \tag{12.2}$$

for all $t$.

Proof. Consider any time $t$. Considering any time $\tau$ in the interval bounded by $t_0$ and $t$, we have

$$\dot{x}(\tau) = Ax(\tau) + Bu(\tau)$$

Pre-multiply both sides by $e^{-A\tau}$ and rearrange to get:

$$e^{-A\tau}\dot{x}(\tau) - e^{-A\tau}Ax(\tau) \quad = \quad e^{-A\tau}Bu(\tau)$$

Hence

$$\frac{d}{d\tau}(e^{-A\tau}x(\tau)) \quad = \quad e^{-A\tau}Bu(\tau)$$

Integrate with respect to to $\tau$ from $t_0$ to $t$ and use $x(t_0) = x_0$:

$$e^{-At}x(t) - e^{-At_0}x_0 \quad = \quad \int_{t_0}^{t} e^{-A\tau}Bu(\tau)d\tau$$

Now premultiply by $e^{At}$ and use $e^{At}e^{-At} = I$ and $e^{At}e^{-At_0} = e^{A(t-t_0)}$ to obtain

$$x(t) - e^{A(t-t_0)}x_0 \quad = \quad \int_{t_0}^{t} e^{A(t-\tau)}Bu(\tau)d\tau$$

and, hence, the desired result. ■

• Note that the solution depends linearly on the pair $(x_0, u(\cdot))$.

• Suppose that for some input function $u(\cdot)$, the function $x^p(\cdot)$ is a particular solution to (12.1) with $x^p(t_0) = x_0^p$. Then all other solutions are given by

$$x(t) = e^{A(t-t_0)}(x_0 - x_0^p) + x^p(t)$$

(Why?)

**Impulse response**

Suppose $t_0 = 0$ and

$$u(t) = \delta(t)v$$

where $\delta(\cdot)$ is the Dirac delta function or the unit impulse function and $v$ is some constant vector. (Don't confuse with the $\delta$ used in linearization.) Recall that if zero is in the interval bounded by $t_0$ and $t$ and $g$ is any continuous function then,

$$\int_{t_0}^{t} g(t)\delta(t)dt = g(0)$$

Hence, the solution to (12.1) is given by

$$x(t) = e^{At}x_0 + e^{At}Bv$$

From this we see that the zero initial state response (that is, $x_0 = 0$) to the above impulse input is given by

$$x(t) = e^{At}Bv$$

*This is the same as the system response to zero input and initial state, $x_0 = Bv$.*

**Response to exponentials and sinusoids**

Suppose $t_0 = 0$ and

$$u(t) = e^{\lambda t} v$$

where $v$ is a constant $m$-vector and $\lambda$ is not an eigenvalue of $A$. Then the solution to (12.1) is

$$x(t) = \underbrace{e^{At}(x_0 - x_0^p)}_{\text{transient part}} \quad + \quad \underbrace{e^{\lambda t} x_0^p}_{\text{steady state part}}$$

where

$$x_0^p = (\lambda I - A)^{-1} B v$$

We need only prove the result for $x_0 = x_0^p$. So considering

$$x(t) = e^{\lambda t}(\lambda I - A)^{-1} B v$$

we have

$$\dot{x}(t) = \lambda e^{\lambda t}(\lambda I - A)^{-1} B v = \lambda x(t)$$

and

$$(\lambda I - A)x(t) = e^{\lambda t} B v = B u(t)$$

Hence,

$$\lambda x(t) = A x(t) + B u(t)$$

and

$$\dot{x}(t) = \lambda x(t) = A x(t) + B u(t)$$

■

- If $A$ is asymptotically stable, the transient part goes to zero as $t$ goes to infinity.

### 12.1.2 Output response

The output response of the system

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) + Du(t) \end{aligned}$$

subject to initial condition $x(t_0) = x_0$ is unique and is given by

$$\boxed{y(t) = Ce^{A(t-t_0)} x_0 + \int_{t_0}^{t} G(t-\tau) u(\tau) d\tau} \tag{12.3}$$

where the impulse response matrix $G$ is defined by

$$\boxed{G(t) = Ce^{At} B + \delta(t) D}$$

where $\delta(\cdot)$ is the Dirac delta function or the unit impulse function. (Don't confuse with the $\delta$ used in linearization.)

To see this, recall that if zero is in the interval bounded by $t_0$ and $t$ and $g$ is any continuous function then,

$$\int_{t_0}^{t} \delta(t)g(t)dt = g(0)$$

Recalling that

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

we see that

$$\begin{aligned}
y(t) &= Cx(t) + Du(t) \\
&= Ce^{A(t-t_0)}x_0 + C\int_{t_0}^{t} e^{A(t-\tau)}Bu(\tau)d\tau + Du(t) \\
&= Ce^{A(t-t_0)}x_0 + \int_{t_0}^{t} Ce^{A(t-\tau)}Bu(\tau) + \delta(t-\tau)Du(\tau)d\tau \\
&= Ce^{A(t-t_0)}x_0 + \int_{t_0}^{t} G(t-\tau)u(\tau)\,d\tau
\end{aligned}$$

**Zero initial conditions response:** $x(0) = 0$.

$$\boxed{y(t) = \int_{0}^{t} G(t-\tau)u(\tau)d\tau} \tag{12.4}$$

Sometimes this is represented by

$$y \longleftarrow \boxed{G} \longleftarrow u$$

This defines a linear map from the space of input functions to the space of output functions.

**Example 157**

$$\begin{aligned}
\dot{x} &= -x + u \\
y &= x
\end{aligned}$$

Here

$$\begin{aligned}
G(t) &= Ce^{At}B + \delta(t)D \\
&= e^{-t}
\end{aligned}$$

So, for zero initial state,

$$y(t) = \int_{0}^{t} e^{-(t-\tau)}u(\tau)\,d\tau$$

**Example 158** Unattached mass

$$\begin{aligned}\dot{x}_1 &= x_2\\ \dot{x}_2 &= u\\ y &= x_1\end{aligned}$$

Here

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad B = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \qquad C = \begin{pmatrix} 1 & 0 \end{pmatrix} \qquad D = 0$$

Hence

$$e^{At} = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$$

and

$$\begin{aligned} G(t) &= Ce^{At}B \\ &= t \end{aligned}$$

So,

$$y(t) = \int_0^t (t-\tau)u(\tau)\,d\tau$$

**Remark 5** Description (12.4) is a very general representation of a linear input-output system. It can also be used for infinite-dimensional systems such as distributed parameter systems and systems with delays.

**Example 159** Consider the system with input delay described by

$$\begin{aligned}\dot{x}(t) &= -x(t) + u(t-1)\\ y(t) &= \phantom{-}x(t)\,.\end{aligned}$$

Since

$$y(t) = \int_0^t e^{-(t-\hat{\tau})}u(\hat{\tau}-1)\,d\hat{\tau}$$

(why?) we have (upon changing variable of integration)

$$y(t) = \int_{-1}^{t-1} e^{-(t-1-\tau)}u(\tau)\,d\tau$$

Considering only inputs for which

$$u(t) = 0 \qquad \text{for} \quad t < 0$$

and defining

$$G(t) = \begin{cases} 0 & \text{for} \quad t < 1 \\ e^{-(t-1)} & \text{for} \quad t \geq 1 \end{cases}$$

the input-output response of this system is described by (12.4).

### Impulse response

Consider first a SISO (single-input single-output) system with impulse response function $g$ and subject to zero initial conditions, $x(0) = 0$. Suppose such a system is subject to a unit impulse input, that is, $u(t) = \delta(t)$. Then, using (12.3), its output is given by

$$y(t) = \int_0^t g(t-\tau)\delta(\tau)\,\tau = g(t)\,.$$

Thus, as its name suggests, *the impulse response function $g$ is the output response of the system to zero initial conditions and a unit impulse input.*

Consider now a MIMO (multi-input multi-output) system with impulse response function $G$ and subject to zero initial conditions. Suppose this system is subject to an input of the form

$$u(t) = \delta(t)v$$

where $v$ is constant $m$-vector. Then, from (12.3),

$$y(t) = \int_0^t G(t-\tau)v\delta(\tau)\,\tau = G(t)v\,.$$

By considering $v = e^j$, where each component of $e^j$ is zero except for its $j$-th component which equals one, we see that $G_{ij}$ is the zero initial state response of $y_i$ when $u_j = \delta$ and all other inputs are zero, that is,

$$\left.\begin{array}{rcl} u_1(t) & = & 0 \\ & \vdots & \\ u_j(t) & = & \delta(t) \\ & \vdots & \\ u_m(t) & = & 0 \end{array}\right\} \qquad \Longrightarrow \qquad \left\{\begin{array}{rcl} y_1(t) & = & G_{1j}(t) \\ & \vdots & \\ y_i(t) & = & G_{ij}(t) \\ & \vdots & \\ y_p(t) & = & G_{pj}(t) \end{array}\right.$$

### Impulse response and transfer function

Recall that the transfer function (matrix) is given by

$$\hat{G}(s) = C(sI - A)^{-1}B + D$$

Also, the impulse response (matrix) is given by

$$G(t) = Ce^{At}B + \delta(t)D$$

Using the following Laplace transform results:

$$\mathcal{L}(e^{At})(s) = (sI - A)^{-1}, \qquad \mathcal{L}(\delta) = 1$$

we obtain

$$\boxed{\hat{G} = \mathcal{L}(G)}$$

that is,

*the transfer function $\hat{G}$ is the Laplace transform of the impulse response $G$.*

**Response to exponentials and sinusoids**

Suppose $t_0 = 0$ and

$$u(t) = e^{\lambda t} v$$

where $v$ is a constant $m$ vector and $\lambda$ is not an eigenvalue of $A$. Then

$$y(t) = \underbrace{Ce^{At}(x_0 - x_0^p)}_{\text{transient part}} + \underbrace{e^{\lambda t}\hat{G}(\lambda)v}_{\text{steady state part}}$$

where

$$x_0^p = (\lambda I - A)^{-1}Bv$$

and

$$\boxed{\hat{G}(\lambda) := C(\lambda I - A)^{-1}B + D}$$

If $A$ is asymptotically stable, then the transient part goes to zero as $t$ goes to infinity.

## 12.2 Discretization and zero-order hold

Here we consider the discretization of a continuous-time system described by

$$\boxed{\dot{x}_c(t) = A_c x_c(t) + B_c u_c(t)} \tag{12.5}$$

Let $T > 0$ be the sampling time for discretization.

• **Zero-order hold.** Suppose that the input to the continuous-time system comes from a zero-order hold with sampling time $T$, that is,

$$u_c(t) = u_d(k) \qquad \text{for} \quad kT \le t < (k+1)T$$

Thus the input $u_c$ is constant over each sampling interval, $[kT, (k+1)T)$.

Suppose also that we sample the state of the continuous-time system at multiples of the sampling time and let

$$x_d(k) := x_c(kT)$$

Then

$$\begin{aligned}
x_d(k+1) &= x_c((k+1)T) \\
&= e^{A_cT}x_c(kT) + \int_{kT}^{(k+1)T} e^{A_c((k+1)T-t)}B_c u_c(t)\,dt \\
&= e^{A_cT}x_d(k) + \int_{kT}^{(k+1)T} e^{A_c((k+1)T-t)}B_c u_d(k)\,dt \\
&= e^{A_cT}x_d(k) + \int_{kT}^{(k+1)T} e^{A_c((k+1)T-t)}\,dt\,B_c u_d(k)
\end{aligned}$$

Changing the variable of integration from $t$ to $\tau := (k+1)T - t$ yields

$$\int_{kT}^{(k+1)T} e^{A_c((k+1)T-t)}\,dt \;=\; -\int_{T}^{0} e^{A_c\tau}\,d\tau = \int_0^T e^{A_c\tau}\,d\tau$$

Hence, we obtain the following discretization of the continuous-time system (12.5):

$$\boxed{x_d(k+1) = A_d x_d(k) + B_d u_d(k)} \tag{12.6}$$

where

$$\boxed{\begin{aligned} A_d &= e^{A_cT} \\ B_d &= \textstyle\int_0^T e^{A_c\tau}d\tau\, B_c \end{aligned}}$$

Recall that

$$e^{A_ct} = I + A_ct + \frac{1}{2!}A_c^2t^2 + \frac{1}{3!}A_c^3t^3 + \cdots$$

Thus,

$$\int_0^T e^{A_c\tau}d\tau = TI + \frac{T^2}{2!}A_c + \frac{T^2}{3!}A_c^2T^3 + \cdots .$$

and we have the following power series for $A_d$ and $B_d$:

$$\begin{aligned}
A_d &= I + TA_c + \frac{T^2}{2!}A_c^2 + \frac{T^3}{3!}A_c^3 + \cdots \\
B_d &= TB_c + \frac{T^2}{2!}A_cB_c + \frac{T^3}{3!}A_c^2B_c + \cdots .
\end{aligned}$$

If $A_c$ is invertible, then

$$\int_0^T e^{A_c\tau}d\tau = (e^{A_cT} - I)A_c^{-1} = (A_d - I)A_c^{-1}$$

and

$$B_d = (A_d - I)A_c^{-1}B_c \tag{12.7}$$

If we approximate $A_d$ and $B_d$ by considering only terms up to first order in $T$ in their power series, we obtain

$$A_d \approx I + TA_c \qquad \text{and} \qquad B_d \approx TB_c\,.$$

**Example 160** (Discretization of forced undamped oscillator)

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

$$A_c = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \qquad B_c = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Here,

$$e^{A_c t} = \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix}$$

Hence,

$$A_d = e^{A_c T} = \begin{pmatrix} \cos T & \sin T \\ -\sin T & \cos T \end{pmatrix} \qquad B_d = \int_0^T e^{A_c \tau}\, d\tau\, B_c = \begin{pmatrix} 1 - \cos T \\ \sin T \end{pmatrix}$$

Consider $T = 2\pi$. Then,

$$A_d = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad B_d = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

So,

$$\begin{aligned} x_{1d}(k+1) &= x_{1d}(k) \\ x_{2d}(k+1) &= x_{2d}(k) \end{aligned}$$

Where did the input go?

What about $T = \pi/2$?

$$A_d = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \qquad B_d = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

```
>> help c2d

 C2D    Conversion of state space models from continuous to discrete time.
        [Phi, Gamma] = C2D(A,B,T)  converts the continuous-time system:
                .
                x = Ax + Bu
```

```
        to the discrete-time state-space system:

                x[n+1] = Phi * x[n] + Gamma * u[n]

        assuming a zero-order hold on the inputs and sample time T.

        See also: C2DM, and D2C.

>> a=[0 1 ; -1 0] ; b=[0; 1] ; T= pi/2;
>> [phi, gamma]= c2d(a,b,T)

phi =
   -0.0000    1.0000
   -1.0000    0.0000

gamma =
    1.0000
    1.0000
```

### 12.2.1 Input-output discretization

Consider a continuous-time system:

$$\begin{aligned}\dot{x}_c(t) &= A_c x_c(t) + B_c u_c(t)\\ y_c(t) &= C_c x_c(t) + D_c u_c(t)\end{aligned}$$

Suppose $T > 0$ is the *sampling time*.

• *Zero order hold:*

$$u_c(t) = u_d(k) \qquad \text{for} \quad kT \leq t < (k+1)T$$

Letting

$$x_d(k) := x_c(kT) \qquad y_d(k) := y(kT)$$

we obtain

$$\begin{aligned}x_d(k+1) &= A_d x_d(k) + B_d u_d(k)\\ y_d(k) &= C_d x_d(k) + D_d u_d(k)\end{aligned}$$

where

$$\boxed{A_d = e^{A_c T} \qquad B_d = \int_0^T e^{A_c\tau}\, d\tau\, B_c \qquad C_d = C_c \qquad D_d = D_c}$$

## 12.3 Response of discrete-time systems*

A general linear time-invariant discrete-time system is described by

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) \\ y(k) &= Cx(k) + Du(k) \end{aligned}$$

where $k \in \mathbb{R}$ is 'time', $x(k) \in \mathbb{R}^n$ in the state, $u(k) \in \mathbb{R}^m$ the input, and $y(k) \in \mathbb{R}^p$ is the output.

### 12.3.1 State response and the convolution sum

The solution to

$$x(k+1) = Ax(k) + Bu(k) \qquad x(k_0) = x_0 \tag{12.8}$$

is unique and is given by

$$\boxed{x(k) \quad = \quad A^{(k-k_0)}x_0 + \sum_{j=k_0}^{k-1} A^{(k-1-j)}Bu(j)} \tag{12.9}$$

Proof. Use induction.

### 12.3.2 Input-output response and the pulse response matrix

The solution to

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) \\ y(k) &= Cx(k) + Du(k) \end{aligned}$$

subject to initial condition

$$x(k_0) = x_0$$

is unique and is given by

$$\boxed{y(k) = CA^{(k-k_0)}x_0 + \sum_{j=k_0}^{k} G(k-j)u(j)}$$

where the pulse response matrix $G$ is defined by

$$\boxed{\begin{aligned} G(0) &= D \\ G(k) &= CA^{k-1}B \qquad \text{for} \quad k = 1, 2, \ldots \end{aligned}}$$

Clearly, the above holds for $k = k_0$. For $k > k_0$, recall that

$$x(k) = A^{(k-k_0)}x_0 + \sum_{j=k_0}^{k-1} A^{(k-1-j)}Bu(j)$$

Hence,

$$
\begin{aligned}
y(k) &= CA^{(k-k_0)}x_0 + \sum_{j=k_0}^{k-1} CA^{(k-1-j)}Bu(j) + Du(k) \\
&= CA^{(k-k_0)}x_0 + \sum_{j=k_0}^{k-1} G(k-j)u(j) + G(0)u(k) \\
&= CA^{(k-k_0)}x_0 + \sum_{j=k_0}^{k} G(k-j)u(j)
\end{aligned}
$$

**Zero initial conditions response:** $x(0) = 0$.

$$\boxed{y(k) = \sum_{j=0}^{k} G(k-j)u(j)}$$

Sometimes this is represented by

$$y \longleftarrow \boxed{G} \longleftarrow u$$

**Pulse response:** Suppose $x(0) = 0$ and

$$u(0) = p(k)v \qquad \text{for} \quad k = 0, 1, \ldots$$

where $v$ is constant $m$ vector and $p$ is the unit pulse function defined by

$$
\begin{aligned}
p(0) &= 1 \\
p(k) &= 0 \qquad \text{for} \quad k = 1, 2, \ldots
\end{aligned}
$$

Then,

$$y(k) = G(k)v \qquad \text{for} \quad k = 0, 1, \ldots$$

From this we can see that $G_{ij}$ is the zero initial condition response of output component $y_i$ when input component $u_j$ is a unit pulse and all other inputs are zero, that is,

$$
\left.
\begin{aligned}
u_1(k) &= 0 \\
&\vdots \\
u_j(k) &= p(k) \\
&\vdots \\
u_m(k) &= 0
\end{aligned}
\right\}
\quad \Longrightarrow \quad
\left\{
\begin{aligned}
y_1(k) &= G_{1j}(k) \\
&\vdots \\
y_i(k) &= G_{ij}(k) \\
&\vdots \\
y_p(k) &= G_{pj}(k)
\end{aligned}
\right.
$$

• So, for a SISO (single-input single-output) system, $G$ is the output response due a unit pulse input and zero initial conditions.

# Exercises

**Exercise 107** Obtain (by hand) the response of each of the following systems due to a unit impulse input and the zero initial conditions. For each case, determine whether the response contains all the system modes.

a)

$$\begin{aligned}\dot{x}_1 &= -5x_1 + 2x_2 + u \\ \dot{x}_2 &= -12x_1 + 5x_2 + u\end{aligned}$$

b)

$$\begin{aligned}\dot{x}_1 &= -5x_1 + 2x_2 + u \\ \dot{x}_2 &= -12x_1 + 5x_2 + 2u\end{aligned}$$

c)

$$\begin{aligned}\dot{x}_1 &= -5x_1 + 2x_2 + u \\ \dot{x}_2 &= -12x_1 + 5x_2 + 3u\end{aligned}$$

**Exercise 108** Consider the system with disturbance input $w$ and performance output $z$ described by

$$\begin{aligned}\dot{x}_1 &= -x_1 + x_2 + w \\ \dot{x}_2 &= -x_1 + 4x_2 + 2w \\ z &= x_1\,.\end{aligned}$$

Using an appropriate Lyapunov equation, determine

$$\int_0^\infty \|z(t)\|^2\,dt$$

for each of the following situations.

(a)

$$w = 0 \qquad \text{and} \qquad x(0) = (1, 0)\,.$$

(b)

$$w(t) = \delta(t) \qquad \text{and} \qquad x(0) = 0\,.$$

# Chapter 13

# Observability

## 13.1 Observability

We are concerned here with the problem of determining the state of a system based on measurements of its inputs and outputs. Consider the system

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y &= Cx + Du \end{aligned} \tag{13.1}$$

with initial condition

$$x(0) = x_0$$

where the $n$-vector $x(t)$ is the state, the $m$-vector $u(t)$ is the input and the $p$-vector $y(t)$ is a vector of measured outputs; we call $y$ the measured output (vector).

Consider any time interval $[0, T]$ with $T > 0$. The basic observability problem is as follows. Suppose we have knowledge of the input $u(t)$ and the output $y(t)$ over the interval $[0, T]$; can we uniquely determine the initial state $x_0$? If we can do this for all input histories and initial states $x_0$ we say that the system is observable. So, for observability, we require that for each input history, different initial states produce different output histories, or, equivalently, if two output histories are identical, then the corresponding initial states must be the same. A formal definition of observability follows the next example. Note that knowledge of the initial state and the input history over the interval $[0, T]$ allows one to compute the state $x(t)$ for $0 \leq t \leq T$. This follows from the relationship

$$x(t) = e^{At}x_0 + \int_0^t e^{A(t-\tau)}Bu(\tau)\, d\tau \,.$$

**Example 161** Two unattached masses

$$\begin{aligned} \dot{x}_1 &= 0 \\ \dot{x}_2 &= 0 \\ y &= x_1 \end{aligned}$$

Clearly, this system is unobservable.

**DEFN.** *System (13.1) is* observable *over an interval* $[0, T]$ *if the following holds for each input history* $u(\cdot)$. *Suppose* $y^a(\cdot)$ *and* $y^b(\cdot)$ *are any two output histories of system (13.1) due to initial states* $x_0^a$ *and* $x_0^b$*, respectively, and* $y^a(t) = y^b(t)$ *for* $0 \le t \le T$*; then* $x_0^a = x_0^b$.

## 13.2 Main result

To describe the main results associated with observability of LTI systems, we introduce the observability matrix $Q_o$ associated with system (13.1). This is the following $pn \times n$ matrix

$$\boxed{Q_o = \begin{pmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{pmatrix}}$$

The following result is the main observability result for continuous-time LTI systems.

**Theorem 19 (Main observability result)** *For any* $T > 0$*, system (13.1) is observable over the interval* $[0, T]$ *if and only if the its observability matrix has maximum rank, that is,*

$$\boxed{rank \begin{pmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{pmatrix} = n}$$

*where* $n$ *is the number of state variables.*

An immediate consequence of the above theorem is that observability does not depend on the interval $[0, T]$. So, from now on we drop reference to the interval. Since observability only depends on the matrix pair $(C,\ A)$, we say that this pair is observable if system (13.1) is observable.

**SO systems.** For scalar output ($p = 1$) systems, $Q_o$ is a square $n \times n$ matrix; hence it has rank $n$ if and only if its determinant is nonzero. So, the above theorem has the following corollary.

**Corollary 1** *A scalar output system of the form (13.1) is observable if and only if its observability matrix* $Q_o$ *has non-zero determinant.*

**Example 162** (Two unattached masses.) Recall example 161. Here $n = 2$,

$$A = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \qquad C = \begin{pmatrix} 1 & 0 \end{pmatrix}$$

Hence,

$$Q_o = \begin{pmatrix} C \\ CA \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

Since $\operatorname{rank} Q_o = 1 < n$, this system is unobservable.

**Example 163** (The unattached mass)

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= 0 \end{aligned}$$

We consider two cases:

(a) (Velocity measurement.) Suppose $y = x_2$. Then

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad C = \begin{pmatrix} 0 & 1 \end{pmatrix}$$

So,

$$Q_o = \begin{pmatrix} C \\ CA \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

Since $\operatorname{rank} Q_0 = 1 < n$, we have unobservability.

(b) (Position measurement.) Suppose $y = x_1$. Then

$$Q_o = \begin{pmatrix} C \\ CA \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Since $\operatorname{rank} Q_0 = 2 = n$, we have observability.

**Example 164 (Beavis and Butthead: mass center measurement)**

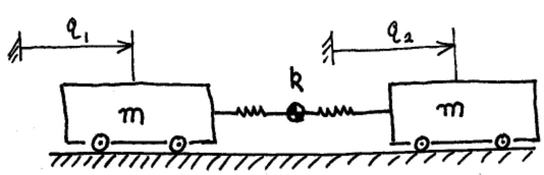



Figure 13.1: Example 164

$$\begin{aligned} m\ddot{q}_1 &= k(q_2 - q_1) \\ m\ddot{q}_2 &= -k(q_2 - q_1) \\ y &= q_1 + q_2 \end{aligned}$$

Actually, $y$ is not the location of the mass center of the system, however there is a one-to-one correspondence between $y$ and the mass center location.

With

$$x_1 = q_1 \qquad x_2 = q_2 \qquad x_3 = \dot{q}_1 \qquad x_4 = \dot{q}_2$$

we have

$$A = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -k/m & k/m & 0 & 0 \\ k/m & -k/m & 0 & 0 \end{pmatrix} \qquad\qquad C = \begin{pmatrix} 1 & 1 & 0 & 0 \end{pmatrix}$$

Since,

$$\begin{array}{rcccl} CA & & & = & \begin{pmatrix} 0 & 0 & 1 & 1 \end{pmatrix} \\ CA^2 & = & (CA)A & = & \begin{pmatrix} 0 & 0 & 0 & 0 \end{pmatrix} \\ CA^3 & = & (CA^2)A & = & \begin{pmatrix} 0 & 0 & 0 & 0 \end{pmatrix} \end{array}$$

the observability matrix for this system is

$$Q_o = \begin{pmatrix} C \\ CA \\ CA^2 \\ CA^3 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Since $\operatorname{rank} Q_o = 2 < n$, this system is not observable.

Physically, one can see why this system is not observable. The system can oscillate with the mass center fixed, that is, a non-zero state motion results in zero output history.

**Example 165 (Beavis and Butthead: observable)** Here we take the system from the last example and consider the measured output to be the position of the first mass.

$$\begin{array}{rcl} m\ddot{q}_1 & = & k(q_2 - q_1) \\ m\ddot{q}_2 & = & -k(q_2 - q_1) \\ y & = & q_1 \end{array}$$

With

$$x_1 = q_1 \qquad x_2 = q_2 \qquad x_3 = \dot{q}_1 \qquad x_4 = \dot{q}_2$$

we have

$$A = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -k/m & k/m & 0 & 0 \\ k/m & -k/m & 0 & 0 \end{pmatrix} \qquad\qquad C = \begin{pmatrix} 1 & 0 & 0 & 0 \end{pmatrix}$$

Hence,

$$\begin{array}{rcl} CA & = & \begin{pmatrix} 0 & 0 & 1 & 0 \end{pmatrix} \\ \\ CA^2 = (CA)A & = & \begin{pmatrix} -k/m & k/m & 0 & 0 \end{pmatrix} \\ \\ CA^3 = (CA^2)A & = & \begin{pmatrix} 0 & 0 & -k/m & k/m \end{pmatrix} \end{array}$$

So,

$$
Q_o = \begin{pmatrix} C \\ CA \\ CA^2 \\ CA^3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -k/m & k/m & 0 & 0 \\ 0 & 0 & -k/m & k/m \end{pmatrix}.
$$

Since $\operatorname{rank} Q_o = 4 = n$; this system is observable.

**Useful facts for checking observability.** In checking observability, the following facts are useful.

(a) Suppose there exists $n^*$ such that

$$
\mathcal{N}\left(\begin{pmatrix} C \\ CA \\ \vdots \\ CA^{n^*-1} \\ CA^{n^*} \end{pmatrix}\right) = \mathcal{N}\left(\begin{pmatrix} C \\ CA \\ \vdots \\ CA^{n^*-1} \end{pmatrix}\right)
$$

or, equivalently,

$$
\operatorname{rank} \begin{pmatrix} C \\ CA \\ \vdots \\ CA^{n^*-1} \\ CA^{n^*} \end{pmatrix} = \operatorname{rank} \begin{pmatrix} C \\ CA \\ \vdots \\ CA^{n^*-1} \end{pmatrix}
$$

Then

$$
\mathcal{N}(Q_o) = \mathcal{N}\left(\begin{pmatrix} C \\ CA \\ \vdots \\ CA^{n^*-1} \end{pmatrix}\right)
$$

and

$$
\operatorname{rank}(Q_o) = \operatorname{rank} \begin{pmatrix} C \\ CA \\ \vdots \\ CA^{n^*-1} \end{pmatrix}
$$

(b) Suppose $V^1, \dots, V^k$ is a sequence of matrices which satisfy

$$\mathcal{N}(V^1) \;=\; \mathcal{N}(C)$$

$$\mathcal{N}\left(\begin{pmatrix} V^1 \\ \vdots \\ V^{j-1} \\ V^j \end{pmatrix}\right) \;=\; \mathcal{N}\left(\begin{pmatrix} V^1 \\ \vdots \\ V^{j-1} \\ V^{j-1}A \end{pmatrix}\right) \qquad \text{for} \quad j = 2, \dots, k$$

Then

$$\mathcal{N}\left(\begin{pmatrix} C \\ \vdots \\ CA^k \\ CA^{k+1} \end{pmatrix}\right) = \mathcal{N}\left(\begin{pmatrix} V^1 \\ \vdots \\ V^k \\ V^kA \end{pmatrix}\right)$$

and

$$\text{rank}\begin{pmatrix} C \\ \vdots \\ CA^k \\ CA^{k+1} \end{pmatrix} \;=\; \text{rank}\begin{pmatrix} V^1 \\ \vdots \\ V^k \\ V^kA \end{pmatrix}$$

See next two examples.

**Example 166 (Beavis and Butthead: mass center measurement revisited)**

$$\begin{aligned} m\ddot{q}_1 &= k(q_2 - q_1) \\ m\ddot{q}_2 &= -k(q_2 - q_1) \\ y &= q_1 + q_2 \end{aligned}$$

With

$$x_1 = q_1 \qquad x_2 = q_2 \qquad x_3 = \dot{q}_1 \qquad x_4 = \dot{q}_2$$

we have

$$A = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -k/m & k/m & 0 & 0 \\ k/m & -k/m & 0 & 0 \end{pmatrix} \qquad\qquad C = \begin{pmatrix} 1 & 1 & 0 & 0 \end{pmatrix}$$

Hence,

$$CA \;=\; \begin{pmatrix} 0 & 0 & 1 & 1 \end{pmatrix}$$

$$CA^2 \;=\; \begin{pmatrix} 0 & 0 & 0 & 0 \end{pmatrix}$$

So, using useful fact (b) above,

$$\text{rank}\,(Q_o) = \text{rank}\begin{pmatrix} C \\ CA \end{pmatrix} = \text{rank}\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} = 2$$

Since $\text{rank}\,(Q_o) < n$, this system is not observable.

**Example 167** (Beavis and Butthead: observable)

$$
\begin{array}{rcl}
m\ddot{q}_1 &=& k(q_2 - q_1) \\
m\ddot{q}_2 &=& -k(q_2 - q_1) \\
y &=& q_1
\end{array}
$$

With

$$
x_1 = q_1 \qquad x_2 = q_2 \qquad x_3 = \dot{q}_1 \qquad x_4 = \dot{q}_2
$$

we have

$$
A = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -k/m & k/m & 0 & 0 \\ k/m & -k/m & 0 & 0 \end{pmatrix} \qquad\qquad C = \begin{pmatrix} 1 & 0 & 0 & 0 \end{pmatrix}
$$

Hence,

$$
\begin{array}{rcl}
CA &=& \begin{pmatrix} 0 & 0 & 1 & 0 \end{pmatrix} \\[2ex]
CA^2 &=& \begin{pmatrix} -k/m & k/m & 0 & 0 \end{pmatrix}
\end{array}
$$

So,

$$
\text{rank} \begin{pmatrix} C \\ CA \\ CA^2 \end{pmatrix} = \text{rank} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}
$$

and

$$
\text{rank}\,(Q_o) = \text{rank} \begin{pmatrix} C \\ CA \\ CA^2 \\ CA^3 \end{pmatrix} = \text{rank} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad = \quad 4
$$

Since $\text{rank}\, Q_o = n$; this system is observable.

**Example 168** *State feedback can destroy observability.*

$$
A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad C = \begin{pmatrix} 1 & 0 \end{pmatrix}
$$

$(C, A)$ is observable. Consider

$$
BK = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}
$$

Then

$$
A + BK = 0
$$

and $(C,\, A + BK)$ is unobservable.

## 13.3 Development of main result

To develop the main result presented in the last section, we first demonstrate the following result:

*System (13.1) is observable over an interval* $[0, T]$ *if and only if the uncontrolled system* ($u(t) \equiv 0$)

$$\begin{aligned} \dot{x} &= Ax \\ y &= Cx \end{aligned} \tag{13.2}$$

*has the property that the zero state is the only initial state which results in an output history which is identically zero; that is, it has the following property:*

$$y(t) = 0 \quad \textit{for} \quad 0 \le t \le T \qquad \textit{implies} \qquad x(0) = 0$$

To see that the above property is necessary for observability, simply note that the zero initial state produces a zero output history; hence if some nonzero initial state $x_0$ produces a zero output history, then there are two different initial states producing the same output history and the system is not observable.

To see that the above property is sufficient, consider any input history $u(\cdot)$ and suppose $y^a(\cdot)$ and $y^b(\cdot)$ are two output histories due to initial states $x_0^a$ and $x_0^b$, repectively, and $y^a(\cdot) = y^b(\cdot)$. Let $x^a(\cdot)$ and $x^b(\cdot)$ be the state histories corresponding the $x_0^a$ and $x_0^b$, respectively, that is,

$$\begin{aligned} \dot{x}^a &= Ax^a + Bu\,, \qquad x^a(0) = x_0^a \\ y^a &= Cx^a + Du \end{aligned}$$

and

$$\begin{aligned} \dot{x}^b &= Ax^b + Bu\,, \qquad x^b(0) = x_0^b \\ y^b &= Cx^b + Du \end{aligned}$$

Then, letting

$$\begin{aligned} x(t) &:= x^b(t) - x^a(t) \\ y(t) &:= y^b(t) - y^a(t) \end{aligned}$$

we obtain

$$\begin{aligned} \dot{x} &= Ax\,, \quad x(0) = x_0^b - x_0^a \\ y &= Cx \end{aligned}$$

Since $y(t) \equiv 0$ we must have $x_0^b - x_0^a = 0$; hence $x_0^b = x_0^a$. So, system (13.1) is observable.

**Example 169** (The unattached mass)

$$\begin{aligned}\dot{x}_1 &= x_2 \\ \dot{x}_2 &= 0\end{aligned}$$

If $y = x_2$, the system is unobservable: consider nonzero initial state $x_0 = [1 \quad 0]^T$; the corresponding output history satisfies $y(t) \equiv 0$. If $y = x_1$, the system is observable. To see this note that $x = [y,\ \dot{y}]^T$. If $y(t) \equiv 0$ over a nonzero interval, then $\dot{y}(0) = 0$ and hence, $x(0) = 0$.

We have just shown that to investigate the observability of system (13.1), we need only determine the initial states which result in an identically zero output for the system with zero input, that is system (13.2). So, let us look at the consequences of a zero output for the above system. So, suppose $x_0$ is an initial state which, for some $T > 0$, results in $y(t) = 0$ for $0 \le t \le T$, that is,

$$Cx(t) = 0 \quad \text{for} \quad 0 \le t \le T \tag{13.3}$$

where

$$\dot{x}(t) = Ax(t), \quad x(0) = x_0\,.$$

Differentiating the equation in (13.3) with respect to $t$ results in

$$0 = C\dot{x}(t) = CAx(t) = 0 \qquad \text{for} \quad 0 \le t \le T\,.$$

Repeated differentiation of (13.3) yields:

$$CA^k x(t) = 0 \qquad \text{for} \quad 0 \le t \le T \quad \text{and} \quad k = 0, 1, 2, \ldots$$

Letting $t = 0$ and using the initial condition $x(0) = x_0$ results in

$$CA^k x_0 = 0 \qquad \text{for} \quad k = 0, 1, 2, \ldots\,. \tag{13.4}$$

So, we have shown that if an initial state $x_0$ results in a zero output over some nonzero interval then, (13.4) holds.

Let us now demonstrate the converse, that is, if (13.4) holds, then the resulting output is zero over any nonzero interval. So, suppose (13.4) holds for some initial state $x_0$. Recall that $x(t) = e^{At}x_0$; hence

$$y(t) = Ce^{At}x_0$$

Since $e^{At}$ can be expressed as $e^{At} = \sum_{k=0}^{\infty} \beta_k(t)A^k$, we have

$$y(t) \;\; = \;\; Ce^{At}x_0 = \sum_{k=0}^{\infty} \beta(t) CA^k x_0 = 0\,.$$

So, we now conclude that an initial state $x_0$ results in a zero output over some nonzero interval if and only if (13.4) holds. Since, for each $k = 0, 1, 2\ldots$, the matrix $A^k$ can be expressed as a linear combination of $I, A, A^2, \dots, A^{n-1}$, it follows that (13.4) holds if and only if

$$CA^k x_0 = 0 \qquad \text{for} \quad k = 0, 1, 2, n{-}1\,. \tag{13.5}$$

Thus an initial state $x_0$ results in a zero output over some nonzero interval if and only if (13.5) holds. Recalling the definition of the observability matrix $Q_o$ associated with system (13.1) we obtain the following result:

| |
|---|
| *An initial state $x_0$ results in zero output over some nonzero interval if and only if $Q_o x_0 = 0$* |

Hence, system (13.1) is observable if and only if the only vector satisfying $Q_o x_0 = 0$ is the zero vector. This immediately leads to the following result on observability. Recall that the nullity of a matrix is the dimension of its null space.

*For any $T > 0$, system (13.1) is observable over the interval $[0, T]$ if and only if the nullity of its observability matrix is zero.*

Since the rank and nullity of $Q_o$ must sum to $n$ which is the maximum of $Q_o$, we now obtain main result, that is Theorem 19.

## 13.4 The unobservable subspace

**Example 170** At first sight, the system

$$\begin{aligned}\dot{x}_1 &= x_1\\ \dot{x}_2 &= x_2\\ y &= x_1 + x_2\end{aligned}$$

might look observable. However, it is not. To see this, introduce new state variables

$$x_o := x_1 + x_2 \qquad \text{and} \qquad x_{uo} := x_1 - x_2$$

Then

$$\begin{aligned}\dot{x}_o &= x_o\\ \dot{x}_u &= x_u\\ y &= x_o\end{aligned}$$

Hence, $x_o(0) = 0$ implies $x_o(t) = 0$ and, hence, $y(t) = 0$ for all $t$. In other words if the initial state $x(0)$ is in the set

$$\{x \in \mathbb{R}^2 \, : \, x_1 + x_2 = 0\}$$

The resulting output history is zero.

Figure 13.2: Example 170

Recall that observability matrix $Q_o$ associated with $(C,\, A)$ is defined to be the following $pn \times n$ matrix

$$Q_o = \begin{pmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{pmatrix}$$

Recall also that an initial state $x_0$ results in a zero output if and only if $Q_o x_0 = 0$, that is $x_o$ is in the null space of $Q_o$. So, we define the unobservable subspace $\mathcal{X}_{uo}$ associated with system (13.1) or $(C,\, A)$ to be the null space of the observability matrix, that is,

$$\boxed{\mathcal{X}_{uo} = \mathcal{N}(Q_o)}$$

and we have the following result.

**Lemma 10 (Unobservable subspace lemma)** *For any nonzero time interval* $[0, T]$, *the following statements are equivalent for system (13.2).*

*(i)* $y(t) = 0$ *for* $0 \leq t \leq T$.

*(ii)* $Q_o x(0) = 0$, *or, equivalently,* $x(0)$ *is in the unobservable subspace* $\mathcal{X}_{uo}$.

**Example 171** Recalling the last example we have $n = 2$ and

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad C = \begin{pmatrix} 1 & 1 \end{pmatrix}$$

Hence,

$$Q_o = \begin{pmatrix} C \\ CA \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

Clearly,

$$\mathcal{X}_{uo} = \mathcal{N}(Q_o) = \mathcal{N}\left(\begin{pmatrix} 1 & 1 \end{pmatrix}\right) = \mathcal{N}(C)$$

Hence

$$\mathcal{N}(Q_o) = \{x \in \mathbb{R}^2 \, : \, x_1 + x_2 = 0\}$$

## 13.5 Unobservable modes

Suppose we are interested in the observability of the the input-output system

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y &= Cx + Du \end{aligned}$$

Since observability is independent of $u$, we only need to consider the corresponding zero input system:

$$\begin{aligned} \dot{x} &= Ax \\ y &= Cx \end{aligned} \tag{13.6}$$

Recall that a mode of this system is a special solution of the form $x(t) = e^{\lambda t}v$. Such a solution results when $x(0) = v$ and $v$ is an eigenvector of $A$ with eigenvalue $\lambda$. We say that $e^{\lambda t}v$ is an unobservable mode of the above system if, the above system has a solution of the form $x(t) = e^{\lambda t}v$ and the corresponding output is zero, that is,

$$\boxed{x(t) \equiv e^{\lambda t}v \qquad \text{and} \quad y(t) \equiv 0}$$

The eigenvalue $\lambda$ is called an unobservable eigenvalue of the system. We also refer to $\lambda$ as an unobservable eigenvalue of $(C, A)$.

Clearly, if a system has an unobservable mode then, the system is not observable. In the next section, we show that if an LTI system is not observable, then it must have an unobservable mode. Thus:

*An LTI system is observable if and only if it has no unobservable modes.*

This is illustrated in the following example.

**Example 172**

$$\begin{aligned} \dot{x}_1 &= \lambda_1 x_1 + b_1 u \\ \dot{x}_2 &= \lambda_2 x_2 + b_2 u \\ y &= c_1 x_1 + c_2 x_2 + du \end{aligned}$$

Since

$$Q_o = \begin{pmatrix} C \\ CA \end{pmatrix} = \begin{pmatrix} c_1 & c_2 \\ \lambda_1 c_1 & \lambda_2 c_2 \end{pmatrix}$$

This system is observable if and only if

$$c_1 \neq 0,\ c_2 \neq 0\,, \lambda_1 \neq \lambda_2$$

Note that if $c_1 = 0$ then considering $u(t) \equiv 0$ this system has a solution

$$x(t) = e^{\lambda_1 t} \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

and the corresponding output satisfies

$$y(t) \equiv 0$$

### 13.5.1 PBH Test

The next lemma provides a useful characterization of unobservable eigenvalues.

**Lemma 11** *A complex number $\lambda$ is an unobservable eigenvalue of the pair $(C, A)$ if and only if the matrix*

$$\begin{pmatrix} A - \lambda I \\ C \end{pmatrix} \tag{13.7}$$

*does not have maximum rank (that is $n$, where $n$ is the number of columns in $A$.)*

Proof. Recall that $\lambda$ is an unobservable eigenvalue of $(C, A)$ if the system $\dot{x} = Ax$ has a non-zero solution of the form $x(t) = e^{\lambda t}v$ which satisfies $Cx(t) \equiv 0$. We have already seen that the existence of a nonzero solution of the form $e^{\lambda t}v$ is equivalent to $v$ being an eigenvector of $A$ with eigenvalue $\lambda$, that is, $v \neq 0$ and $Av = \lambda v$; this is equivalent to

$$(A - \lambda I)v = 0$$

and $v \neq 0$. Since $x(t) = e^{\lambda t}v$ and $e^{\lambda t} \neq 0$, the statement that $Cx(t) \equiv 0$ is equivalent to

$$Cv = 0\,.$$

So, we have shown that $\lambda$ is an unobservable eigenvalue of $(C, A)$ if and only if there is a nonzero vector satisfying

$$\begin{pmatrix} A - \lambda I \\ C \end{pmatrix} v = 0\,.$$

The existence of a nonzero $v$ satisfying the above equality is equivalent to saying that the columns of the matrix in the equality are linearly dependent, that is, the rank of the matrix is less than the number of its columns. Since the number of columns in the above matrix is less than or equal to the number of its rows, the maximum rank of this matrix equals the number of its columns. Hence, $\lambda$ being an unobservable eigenvalue of $(C, A)$ is equivalent to the matrix in (13.7) not having maximum rank. ■

**Example 173** Consider a system with

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad B = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \qquad C = \begin{pmatrix} -1 & 1 \end{pmatrix} \qquad D = 0\,.$$

Here, the observability matrix is

$$Q_o = \begin{pmatrix} C \\ CA \end{pmatrix} = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Since $Q_o$ has zero determinant, this system is not observable. Hence, the system has some unobservable modes. The eigenvalues of $A$ are 1 and $-1$. To determine the unobservable eigenvalues, we carry out the PBH test:

$$\begin{pmatrix} A - \lambda I \\ C \end{pmatrix} = \begin{pmatrix} -\lambda & 1 \\ 1 & -\lambda \\ -1 & 1 \end{pmatrix}$$

For $\lambda = 1$, the above matrix is rank deficient; hence 1 is an unobservable eigenvalue. For $\lambda = -1$, the PBH matrix has maximum rank. Hence, this eigenvalue is observable.

Note that the transfer function for this system is given by

$$\hat{G}(s) = \frac{s-1}{s^2-1} = \frac{1}{s+1}\,.$$

So, *the unobservable eigenvalue is not a pole of the transfer function.* This transfer function is also the transfer function of the stable system with with $A = -1$ and $B = C = 1$ and $D = 0$.

Since we have already stated that an LTI system is observable if and only if it has no unobservable eigenvalues, we obtain the following corollary to Lemma 11. This new result provides another useful test for observability.

**Corollary 1** (PBH observability test.) *The pair $(C, A)$ is observable if and only if the matrix*

$$\begin{pmatrix} A - \lambda I \\ C \end{pmatrix}$$

*has maximum rank for every eigenvalue $\lambda$ of $A$.*

**Example 174 (Unobservable modes)**

```
A
    -2     2     1     2
    -1     1     1     2
    -3     2     2     2
    -1     0    -1     1

B =

     2
     0
     0
     2

C
     1     0     0    -1

Qo= obsv(A, C)
     1     0     0    -1
    -1     2     2     1
    -7     4     4     7
    -9    -2    -2     9

rank(Qo)
     2                     %System is unobservable
```

```
eig(A)
  1.0000 + 2.0000i
  1.0000 - 2.0000i
 -1.0000
  1.0000


I=eye(4);

rank([A-(-1)*I; C])                    %PBH Time
    3

rank([A-(1)*I; C])
    3

rank([A-(1+2i)*I; C])
    4

rank([A-(1-2i)*I; C])
    4
```

So the eigenvalues $-1$ and $+1$ are unobservable. The eigenvalues $1+\imath 2$ and $1-\imath 2$ are not unobservable.

```
 [num den] = ss2tf(A,B,C,0)

num =  1.0e-013 *
        0   -0.0400    0.0933    0.0044   -0.1954


den =
   1.0000   -2.0000    4.0000    2.0000   -5.0000


        %Actually, the transfer function is ZERO!  Why?
```

Here we observe that unobservable eigenvalues are invariant under a state transformation. To see this consider any state transformation of the form $x = T\tilde{x}$; then the new PBH observability matrix is given by

$$\begin{pmatrix} \tilde{A}-\lambda I \\ \tilde{C} \end{pmatrix} = \begin{pmatrix} T^{-1}AT-\lambda I \\ CT \end{pmatrix} = \begin{pmatrix} T^{-1} & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} A-\lambda I \\ C \end{pmatrix} T$$

Since $T$ is invertible, we see that the new and the old PBH matrices have the same rank.

### 13.5.2 Existence of unobservable modes*

Here we will show that, if a system is unobservable then, is must have an unobservable mode. So, consider a linear time-invariant input-output system described

$$\dot{x} = Ax + Bu \tag{13.8a}$$

$$y = Cx + Du \tag{13.8b}$$

where $x(t)$ is an $n$-vector, $u(t)$ is an $m$-vector, and $y(t)$ is a $p$-vector. Suppose $(C, A)$ is not observable. In the following section, we show that there is a state transformation $x = T\tilde{x}$ so that the corresponding transformed system has the following special structure:

$$\boxed{\begin{array}{rcl} \dot{x}_o &=& A_{oo}x_o + B_o u \\ \dot{x}_u &=& A_{uo}x_o + A_{uu}x_u + B_u u \\ y &=& C_o x_o + Du \end{array}}$$

with

$$\begin{pmatrix} x_o \\ x_u \end{pmatrix} = \tilde{x}\,.$$

Also, the pair $(C_o, A_{oo})$ is observable. Roughly speaking, this transformation splits the state into an observable part $x_o$ and an unobservable part $x_u$; the subsystem

$$\begin{array}{rcl} \dot{x}_o &=& A_{oo}x_o + B_o u \\ y &=& C_o x_o + Du \end{array}$$

is observable; one can obtain $x_o$ from knowledge of the histories of $u$ and $y$.

Notice that, in the transformed system description, there is no connection from the state $x_u$ to the output $y$. We also note that, when $u(t) \equiv 0$ and $x_o(0) = 0$ we must have $x_o(t) \equiv 0$. Also,

$$\begin{array}{rcl} \dot{x}_u &=& A_{uu}x_u \\ y &=& 0 \end{array}$$

From this it should be clear that, if $\lambda$ is any eigenvalue of $A_{uu}$ then the transformed system (with zero input) has a solution which satisfies

$$\tilde{x}(t) = e^{\lambda t}\tilde{v} \qquad \text{and} \quad y(t) \equiv 0$$

for some nonzero $\tilde{v}$, that is $\lambda$ is an unobservable eigenvalue of the transformed system.

Hence, the system

$$\dot{x} = Ax\,, \qquad y = Cx$$

has a nonzero solution $x$ which satisfies

$$x(t) \equiv e^{\lambda t}v \qquad \text{and} \quad y(t) \equiv 0$$

Since the matrices $A$ and
$$\tilde{A} = \begin{pmatrix} A_{oo} & 0 \\ A_{uo} & A_{uu} \end{pmatrix}$$
are similar, they have the same characteristic polynomial. Moreover the characteristic polynomial of $\tilde{A}$ is given by

$$\det(sI - \tilde{A}) = \det(sI - A_{oo})\det(sI - A_{uu})$$

Hence the characteristic polynomial of $A$ is the product of the characteristic polynomials of $A_{oo}$ and $A_{uu}$. Thus, the eigenvalues of $A$ are the union of those of $A_{oo}$ and $A_{uu}$. Hence, the eigenvalues of $A_{uu}$ are the unobservable eigenvalues of $(C, A)$. The corresponding modes are the unobservable modes.

**Transfer function considerations.** Consider the reduced order observable system:

$$\begin{aligned} \dot{x}_o &= A_{oo}x_o + B_o u \\ y &= C_o x_o + Du \end{aligned}$$

One can readily show that the transfer function of this system is exactly the same as that of the original system, that is,

$$C(sI - A)^{-1}B + D = C_o(sI - A_o)^{-1}B_o + D$$

### 13.5.3 A nice transformation*

Consider a LTI input-output system described by

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y &= Cx + Du \end{aligned}$$

where $x(t)$ is an $n$-vector, $u(t)$ is an $m$-vector, and $y(t)$ is a $p$-vector. Suppose $(C, A)$ is not observable; then, its unobservable subspace $\mathcal{X}_{uo}$ is nontrivial. The next lemma states that $\mathcal{X}_{uo}$ is an invariant subspace of $A$.

**Lemma 12** *For any $n \times n$ matrix $A$ and any $p \times n$ matrix, $C$, the subspace*

$$\mathcal{X}_{uo} = \mathcal{N}\left(\begin{pmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{pmatrix}\right)$$

*is an invariant subspace for $A$.*

Proof. Consider any element $x$ of $\mathcal{X}_{uo}$. It satisfies

$$CA^k x = 0 \qquad \text{for} \quad k = 0, \cdots, n-1$$

Since

$$A^n = -\alpha_0 I - \alpha_1 A - \cdots - \alpha_{n-1} A^{n-1}$$

we obtain

$$CA^n x = 0$$

from which it follows that

$$CA^k(Ax) = 0 \qquad \text{for} \quad k = 0, \cdots, n-1$$

hence $Ax$ is contained in $\mathcal{X}_{uo}$. Since this holds for every element $x$ of $\mathcal{X}_{uo}$, it follows that $\mathcal{X}_{uo}$ is $A$-invariant. ■

**Transformation.** Suppose $n_u$ is the dimension of the unobservable subspace $\mathcal{X}_{uo}$, that is, $n_u$ is the nullity of the observability matrix, and let

$$n_o := n - n_u \,.$$

Choose any basis

$$t^{n_o+1}, \ldots, t^n$$

for the unobservable subspace, that is, the null space of the observability matrix $Q_o$. Extend it to a basis

$$t^1, \ldots, t^{n_o}, t^{n_o+1}, \ldots, t^n$$

for the whole state space $\mathbb{R}^n$ and let

$$T := \begin{pmatrix} t^1 & t^2 & \ldots & t^n \end{pmatrix}$$

(One could do this by choosing $T = V$ where $USV^*$ is a singular value decomposition of $Q_o$). Introduce the state transformation

$$x = T\tilde{x}$$

If we let

$$\begin{pmatrix} x_o \\ x_u \end{pmatrix} = \tilde{x}$$

where $x_o$ is an $n_o$-vector and $x_u$ is an $n_u$-vector, then

$$x \text{ is in } \mathcal{X}_{uo} \qquad \text{if and only if} \qquad x_o = 0$$

The transformed system description is given by

$$\begin{aligned} \dot{\tilde{x}} &= \tilde{A}\tilde{x} + \tilde{B}u \\ y &= \tilde{C}\tilde{x} + Du \end{aligned}$$

where

$$\tilde{A} = T^{-1}AT \qquad \tilde{B} = T^{-1}B \qquad \tilde{C} = CT$$

Since $\mathcal{X}_{uo}$ is an invariant subspace of $A$, the matrix $\tilde{A}$ has the following structure:

$$\tilde{A} = \begin{pmatrix} A_{oo} & 0 \\ A_{uo} & A_{uu} \end{pmatrix}$$

where $A_{oo}$ is $n_o \times n_o$. Since the vectors $t^{n_o+1}, \ldots, t^n$ are in the null space of $C$, the matrix $\tilde{C}$ has the following structure

$$\tilde{C} = \begin{pmatrix} C_o & 0 \end{pmatrix}$$

where $C_o$ is $p \times n_o$. If we partition $\tilde{B}$ as

$$\tilde{B} = \begin{pmatrix} B_o \\ B_u \end{pmatrix}$$

where $B_o$ is $n_o \times m$. then the transformed system is described by

$$\boxed{\begin{array}{rcll} \dot{x}_o & = & A_{oo}x_o & +B_o u \\ \dot{x}_u & = & A_{uo}x_o + A_{uu}x_u & +B_u u \\ y & = & C_o x_o & +Du \end{array}}$$

- The pair $(C_o, A_{oo})$ is observable.

**Example 175 (Unobservable modes)**

```
A
    -2     2     1     2
    -1     1     1     2
    -3     2     2     2
    -1     0    -1     1

C
     1     0     0    -1

Qo= obsv(A, C)
     1     0     0    -1
    -1     2     2     1
    -7     4     4     7
    -9    -2    -2     9

rank(Qo)
     2                       %System is unobservable



eig(A)
   1.0000 + 2.0000i
```

```
  1.0000 - 2.0000i
 -1.0000
  1.0000

I=eye(4);


rank([A-(-1)*I; C])                  %PBH Time
    3

rank([A-(1)*I; C])
    3

rank([A-(1+2i)*I; C])
    4

rank([A-(1-2i)*I; C])
    4
```

So the eigenvalues $-1$ and $+1$ are unobservable. The eigenvalues $1 + \imath 2$ and $1 - \imath 2$ are not unobservable.

```
rref(Qo)
    1     0     0    -1
    0     1     1     0
    0     0     0     0
    0     0     0     0

N=[1 0 0 1; 0 1 -1 0]'          %Basis for unobservable subspace
    1     0
    0     1
    0    -1
    1     0

Qo*N
    0     0                     %Just checking!
    0     0
    0     0
    0     0


rref([N I])
    1     0     0     0     0     1
    0     1     0     0    -1     0
    0     0     1     0     0    -1
```

```
     0     0     0     1     1     0

 T=[E(:,1:2) N ]                    % A nice transformation
     1     0     1     0
     0     1     0     1
     0     0     0    -1
     0     0     1     0

 Anew = inv(T)*A*T
    -1     2     0     0
    -4     3     0     0
    -1     0     0     1
     3    -2     1     0

 Cnew = C*T
     1     0     0     0

 Aoo= Anew(1:2,1:2)
    -1     2
    -4     3

 Auu= Anew(3:4,3:4)
     0     1
     1     0

      eig(Auu)
   -1.0000                           %Unobservable eigenvalues
    1.0000

 eig(Aoo)
   1.0000 + 2.0000i                  %The remaining eigenvalues
   1.0000 - 2.0000i
 B
     2
     0
     0
     2

 [num den] = ss2tf(A,B,C,0) num =  1.0e-013 *
         0   -0.0400    0.0933    0.0044   -0.1954


den =
    1.0000   -2.0000    4.0000    2.0000   -5.0000
```

```
%Actually, the transfer function is ZERO!  Why?
```

## 13.6 Observability grammians*

For each $T > 0$, we define the finite time observability grammian associated with $(C, A)$:

$$\boxed{W_o(T) = \int_0^T e^{A^*t} C^* C e^{At}\, dt}$$

To obtain an interpretation of $W_o(T)$, consider the output $y$ due to zero input and any initial state $x_0$, that is,

$$\begin{aligned} \dot{x} &= Ax \qquad x(0) = x_0 \\ y &= Cx \end{aligned}$$

Then $y(t) = Ce^{At}x_0$, and

$$\begin{aligned} \int_0^T ||y(t)||^2\, dt &= \int_0^T ||Ce^{At}x_0||^2\, dt = \int_0^T \left(Ce^{At}x_0\right)^* \left(Ce^{At}x_0\right)\, dt \\ &= \int_0^T x_0^* e^{A^*t} C^* C e^{At} x_0\, dt = x_0^* \left(\int_0^T e^{A^*t} C^* C e^{At}\, dt\right) x_0 \\ &= x_0^* W_o(T) x_0\,; \end{aligned}$$

thus

$$\int_0^T ||y(t)||^2\, dt = x_0^* W_o(T) x_0\,.$$

**Remark 6** Note that $W_o(T)$ is the solution at $t = T$ to the following initial value problem:

$$\dot{W}_o = W_o A + A^* W_o + C^* C \qquad \text{and} \qquad W_o(0) = 0$$

PROOF. With $s = t - \tau$, we have $\tau = t - s$ and

$$W_c(t) = \int_0^t e^{A\tau} C^* C e^{A^*\tau}\, d\tau = \int_0^t e^{A(t-s)} C^* C e^{A^*(t-s)}\, ds$$

Hence, $W_c(0) = 0$ and

$$\dot{W}_c = AW_c + W_c A^* + C^* C$$

■

**Lemma 13** *For each $T > 0$, the matrices $W_o(T)$ and $Q_o$ have the same null space.*

PROOF. We need to show that $\mathcal{N}(Q_o) = \mathcal{N}(W_o(T))$. Since

$$Q_o = \begin{pmatrix} C \\ CA \\ \vdots \\ CA^{(n-1)} \end{pmatrix}$$

it follows that $x$ is in $\mathcal{N}(Q_o)$ if and only if

$$CA^k x = 0 \qquad \text{for} \quad k = 0, 1, \ldots, n-1 \tag{13.9}$$

We first show that $\mathcal{N}(Q_o) \subset \mathcal{N}(W_o(T))$: Consider any $x$ in $\mathcal{N}(Q_o)$. Then (13.9) holds. Recalling the application of the Cayley-Hamilton theorem to functions of a matrix, it follows that for $0 \le t \le T$, there are scalars $\beta_0(t), \ldots, \beta_{n-1}(t)$ such that

$$e^{At} = \sum_{i=0}^{n-1} \beta_i(t) A^i$$

Hence,

$$Ce^{At}x = \sum_{i=0}^{n-1} \beta_i(t) CA^i x = 0$$

and

$$\begin{aligned} W_o(T)x &= \int_0^T e^{A^*t}C^*Ce^{At}\,dt\,x \\ &= \int_0^T e^{A^*t}C^*Ce^{At}x\,dt \\ &= 0 \end{aligned}$$

that is, $x$ is in $\mathcal{N}(W_o(T))$. Since the above holds for any $x \in \mathcal{N}(Q_o)$, we must have $\mathcal{N}(Q_o) \subset \mathcal{N}(W_o(T))$.

We now show that $\mathcal{N}(W_o(T)) \subset \mathcal{N}(Q_o)$ and hence $\mathcal{N}(Q_o) = \mathcal{N}(W_o(T))$. To this end, consider any $x$ in $\mathcal{N}(W_o(T))$. Then,

$$\begin{aligned} 0 &= x^*W_o(T)x \\ &= x^*\left(\int_0^T e^{A^*t}C^*Ce^{At}\,dt\right)x \\ &= \int_0^T x^*e^{A^*t}C^*Ce^{At}x\,dt \\ &= \int_0^T \left(Ce^{At}x\right)^*\left(Ce^{At}x\right)\,dt \\ &= \int_0^T ||Ce^{At}x||^2\,dt\,. \end{aligned}$$

Since the last integrand is non-negative for all $t$ and the integral is zero, the integrand must be zero for all $t$, that is,

$$Ce^{At}x \equiv 0\,.$$

If a function is zero for all $t$, all its derivatives must be zero for all $t$; hence

$$CA^k e^{At}x \equiv 0 \qquad \text{for} \quad k = 0, 1, \ldots, n-1\,.$$

Considering $t = 0$ yields

$$CA^k x = 0 \qquad \text{for} \quad k = 0, 1, \ldots, n-1\,,$$

that is, $x$ is in $\mathcal{N}(Q_o)$. Since the above holds for any $x$ in $\mathcal{N}(W_o(T))$, we must have $\mathcal{N}(W_o(T)) \subset \mathcal{N}(Q_o)$. ■

**Example 176** (The unattached mass)

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= 0 \end{aligned}$$

Here

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad e^{At} = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$$

(a) (Position measurement.) $y = x_1$. Here

$$C = \begin{pmatrix} 1 & 0 \end{pmatrix}$$

Hence,

$$W_o(T) = \int_0^T e^{A^*t}C^*Ce^{At}\,dt = \begin{pmatrix} T & T^2/2 \\ T^2/2 & T^3/3 \end{pmatrix}$$

and $W_o(T)$ and $Q_o$ have the same null space; that is, $\{0\}$.

(b) (Velocity measurement.) $y = x_2$. Here

$$C = \begin{pmatrix} 0 & 1 \end{pmatrix}$$

Hence,

$$W_o(T) = \int_0^T e^{A^*t}C^*Ce^{At}\,dt = \begin{pmatrix} 0 & 0 \\ 0 & T \end{pmatrix}$$

and $W_o(T)$ and $Q_o$ have the same null space.

The above lemma has the following consequence.

*The pair $(C, A)$ is observable if and only if for every $T > 0$ the corresponding observability grammian is invertible.*

**Computation of initial state.** One reason for introducing the observability grammian is as follows: Suppose $(C, A)$ is observable; then from the above lemma, $W_o(T)$ is invertible. The output for

$$\begin{aligned} \dot{x} &= Ax + Bu \qquad x(0) = x_0 \\ y &= Cx + Du \end{aligned}$$

is given by

$$y(t) = Ce^{At}x_0 + \int_0^t Ce^{A(t-\tau)}Bu(\tau)\,d\tau + Du(t)$$

Letting

$$\tilde{y}(t) = y(t) - \int_0^t Ce^{A(t-\tau)}Bu(\tau)\,d\tau - Du(t)$$

we have $\tilde{y}(t) = Ce^{At}x_0$ and

$$\begin{aligned} \int_0^T e^{A^*t}C^*\tilde{y}(t)\,dt &= \int_0^T e^{A^*t}C^*Ce^{At}x_0\,dt \\ &= W_o(T)x_0\,. \end{aligned}$$

Since $W_o(T)$ is invertible, we obtain the following explicit expression for the initial state in terms of the input and output histories:

$$\boxed{x_0 = W_o(T)^{-1}\int_0^T e^{A^*t}C^*\tilde{y}(t)\,dt}$$

The above expression explicitly shows us that if a system is observable, we can obtain the initial state from measurements of the input and output over any time interval. However, this formula is not usually used to obtain the state. One usually uses asymptotic observers or state estimators to asymptotically obtain the state; these are covered in a later chapter.

### 13.6.1 Infinite time observability grammian

Suppose $A$ is asymptotically stable, that is, all its eigenvalues have negative real parts, and consider the linear matrix equation

$$\boxed{W_oA + A^*W_o + C^*C = 0}$$

This is a Lyapunov equation of the form $PA + A^*P + Q = 0$ with $Q = C^*C$. Since $A$ is asymptotically stable the above equation has a unique solution for $W_o$ and this solution is given by

$$W_o = \int_0^\infty e^{A^*t}C^*Ce^{At}dt$$

We call this matrix the (infinite-time) observability grammian associated with $(C, A)$. One may show that $W_o$ and the observability matrix $Q_o$ have the same null-space. Since $W_o$ is a symmetric positive definite matrix, it follows that the null space of $W_o$ is $\{0\}$ if and only if $W_o$ is positive definite. Hence, we have the following result.

**Lemma 14** *Suppose all the eigenvalues of $A$ have negative real parts. Then $(C, A)$ is observable if and only if $W_o$ is positive definite.*

Note that if $y$ is the output of the system

$$\begin{aligned} \dot{x} &= Ax \\ y &= Cx \end{aligned}$$

with initial condition $x(0) = x_0$ then

$$\boxed{\int_0^\infty \|y(t)\|^2 \, dt = x_0^* W_0 x_0}$$

This follows from $y(t) = Ce^{At}x_0$ and

$$\begin{aligned} \int_0^\infty \|y(t)\|^2 \, dt &= \int_0^\infty y(t)^* y(t) \, dt \\ &= \int_0^\infty (Ce^{At}x_0)^* Ce^{At} x_0 \, dt \\ &= x_0^* \int_0^\infty e^{A^*t} C^* C e^{At} \, dt \, x_0 \\ &= x_0^* W_0 x_0 \, . \end{aligned}$$

## 13.7 Discrete time

Consider the discrete-time system described by

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) \\ y(k) &= Cx(k) + Du(k) \end{aligned} \tag{13.10}$$

where $k \in \mathbb{Z}$ is time, $x(k) \in \mathbb{R}^n$ is the state, $u(k) \in \mathbb{R}^m$ is the input, and $y(k) \in \mathbb{R}^p$ is the *measured output.*

Consider any fixed time $N > 0$. The basic observability problem is as follows. Suppose we have knowledge of the input sequence $\{ u(0), u(1), \cdots, u(N) \}$ and output sequence $\{ y(0), y(1), \cdots, y(N) \}$ over the interval $[0, N]$. Can we uniquely determine the initial state

$$x_0 = x(0)$$

If we can do this for input histories and all initial states we say the system is observable.

**Defn.** *System (13.10) is* observable *over the interval* $[0, N]$ *if the following holds for each input history* $u(\cdot)$. *Suppose* $y^a(\cdot)$ *and* $y^b(\cdot)$ *are any two output histories of system (13.1) due to initial states* $x_0^a$ *and* $x_0^b$*, respectively, and* $y^a(k) = y^b(k)$ *for* $0 \le k \le N$*; then* $x_0^a = x_0^b$.

• Our first observation is that system (13.10) is observable over $[0, N]$ if and only if the system

$$\begin{aligned} x(k+1) &= Ax(k) \\ y(k) &= Cx(k) \end{aligned} \tag{13.11}$$

has the property that the zero state is the only initial state resulting in an output history which is identically zero; that is, it has the the following property:

$$y(k) = 0 \quad \text{for} \quad 0 \le k \le N \qquad \Longrightarrow \qquad x(0) = 0$$

### 13.7.1 Main observability result

**Theorem 20 (Main observability theorem)** *For each* $\boxed{N \ge n-1}$*, system (13.10) is observable over* $[0, N]$ *if and only if*

$$\text{rank} \begin{pmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{pmatrix} = n$$

Proof. The proof can be obtained by showing that, for system (13.11) with $x_0 = x(0)$,

$$\begin{pmatrix} y(0) \\ y(1) \\ \vdots \\ y(n-1) \end{pmatrix} = \begin{pmatrix} Cx_0 \\ CAx_0 \\ \vdots \\ CA^{n-1}x_0 \end{pmatrix} = Q_0 x_0$$

Also use Cayley Hamilton. ∎

## 13.8 Exercises

**Exercise 109** Determine whether or not each of the following systems are observable.

$$
\begin{array}{lll}
\dot{x}_1 &=& -x_1 \\
\dot{x}_2 &=& x_2 + u \\
y &=& x_1 + x_2
\end{array}
\qquad
\begin{array}{lll}
\dot{x}_1 &=& -x_1 \\
\dot{x}_2 &=& x_2 + u \\
y &=& x_2
\end{array}
\qquad
\begin{array}{lll}
\dot{x}_1 &=& x_1 \\
\dot{x}_2 &=& x_2 + u \\
y &=& x_1 + x_2
\end{array}
\qquad
\begin{array}{lll}
\dot{x}_1 &=& x_2 \\
\dot{x}_2 &=& 4x_1 + u \\
y &=& -2x_1 + x_2
\end{array}
$$

**Exercise 110** Consider the system with input $u$, output $y$, and state variables $x_1, x_2$ described by

$$
\begin{aligned}
\dot{x}_1 &= -3x_1 + 2x_2 + u \\
\dot{x}_2 &= -4x_1 + 3x_2 + 2u \\
y &= -2x_1 + x_2
\end{aligned}
$$

(a) Is this system observable?
(b) If the system is unobservable, determine the unobservable eigenvalues.

**Exercise 111** Determine (by hand) whether or not the following system is observable.

$$
\begin{aligned}
\dot{x}_1 &= 5x_1 - x_2 - 2x_3 \\
\dot{x}_2 &= x_1 + 3x_2 - 2x_3 \\
\dot{x}_3 &= -x_1 - x_2 + 4x_3 \\
y_1 &= x_1 + x_2 \\
y_2 &= x_2 + x_3
\end{aligned}
$$

If the system is unobservable, compute the unobservable eigenvalues.

**Exercise 112** Determine the unobservable eigenvalues for each of the systems of Exercise 109.

**Exercise 113** For each system in Exercise 109 which is not observable, obtain a reduced order system which is observable and has the same transfer function as the original system.

**Exercise 114** (Damped linear oscillator.) Recall

$$
\begin{aligned}
\dot{x}_1 &= x_2 \\
\dot{x}_2 &= -(k/m)\,x_1 - (d/m)\,x_2
\end{aligned}
$$

Show the following:

(a) (Position measurement.) If $y = x_1$, we have observability.

(b) (Velocity measurement.) If $y = x_2$, we have observability if and only if $k \neq 0$.

(c) (Acceleration measurement.) If $y = \dot{x}_2$, we have observability if and only if $k \neq 0$.

**Exercise 115** (BB in laundromat) Obtain a state space representation of the following system.

$$\begin{aligned} m\ddot{q}_1 - m\Omega^2 q_1 + k(q_1 - q_2) &= 0 \\ m\ddot{q}_2 - m\Omega^2 q_2 - k(q_1 - q_2) &= 0 \\ y &= q_1 \end{aligned}$$

Determine whether or not your state space representation is observable.

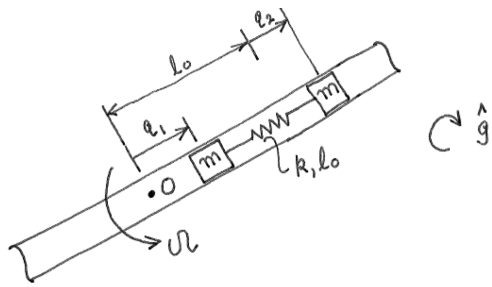



Figure 13.3: Beavis and Butthead in the laundromat

**Exercise 116** (BB in laundromat: mass center observations.) Obtain a state space representation of the following system.

$$\begin{aligned} m\ddot{q}_1 - m\Omega^2 q_1 + k(q_1 - q_2) &= 0 \\ m\ddot{q}_2 - m\Omega^2 q_2 - k(q_1 - q_2) &= 0 \\ y &= \tfrac{1}{2}(q_1 + q_2) \end{aligned}$$

(a) Obtain a basis for its unobservable subspace.

(b) Determine the unobservable eigenvalues. Consider $\omega := \sqrt{k/2m} > \Omega$.

**Exercise 117** (BB in laundromat: internal observations.) Obtain a state representation of the following system.

$$\begin{aligned} m\ddot{q}_1 - m\Omega^2 q_1 + k(q_1 - q_2) &= 0 \\ m\ddot{q}_2 - m\Omega^2 q_2 - k(q_1 - q_2) &= 0 \\ y &= q_2 - q_1 \end{aligned}$$

(a) Obtain a basis for its unobservable subspace.

(b) By using a nice transformation, obtain a reduced order observable system which has the same transfer function as the original system.

(c) Determine the unobservable eigenvalues. Consider $\omega := \sqrt{k/2m} > \Omega$.

**Exercise 118** Consider a system described by

$$
\begin{aligned}
\dot{x}_1 &= \lambda_1 x_1 + b_1 u \\
\dot{x}_2 &= \lambda_2 x_2 + b_2 u \\
&\vdots \\
\dot{x}_n &= \lambda_n x_n + b_n u \\
y &= c_1 x_1 + c_2 x_2 + \cdots + c_n x_n
\end{aligned}
$$

where all quantities are scalar. Obtain conditions on the numbers $\lambda_1, \cdots, \lambda_n$ and $c_1, \cdots, c_n$ which are necessary and sufficient for the observability of this system.

**Exercise 119** Compute the observability grammian for the damped linear oscillator with position measurement and $d, k > 0$.

**Exercise 120** Carry out the following for linearizations L1-L8 of the two pendulum cart system.

(a) Determine which linearizations are observable?

(b) Compute the singular values of the observability matrix.

(b) Determine the unobservable eigenvalues for the unobservable linearizations.

# Chapter 14

# Controllability

## 14.1 Controllability

Consider a system in which $x(t)$ is the system state at time $t$ and $u(t)$ is the control input at time $t$. Consider now any non-trivial time interval $[0, T]$; by non-trivial we mean that $T > 0$. We say that the system is controllable over the interval if it can be 'driven' from any one state to any other state over the time interval by appropriate choice of control input over the interval. A formal definition follows the next two examples.

**Example 177** (Two unattached masses)

$$\begin{aligned} \dot{x}_1 &= u \\ \dot{x}_2 &= 0 \end{aligned}$$

Since the behavior of $x_2$ is completely unaffected by the control input $u$, this system is clearly uncontrollable over any time interval.

**Example 178** (Integrator) Consider the scalar system described by

$$\dot{x} = u$$

Consider any initial state $x_0$ and any desired final state $x_f$. For any any $T > 0$, consider the constant input:

$$u(t) = (x_f - x_0)/T \qquad \text{for} \qquad 0 \le t \le T$$

Then the state trajectory $x(\cdot)$ of this system with $x(0) = x_0$ satisfies $x(T) = x_f$. We consider this system to be controllable over any time interval $[0, T]$.

**Defn.**[Controllability] *A system is* controllable *over a non-trivial time interval* $[0, T]$ *if, for every pair of states* $x_0$ *and* $x_f$*, there is a continuous control function* $u(\cdot)$ *defined on the interval such that the solution* $x(\cdot)$ *of the system with* $x(0) = x_0$ *satisfies* $x(T) = x_f$*.*

Note that the definition says nothing about keeping the state at $x_f$ after $T$.

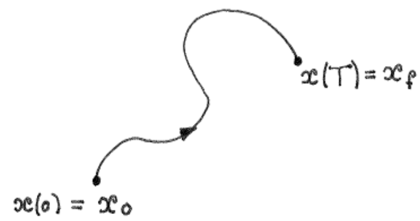



Figure 14.1: Controllability

## 14.2 Main controllability result

The concept of controllability also applies to linear and nonlinear systems. However, we are concerned here with linear time-invariant systems described by

$$\dot{x} = Ax + Bu \tag{14.1}$$

where the $n$-vector $x(t)$ is the system state at time $t$ and the $m$-vector $u(t)$ is the control input at time $t$. In this section, we introduce a simple algebraic condition which is necessary and sufficient for the controllability of a linear time invariant system over *any* time interval. This condition involves the controllability matrix $Q_c$ associated with the above system or the pair $(A,\ B)$. This matrix is defined to be the following $n \times nm$ partitioned matrix:

$$\boxed{Q_c = \begin{pmatrix} B & AB & \cdots & A^{n-1}B \end{pmatrix}} \tag{14.2}$$

**Theorem 21 (Main controllability result)** *For any $T > 0$, system (14.1) is controllable over the interval $[0, T]$ if and only if its controllability has maximum rank, that is,*

$$\boxed{rank \begin{pmatrix} B & AB & \cdots & A^{n-1}B \end{pmatrix} = n}$$

Proof. A proof is given at the end of Section 14.3.

An immediate consequence of the above theorem is that controllability does not depend on the interval $[0, T]$. So, from now on we drop reference to the interval. Since controllability only depends on the matrix pair $(A, B)$, we say that this pair is controllable if system (14.1) is controllable.

For scalar input ($m = 1$) systems, the controllability matrix is a square matrix; hence it has maximum rank if and only if its determinant is nonzero. So, the above theorem has the following corollary.

**Corollary 2** *A scalar input system of the form (14.1) is controllable if and only if its controllability matrix has non-zero determinant.*

**Example 179** (Two unattached masses.) This is a scalar input system with

$$A = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \qquad \text{and} \qquad B = \begin{pmatrix} 1 \\ 0 \end{pmatrix} .$$

Hence,

$$Q_c = \begin{pmatrix} B & AB \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

Clearly $Q_c$ has zero determinant, hence this system is not controllable.

**Example 180**

$$\begin{aligned} \dot{x}_1 &= x_3 + u_1 \\ \dot{x}_2 &= x_3 + u_1 \\ \dot{x}_3 &= u_2 \end{aligned}$$

Since there is a control input entering every state equation, this system looks controllable. Is it?

Here $n = 3$ and

$$A = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \qquad B = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$$

So,

$$Q_c = \begin{pmatrix} B & AB & A^2B \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix}$$

and

$$\text{rank } (Q_c) \;=\; \text{rank} \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} = 2$$

Since $\operatorname{rank} Q_c = 2$, we have $\operatorname{rank} Q_c \neq n$; hence this system is not controllable.

**MATLAB**

```
>> help ctrb

 CTRB   Form controllability matrix.
        CTRB(A,B) returns the controllability matrix
        Co = [B AB A^2B ...]

>> A=[0 0 1; 0 0 1; 0 0 0];
>> B=[1 0; 1 0; 0 1];

>> ctrb(A,B)
ans =
     1     0     0     1     0     0
     1     0     0     1     0     0
     0     1     0     0     0     0
```

**Useful fact for computing rank $Q_c$.** In computing rank $Q_c$, the following fact is useful.

Suppose there exists a positive integer $k$ such that

$$\text{rank}\begin{pmatrix} B & AB & \dots & A^kB \end{pmatrix} = \text{rank}\begin{pmatrix} B & AB & \dots & A^{k-1}B \end{pmatrix}$$

Then

$$\text{rank}\, Q_c = \text{rank}\begin{pmatrix} B & AB & \dots & A^{k-1}B \end{pmatrix}$$

The next example illustrates the above fact.

**Example 181 (Beavis and Butthead: self-controlled)** Consider the situation in which the system generates a pair of control forces which are equal in magnitude and opposite in direction.

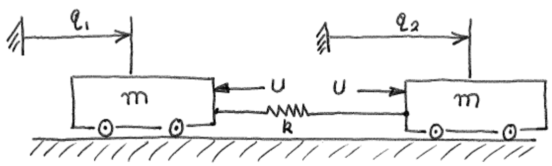



Figure 14.2: B&B: self-controlled

The motion of this system can be described by

$$\begin{aligned} m\ddot{q}_1 &= k(q_2 - q_1) - u \\ m\ddot{q}_2 &= -k(q_2 - q_1) + u \end{aligned}$$

Based on simple mechanical considerations, we can see that this system is not controllable.

Now lets do the controllability test. With

$$x_1 = q_1 \qquad x_2 = q_2 \qquad x_3 = \dot{q}_1 \qquad x_4 = \dot{q}_2$$

we have

$$A = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -k/m & k/m & 0 & 0 \\ k/m & -k/m & 0 & 0 \end{pmatrix} \qquad\qquad B = \begin{pmatrix} 0 \\ 0 \\ -1/m \\ 1/m \end{pmatrix}$$

Here

$$\begin{pmatrix} B & AB \end{pmatrix} = \begin{pmatrix} 0 & -1/m \\ 0 & 1/m \\ -1/m & 0 \\ 1/m & 0 \end{pmatrix}$$

and

$$\begin{pmatrix} B & AB & A^2B \end{pmatrix} = \begin{pmatrix} 0 & -1/m & 0 \\ 0 & 1/m & 0 \\ -1/m & 0 & 2k/m^2 \\ 1/m & 0 & -2k/m^2 \end{pmatrix}$$

Clearly,

$$\operatorname{rank}\begin{pmatrix} B & AB & A^2B \end{pmatrix} = \operatorname{rank}\begin{pmatrix} B & AB \end{pmatrix}$$

Hence

$$\operatorname{rank} Q_c = \operatorname{rank}\begin{pmatrix} B & AB \end{pmatrix} = 2$$

MATLAB time.

```
>> a = [0      0      1      0
     0      0      0      1
    -1      1      0      0
     1     -1      0      0];
>> b = [0; 0; -1; 1];

>> rank(ctrb(a,b))
ans = 2
```

**Example 182 (Beavis and Butthead with external control)** Here we consider the system of the previous example and replace the pair of internal forces with a single external control force $u$. This system is described by

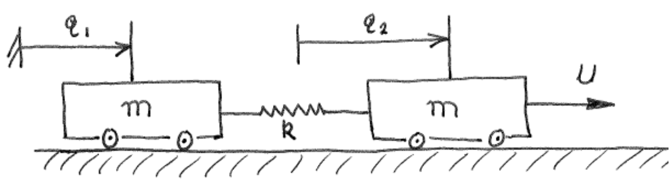



Figure 14.3: B&B: external control

$$\begin{aligned} m\ddot{q}_1 &= k(q_2 - q_1) \\ m\ddot{q}_2 &= -k(q_2 - q_1) \quad +u \end{aligned}$$

With

$$x_1 = q_1 \qquad x_2 = q_2 \qquad x_3 = \dot{q}_1 \qquad x_4 = \dot{q}_2$$

we have

$$A = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -k/m & k/m & 0 & 0 \\ k/m & -k/m & 0 & 0 \end{pmatrix} \qquad\qquad B = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1/m \end{pmatrix}$$

Hence,

$$Q_c = \begin{pmatrix} B & AB & A^2B & A^3B \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & k/m^2 \\ 0 & 1/m & 0 & -k/m^2 \\ 0 & 0 & k/m^2 & 0 \\ 1/m & 0 & -k/m^2 & 0 \end{pmatrix}$$

Clearly, $\operatorname{rank} Q_c = 4 = n$; hence this system is controllable.

**Exercise 121** Suppose $A$ is $n \times n$, $B$ is $n \times m$, $T$ is $n \times n$ and nonsingular and let

$$\tilde{A} = T^{-1}AT \qquad \tilde{B} = T^{-1}B$$

Then $(A, B)$ is controllable if and only if $(\tilde{A}, \tilde{B})$ is controllable.

## 14.3 The controllable subspace

Let us now consider the situation when the system $\dot{x} = Ax + Bu$ is not controllable.

**Example 183** At first sight the system

$$\begin{aligned}\dot{x}_1 &= x_1 + u \\ \dot{x}_2 &= x_2 + u\end{aligned}$$

might look controllable. However, it is not. To see this, consider the following change of state variables:

$$x_c = x_1 + x_2 \qquad \text{and} \qquad x_u = x_1 - x_2$$

These variables satisfy

$$\begin{aligned}\dot{x}_c &= x_c + 2u \\ \dot{x}_u &= x_u\end{aligned}$$

Hence, $x_u(0) = 0$ implies $x_u(t) = 0$ for all $t$. In other words if the initial $x(0)$ is in the set

$$\{(x_1, x_2) \in \mathbb{R}^2 \,:\, x_1 - x_2 = 0\}$$

then, regardless of the control input, the state will always stay in this set; it cannot leave this set. So, the system is not controllable.

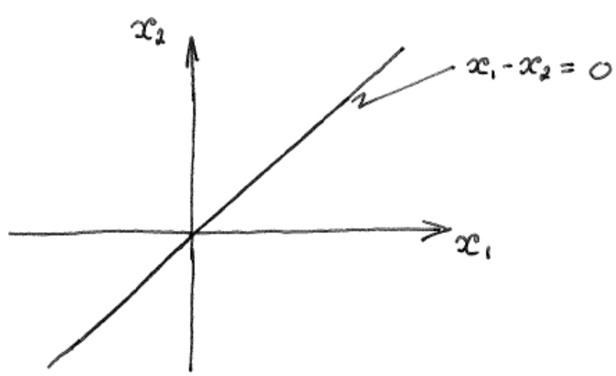



Figure 14.4: Controllable subspace in Example 183

Recall that the controllability matrix associated with the pair $(A, B)$ is defined to be the following $n \times nm$ matrix

$$Q_c = \begin{pmatrix} B & AB & \dots & A^{n-1}B \end{pmatrix}$$

*The* controllable subspace *is defined as the range of the controllability matrix, that is,*

$$\mathcal{R}(Q_c)$$

So, according to the Main Controllability Theorem (Theorem 21), the system $\dot{x} = Ax+Bu$ is controllable if and only if its controllable subspace is the whole state space. It should be clear that a vector $x$ is in the controllable subspace associated with $(A, B)$ if and only if it can be expressed as

$$x = \sum_{i=0}^{n-1} A^i B \nu^i$$

where $\nu^0, \dots, \nu^{n-1}$ are $m$-vectors.

The following lemma states that when the state of system (14.1) starts at zero, then regardless of the control input, the state always stays in the controllable subspace; that is, it cannot leave this subspace.

**Lemma 15** *If $x(0) = 0$ then, for every continuous control input function $u(\cdot)$, the solution $x(\cdot)$ of (14.1) lies in the controllable subspace, that is, $x(t)$ is in the range of $Q_c$ for all $t \geq 0$.*

PROOF. For any time $t > 0$ we have

$$x(t) = \int_0^t e^{A(t-\tau)} B u(\tau)\, d\tau$$

Recalling the application of the Cayley-Hamilton theorem to functions of a matrix, it follows that for $0 \leq \tau \leq t$, there are scalars $\beta_0(\tau), \dots, \beta_{n-1}(\tau)$ such that

$$e^{A(t-\tau)} = \sum_{i=0}^{n-1} \beta_i(\tau) A^i$$

Hence,

$$e^{A(t-\tau)} B u(\tau) = \sum_{i=0}^{n-1} A^i B \nu_i(\tau) \qquad \nu_i(\tau) := \beta_i(\tau) u(\tau)$$

From this it should be clear that for each $\tau$, the $n$ vector $e^{A(t-\tau)} B u(\tau)$ is in $\mathcal{R}(Q_c)$, the range of $Q_c$. From this one obtains that $x(t) = \int_0^t e^{A(t-\tau)} B u(\tau)\, d\tau$ is in $\mathcal{R}(Q_c)$. ■

**Example 184** For the system of example 183, we have

$$Q_c = \begin{pmatrix} B & AB \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

Clearly,

$$\mathcal{R}(Q_c) = \mathcal{R}\left(\left(\begin{array}{c} 1 \\ 1 \end{array}\right)\right) = \mathcal{R}(B)$$

If a vector $x$ is in this controllable subspace, then $x = c\,[\ 1 \quad 1\ ]^T$ for some scalar $c$. This is equivalent to the requirement that $x_1 - x_2 = 0$.

Actually, if the state starts at any point in $\mathcal{R}(Q_c)$ then, regardless of the control input, the state always stays in $\mathcal{R}(Q_c)$; see next exercise. In other words, *the control input cannot drive the state out of the controllable subspace.*

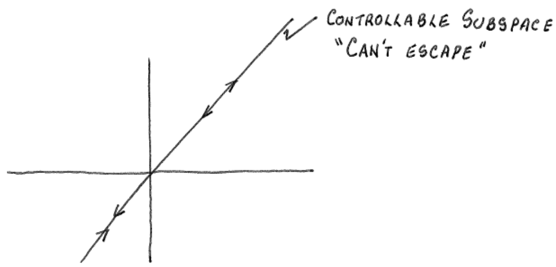



Figure 14.5: Invariance of the controllable subspace

**Exercise 122** Prove the following statement. If $x(0)$ is in $\mathcal{R}(Q_c)$, then for every continuous control input function $u(\cdot)$, the solution of (14.1) satisfies $x(t) \in \mathcal{R}(Q_c)$ for all $t \geq 0$.

The next result states that over any time interval, system (14.1) can be driven between any two states in the controllable subspace $\mathcal{R}(Q_c)$ by appropriate choice of control input.

**Lemma 16** *[Controllable subspace lemma] Consider any pair of states $x_0$, $x_f$ in the controllable subspace $\mathcal{R}(Q_c)$. Then for each $T > 0$, there is a continuous control function $u(\cdot) : [0, T] \to \mathbb{R}^m$ such that the solution $x(\cdot)$ of (14.1) with $x(0) = x_0$ satisfies $x(T) = x_f$*

PROOF. A proof is contained in section 14.4.

**Example 185**

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

Here,

$$A = \left(\begin{array}{cc} -1 & 1 \\ -1 & 1 \end{array}\right) \qquad B = \left(\begin{array}{c} 1 \\ 1 \end{array}\right)$$

Hence

$$Q_c = \left(\begin{array}{cc} B & AB \end{array}\right) = \left(\begin{array}{cc} 1 & 0 \\ 1 & 0 \end{array}\right)$$

and

$$\mathcal{R}(Q_c) = \mathcal{R}\left(\left(\begin{array}{c} 1 \\ 1 \end{array}\right)\right)$$

To illustrate the controllable subspace lemma, introduce new states

$$x_c = x_1 \qquad x_u = -x_1 + x_2$$

Then

$$x = x_c \left(\begin{array}{c} 1 \\ 1 \end{array}\right) + x_u \left(\begin{array}{c} 0 \\ 1 \end{array}\right)$$

and

$$\begin{aligned} \dot{x}_c &= x_u + u \\ \dot{x}_u &= 0 \end{aligned}$$

It should be clear that $x$ is in the controllable subspace if and only if $x_u = 0$. In this case $x$ is completely specified by the scalar $x_c$ and

$$\dot{x}_c = u$$

Recalling example 178, we see that given any $x_{c0}$, $x_{cf}$ and any $T > 0$, one can always find a control input which 'drives' $x_c$ from $x_{c0}$ to $x_{cf}$ over the interval $[0, T]$. Hence, the state $x$ can be driven between any two points in the controllable subspace.

**Proof of main controllability result.** First note that $\operatorname{rank} Q_c = n$ if and only if $\mathcal{R}(Q_c) = \mathbb{R}^n$, that is, the controllability matrix has rank $n$ if and only if the controllable subspace is the whole state space. If $\operatorname{rank} Q_c = n$, then Lemma 16 implies controllability. If $\operatorname{rank} Q_c < n$, then the controllable subspace is not the whole state space and from Lemma 15 any state which is not in the controllable subspace cannot be reached from the origin. ■

## 14.4 Proof of the controllable subspace lemma*

### 14.4.1 Finite time controllability grammian

Before proving the controllable subspace lemma, we first establish a preliminary lemma. For each $T > 0$, we define the finite time controllability grammian associated with $(A, B)$:

$$W_c(T) := \int_0^T e^{At} BB^* e^{A^*t} \, dt \tag{14.3}$$

Note that this matrix is hermitian and positive semi-definite.

**Remark 7** Note that $W_c(T)$ is the solution at $t = T$ of the following initial value problem:

$$\dot{W}_c = AW_c + W_c A^* + BB^* \qquad \text{and} \qquad W_c(0) = 0$$

PROOF.

$$W_c(t) = \int_0^t e^{A\tau} BB^* e^{A^*\tau} \, d\tau = \int_0^t e^{A(t-s)} BB^* e^{A^*(t-s)} \, ds$$

Hence, $W_c(0) = 0$ and

$$\dot{W}_c = AW_c + W_c A^* + BB^*$$

■

**Lemma 17** *For each $T > 0$,*

$$\mathcal{R}(W_c(T)) = \mathcal{R}(Q_c)$$

PROOF. We will show that $\mathcal{N}(Q_c^*) = \mathcal{N}(W_c(T))$. From this it will follow that $\mathcal{R}(Q_c) = \mathcal{R}(W_c(T)^*) = \mathcal{R}(W_c(T))$.

Since

$$Q_c^* = \begin{pmatrix} B^* \\ B^*A^* \\ \vdots \\ B^*A^{*(n-1)} \end{pmatrix}$$

it follows that $x$ is in $\mathcal{N}(Q_c^*)$ if and only if

$$B^* A^{*k} x = 0 \qquad \text{for} \quad k = 0, 1, \ldots, n-1 \tag{14.4}$$

We first show that $\mathcal{N}(Q_c^*) \subset \mathcal{N}(W_c(T))$: Consider any $x \in \mathcal{N}(Q_c^*)$. Then (14.4) holds. Recalling the application of the Cayley-Hamilton theorem to functions of a matrix, it follows that for $0 \le t \le T$, there are scalars $\beta_0(t), \ldots, \beta_{n-1}(t)$ such that

$$e^{A^*t} = \sum_{i=0}^{n-1} \beta_i(t) A^{*i}$$

Hence,

$$B^* e^{A^*t} x = \sum_{i=0}^{n-1} \beta_i(t) B^* A^{*i} x = 0$$

and

$$\begin{aligned} W_c(T)x &= \int_0^T e^{At} BB^* e^{A^*t} \, dt \, x \\ &= \int_0^T e^{At} BB^* e^{A^*t} x \, dt \\ &= 0 \end{aligned}$$

that is, $x \in \mathcal{N}(W_c(T))$. Since the above holds for any $x \in \mathcal{N}(Q_c^*)$, we must have $\mathcal{N}(Q_c^*) \subset \mathcal{N}(W_c(T))$.

We now show that $\mathcal{N}(W_c(T)) \subset \mathcal{N}(Q_c^*)$ and hence $\mathcal{N}(Q_c^*) = \mathcal{N}(W_c(T))$: Consider any $x \in \mathcal{N}(W_c(T))$. Then,

$$\begin{aligned} 0 &= x^* W_c(T) x \\ &= x^* \left( \int_0^T e^{At} BB^* e^{A^*t} \, dt \right) x \\ &= \int_0^T x^* e^{At} BB^* e^{A^*t} x \, dt \\ &= \int_0^T \left(B^* e^{A^*t} x\right)^* \left(B^* e^{A^*t} x\right) \, dt \\ &= \int_0^T ||B^* e^{A^*t} x||^2 \, dt \end{aligned}$$

Since the integrand is non-negative for all $t$ and the integral is zero, the integrand must be zero for all $t$, that is,

$$B^* e^{A^*t} x \equiv 0$$

If a function is zero for all $t$, all its derivatives must be zero for all $t$; hence

$$B^* A^{*k} e^{A^*t} x \equiv 0 \qquad \text{for} \quad k = 0, 1, \ldots, n-1$$

Considering $t = 0$ yields

$$B^* A^{*k} x = 0 \qquad \text{for} \quad k = 0, 1, \ldots, n-1$$

that is $x \in \mathcal{N}(Q_c^*)$. Since the above holds for any $x \in \mathcal{N}(W_c(T))$, we must have $\mathcal{N}(W_c(T)) \subset \mathcal{N}(Q_c^*)$. ■

From our main controllability result and the above lemma, we see that $(A, B)$ is controllable if and only if , for any $T > 0$, its finite time controllability grammian $W_c(T)$ has full rank. Since $W_c(T)$ is square, it has full rank if and only if it is invertible. Moreover, since $W_c(T)$ is symmetric and positive semidefinite, it has full rank if and only if it is positive definite. This yields the following result.

**Corollary 2** *The system $\dot{x} = Ax + Bu$ is controllable if and only if its controllability grammian over any non-trivial interval is positive definite.*

### 14.4.2 Proof of the controllable subspace lemma

Consider any $T > 0$ and any pair of states $x_0, x_f$ in $\mathcal{R}(Q_c)$. The solution of (14.1) at time $T$ is:

$$x(T) = e^{AT}x_0 + \int_0^T e^{A(T-\tau)}Bu(\tau)\,d\tau$$

Hence, a control input drives the state from $x_0$ to $x_f$ if and only if

$$\int_0^T e^{A(T-\tau)}Bu(\tau)\,d\tau \;=\; z := x_f - e^{AT}x_0$$

We first show that

$$z \in \mathcal{R}(Q_c)$$

Since $x_0 \in \mathcal{R}(Q_c)$,

$$x_0 = \sum_{i=0}^{n-1} A^i B \nu^i$$

hence

$$e^{AT}x_0 = \sum_{i=0}^{n-1} e^{AT}A^i B \nu^i$$

By application of Cayley Hamilton to functions of a matrix

$$e^{AT}A^i = \sum_{k=0}^{n-1} \beta_{ik}A^k$$

so,

$$e^{AT}x_0 = \sum_{i=0}^{n-1}\sum_{k=0}^{n-1} \beta_{ik}A^k B \nu^i = \sum_{k=0}^{n-1} A^k B \tilde{\nu}^k$$

where $\tilde{\nu}^k := \sum_{i=0}^{n-1} \beta_{ik}\nu^i$. From this it should be clear that $e^{AT}x_0$ is in $\mathcal{R}(Q_c)$. Since $x_f$ is also in $\mathcal{R}(Q_c)$ we have $z \in \mathcal{R}(Q_c)$.

Since $z$ is in $\mathcal{R}(Q_c)$, it follows from the previous lemma that $z$ is also in $\mathcal{R}(W_c(T))$. Hence there exists an $n$-vector $y$ such that

$$z = W_c(T)y$$

Considering the control input

$$u(\tau) = B^* e^{A^*(T-\tau)}y$$

we obtain

$$
\begin{aligned}
\int_0^T e^{A(T-\tau)}Bu(\tau)\,d\tau &= \int_0^T e^{A(T-\tau)}BB^*e^{A^*(T-\tau)}y\,d\tau \\
&= \int_0^T e^{A(T-\tau)}BB^*e^{A^*(T-\tau)}\,d\tau\,y \\
&= \int_0^T e^{At}BB^*e^{A^*t}\,dt\,y \\
&= W_c(T)y \\
&= z
\end{aligned}
$$

hence, this control input drives the state from $x_0$ to $x_f$. ■

**Remark 8** Suppose $(A, B)$ is controllable. Then $\operatorname{rank} W_c(T) = \operatorname{rank} Q_c = n$. Hence $W_c(T)$ is invertible and the control which drives $x_0$ to $x_f$ over the interval $[0, T]$ is given by

$$
\boxed{u(t) = B^*e^{A^*(T-t)}\tilde{x} \qquad \tilde{x} = W_c(T)^{-1}(x_f - e^{AT}x_0)}
$$

# 14.5 Uncontrollable modes

We shall see that if a system is uncontrollable, it has modes which cannot be affected by the input. To define exactly what we mean by an uncontrollable mode, we need to look at the left eigenvectors of the system $A$ matrix.

## 14.5.1 Eigenvectors revisited: left eigenvectors

Suppose that $\lambda$ is an eigenvalue of a square matrix $A$. Then an eigenvector corresponding to $\lambda$ is any nonzero vector satisfying $Av = \lambda v$. We say that *a nonzero vector $w$ is a* left eigenvector *of $A$ corresponding to $\lambda$ if it satisfies*

$$
\boxed{w'A = \lambda w'} \tag{14.5}
$$

To see that left eigenvectors exist, note that the above equation is equivalent to $A'w = \bar{\lambda}w$, that is, $\bar{\lambda}$ is an eigenvalue of $A'$ and $w$ is an eigenvector of $A'$ corresponding to $\bar{\lambda}$. We now recall that $\lambda$ is an eigenvalue of $A$ if and only if $\bar{\lambda}$ is an eigenvalue of $A'$ . Thus, we can say that $\lambda$ is an eigenvalue of $A$ if and only if it satisfies (14.5) for some non-zero vector $w$. Moreover, we can also say that every eigenvalue $\lambda$ of $A$ has left eigenvectors and these vectors are simply the eigenvectors of $A'$ corresponding to $\bar{\lambda}$.

Consider now any two eigenvectors $\lambda_1$ and $\lambda_2$ of $A$ with $\lambda_1 \neq \lambda_2$. Suppose that $v_1$ is an eigenvector for $\lambda_1$ and $w_2$ is a left eigenvector for $\lambda_2$. Then

$$
w_2'v_1 = 0\,.
$$

To see this, we first use the eigenvector definitions to obtain

$$Av_1 = \lambda_1 v_1 \qquad \text{and} \qquad w_2' A = \lambda_2 w_2' \,.$$

Pre-multiplying the first equation by $w_2'$ we obtain that $w_2' A v_1 = \lambda_1 w_2' v_1$. Post-multiplying the section equation by $v_1$ we obtain that $w_2' A v_1 = \lambda_2 w_2' v_1$. This implies that $\lambda_1 w_2' v_1 = \lambda_2 w_2' v_1$, that is $(\lambda_1 - \lambda_2) w_2' v_1 = 0$. Since $\lambda_1 - \lambda_2 \neq 0$, we must have $w_2' v_1 = 0$.

**System significance of left eigenvectors.** Consider now a system described by

$$\dot{x} = Ax \,. \tag{14.6}$$

Recall that $\lambda$ is an eigenvalue of $A$ if and only if the above system has a solution of the form $x(t) = e^{\lambda t} v$ where $v$ is a nonzero constant vector. When this occurs, the vector $v$ can be any eigenvector of $A$ corresponding to $\lambda$ and the above special solution is described as a mode of the system. We now demonstrate the following result:

*A scalar $\lambda$ is an eigenvalue of $A$ if and only if there is a nonzero vector $w$ such that* every *solution $x(\cdot)$ of system (14.6) satisfies*

$$\boxed{w' x(t) = e^{\lambda t} w' x(0)} \tag{14.7}$$

*When this occurs, the vector $w$ can be any left eigenvector of $A$ corresponding to $\lambda$.*

In particular, if $x(0) = w/||w||^2$ then, $w'x(t) = e^{\lambda t}$. If $w'x(0) = 0$ then $w'x(t) = 0$ for all $t$.

To gain further insight into the above result, introduce the scalar variable $\xi(t) = w'x(t)$. Then $\xi(0) = w'x(0)$ and (14.7) is equivalent to

$$\xi(t) = e^{\lambda t} \xi(0) \,,$$

that is, the behavior of $\xi$ is governed by

$$\dot{\xi} = \lambda \xi \,.$$

To demonstrate the above result, we first suppose that $\lambda$ is an eigenvalue of $A$. Then there is a nonzero vector $w$ such that (14.5) holds; also $w$ is a left eigenvector of $A$ corresponding to $\lambda$. Consider now any solution $x(\cdot)$ of system (14.6). Since it satisfies $\dot{x} = Ax$, it follows from (14.5) that

$$\frac{d(w'x)}{dt} = w'\dot{x} = w'Ax = \lambda (w'x)$$

It now follows that (14.7) holds for all $t$.

To demonstrate the converse, suppose that there is a nonzero vector $w$ so that (14.7) holds for every solution $x(\cdot)$. This implies that

$$w'Ax(t) = w'\dot{x}(t) = \frac{d(w'x)}{dt} = \lambda e^{\lambda t} w' x(0) = \lambda w' x(t) \,;$$

hence,

$$w'(A - \lambda I) x(t) = 0 \,.$$

Since the above holds for any $x(t)$, we must have $w'(A - \lambda I) = 0$. Thus, $w'A = \lambda w'$. This implies that $\lambda$ is an eigenvalue of $A$ with left eigenvector $w$.

### 14.5.2 Uncontrollable eigenvalues and modes

Consider now a system with input $u$ described by

$$\dot{x} = Ax + Bu \tag{14.8}$$

We say that $\lambda$ *is an* uncontrollable eigenvalue *of this system or* $\lambda$ *is an uncontrollable eigenvalue of* $(A, B)$ *if there is a nonzero vector* $w$ *such that for* every *input* $u(\cdot)$, every *solution* $x(\cdot)$ *of the system satisfies*

$$\boxed{w'x(t) = e^{\lambda t} w'x(0)} \tag{14.9}$$

To gain further insight into the above concept, introduce the scalar variable $\xi(t) = w'x(t)$. Then, as we have shown above, (14.9) is equivalent to the statement that the behavior of $\xi$ is governed by

$$\dot{\xi} = \lambda \xi\,.$$

Thus the behavior of $\xi$ is completely unaffected by the input $u$. So, the mode $e^{\lambda t}$ is completely unaffected by the input $u$. We say that this is an uncontrollable mode.

*When system (14.8) has an uncontrollable mode, then the system is not controllable.* To see this choose any initial state $x(0) = x_0$ such that $w'x_0 \neq 0$. Then, regardless of the control input, we have $w'x(t) = e^{\lambda t} w'x_0 \neq 0$ for all $t$. Hence, regardless of the control input, $x(t) \neq 0$ for all $t$. This means that over any time interval $[0, T]$ we cannot obtain a control history which drives the system from $x_0$ to 0. Hence the system is not controllable. In Section 14.5.3, we demonstrate the following result.

*If a system is not controllable then, it must have some uncontrollable modes.*

### 14.5.3 Existence of uncontrollable modes*

In this section, we show that if a LTI system is not controllable, then it has uncontrollable modes. Consider a linear time-invariant input-output system described

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y &= Cx + Du \end{aligned} \tag{14.10}$$

where $x(t)$ is an $n$-vector, $u(t)$ is an $m$-vector, and $y(t)$ is a $p$-vector. Suppose $(A, B)$ is not controllable; then, its controllable subspace $\mathcal{X}_c$ is not equal to the whole state space. In the following section, we show that there is a state transformation $x = T\tilde{x}$ so that the corresponding transformed system is described by

$$\begin{aligned} \dot{x}_c &= A_{cc}x_c + A_{cu}x_u + B_c u \\ \dot{x}_u &= A_{uu}x_u \\ y &= C_c x_c + C_u x_u + Du \end{aligned} \tag{14.11}$$

with

$$\begin{pmatrix} x_c \\ x_u \end{pmatrix} = \tilde{x}\,.$$

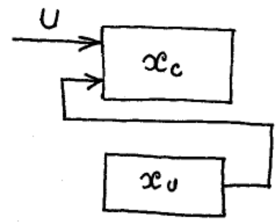



Figure 14.6: Controllable/uncontrollable decomposition

Also, the pair $(A_{cc}, B_c)$ is controllable.

Clearly, the state $x_u$ of the subsystem

$$\dot{x}_u = A_{uu} x_u \tag{14.12}$$

is completely unaffected by the input $u$. Also, if $x_u(0) = 0$ then $x_u(t) = 0$ for all $t$ and the input-output system in (14.11) can be described by the controllable subsystem:

$$\begin{aligned} \dot{x}_c &= A_{cc} x_c + B_c u \\ y &= C_c x_c + Du \end{aligned} \tag{14.13}$$

Hence, when the initial state $x(0)$ of the input-output system (14.10) is zero, its behavior can be described by the lower order system (14.13).

Since the matrices $A$ and

$$\tilde{A} = \begin{pmatrix} A_{cc} & A_{cu} \\ 0 & A_{uu} \end{pmatrix}$$

are similar, they have the same characteristic polynomial. In particular, the characteristic polynomial of $\tilde{A}$ is given by

$$\det(sI - \tilde{A}) = \det(sI - A_{cc}) \det(sI - A_{uu})$$

Hence the characteristic polynomial of $A$ is the product of the characteristic polynomials of $A_{cc}$ and $A_{uu}$. Thus, the eigenvalues of $A$ are the union of those of $A_{cc}$ and $A_{uu}$. We now make the following claim:

*The eigenvalues of $A_{uu}$ are the uncontrollable eigenvalues of $A$.*

Proof. To prove this claim, let $\lambda$ be any eigenvalue of $A_{uu}$. We will show that $\lambda$ is an uncontrollable eigenvalue of $(A, B)$. Since $\lambda$ is an eigenvalue of $A_{uu}$ there is a nonzero vector $w_u$ such that for every solution $x_u(\cdot)$ of (14.12), we have

$$w_u' x_u(t) = e^{\lambda t} w_u' x_u(0) \,.$$

Let $\tilde{w}' = [0 \quad w_u']$. Then every solution $\tilde{x}$ of (14.11) satisfies

$$\tilde{w}' \tilde{x}(t) = e^{\lambda t} \tilde{w}' \tilde{x}(0)$$

With $w = T\tilde{w}$, we now obtain that every solution of (14.10) satisfies $w' x(t) = e^{\lambda t} w' x(0)$. Thus, $\lambda$ is an uncontrollable eigenvalue of $A$.

We now prove the converse, that is, if $\lambda$ is an uncontrollable eigenvalue of $(A, B)$, then $\lambda$ must be an eigenvalue of $A_{uu}$. So, suppose that $\lambda$ is an uncontrollable eigenvalue of $(A, B)$. Then $(A, B)$ is not controllable and there is a state transformation $x = T[x_c'\ x_u']'$ such that the system can be represented as in (14.11) with a nontrivial $x_u$ and with $(A_{cc}, B_c)$ controllable. Suppose, on the contrary, that $\lambda$ is not an eigenvalue of $A_{uu}$. Then $\lambda$ must be an eigenvalue of $A_{cc}$. Let $w$ be such that (14.9) holds and let $[w_u'\ w_c']' = T^{-1}w$. Then we must have

$$w_c'x_c(t) + w_u'x_u(t) = e^{\lambda t}[w_c'x_c(0) + w_u'x_u(0)]$$

for every input $u$ and every solution $\tilde{x}(\cdot)$ of (14.11). Considering those solutions with $x_u(0) = 0$, we obtain that $x_u(t) = 0$ for all $t$ and

$$\dot{x}_c = A_c x_c + B_c u\,. \tag{14.14}$$

Thus $w_c'x_c(t) = e^{\lambda t}w_c'x_c(0)$ for every input and every solution $x_c(\cdot)$ of (14.14); hence $\lambda$ is an uncontrollable eigenvalue of $(A_c, B_c)$. This implies that $(A_c, B_c)$ is not controllable and we have a contradiction. So, $\lambda$ must be an eigenvalue of $A_{uu}$ ■

**Transfer function considerations.** Consider the reduced order controllable subsystem:

$$\begin{aligned}\dot{x}_c &= A_{cc}x_c + B_c u\\ y &= C_c x_c + Du\end{aligned}$$

One can readily show that the transfer function of this system is exactly the same as that of the original system, that is

$$C(sI - A)^{-1}B + D = C_c(sI - A_c)^{-1}B_c + D\,.$$

An immediate consequence of this result is that a minimal state space realization of a transfer function must be controllable.

### 14.5.4 A nice transformation*

We first show that the controllability subspace $\mathcal{X}_c$ is an invariant subspace for $A$, that is, is $x$ is in $\mathcal{X}_c$, then $Ax$ is in $\mathcal{X}_c$.

**Lemma 18** *For any pair of matrices $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{n\times m}$, the controllable subspace*

$$\mathcal{X}_c = \mathcal{R}\left(\begin{pmatrix} B & AB & \cdots & A^{n-1}B \end{pmatrix}\right)$$

*is an invariant subspace for $A$.*

Proof. Consider any element $x$ of $\mathcal{X}_c$. It can be expressed as

$$x = B\nu_1 + AB\nu_2 + \cdots + A^{n-1}B\nu_n$$

Hence

$$Ax = AB\nu_1 + A^2B\nu_2 + \cdots + A^nB\nu_n$$

Since

$$A^n = -\alpha_0 I - \alpha_1 A - \cdots - \alpha_{n-1} A^{n-1}$$

we obtain

$$Ax = B(-\alpha_0 \nu_n) + AB(\nu_1 - \alpha_1 \nu_n) + \cdots + A^{n-1} B(\nu_{n-1} - \alpha_{n-1} \nu_n)$$

hence $Ax$ is contained in $\mathcal{X}_c$. Since this holds for every element $x$ of $\mathcal{X}_c$, it follows that $\mathcal{X}_c$ is $A$-invariant. ■

**A nice transformation.** Choose any basis

$$t^1, \ldots, t^{n_c}$$

for the controllable subspace $\mathcal{X}_c$ of $(A, B)$. Extend it to a basis

$$t^1, \ldots, t^{n_c}, t^{n_c+1}, \ldots, t^n$$

for $\mathbb{R}^n$ and let

$$T := \begin{pmatrix} t^1 & t^2 & \ldots & t^n \end{pmatrix}$$

One could do this by choosing $T = U$ where $USV^*$ is a singular value decomposition of the controllability matrix, that is

$$USV^* = \begin{pmatrix} B & AB & \cdots & A^{n-1}B \end{pmatrix}$$

Introduce the state transformation

$$x = T\tilde{x}$$

Let

$$\begin{pmatrix} x_c \\ x_u \end{pmatrix} = \tilde{x}$$

where $x_c$ is a $n_c$-vector and $x_u$ is a $(n–n_c)$-vector. Then

$$x \text{ is in } \mathcal{X}_c \qquad \text{if and only if} \qquad x_u = 0$$

The transformed system description is given by

$$\begin{aligned} \dot{\tilde{x}} &= \tilde{A}\tilde{x} + \tilde{B}u \\ y &= \tilde{C}\tilde{x} + Du \end{aligned}$$

where

$$\tilde{A} = T^{-1}AT \qquad \tilde{B} = T^{-1}B \qquad \tilde{C} = CT$$

Since $\mathcal{X}_c$ is an invariant subspace of $A$, the matrix $\tilde{A}$ has the following structure:

$$\tilde{A} = \begin{pmatrix} A_{cc} & A_{cu} \\ 0 & A_{uu} \end{pmatrix}$$

where $A_{cc}$ is $n_c \times n_c$. Since the columns of $B$ are contained in the $\mathcal{X}_c$, the matrix $\tilde{B}$ has the following structure

$$\tilde{B} = \begin{pmatrix} B_c \\ 0 \end{pmatrix}$$

where $B_c$ is $n_c \times m$. If we partition $\tilde{C}$ as

$$\tilde{C} = \begin{pmatrix} C_c & C_u \end{pmatrix}$$

where $C_c$ is $p \times n_c$, then the transformed system is described by

$$\boxed{\begin{array}{rcl} \dot{x}_c &=& A_{cc}x_c + A_{cu}x_u + B_c u \\ \dot{x}_u &=& A_{uu}x_u \end{array}}$$
$$y = C_c x_c + C_u x_u + Du$$

• The pair $(A_{cc}, B_c)$ is controllable.

**Example 186** Consider

$$\begin{array}{rcl} \dot{x}_1 &=& x_2 - u \\ \dot{x}_2 &=& x_1 + u \\ y &=& x_2 \end{array}$$

Here,

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad B = \begin{pmatrix} -1 \\ 1 \end{pmatrix} \qquad C = \begin{pmatrix} 0 & 1 \end{pmatrix} \qquad D = 0$$

and the system transfer function is given by

$$\hat{G}(s) = \frac{1}{s+1}$$

Since the controllability matrix

$$Q_c = \begin{pmatrix} B & AB \end{pmatrix} \quad = \quad \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}$$

has rank less than two, this system is not controllable. The controllable subspace is spanned by the vector

$$\begin{pmatrix} -1 \\ 1 \end{pmatrix} = B$$

Introducing the state transformation

$$x = T\tilde{x} \qquad \tilde{x} = \begin{pmatrix} x_c \\ x_u \end{pmatrix} \qquad T = \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}$$

we obtain an equivalent system description with

$$\tilde{A} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \tilde{B} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \tilde{C} = \begin{pmatrix} 1 & 1 \end{pmatrix} \qquad \tilde{D} = 0$$

that is

$$
\begin{aligned}
\dot{x}_c &= -x_c + u \\
\dot{x}_u &= x_u \\
y &= x_c + x_u
\end{aligned}
$$

The transfer function of the reduced order controllable system

$$
\begin{aligned}
\dot{x}_c &= -x_c + u \\
y &= x_c
\end{aligned}
$$

is clearly the same as that of the original system.

## 14.6 PBH test

### 14.6.1 PBH test for uncontrollable eigenvalues

The next lemma provides a useful characterization of uncontrollable eigenvalues.

**Lemma 19** *A complex number $\lambda$ is an uncontrollable eigenvalue of the pair $(A, B)$ if and only if*

$$\text{rank}\,(A - \lambda I \quad B) < n$$

*where $n$ is the number of rows in $A$.*

PROOF. When the above rank condition holds, there exists a non-zero $n$-vector $w$ such that

$$w'\,(A - \lambda I \quad B) = 0\,,$$

that is

$$w'A = \lambda w' \qquad \text{and} \qquad w'B = 0\,.$$

So, if $x(\cdot)$ is any solution of the system $\dot{x} = Ax + Bu$ corresponding to any input $u(\cdot)$, we obtain that

$$\frac{d(w'x)}{dt} = w'Ax + w'Bu = \lambda(w'x)\,.$$

Hence,

$$w'x(t) = e^{\lambda t}w'x(0)\,. \tag{14.15}$$

So $\lambda$ is an uncontrollable eigenvalue.

To prove the converse, suppose now that $\lambda$ is an uncontrollable eigenvalue of $(A, B)$. Then, (14.15) holds for any solution $x(\cdot)$ of the system $\dot{x} = Ax + Bu$ corresponding to any input $u(\cdot)$. Since (14.15) implies that

$$w'\dot{x}(t) = \lambda e^{\lambda t}w'x(0) = \lambda w'x(t),,$$

it follows that

$$w'Ax(t) + w'Bu(t) = \lambda w'x(t)\,;$$

hence

$$w'\begin{pmatrix} A - \lambda I & B \end{pmatrix}\begin{pmatrix} x(t) \\ u(t) \end{pmatrix} = 0$$

Since $x(t)$ and $u(t)$ can be arbitrary chosen, we must have

$$w'\,(A - \lambda I \quad B) = 0\,,$$

Since $w$ is non zero, it follows that the rank of $(A - \lambda I \quad B)$ is less than $n$.

■

ALTERNATE PROOF*. We first show that when the above rank condition holds, the pair $(A, B)$ is not controllable. When this rank condition holds, there exists a non-zero $n$ vector $x$ such that

$$x^*\,(A - \lambda I \quad B) = 0$$

hence,

$$x^*A = \lambda x^*, \qquad x^*B = 0$$

It now follows that

$$x^*AB = \lambda x^*B = 0$$

and, by induction, one can readily prove that

$$x^*A^kB = 0 \qquad \text{for} \quad k = 0, \cdots, n-1$$

Hence

$$x^*Q_c = 0$$

where $Q_c$ is the controllability matrix associated with $(A, B)$. This last condition implies that $Q_c$ has rank less than $n$, hence, $(A, B)$ is not controllable.

So, if either $(A, B)$ has an uncontrollable eigenvalue or the above rank condition holds then, $(A, B)$ is not controllable and there exists a nonsingular matrix $T$ such that the matrices $\tilde{A} = T^{-1}AT$, $\tilde{B} = T^{-1}B$ have the following structure;

$$\tilde{A} = \begin{pmatrix} A_{cc} & A_{cu} \\ 0 & A_{uu} \end{pmatrix} \qquad \qquad \tilde{B} = \begin{pmatrix} B_c \\ 0 \end{pmatrix}$$

with $(A_{cc}, B_c)$ controllable. Noting that

$$\begin{pmatrix} \tilde{A} - \lambda I & \tilde{B} \end{pmatrix} = T^{-1} \begin{pmatrix} A - \lambda I & B \end{pmatrix} \begin{pmatrix} T & 0 \\ 0 & I \end{pmatrix}$$

it follows that

$$\text{rank} \begin{pmatrix} \tilde{A} - \lambda I & \tilde{B} \end{pmatrix} = \text{rank} \begin{pmatrix} A - \lambda I & B \end{pmatrix}$$

Since $(A_{cc}, B_c)$ is controllable,

$$\text{rank} \begin{pmatrix} A_{cc} - \lambda I & B_c \end{pmatrix} = n_c$$

where $n_c$ is the number of rows of $A_c$; hence

$$\begin{aligned}
\text{rank} \begin{pmatrix} \tilde{A} - \lambda I & \tilde{B} \end{pmatrix} &= \text{rank} \begin{pmatrix} A_{cc} - \lambda I & A_{cu} & B_c \\ 0 & A_{uu} - \lambda I & 0 \end{pmatrix} \\
&= \text{rank} \begin{pmatrix} A_{cc} - \lambda I & B_c & A_{cu} \\ 0 & 0 & A_{uu} - \lambda I \end{pmatrix} \\
&= \text{rank} \begin{pmatrix} A_{cc} - \lambda I & B_c & 0 \\ 0 & 0 & A_{uu} - \lambda I \end{pmatrix} \\
&= n_c + \text{rank}\,[A_{uu} - \lambda I]
\end{aligned}$$

It now follows that

$$\text{rank} \begin{pmatrix} A - \lambda I & B \end{pmatrix} < n$$

if and only if

$$\text{rank}\,[A_{uu} - \lambda I] < n_u$$

where $n_u = n - n_c$ is the number of rows of $A_{uu}$. This last condition is equivalent to $\lambda$ being an uncontrollable eigenvalue of $(A, B)$. ■

### 14.6.2 PBH controllability test

Since a system is controllable if and only if it has no uncontrollable eigenvalues, the PBH test for uncontrollable eigenvalues gives us the following test for controllability.

**Corollary 3 (PBH controllability test)** *The pair $(A, B)$ is controllable if and only if*

$$\text{rank}\,(A - \lambda I \quad B) = n \qquad \text{for all} \quad \lambda \in \mathbb{C}$$

**Example 187** Two unattached masses. There,

$$(A - \lambda I \quad B) = \begin{pmatrix} -\lambda & 0 & 0 \\ 0 & -\lambda & 1 \end{pmatrix}$$

The above rank test fails for $\lambda = 0$.

**Example 188** Recall example 186. There

$$(A - \lambda I \quad B) = \begin{pmatrix} -\lambda & 1 & -1 \\ 1 & -\lambda & 1 \end{pmatrix}$$

The above rank test fails for $\lambda = 1$

**Exercise 123** If $\alpha$ is any complex number, prove the following result. The pair $(A+\alpha I, B)$ is controllable if and only if $(A, B)$ is controllable.

**Example 189 (Beavis and Butthead: self-controlled)** Recall the system of Example 181.

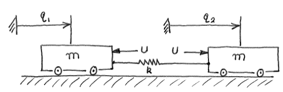


Figure 14.7: B&B: self-controlled

Here

$$A = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -k/m & k/m & 0 & 0 \\ k/m & -k/m & 0 & 0 \end{pmatrix} \qquad B = \begin{pmatrix} 0 \\ 0 \\ -1/m \\ 1/m \end{pmatrix}$$

So, the PBH controllability matrix for the zero eigenvalue is

$$(A \quad B) = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ -k/m & k/m & 0 & 0 & -1/m \\ k/m & -k/m & 0 & -0 & 1/m \end{pmatrix}$$

Since the last two rows of this matrix are not linearly independent, this matrix does not have the maximum rank of 4; hence zero is an uncontrollable eigenvalue. This eigenvalue corresponds to the rigid body mode of this system.

## 14.7 Other characterizations of controllability*

### 14.7.1 Infinite time controllability grammian

Suppose $A$ is asymptotically stable, that is all its eigenvalues have negative real parts, and consider the linear matrix equation

$$\boxed{AW_c + W_cA^* + BB^* = 0}$$

This is a Lyapunov equation of the form $AS + SA^* + Q = 0$ with $Q = BB^*$. Since $A$ is asymptotically stable the above equation has a unique solution for $W_c$ and this solution is given by

$$W_c = \int_0^\infty e^{At}BB^*e^{A^*t}dt\,.$$

We call this matrix the (infinite-time) controllability grammian associated with $(A, B)$. For asymptotically stable systems, this matrix provides another controllability test.

**Lemma 20** *Suppose all the eigenvalues of $A$ have negative real parts. Then $(A, B)$ is controllable if and only if $W_c > 0$*

PROOF. When $A$ is asymptotically stable, one can readily generalize Lemma 17 to include the case $T = \infty$; hence

$$\mathcal{R}(W_c) = \mathcal{R}(Q_c)$$

Since $W_c > 0$ if and only if it has rank $n$, the lemma now follows from the main controllability theorem. ∎

**MATLAB**

```
>> help gram
 GRAM   Controllability and observability gramians.
        GRAM(A,B) returns the controllability gramian:

                Gc = integral {exp(tA)BB'exp(tA')} dt

        GRAM(A',C') returns the observability gramian:

                Go = integral {exp(tA')C'Cexp(tA)} dt

        See also DGRAM, CTRB and OBSV.
```

## 14.8 Controllability, observability, and duality

Recall that for any matrix $M$, we have $\operatorname{rank} M^* = \operatorname{rank} M$. It now follows from the main controllability theorem that a pair $(A,\, B)$ is controllable if and only if the rank of $Q_c^*$ is $n$, that is,

$$\operatorname{rank} \begin{pmatrix} B^* \\ B^* A^* \\ \vdots \\ B^* A^{*(n-1)} \end{pmatrix} = n$$

Note that the above matrix is the observability matrix associated with the pair $(B^*,\, A^*)$. Recalling the main observability theorem, we obtain the following result:

- $(A,\, B)$ is controllable if and only if $(B^*,\, A^*)$ is observable.

  The above statement is equivalent to

- $(C,\, A)$ is observable if and only if $(A^*,\, C^*)$ is controllable.

  This is what is meant by the duality between controllability and observability.

## 14.9 Discrete-time

With one exception, all the results we have developed for continuous time also hold in an analogous fashion for discrete time. In discrete-time we need to consider time intervals $[0, N]$ where $N$ is bigger then $n$, the number state variables. Consider the system described by

$$x(k+1) = Ax(k) + Bu(k) \tag{14.16}$$

where $k \in \mathbb{Z}$ is time, $x(k) \in \mathbb{R}^n$ is the state, and $u(k) \in \mathbb{R}^m$ is the *control input*.

Consider any fixed time $N > 0$. We say that this system is controllable if it can be 'driven' from any state to any other state over the time interval $[0, N]$ by appropriate choice of control input.

**DEFN.** *System (14.16) is* controllable *over the interval* $[0, N]$ *if for every pair of states* $x_0, x_f \in \mathbb{R}^n$, *there is a control function* $u(\cdot) : \{0, 1 \ldots, N\} \to \mathbb{R}^m$ *such that the solution* $x(\cdot)$ *of*

$$x(k+1) = Ax(k) + Bu(k) \qquad x(0) = x_0 \tag{14.17}$$

*satisfies* $x(N) = x_f$.

### 14.9.1 Main controllability theorem

**Theorem 22** (Main controllability theorem.) *For each* $\boxed{N \geq n}$, *system (14.16) is controllable over* $[0, N]$ *if and only if*

$$\boxed{\mathit{rank}\begin{pmatrix} B & AB & \ldots & A^{n-1}B \end{pmatrix} = n}$$

# Part IV

# Control Design

# Chapter 15

# Stabilizability and state feedback

In this chapter we look at the problem of changing the dynamics of a system using state feedback. In particular, we look at the problem of stabilization via state feedback.

## 15.1 Stabilizability and state feedback

Consider a linear time-invariant system with control input described by

$$\dot{x} = Ax + Bu\,. \tag{15.1}$$

For now, we consider linear static state feedback controllers, that is, at each instant $t$ of time the current control input $u(t)$ depends linearly on the current state $x(t)$, that is, we consider the control input to be given by

$$\boxed{u(t) = Kx(t)} \tag{15.2}$$

where $K$ is a constant $m \times n$ matrix, sometimes called a state feedback gain matrix. When system (15.1) is subject to such a controller, its behavior is governed by

$$\boxed{\dot{x} = (A + BK)x} \tag{15.3}$$

We call this the closed loop system resulting from controller (15.2). The following result is

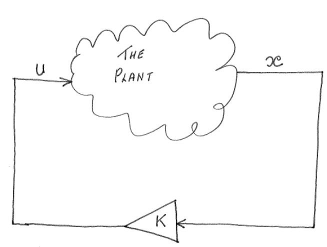



Figure 15.1: State feedback

easily shown using the PBH test for controllability.

**Fact 9** *The pair $(A + BK,\, B)$ is controllable if and only if $(A,\, B)$ is controllable.*

Suppose the open loop system, that is, the system with zero feedback, or, $\dot{x} = Ax$ is unstable or at most marginally stable. A natural question is the following. Can we choose the gain matrix $K$ so that $A + BK$ is asymptotically stable? If yes, we say that system (15.1) is stabilizable.

**DEFN.**[Stabilizability] *System (15.1) is* stabilizable *if there exists a matrix $K$ such that $A+BK$ is asymptotically stable.*

We sometimes say that the pair $(A, B)$ is stabilizable if system (15.1) is stabilizable.

The big question is: under what conditions is a given pair $(A, B)$ stabilizable? We shall see in Section 15.4.2 that controllability is a sufficient condition for stabilizability; that is, if a system is controllable it is stabilizable. However, controllability is not necessary for stabilizability, that is, it is possible for a system to be stabilizable but not controllable. To see this, consider a system with $A$ asymptotically stable and $B$ equal to zero; see also the next two examples.

**Example 190** *(Not stabilizable and not controllable)*

$$\begin{aligned} \dot{x}_1 &= x_1 \\ \dot{x}_2 &= u \end{aligned}$$

For any gain matrix $K$, the matrix

$$A + BK = \begin{pmatrix} 1 & 0 \\ k_1 & k_2 \end{pmatrix}$$

has eigenvalues 1 and $k_2$. Hence, $A + BK$ is unstable for all $K$. So, this system is not stabilizable.

**Example 191** *(Stabilizable but not controllable.)*

$$\begin{aligned} \dot{x}_1 &= -x_1 \\ \dot{x}_2 &= u \end{aligned}$$

For any gain matrix $K$, the matrix

$$A + BK = \begin{pmatrix} -1 & 0 \\ k_1 & k_2 \end{pmatrix}$$

has eigenvalues $-1$ and $k_2$. Hence, $A + BK$ is asymptotically stable provided $k_2 < 0$. So, this system is stabilizable. However, it is not controllable.

## 15.2 Eigenvalue placement (pole placement) by state feedback

Recall the stabilizability problem for the system

$$\dot{x} = Ax + Bu$$

and suppose $A$ and $B$ are real matrices. Here we consider a more general question: *Can we arbitrarily assign the eigenvalues of $A+BK$ by choice of the gain matrix $K$?* The main result of this section is that the answer is yes if and only if the pair $(A,B)$ is controllable. We consider first single input systems.

### 15.2.1 Single input systems

Here we consider scalar input ($m=1$) systems. We first demonstrate that the characteristic polynomial of $A+BK$ depends in an affine linear fashion on the gain matrix $K$. To achieve this, we need to use the following result whose proof is given in the appendix at the end of this chapter.

**Fact 10** *Suppose $M$ and $N$ are any two matrices of dimensions $n \times m$ and $m \times n$, respectively. Then,*

$$\det(I+MN) = \det(I+NM)\,.$$

*In particular, if $m=1$, we have*

$$\det(I+MN) = 1+NM\,.$$

Whenever $s$ is not an eigenvalue of $A$, we have $sI-A-BK = (sI-A)(I-(sI-A)^{-1}BK)$. Thus, using Fact 10, we obtain that

$$\begin{aligned}
\det(sI-A-BK) &= \det(sI-A)\det(I-(sI-A)^{-1}BK) \\
&= \det(sI-A)\det(I-K(sI-A)^{-1}B) \\
&= \det(sI-A)(1-K(sI-A)^{-1}B)\,,
\end{aligned}$$

that is, the characteristic polynomial of $A+BK$ satisfies

$$\det(sI-A-BK) = \det(sI-A)(1-K(sI-A)^{-1}B)\,. \tag{15.4}$$

From this expression, we can make the following statement:

*For a single input system, the characteristic polynomial of $A+BK$ depends in an affine linear fashion on the gain matrix $K$.*

We now show that we can arbitrarily assign the eigenvalues of $A+BK$ by choice of the gain matrix $K$ if $(A,B)$ is controllable. Consider any bunch of complex numbers $\lambda_1,\ldots,\lambda_n$ with the property that if $\lambda$ is in the bunch then so is $\bar{\lambda}$ and suppose we wish to choose a gain matrix $K$ such that the eigenvalues of $A+BK$ are precisely these complex numbers. Let

$$\hat{d}(s) = \prod_{i=1}^{n}(s-\lambda_i)$$

that is, $\hat{d}$ is the unique monic polynomial whose roots are the desired closed loop eigenvalues, $\lambda_1,\ldots,\lambda_n$. We need to show that there is a gain matrix $K$ so that $\det(sI-A-BK) = \hat{d}(s)$.

Let $d$ be the characteristic polynomial of $A$, that is, $d(s) = \det(sI - A)$. Then, recalling (15.4) we need to show that there exists a gain matrix $K$ such that

$$\hat{d}(s) = d(s)(1 - K(sI - A)^{-1}B)$$

that is,

$$d(s)K(sI - A)^{-1}B = d(s) - \hat{d}(s)\,. \tag{15.5}$$

Let

$$\begin{aligned} d(s) &= a_0 + a_1 s + \cdots + a_{n-1}s^{n-1} + s^n \\ \hat{d}(s) &= \hat{a}_0 + \hat{a}_1 s + \cdots + \hat{a}_{n-1}s^{n-1} + s^n\,. \end{aligned}$$

Recall the power series expansion for $(sI - A)^{-1}$:

$$(sI - A)^{-1} = \frac{1}{s}I + \frac{1}{s^2}A + \cdots + \frac{1}{s^k}A^{k-1} + \cdots$$

Substituting this expression into (15.5) and equating the coefficients of like powers of $s$ we obtain that (15.5) holds if and only if

$$\begin{array}{lrcl} s^0: & a_1KB + a_2KAB + \cdots + a_{n-1}KA^{n-2}B + KA^{n-1}B & = & a_0 - \hat{a}_0 \\ s^1: & a_2KB + \cdots + a_{n-1}KA^{n-3}B + KA^{n-2}B & = & a_1 - \hat{a}_1 \\ & \vdots & & \\ s^{n-2} & a_{n-1}KB + KAB & = & a_{n-2} - \hat{a}_{n-2} \\ s^{n-1}: & KB & = & a_{n-1} - \hat{a}_{n-1} \end{array}$$

that is,

$$\begin{pmatrix} KB & KAB & \cdots & KA^{n-1}B \end{pmatrix} \Upsilon = \begin{pmatrix} a_0 - \hat{a}_0 & a_1 - \hat{a}_1 & \cdots & a_{n-1} - \hat{a}_{n-1} \end{pmatrix} \tag{15.6}$$

where $\Upsilon$ is the invertible matrix given by

$$\Upsilon = \begin{pmatrix} a_1 & a_2 & \cdots & a_{n-2} & a_{n-1} & 1 \\ a_2 & a_3 & \cdots & a_{n-1} & 1 & 0 \\ a_3 & a_4 & \cdots & 1 & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots \\ a_{n-1} & 1 & \cdots & 0 & 0 & 0 \\ 1 & 0 & \ldots & 0 & 0 & 0 \end{pmatrix}. \tag{15.7}$$

Let $Q_c$ be the controllability matrix associated with $(A, B)$, that is,

$$Q_c = \begin{pmatrix} B & AB & \cdots & A^{n-1}B \end{pmatrix}.$$

Then condition (15.6) on $K$ reduces to

$$KQ_c\Upsilon = \begin{pmatrix} a_0 - \hat{a}_0 & a_1 - \hat{a}_1 & \cdots & a_{n-1} - \hat{a}_{n-1} \end{pmatrix}.$$

Because the pair $(A, B)$ is controllable, $Q_c$ is invertible. This readily implies that the gain matrix $K$ is uniquely given by

$$K = \begin{pmatrix} a_0 - \hat{a}_0 & a_1 - \hat{a}_1 & \cdots & a_{n-1} - \hat{a}_{n-1} \end{pmatrix} \Upsilon^{-1}Q_c^{-1}\,. \tag{15.8}$$

We have just demonstrated the following result.

**Theorem 23 (Pole placement theorem: SI case)** *Suppose the real matrix pair $(A, B)$ is a single input controllable pair with state dimension $n$ and $\lambda_1, \ldots, \lambda_n$ is any bunch of $n$ complex numbers with the property that if $\lambda$ is in the bunch then so is $\bar{\lambda}$. Then there exists a real matrix $K$ such that $\lambda_1, \ldots, \lambda_n$ are the eigenvalues (multiplicities included) of $A + BK$.*

**MATLAB**

```
>> help acker

 ACKER  Pole placement gain selection using Ackermann's formula.
        K = ACKER(A,B,P)  calculates the feedback gain matrix K such that
        the single input system
                .
                x = Ax + Bu

        with a feedback law of  u = -Kx  has closed loop poles at the
        values specified in vector P, i.e.,  P = eig(A-B*K).

        See also PLACE.

        Note: This algorithm uses Ackermann's formula.  This method
        is NOT numerically reliable and starts to break down rapidly
        for problems of order greater than 10, or for weakly controllable
        systems.  A warning message is printed if the nonzero closed loop
        poles are greater than 10% from the desired locations specified
        in P.
```

**Example 192** Consider

$$A = \begin{pmatrix} 3 & -2 \\ 4 & -3 \end{pmatrix} \qquad \text{and} \qquad B = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$$

Here the controllability matrix

$$Q_c = \begin{pmatrix} B & AB \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 3 & -1 \end{pmatrix}$$

is full rank; hence $(A, B)$ is controllable. Considering any state feedback gain matrix $K = \begin{pmatrix} k_1 & k_2 \end{pmatrix}$, we have

$$A + BK = \begin{pmatrix} 3 + 2k_1 & -2 + 2k_2 \\ 4 + 3k_1 & -3 + 3k_2 \end{pmatrix}$$

and

$$\det(sI - A - BK) = s^2 - (2k_1 + 3k_2)s - 1 + k_2 \,.$$

Suppose we desire the eigenvalues of $A + BK$ to be $-2$ and $-3$. Then the desired characteristic polynomial $\hat{d}$ of $A + BK$ is given by

$$\hat{d}(s) = (s+2)(s+3) = s^2 + 5s + 6\,.$$

By equating the coefficients of like powers of $s$, we obtain

$$\begin{aligned} -1 + k_2 &= 6 \\ -2k_1 - 3k_2 &= 5 \end{aligned}$$

Solving these equations yields (uniquely) $k_1 = -13$ and $k_2 = 7$. Hence the gain matrix is given by

$$K = \begin{pmatrix} -13 & 7 \end{pmatrix}\,.$$

Checking our answer in Matlab, we obtain

```
K = -acker(A,B,[-2 -3])
K =
   -13     7
```

**Example 193 (Beavis and Butthead with external control)** Considering $m = k =$

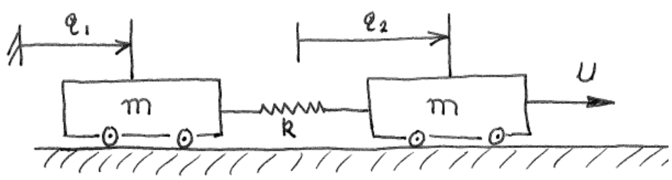


Figure 15.2: B&B: external control

1, a state space description of this system is given by

$$A = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \end{pmatrix} \qquad \text{and} \qquad B = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}\,.$$

Here,

$$\det(sI - A) = s^4 + 2s^2$$

Let

$$\hat{d}(s) = s^4 + \hat{a}_3 s^3 + \hat{a}_2 s^2 + \hat{a}_1 s + \hat{a}_0$$

be the desired characteristic polynomial of $A + BK$.

Since this is a scalar input system, the coefficients of the characteristic polynomial of $A + BK$ depend affinely on $K$ and we can use the following solution approach to find $K$. Let

$$K = \begin{pmatrix} k_1 & k_2 & k_3 & k_4 \end{pmatrix}$$

Then

$$\det(sI - A - BK) = \det\begin{pmatrix} s & 0 & -1 & 0 \\ 0 & s & 0 & -1 \\ 1 & -1 & s & 0 \\ -1-k_1 & 1-k_2 & -k_3 & s-k_4 \end{pmatrix}$$
$$= s^4 - k_4 s^3 + (2-k_2)s^2 - (k_3+k_4)s - (k_1+k_2)\,.$$

Note the affine dependence of the coefficients of the above polynomial on $K$. Letting the above polynomial equal the desired polynomial $\hat{d}$ and equating like coefficients results in the following four linear equations in the four unknowns components of $K$:

$$\begin{aligned} -k_4 &= \hat{a}_3 \\ 2-k_2 &= \hat{a}_2 \\ -k_3-k_4 &= \hat{a}_1 \\ -k_1-k_2 &= \hat{a}_0 \end{aligned}$$

Solving for $k_1, \ldots, k_4$ uniquely yields

$$\begin{aligned} k_1 &= -2-\hat{a}_0+\hat{a}_2 \\ k_2 &= 2-\hat{a}_2 \\ k_3 &= -\hat{a}_1+\hat{a}_3 \\ k_4 &= -\hat{a}_3 \end{aligned}$$

• To illustrate MATLAB, consider desired closed loop eigenvalues: $-1, -2-3, -4$.

```
>> poles=[-1 -2 -3 -4];

>> poly(poles)
ans =
     1    10    35    50    24
```

Hence

$$\hat{a}_0 = 24 \qquad \hat{a}_1 = 50 \qquad \hat{a}_2 = 35 \qquad \hat{a}_3 = 10$$

and the above expressions for the gain matrix yield

$$k_1 = 9 \qquad k_2 = -33 \qquad k_3 = -40 \qquad k_4 = -10$$

Using MATLAB pole placement commands, we obtain:

```
>> k= place(a,b,poles)
place: ndigits= 18

k =
```

```
   -9.0000   33.0000   40.0000   10.0000

>> k= acker(a,b,poles)

k =
    -9    33    40    10
```

Remembering that MATLAB yields $-K$, these results agree with our 'hand' calculated results. Lets check to see if we get the desired closed loop eigenvalues.

```
>> eig(a-b*k)

ans =
   -1.0000
   -2.0000
   -3.0000
   -4.0000
```

YEES!

### 15.2.2 Multi-input systems

The following theorem (which we do not prove here) states that the eigenvalue placement result also holds in the general multi-input case.

> **Theorem 24 (Pole placement theorem )** *Suppose the real matrix pair $(A, B)$ is a controllable pair with state dimension $n$ and $\lambda_1, \ldots, \lambda_n$ is any bunch of $n$ complex numbers with the property that if $\lambda$ is in the bunch then so is $\bar{\lambda}$. Then there exists a real matrix $K$ such that $\lambda_1, \ldots, \lambda_n$ are the eigenvalues (multiplicities included) of $A + BK$.*

• It follows from the above theorem that

$$\boxed{\text{controllability} \quad \Longrightarrow \quad \text{stabilizability}}$$

**MATLAB**

```
>> help place

 PLACE  K = place(A,B,P)  computes the state feedback matrix K such that
        the eigenvalues of  A-B*K  are those specified in vector P.
        The complex eigenvalues in the vector P must appear in consecu-
        tive complex conjugate pairs. No eigenvalue may be placed with
        multiplicity greater than the number of inputs.
```

The displayed "ndigits" is an estimate of how well the eigenvalues were placed. The value seems to give an estimate of how many decimal digits in the eigenvalues of A-B*K match the specified numbers given in the array P.

A warning message is printed if the nonzero closed loop poles are greater than 10% from the desired locations specified in P.

See also: LQR and RLOCUS.

## 15.3 Uncontrollable eigenvalues (you can't touch them)

In the previous section, we saw that, if the pair $(A, B)$ is controllable, then one could arbitrarily place the eigenvalues of $A + BK$ by appropriate choice of the gain matrix $K$; hence controllability is a sufficient condition for arbitrary eigenvalue placement. In this section, we show that controllability is also necessary for arbitrary eigenvalue placement. In particular, we show that if $\lambda$ is an uncontrollable eigenvalue of $(A, B)$, then $\lambda$ is an eigenvalue of $A + BK$ for every $K$. Recall Examples 190 and 191 in which the matrix $A$ had an eigenvalue $\lambda$ with the property that for every gain matrix $K$, it is an eigenvalue of $A + BK$. We now demonstrate the following result:

*If $\lambda$ is an uncontrollable eigenvalue of $(A, B)$, then $\lambda$ is an eigenvalue of $A+BK$ for every gain matrix $K$.*

To see this, recall that $\lambda$ is an uncontrollable eigenvalue of $(A, B)$ if and only if

$$\operatorname{rank} \begin{pmatrix} A - \lambda I & B \end{pmatrix} < n \tag{15.9}$$

where $A$ is $n \times n$. Hence, there exists a nonzero $n$-vector $w$ such that

$$w' \begin{pmatrix} A - \lambda I & B \end{pmatrix} = 0\,.$$

that is

$$w'(A - \lambda I) = 0 \qquad \text{and} \qquad w'B = 0$$

So, for any gain matrix $K$, we have

$$w'(A + BK) = \lambda w'$$

that is, $\lambda$ is an eigenvalue of $A + BK$ with left eigenvector $w$. So, regardless of the feedback gain matrix, $\lambda$ is always an eigenvalue of $A + BK$. We cannot alter this eigenvalue by feedback. In particular, if $\lambda$ has a non-negative real part, then the pair $(A, B)$ is not stabilizable.

Another way to see the above result is to consider the nice transformation introduced earlier, namely,

$$x = T\tilde{x} \qquad \text{and} \qquad \tilde{x} = \begin{pmatrix} x_c \\ x_u \end{pmatrix}$$

which results in the equivalent system description,

$$\dot{\tilde{x}} = \tilde{A}\tilde{x} + \tilde{B}u$$

where

$$\tilde{A} = T^{-1}AT = \begin{pmatrix} A_{cc} & A_{cu} \\ 0 & A_{uu} \end{pmatrix} \qquad \text{and} \qquad \tilde{B} = T^{-1}B = \begin{pmatrix} B_c \\ 0 \end{pmatrix}$$

and the pair $(A_{cc}, B_c)$ is controllable. If $u = Kx$, then $u = \tilde{K}\tilde{x}$ where $\tilde{K} = KT$. Letting $\tilde{K} = \begin{pmatrix} K_c & K_u \end{pmatrix}$, we obtain

$$\tilde{A} = \tilde{B}\tilde{K} = \begin{pmatrix} A_{cc} + B_c K_c & A_{cu} + B_c K_u \\ 0 & A_{uu} \end{pmatrix}$$

Since

$$\tilde{A} + \tilde{B}\tilde{K} = T^{-1}[A + BK]T$$

the eigenvalues of $A + BK$ are exactly the same as those of $\tilde{A} + \tilde{B}\tilde{K}$ which are equal to the union of the eigenvalues of $A_{cc} + B_c K_c$ and $A_{uu}$. From this we see that, regardless of $K$, the eigenvalues of $A_{uu}$ are always eigenvalues of $A + BK$. Now recall that the eigenvalues of $A_{uu}$ are the uncontrollable eigenvalues of $(A, B)$.

Furthermore, since the pair $(A_{cc}, B_c)$ is controllable, the eigenvalues of $A_{cc} + B_c K_c$ can be arbitrarily placed by appropriate choice of $K_c$. Hence, except for the eigenvalues of $A_{uu}$, all eigenvalues of $A + BK$ can be arbitrarily placed. In particular, if all the eigenvalues of $A_{uu}$ have negative real parts, then $K$ can be chosen so that all the eigenvalues of $A + BK$ have negative real parts. So, we have the following conclusion.

*A pair $(A, B)$ is stabilizable if and only if all the uncontrollable eigenvalues of $(A, B)$ have negative real parts.*

Recalling the PBH test for uncontrollable eigenvalues, we now have the following PBH result on stabilizability.

**Theorem 25 (PBH stabilizability theorem)** *A pair $(A, B)$ is stabilizable if and only if*

$$\mathrm{rank}\begin{pmatrix} A - \lambda I & B \end{pmatrix} = n$$

*for every eigenvalue $\lambda$ of $A$ with nonnegative real part.*

**Example 194**

```
A =
     0    -1     0
     1    -1    -1
    -1     0     0


B =
     1
    -1
     1

rank(ctrb(A,B))
     2                  %Uncontrollable

eig(A)
```

```
        0 + 1.0000i
        0 - 1.0000i
         -1.0000

rank([ A-i*eye(3) B])
     3                      %Stabilizable

rank([ A+eye(3) B])
     2                      %-1 is an uncontrollable  eigenvalue
```

Considering $K = \begin{pmatrix} k_1 & k_2 & k_3 \end{pmatrix}$ we obtain that

$$\det(sI - A - BK) = s^3 + (1 - k_1 + k_2 - k_3)s^2 + (1 - 2k_1)s + (1 - k_1 - k_2 + k_3)\,.$$

As expected, we can write this as

$$\det(sI - A - BK) = (s+1)\left(s^2 + (-k_1 + k_2 - k_3)s + (1 - k_1 - k_2 + k_3)\right)$$

Suppose that the desired eigenvalues of $A + BK$ are $-1 - 2, -3$. Then

$$\hat{d}(s) = (s+1)(s+2)(s+3) = s^3 + 6s^2 + 11s + 6\,.$$

Equating coefficients, we obtain

$$\begin{aligned} 1 - k_1 - k_2 + k_3 &= 6 \\ 1 - 2k_1 &= 11 \\ 1 - k_1 + k_2 - k_3 &= 6 \end{aligned}$$

Solving yields

$$\begin{array}{rcl} k_1 & = & -5 \\ k_2 & = & k_3 \\ k_3 & \text{is} & \text{arbitrary} \end{array}$$

We choose

$$K = \begin{pmatrix} -5 & 0 & 0 \end{pmatrix}\,.$$

```
 eig(A+B*K)
   -3.0000
   -2.0000                      %Just checking!
   -1.0000
```

**Remark 9** *If system (15.1) is not stabilizable via a linear static state controller of the form $u = Kx$ then, it is not stabilizable with any controller.* This can be seen as follows. The fact that the system is not stabilizable via a linear controller implies that the system has an uncontrollable eigenvalue $\lambda$ with non-negative real part. Hence there exists a nonzero vector $w$ such that, regardless of the input $u$, we have

$$w'x(t) = e^{\lambda t} w'x(0)$$

for all $t$. Considering any initial state $x(0)$ for which $w'x(0) \neq 0$ (for example, let $x(0)$ be the real part of $w$), we see that, regardless of the input $u$, the resulting solution $x(t)$ does not go to zero as $t$ goes to infinity.

## 15.4 Controllable canonical form

Consider a scalar input system of the form $\dot{x} = Ax + Bu$ where $A$ and $B$ have the following structure:

$$A = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & & \ddots & & \vdots \\ \vdots & \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 1 \\ -a_0 & -a_1 & -a_2 & \dots & -a_{n-2} & -a_{n-1} \end{pmatrix} \qquad \text{and} \qquad B = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

We say that the pair $(A, B)$ is in controllable canonical form. We saw this structure when we constructed realizations for scalar transfer functions. Note that $\dot{x} = Ax + Bu$ looks like:

$$\begin{aligned} \dot{x}_1 &= x_2 \\ &\vdots \\ \dot{x}_{n-1} &= x_n \\ \dot{x}_n &= -a_0 x_1 + a_1 x_2 + \ldots + a_{n-1} x_n + u \end{aligned}$$

Also,

$$\det(sI - A) = a_0 + a_1 s + \cdots + a_{n-1} s^{n-1} + s^n$$

**Example 195** (*Controlled unattached mass*)

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= u \end{aligned}$$

Here,

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad \text{and} \qquad B = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

**Fact 11** *If the pair $(A, B)$ is in controllable canonical form, then it is controllable.*

To prove the above result, let $K = \begin{pmatrix} a_0 & a_1 & \cdots & a_{n-1} \end{pmatrix}$. Then

$$A + BK = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & & \ddots & & \vdots \\ \vdots & \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & \dots & 0 & 0 \end{pmatrix}$$

The controllability matrix $Q_c$ for the pair $(A + BK, B)$ is simply given by

$$Q_c = \begin{pmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 1 & 0 \\ \vdots & \vdots & \ddots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \end{pmatrix}$$

This matrix has rank $n$; hence $(A + BK, B)$ is controllable. It now follows that $(A, B)$ is controllable.

### 15.4.1 Transformation to controllable canonical form*

Here we show that any controllable single-input system

$$\dot{x} = Ax + Bu$$

can be transformed via a state transformation to a system which is in controllable canonical form.

**State transformations.** Suppose $T$ is a nonsingular matrix and consider the state transformation

$$x = T\tilde{x}$$

Then the evolution of $\tilde{x}$ is governed by

$$\dot{\tilde{x}} = \tilde{A}\tilde{x} + \tilde{B}u$$

where $\tilde{A} = T^{-1}AT$ and $\tilde{B} = T^{-1}B$.

Note that *controllability is invariant under a state transformation*, that is, $(A, B)$ is controllable if and only if $(T^{-1}AT,\ T^{-1}B)$ is controllable. This can readily seen by noting the controllability matrix $\tilde{Q}_c$ for the pair $(T^{-1}AT,\ T^{-1}B)$ is given by $\tilde{Q}_c = T^{-1}Q_c$ where $Q_c$ is the controllability matrix for $(A, B)$; hence $\tilde{Q}_c$ and $Q_c$ have the same rank.

**Transformation to controllable canonical form.** Suppose $(A, B)$ is controllable and $m = 1$. Then the following algorithm yields a state transformation which transforms $(A,\ B)$ into controllable canonical form.

**Algorithm.** Let

$$a_0 + \ldots + a_{n-1}s^{n-1} + s^n = \det(sI - A)$$

Recursively define the following sequence of $n$-vectors:

$$\boxed{\begin{array}{lcl} t^n & = & B \\ t^j & = & At^{j+1} + a_j B\,, \qquad j = n-1,\ \ldots,\ 1 \end{array}}$$

and let

$$T = \begin{pmatrix} t^1 & \ldots & t^n \end{pmatrix} \qquad \blacksquare$$

**Fact 12** *If $(A, B)$ is controllable then, the matrix $T$ is generated by the above algorithm is nonsingular and the pair $(T^{-1}AT,\ T^{-1}B)$ is in controllable canonical form.*

Proof. We first need to prove that $T$ is nonsingular. By applying the algorithm, we obtain

$$\begin{array}{rcl} t^n & = & B \\ t^{n-1} & = & AB + a_{n-1}B \\ t^{n-2} & = & A(AB + a_{n-1}B) + a_{n-2}B \quad = \quad A^2B + a_{n-1}AB + a_{n-2}B \end{array}$$

One may readily show by induction that for $j = n-1, n-2 \cdots, 1$, one has

$$t^j = A^{n-j}B + a_{n-1}A^{n-j-1}B + \cdots + a_{j+1}AB + a_jB \tag{15.10}$$

From this it should be clear that for any $j = n, n-1, \cdots 1$,

$$\text{rank}\begin{pmatrix} t^n & t^{n-1} & \cdots & t^j \end{pmatrix} = \text{rank}\begin{pmatrix} B & AB & \cdots & A^{n-j}B \end{pmatrix}$$

Hence

$$\text{rank}\, T = \text{rank}\begin{pmatrix} t^n & t^{n-1} & \cdots & t^1 \end{pmatrix} = \text{rank}\begin{pmatrix} B & AB & \cdots & A^nB \end{pmatrix} = \text{rank}\, Q_c$$

Since $(A, B)$ is controllable, $Q_c$ has rank $n$. Hence $T$ has rank $n$ and is invertible.

We now show that $T^{-1}AT$ and $T^{-1}B$ have the required structure. It follows from (15.10) that

$$t^1 = A^{n-1}B + a_{n-1}A^{n-2}B + \cdots + a_2AB + a_1B$$

Using the Cayley Hamilton Theorem, we obtain

$$At^1 = [A^n + a_{n-1}A^{n-1} + \cdots + a_2A^2 + a_1A]B = -a_0B\,,$$

that is $At^1 = -a_0t^n$. Also from the algorithm, we have

$$At^j = t^{j-1} - a_{j-1}t^n \qquad \text{for} \quad j = 2, 3, \cdots, n$$

Recalling a useful fact on similarity transformations, it follows that $T^{-1}AT$ has the required structure. Since $B = t^n$, it follows that $T^{-1}B$ also has the required structure. $\blacksquare$

We can now state the following result.

**Theorem 26** *A single input system $(A, B)$ is controllable if and only if there is a nonsingular matrix $T$ such that $(T^{-1}AT,\ T^{-1}B)$ is in controllable canonical form.*

**Example 196** *(Beavis and Butthead with external control.)* Consider $m = k = 1$. Then,

$$A = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \end{pmatrix} \qquad \text{and} \qquad B = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Hence,

$$\det(sI - A) = s^4 + 2s^2$$

and

$$\begin{aligned} t^4 &= B \\ t^3 &= At^4 + a_3 B = At^4 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \\ t^2 &= At^3 + a_2 B = At^3 + 2B = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} \\ t^1 &= At^2 + a_1 B = At^2 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \end{aligned}$$

So, the transformation matrix is given by:

$$T = \begin{pmatrix} t^1 & t^2 & t^3 & t^4 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}$$

```
>> inv(t)
ans =
     1     0     0     0
     0     0     1     0
    -1     1     0     0
     0     0    -1     1

>> inv(t)*a*t
ans =
     0     1     0     0
     0     0     1     0
```

```
0    0    0    1
0    0   -2    0
```

### 15.4.2 Eigenvalue placement by state feedback*

Recall that if $(A, B)$ is controllable then, we arbitrarily assign the eigenvalues of $A + BK$ by choice of the gain matrix $K$? We can demonstrate this result for single input systems using the controllable canonical form. Consider any bunch of complex numbers $\lambda_1, \ldots, \lambda_n$ with the property that if $\lambda$ is in the bunch then so is $\bar{\lambda}$ and suppose we wish to choose a gain matrix $K$ such that the eigenvalues of $A + BK$ are precisely these complex numbers.

Let

$$\hat{d}(s) = s^n + \hat{a}_{n-1}s^{n-1} + \ldots + \hat{a}_0 = \prod_{i=1}^{n}(s - \lambda_i)$$

that is, the real numbers, $\hat{a}_0, \ldots, \hat{a}_{n-1}$, are the coefficients of the unique monic polynomial $\hat{d}$ whose roots are the desired closed loop eigenvalues, $\lambda_1, \ldots, \lambda_n$. If $(A,\ B)$ is controllable and $m = 1$, then there exists a nonsingular matrix $T$ such that the pair $(\tilde{A},\ \tilde{B})$ is in controllable canonical form, where $\tilde{A} = T^{-1}AT$ and $\tilde{B} = T^{-1}B$, that is,

$$\tilde{A} = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & & \ddots & & \vdots \\ \vdots & \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 0 & 1 \\ -a_0 & -a_1 & -a_2 & \ldots & -a_{n-2} & -a_{n-1} \end{pmatrix}, \qquad \tilde{B} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

where

$$s^n + a_{n-1}s^{n-1} + \ldots + a_0 = \det(sI - A)\,.$$

Considering the $1 \times n$ real gain matrix

$$\tilde{K} = \begin{pmatrix} a_0 - \hat{a}_0 & a_1 - \hat{a}_1 & \ldots & a_{n-1} - \hat{a}_{n-1} \end{pmatrix}$$

we obtain

$$\tilde{A} + \tilde{B}\tilde{K} = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 0 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & & \ddots & & \vdots \\ \vdots & \vdots & & & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 0 & 1 \\ -\hat{a}_0 & -\hat{a}_1 & -\hat{a}_2 & \ldots & -\hat{a}_{n-2} & -\hat{a}_{n-1} \end{pmatrix}.$$

Clearly, the matrix $\tilde{A} + \tilde{B}\tilde{K}$ has the desired characteristic polynomial $\hat{d}$. Let

$$K = \tilde{K}T^{-1}\,.$$

Then

$$\begin{aligned} A+BK &= T\tilde{A}T^{-1}+T\tilde{B}\tilde{K}T^{-1} \\ &= T[\tilde{A}+\tilde{B}\tilde{K}]T^{-1} \end{aligned}$$

and hence

$$\det(sI-A-BK)=\det(sI-\tilde{A}-\tilde{B}\tilde{K})=\hat{d}(s)$$

in other words, the eigenvalues of $A+BK$ are as desired.

**Remark 10** For scalar input systems, the coefficients of the characteristic polynomial of $A+BK$ depend affinely on $\tilde{K}$. Since $K=\tilde{K}T^{-1}$, it follows that the coefficients of the characteristic polynomial of $A+BK$ depend affinely on $K$.

**Example 197** *(Beavis and Butthead with external control)* Here

$$A=\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \end{pmatrix} \qquad \text{and} \qquad B=\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

We have already shown that this system is controllable and it can be transformed to controllable canonical form with the matrix

$$T=\begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}$$

Also,

$$\det(sI-A)=s^4+2s^2$$

Let

$$\hat{p}(s)=s^4+\hat{a}_3s^3+\hat{a}_2s^2+\hat{a}_1s+\hat{a}_0$$

be the desired characteristic polynomial of $A+BK$. Then

$$\tilde{K}=\begin{pmatrix} -\hat{a}_0 & -\hat{a}_1 & 2-\hat{a}_2 & -\hat{a}_3 \end{pmatrix}$$

and

$$\begin{aligned} K &= =\tilde{K}T^{-1}=\begin{pmatrix} -\hat{a}_0 & -\hat{a}_1 & 2-\hat{a}_2 & -\hat{a}_3 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} -2-\hat{a}_0+\hat{a}_2 & 2-\hat{a}_2 & -\hat{a}_1+\hat{a}_3 & -\hat{a}_3 \end{pmatrix} \end{aligned}$$

This is the same as the result we previously obtained for this example.

## 15.5 Discrete-time systems

Basically, everything we have said in this chapter so far also holds for discrete time systems, except that in talking about asymptotic stability of discrete-time systems, we need the eigenvalues to have magnitude strictly less than one.

### 15.5.1 Dead beat controllers

One cannot drive the state of a LTI continuous-time system to zero in a finite time using a static linear state-feedback controller. However, this is possible for a controllable discrete-time system. To see this, consider

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

where $x(k)$ is an $n$-vector and $u(k)$ is an $m$-vector and suppose that the pair $(A, B)$ is controllable. Then there exists a gain matrix K so that all the eigenvalues of the corresponding closed loop system matrix $A_{cl} := A + BK$ are all zero. Hence the characteristic polynomial of $A_{cl}$ is $p(s) = s^n$. From the Cayley-Hamilton Theorem, we obtain that

$$A_{cl}^n = 0 \, .$$

Since the solutions $x(\cdot)$ of the closed lop system

$$x(k+1) = A_{cl}x(k)$$

satisfy $x(k) = A_{cl}^k x(0)$ for all $k \geq 0$, we see that

$$x(k) = 0 \qquad \text{for} \qquad k \geq n \, .$$

*Thus, all solutions go to zero in at most $n$ steps.*

**Application to continuous-time systems.** One could use the above result to obtain a controller which drives the state of a continuous-time system to zero. Simply, pick a sampling time $T$ and discretize the continuous-time system. Now design a deadbeat controller for the discrete-time system. Using a zero order hold, apply this controller to the continuous-time system. Note that the resulting controller is not a static controller for the continuous-time system.

## 15.6 Stabilization of nonlinear systems

Consider a nonlinear system described by

$$\dot{x} = F(x,u)$$

and suppose we wish to stabilize this system about some controlled equilibrium state $x^e$. Let $u^e$ be a constant input that achieves the desired controlled equilibrium state $x^e$, that is, $F(x^e,u^e) = 0$. Let

$$\delta\dot{x} = A\delta x + B\delta u \tag{15.11}$$

be the linearization of the nonlinear system about $(x^e, u^e)$. Thus,

$$A = \frac{\partial F}{\partial x}(x^e,u^e) \qquad \text{and} \qquad B = \frac{\partial F}{\partial u}(x^e,u^e)\,.$$

### 15.6.1 Continuous-time controllers

If the pair $(A,B)$ is stabilizable, the following procedure yields a controller which stabilizes the nonlinear system about $x^e$.

*Choose a gain matrix $K$ such that all eigenvalues of $A+BK$ have negative real part and let*

$$\boxed{u = u^e + K(x - x^e)} \tag{15.12}$$

We now show that the above controller results in the closed loop nonlinear system being asymptotically about $x^e$. To this end, note that the closed loop system is described by

$$\dot{x} = f(x) = F(x,\ Kx + u^e - Kx^e)$$

Note that

$$f(x^e) = F(x^e,u^e) = 0\,;$$

hence $x^e$ is an equilibrium state for the closed loop system. Linearizing the closed loop system about $x^e$, we obtain

$$\delta\dot{x} = (A + BK)\delta x$$

Since all the eigenvalues of $A+BK$ have negative real parts, the closed loop nonlinear system is asymptotically stable about $x^e$.

### 15.6.2 Discrete-time controllers

Linearize the continuous time nonlinear system to obtain the LTI continuous-time system (15.11).

Choose a sample time $T > 0$ and discretize (15.11) to obtain

$$\delta x_d(k+1) = A_d\delta x_d(k) + B_d\delta u_d(k) \tag{15.13}$$

If the pair $(A_d, B_d)$ is stabilizable, choose matrix $K$ so that all the eigenvalues of $A_d + B_d K$ have magnitude less than one.

Let

$$u_d(k) = K(x_d(k) - x^e) + u^e = K[x(kT) - x^e] + u^e\,. \tag{15.14}$$

The continuous control input $u$ is obtained form $u_d$ via a zero order hold.

# Exercises

**Exercise 124** Consider the system with input $u$, output $y$, and state variables $x_1, x_2, x_3, x_4$ described by

$$\begin{aligned}
\dot{x}_1 &= -x_1 \\
\dot{x}_2 &= x_1 - 2x_2 \\
\dot{x}_3 &= x_1 - 3x_3 + u \\
\dot{x}_4 &= x_1 + x_2 + x_3 + x_4 \\
y &= x_3
\end{aligned}$$

(a) Is this system controllable?
(b) Is this system stabilizable using state feedback? Justify your answers.

**Exercise 125** *(BB in laundromat: external excitation.)* Obtain a state space representation of the following system.

$$\begin{aligned}
m\ddot{\phi}_1 - m\Omega^2\phi_1 + \tfrac{k}{2}(\phi_1 - \phi_2) &= 0 \\
m\ddot{\phi}_2 - m\Omega^2\phi_2 - \tfrac{k}{2}(\phi_1 - \phi_2) &= u
\end{aligned}$$

Show that this system is controllable.

Considering

$$\Omega = 1,\ m = 1,\ k = 2$$

obtain a linear state feedback controller which results in a closed loop system with eigenvalues

$$-1, -2, -3, -4$$

Use the following methods.

(a) Method based on computing the transformation matrix which transforms to controllable canonical form.

(b) Obtaining an expression for the closed loop characteristic polynomial in terms of the components of the gain matrix.

(c) `acker`

(d) `place`

Numerically simulate the closed loop system with the same initial conditions you used in to simulate the open-loop system.

**Exercise 126** Consider the system described by

$$\begin{aligned}\dot{x}_1 &= x_1^3 + \sin x_2 - 1 + u \\ \dot{x}_2 &= -e^{x_1} + (\cos x_2)u\end{aligned}$$

Obtain a state feedback controller which results in a closed loop system which is asymptotically stable about the zero state.

**Exercise 127** *Stabilization of 2-link manipulator with linear state feedback.*

$$\begin{aligned}[m_1 lc_1^2 + m_2 l_1^2 + I_1]\ddot{q}_1 + [m_2 l_1 lc_2 \cos(q_1 - q_2)]\ddot{q}_2 + m_2 l_1 lc_2 \sin(q_1 - q_2)\dot{q}_2^2 & \\ - [m_1 lc_1 + m_2 l_1] g \sin(q_1) &= u\end{aligned}$$

$$[m_2 l_1 lc_2 \cos(q_1 - q_2)]\ddot{q}_1 + [m_2 lc_2^2 + I_2]\ddot{q}_2 - m_2 l_1 lc_2 \sin(q_1 - q_2)\dot{q}_1^2 - m_2 g lc_2 \sin(q_2) = 0$$

Linearize this system about

$$q_1 = q_2 = u = 0$$

and obtain a state space description.

For rest of exercise, use MATLAB and the following data: Is the linearization controllable? Design a state feedback controller which stabilizes this linearized system . Apply the controller to the nonlinear system and simulate the nonlinear closed loop system for various initial conditions.

| $m_1$ | $l_1$ | $lc_1$ | $I_1$ | $m_2$ | $l_2$ | $lc_2$ | $I_2$ | $m_{payload}$ |
|---|---|---|---|---|---|---|---|---|
| $kg$ | $m$ | $m$ | $kg.m^2$ | $kg$ | $m$ | $m$ | $kg.m^2$ | $kg$ |
| 10 | 1 | 0.5 | 10/12 | 5 | 1 | 0.5 | 5/12 | 0 |

**Exercise 128** Consider the system described by

$$\begin{aligned}\dot{x}_1 = x_2 + u \\ \dot{x}_2 = x_1 + u\end{aligned}$$

(a) Is this system stabilizable ?

(b) Does there exist a linear state feedback controller which results in closed loop eigenvalues $-1, -2$?

(b) Does there exist a linear state feedback controller which results in closed loop eigenvalues $-2, -3$?

In parts (b) and (c): If no controller exists, explain why; if one does exist, give an example of one.

**Exercise 129** Consider the system described by

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

where all quantities are scalars.

(a) Is this system stabilizable via state feedback?

(b) Does there exist a linear state feedback controller which results is closed loop eigenvalues $-1, -4$?

(c) Does there exist a linear state feedback controller which results is closed loop eigenvalues $-2, -4$?

In parts (b) and (c): If no controller exists, explain why; if one does exist, give an example of one.

**Exercise 130** Consider the system

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

(a) Design a continuous-time linear state feedback controller which stabilizes this system about the zero state. Illustrate the effectiveness of your controller with a simulation.

(b) For various sampling times, discretize the controller found in part (a) and simulate the corresponding closed loop. Is the closed loop system stable for any sampling time?

(c) Design a stabilizing discrete-time controller based on a discretization of the system. Consider sampling times $T = 0.1, 1, 2\pi, 10$. Illustrate the effectiveness of your controller with simulations of the continuous-time system subject to the discrete-time controller.

(d) Design a dead-beat controller based on a discretization of the system. Consider sampling times $T = 0.1, 1, 10$. Illustrate the effectiveness of your controller with simulations.

**Exercise 131** *(BB in laundromat: external excitation.)* Obtain a state space representation of the following system.

$$\begin{aligned}m\ddot{\phi}_1 - m\Omega^2\phi_1 + \tfrac{k}{2}(\phi_1 - \phi_2) &= 0 \\ m\ddot{\phi}_2 - m\Omega^2\phi_2 - \tfrac{k}{2}(\phi_1 - \phi_2) &= u\end{aligned}$$

(a) Show that this system is controllable.
(b) Considering

$$\Omega = 1,\ m = 1,\ k = 2$$

obtain a linear state feedback controller which results in a closed loop system with eigenvalues

$$-1, -2, -3, -4$$

Use the following methods.

(i) Obtain (by hand) an expression for the closed loop characteristic polynomial in terms of the components of the gain matrix.

(ii) `acker`

(iii) `place`

Illustrate the effectiveness of your controllers with numerical simulations.

**Exercise 132 (Stabilization of cart pendulum system via state feedback.)** Carry out the following for parameter sets P2 and P4 and equilibriums $E1$ and $E2$. Illustrate the effectiveness of your controllers with numerical simulations.

Using eigenvalue placement techniques, obtain a state feedback controller which stabilizes the nonlinear system about the equilibrium.

What is the largest value of $\delta$ (in degrees) for which your controller guarantees convergence of the closed loop system to the equilibrium for initial condition

$$(y,\, \theta_1,\, \theta_2,\, \dot{y},\, \dot{\theta}_1,\, \dot{\theta}_2)(0) = (0,\, \theta_1^e{-}\delta,\, \theta_2^e{+}\delta,\, 0,\, 0,\, 0)$$

where $\theta_1^e$ and $\theta_2^e$ are the equilibrium values of $\theta_1$ and $\theta_2$.

**Exercise 133 (Discrete-time stabilization of cart pendulum system via state feedback.)** Carry out the following for parameter set P4 and equilibriums $E1$ and $E2$. Illustrate the effectiveness of your controllers with numerical simulations.

Using eigenvalue placement techniques, obtain a discrete-time state feedback controller which stabilizes the nonlinear system about the equilibrium.

What is the largest value of $\delta$ (in degrees) for which your controller guarantees convergence of the closed loop system to the equilibrium for initial condition

$$(y,\, \theta_1,\, \theta_2,\, \dot{y},\, \dot{\theta}_1,\, \dot{\theta}_2)(0) = (0,\, \theta_1^e{-}\delta,\, \theta_2^e{+}\delta,\, 0,\, 0,\, 0)$$

where $\theta_1^e$ and $\theta_2^e$ are the equilibrium values of $\theta_1$ and $\theta_2$.

## 15.7 Appendix

**Proof of Fact 10.** We need to prove that for any two matrices $M$ and $N$ of dimensions $m \times n$ and $n \times m$, respectively, we have

$$\det(I + MN) = \det(I + NM)\,.$$

Consider the following matrix equation:

$$\begin{pmatrix} I & M \\ -N & I \end{pmatrix} \begin{pmatrix} I & O \\ N & I \end{pmatrix} = \begin{pmatrix} I + MN & M \\ O & I \end{pmatrix} \tag{15.15}$$

Since the matrix on the right of the equality is block diagonal, its determinant is $\det(I)\det(I + MN) = \det(I + MN)$. Hence the product of the determinants of the two matrices on the left of the equation is $\det(I + MN)$. Reversing the order of the two matrices on the left of (15.15), we obtain

$$\begin{pmatrix} I & O \\ N & I \end{pmatrix} \begin{pmatrix} I & M \\ -N & I \end{pmatrix} = \begin{pmatrix} I & M \\ O & I + NM \end{pmatrix}$$

From this we can conclude that the product of the determinants of the two matrices on the left of equation (15.15) is $\det(I+NM)$. Hence we obtain the desired result that $\det(I+MN) = \det(I + NM)$.

# Chapter 16

# Detectability and observers

## 16.1 Observers, state estimators and detectability

Consider the system (we will call it the plant)

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y &= Cx + Du \end{aligned} \tag{16.1}$$

Suppose that at each instant time $t$ we can measure the plant output $y(t)$ and the plant input $u(t)$ and we wish to estimate the plant state $x(t)$. In this section we demonstrate how to obtain an estimate $\hat{x}(t)$ of the plant state $x(t)$ with the property that as $t \to \infty$ the state estimation error $\hat{x}(t) - x(t)$ goes to zero. We do this by constructing an observer or state estimator.

**Observer or state estimator.** An observer or state estimator for a plant is a system whose inputs consist of the plant input and plant output while its output is an estimate of the plant state; see Figure 16.1.

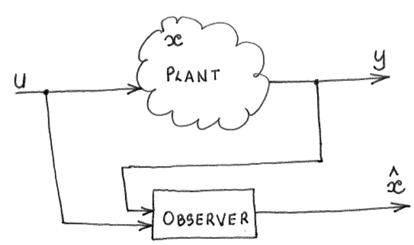



Figure 16.1: Observer

We consider here observers which have the following structure.

$$\boxed{\begin{aligned} \dot{\hat{x}} &= A\hat{x} + Bu + L(\hat{y} - y) \\ \hat{y} &= C\hat{x} + Du \end{aligned}} \qquad \hat{x}(0) = \hat{x}_0 \tag{16.2}$$

where the $n$-vector $\hat{x}(t)$ is the estimated state; the matrix $L$ is called the observer gain matrix and is yet to be determined. The initial state $\hat{x}_0$ for the observer is arbitrary. One can regard $\hat{x}_0$ as an initial guess of the initial plant state $x_0$.

*Note that the observer consists of a copy of the plant plus the "correction term,"* $L(\hat{y}-y)$.

If we rewrite the observer description as

$$\dot{\hat{x}} = (A+LC)\hat{x} + (B+LD)u - Ly\,,$$

it should be clear that we can regard the observer as a linear system whose inputs are the plant input $u$ and plant output $y$. We can regard the estimated state $\hat{x}$ as the observer output.

To discuss the behavior of the above estimator, we introduce the state estimation error

$$\tilde{x} := \hat{x} - x$$

Using (16.1) and (16.2) one can readily show that the evolution of the estimation error is simply governed by

$$\boxed{\dot{\tilde{x}} = (A+LC)\tilde{x}}$$

Hence, if one can choose $L$ such that all the eigenvalues of the matrix $A+LC$ have negative real parts then, we obtain the desired result that

$$\lim_{t\to\infty} \tilde{x}(t) = 0\,.$$

In other words, the state estimation error goes to zero as $t$ goes to infinity. When this happens, we call the observer an asymptotic observer. This leads to the next definition.

**DEFN.** *System (16.1) or the pair* $(C,\, A)$ *is* detectable *if there exists a matrix* $L$ *such that* $A+LC$ *is asymptotically stable.*

Note that the matrix $A+LC$ is asymptotically stable if and only if the matrix

$$(A+LC)^* = A^* + C^*L^*$$

is asymptotically stable. Thus, the problem of choosing a matrix $L$ so that $A+LC$ is asymptotically stable is equivalent to the stabilizability problem of choosing $K$ so that $A^* + C^*K$ is asymptotically stable. If one solves the stabilizability problem for $K$, then $L = K^*$ solves the original problem.

From the above observations, we obtain the following duality between stabilizability and detectability.

- $(C,\, A)$ is detectable if and only if $(A^*,\, C^*)$ is stabilizable.
- $(A,\, B)$ is stabilizable if and only if $(B^*,\, A^*)$ is detectable.
- Also, it follows that

$$\boxed{\text{observability} \qquad \Longrightarrow \qquad \text{detectability}}$$

## 16.2 Eigenvalue placement for estimation error dynamics

Suppose $(C,\, A)$ is a real observable pair. Then $(A^*,\, C^*)$ is a real controllable pair. Now consider any bunch of complex numbers

$$\{\, \lambda_1, \lambda_2, \cdots, \lambda_n \,\}$$

with the property that if $\lambda$ is in the bunch then so is $\bar{\lambda}$. Since $(A^*,\, C^*)$ is real and controllable, there is a real gain matrix $K$ such that $\{\, \lambda_1, \lambda_2, \cdots, \lambda_n \,\}$ are precisely the eigenvalues of $A^* + C^*K$. Letting $L = K^*$ and recalling that a matrix and its transpose have the same eigenvalues, we obtain that the eigenvalues of $A + LC$ are precisely $\{\, \lambda_1, \lambda_2, \cdots, \lambda_n \,\}$. We have just demonstrated the following result.

*If $(C, A)$ is observable, one can arbitrarily place the eigenvalues of $A + LC$ by appropriate choice of the observer gain matrix $L$.*

**Example 198** *(The unattached mass with position measurement and an input.)*

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= u \\ y &= x_1 \end{aligned}$$

Here

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad C = \begin{pmatrix} 1 & 0 \end{pmatrix}$$

Since this system is observable, it is detectable. If

$$L = \begin{pmatrix} l_1 \\ l_2 \end{pmatrix}$$

we have

$$A + LC = \begin{pmatrix} l_1 & 1 \\ l_2 & 0 \end{pmatrix}$$

Hence,

$$\det(sI - A - LC) = s^2 - l_1 s - l_2$$

and $A + LC$ is asymptotically stable if

$$l_1 < 0 \qquad l_2 < 0$$

An asymptotic observer is then given by

$$\begin{aligned} \dot{\hat{x}}_1 &= \hat{x}_2 + l_1(\hat{x}_1 - y) \\ \dot{\hat{x}}_2 &= u + l_2(\hat{x}_1 - y) \end{aligned}$$

and the estimation error $\tilde{x} = \hat{x} - x$ satisfies

$$\begin{aligned} \dot{\tilde{x}}_1 &= l_1\tilde{x}_1 + \tilde{x}_2 \\ \dot{\tilde{x}}_2 &= l_2\tilde{x}_1 \end{aligned}$$

## 16.3 Unobservable modes (you can't see them)

Suppose system (16.1) is unobservable. Then, as we have already seen, this system will have at least one unobservable eigenvalue $\lambda$. This means that, when $u(t) \equiv 0$, the system has a nonzero solution of the form $x^{uo}(t) = e^{\lambda t}v$ which produces zero output, $y(t) \equiv 0$. Thus one cannot distinguish this state motion from the zero motion, $x(t) \equiv 0$. If $\Re(\lambda) < 0$ then $x^{uo}(t)$ asymptotically goes to zero. However, if $\Re(\lambda) \geq 0$ then, $x^{uo}(t)$ does not asymptotically go to zero. Hence, even using nonlinear observers, there is no way that one can asymptotically estimate the state of this system. For linear observers, we have the following result.

*If $\lambda$ an unobservable eigenvalue of $(C, A)$, then $\lambda$ is an eigenvalue of $A + LC$ for every observer gain matrix $L$ .*

To see this, recall that $\lambda$ is an unobservable eigenvalue of $(C, A)$ if and only if

$$\operatorname{rank} \begin{pmatrix} A - \lambda I \\ C \end{pmatrix} < n$$

where $A$ is $n \times n$. Hence, there is a nonzero $n$-vector $v$ such that

$$\begin{pmatrix} A - \lambda I \\ C \end{pmatrix} v = 0$$

or, equivalently,

$$\begin{aligned} Av &= \lambda v \\ Cv &= 0 \end{aligned}$$

Hence, for any matrix $L$

$$(A + LC)v = \lambda v$$

Since $v \neq 0$, it follows that $\lambda$ is an eigenvalue of $A + LC$.

The next result states that a necessary and sufficient condition for stabilizability is that all unobservable eigenvalues are asymptotically stable.

**Theorem 27 (PBH detectability theorem)** *The pair $(C, A)$ is detectable if and only if*

$$\mathit{rank} \begin{pmatrix} A - \lambda I \\ C \end{pmatrix} = n$$

*for every eigenvalue $\lambda$ of $A$ with nonnegative real part.*

Proof. Use duality and the corresponding result for stabilizability.

**Example 199** (The unattached mass with velocity measurement.)

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= 0 \\ y &= x_2 \end{aligned}$$

Here

$$A=\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad C=\begin{pmatrix} 0 & 1 \end{pmatrix}$$

Hence

$$\begin{pmatrix} A-\lambda I \\ C \end{pmatrix}=\begin{pmatrix} -\lambda & 1 \\ 0 & -\lambda \\ 0 & 1 \end{pmatrix}$$

The above matrix has rank $1<n$ for $\lambda=0$; hence $\lambda=0$ is an unobservable eigenvalue and this system in not detectable.

Note that for any observer gain matrix

$$L=\begin{pmatrix} l_1 \\ l_2 \end{pmatrix}$$

we have

$$A+LC=\begin{pmatrix} 0 & 1+l_1 \\ 0 & l_2 \end{pmatrix}$$

Hence,

$$\det(sI-A-LC)=s(s-l_2)$$

So, regardless of choice of $L$, the matrix $A+LC$ always has an eigenvalue at 0.

## 16.4 Observable canonical form*

Suppose $A$ and $C$ have the following structure:

$$A=\begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \cdots & 0 & -a_2 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \cdots & 1 & -a_{n-1} \end{pmatrix} \qquad \text{and} \qquad C=\begin{pmatrix} 0 & 0 & \cdots & 0 & 1 \end{pmatrix}.$$

Then we say that the pair $(C,A)$ is in **observable canonical form.** Note that $(C,A)$ is in observable canonical form if and only if $(A^*,C^*)$ is in controllable canonical form. If $(C,A)$ is in observable canonical form, then $\dot{x}=Ax$ and $y=Cx$ look like

$$\begin{aligned} \dot{x}_1 &= -a_0x_n \\ \dot{x}_2 &= -a_1x_n + x_1 \\ &\vdots \\ \dot{x}_n &= -a_{n-1}x_n + x_{n-1} \\ \\ y &= x_n \end{aligned}$$

Also,

$$\det(sI-A)=s^n+a_{n-1}s^{n-1}+\cdots+a_0\,.$$

If we reverse the ordering of the states, then

$$\begin{array}{rclcl}
\dot{x}_1 & = & -a_{n-1}x_1 & + & x_2 \\
 & \vdots & & & \\
\dot{x}_{n-1} & = & -a_1x_1 & + & x_n \\
\dot{x}_n & = & -a_0x_1 & & \\
 & & & & \\
y & = & x_1 & &
\end{array}$$

and the corresponding $A$ and $C$ matrices have the following structure:

$$A = \begin{pmatrix} -a_{n-1} & 1 & 0 & \cdots & 0 \\ -a_{n-2} & 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & & \vdots \\ -a_1 & 1 & 0 & \cdots & 1 \\ -a_0 & 0 & 0 & \cdots & 0 \end{pmatrix} \qquad \text{and} \qquad C = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \end{pmatrix}$$

## 16.5 Discrete-time systems

The results for discrete-time systems are basically the same as those for continuous-time systems. Consider a discrete-time system described by

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) \\ y(k) &= Cx(k) + Du(k) \end{aligned} \tag{16.3}$$

**Observer or state estimator.** Recalling the observer structure in the continuous-time case, we consider here observers which have the following structure.

$$\boxed{\begin{aligned} \hat{x}(k+1) &= A\hat{x}(k) + Bu(k) + L(\hat{y}(k) - y(k)) \\ \hat{y}(k) &= C\hat{x}(k) + Du(k) \end{aligned}} \qquad \hat{x}(0) = \hat{x}_0 \tag{16.4}$$

where the $n$-vector $\hat{x}(k)$ is the estimated state; the matrix $L$ is called the observer gain matrix.

Introducing the state estimation error

$$\tilde{x} := \hat{x} - x$$

and using (16.3) and (16.4) one can readily show that the evolution of the estimation error is simply governed by

$$\boxed{\tilde{x}(k+1) = (A + LC)\tilde{x}(k)}$$

Hence, if one can choose $L$ such that all the eigenvalues of the matrix $A+LC$ have magnitude less than one then, we obtain the desired result that

$$\lim_{t\to\infty} \tilde{x}(k) = 0\,.$$

**Dead beat observer.** If $(C, A)$ is observable, choose $L$ so that all the eigenvalues of $A+LC$ are zero. Then, the state estimation error goes to zero in at most $n$ steps.

# Exercises

**Exercise 134** *BB in laundromat:* Obtain a state space representation of the following system.

$$
\begin{aligned}
m\ddot{\phi}_1 - m\Omega^2\phi_1 + \tfrac{k}{2}(\phi_1 - \phi_2) &= 0 \\
m\ddot{\phi}_2 - m\Omega^2\phi_2 - \tfrac{k}{2}(\phi_1 - \phi_2) &= u \\
y &= \phi_2
\end{aligned}
$$

Show that this system is observable. Considering

$$\Omega = 1,\ m = 1,\ k = 2$$

obtain a state estimator (observer) which results in an estimation error system with eigenvalues

$$-1, -2, -3, -4$$

Use the following methods.

(a) Obtaining an expression for the characteristic polynomial of the error system in terms of the components of the observer gain matrix.

(b) `acker`

(c) `place`

Illustrate the effectiveness of your observer(s) with numerical simulations.

**Exercise 135** Consider the system described by

$$
\begin{aligned}
\dot{x}_1 &= -x_2 + u \\
\dot{x}_2 &= -x_1 - u \\
y &= x_1 - x_2
\end{aligned}
$$

where all quantities are scalars.

(a) Is this system observable?

(b) Is this system detectable?

(c) Does there exist an asymptotic state estimator for this system? If an estimator does not exist, explain why; if one does exist, give an example of one.

# Chapter 17

# Climax: output feedback controllers

Plant. (The object of your control desires.) Consider a system described by

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y &= Cx \end{aligned} \tag{17.1}$$

where the state $x(t)$ is an $n$-vector, the control input $u(t)$ is an $m$-vector, and the measured output $y(t)$ is a $p$-vector. Initially, we consider $D = 0$ for simplicity of presentation.

## 17.1 Memoryless (static) output feedback

The simplest type of controller is a memoryless (or static) linear output feedback controller; this is of the form

$$\boxed{u = Ky}$$

where $K$ is a real $m \times p$ matrix, sometimes called a gain matrix. This controller results in the following closed loop system:

$$\dot{x} = (A + BKC)x$$

If the open loop system $\dot{x} = Ax$ is not asymptotically stable, a natural question is whether one can choose $K$ so that the closed loop system is asymptotically stable. For full state feedback ($C = I$), we have seen that it is possible to do this if $(A,\, B)$ is controllable or, less restrictively, if $(A, B)$ is stabilizable. If $(C,\, A)$ is observable or detectable, we might expect to be able to stabilize the plant with static output feedback. This is not the case as the following example illustrates.

**Example 200** Unattached mass with position measurement.

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= u \\ y &= x_1 \end{aligned}$$

This system is both controllable and observable. All linear static output feedback controllers are given by

$$u = ky$$

which results in

$$
\begin{aligned}
\dot{x}_1 &= x_2 \\
\dot{x}_2 &= kx_1
\end{aligned}
$$

Such a system is never asymptotically stable.

Note that if the plant has some damping in it, that is,

$$
\begin{aligned}
\dot{x}_1 &= x_2 \\
\dot{x}_2 &= -dx_2 + u \\
y &= x_1
\end{aligned}
$$

where $d > 0$, the closed loop system is given by

$$
\begin{aligned}
\dot{x}_1 &= x_2 \\
\dot{x}_2 &= kx_1 - dx_2
\end{aligned}
$$

This is asymptotically stable provided $k < 0$.

**Example 201** Consider

$$
\begin{aligned}
\dot{x}_1 &= x_2 \\
\dot{x}_2 &= x_1 + u \\
y &= -2x_1 + x_2
\end{aligned}
$$

For a SISO system, one could use root locus techniques to determine whether or not the system is stabilizable via static output feedback. For a general MIMO system, there are currently are no easily verifiable conditions which are both necessary and sufficient for stabilizability via static output feedback. It is a topic of current research. So where do we go now?

## 17.2 Dynamic output feedback

A dynamic output feedback controller is a dynamic input-output system whose input is the measured output of the plant and whose output is the control input to the plant. Thus, a linear, time-invariant, dynamic, output-feedback controller is described by

$$
\boxed{
\begin{aligned}
\dot{x}_c &= A_c x_c + B_c y \\
u &= C_c x_c + D_c y
\end{aligned}
}
\tag{17.2}
$$

where the $n_c$-vector $x_c(t)$ is called the **controller state** and $n_c$ is the **order of the controller.** We regard a memoryless controller as a controller of order zero.

Using transfer functions, the above controller can be described by

$$\hat{u}(s) = \hat{G}_c(s)\hat{y}(s)$$

where

$$\hat{G}_c(s) = C_c(sI - A_c)^{-1}B_c + D_c$$

**Closed loop system.** Application of the dynamic controller (17.2) to the plant (17.1) yields the following closed loop system:

$$\begin{aligned} \dot{x} &= (A + BD_cC)x &+ \; BC_cx_c \\ \dot{x}_c &= B_cCx &+ \; A_cx_c \end{aligned} \tag{17.3}$$

This is an LTI system whose state is

$$\begin{pmatrix} x \\ x_c \end{pmatrix}$$

and whose "A-matrix" is given by

$$\mathcal{A} = \begin{pmatrix} A + BD_cC & BC_c \\ B_cC & A_c \end{pmatrix}$$

So the order of the closed loop system is $n + n_c$.

**Example 202** Consider a SISO system and recall that a PI (proportional integral) controller can be described by

$$u(t) = -k_P y(t) - k_I \int_0^t y(\tau)\, dt$$

Letting

$$x_c(t) = \int_0^t y(\tau)\, dt$$

it can readily be seen that this controller is a first order dynamic system described by

$$\begin{aligned} \dot{x}_c &= y \\ u &= -k_I x_c - k_P y \end{aligned}$$

with initial condition $x_c(0) = 0$.

**Exercise 136** Show that the unattached mass with position feedback can be stabilized with a first order dynamic output feedback controller. What controller parameters place all the eigenvalues of the closed loop system at $-1$?

We now demonstrate the following result:

*If a complex number $\lambda$ is an unobservable eigenvalue of $(C, A)$ then $\lambda$ is an eigenvalue of $\mathcal{A}$.*

To see this, suppose $\lambda$ is an unobservable eigenvalue of $(C,\, A)$. Then

$$\text{rank} \begin{pmatrix} A - \lambda I \\ C \end{pmatrix} < n$$

Since the above matrix has $n$ columns, it must have non-zero nullity; this means there exists a nonzero $n$-vector $v$ with

$$\begin{pmatrix} A - \lambda I \\ C \end{pmatrix} v = 0$$

or, equivalently,

$$\begin{aligned} Av &= \lambda v \\ Cv &= 0 \end{aligned}$$

Letting $\clubsuit$ be the $n + p$ vector

$$\clubsuit = \begin{pmatrix} v \\ 0 \end{pmatrix}$$

it should be clear that

$$\mathcal{A}\clubsuit = \lambda\clubsuit$$

Since $\clubsuit \neq 0$, it follows that $\lambda$ is an eigenvalue of $\mathcal{A}$ ■

**Exercise 137** Prove the following statement. If $\lambda$ is an uncontrollable eigenvalue of $(A, B)$, then $\lambda$ is an eigenvalue of $\mathcal{A}$.

We can now state the following intuitively appealing result.

**Lemma 21** *If either $(A, B)$ is not stabilizable or $(C, A)$ is not detectable, then plant (17.1) is not stabilizable by a LTI dynamic output feedback controller of any order.*

The last lemma states that stabilizability of $(A, B)$ and and detectability of $(C, A)$ is necessary for stabilizability via linear dynamic output feedback control. In the next section, we demonstrate that if these conditions are satisfied then closed loop asymptotic stability can be achieved with a controller of order no more than the plant order.

## 17.3 Observer based controllers

Consider a general LTI plant described by

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y &= Cx + Du \end{aligned} \tag{17.4}$$

An observer based controller has the following structure:

$$\boxed{\begin{aligned} \dot{\hat{x}} &= A\hat{x} + Bu + L(C\hat{x} + Du - y) \\ u &= K\hat{x} \end{aligned}} \tag{17.5}$$

Note that this controller is completely specified by specifying the gain matrices $K$ and $L$; also this controller can be written as

$$\begin{aligned}\dot{\hat{x}} &= [A + BK + L(C + DK)]\hat{x} - Ly \\ u &= K\hat{x}\end{aligned}$$

This is a dynamic output feedback controller with state $x_c = \hat{x}$ (hence $n_c = n$, that is, the controller has same order as plant) and $A_c = A + BK + L(C + DK)$, $B_c = -L$, $C_c = K$, and $D_c = 0$.

*Closed loop system.* Combining the plant description (17.1) with the controller description (17.5), the closed loop system can be described by

$$\begin{aligned}\dot{x} &= Ax \quad + \quad BK\hat{x} \\ \dot{\hat{x}} &= -LCx \quad + \quad (A + BK + LC)\hat{x}\end{aligned}$$

Let $\tilde{x}$ be the state estimation error, that is,

$$\tilde{x} = \hat{x} - x\,. \tag{17.6}$$

Then the closed loop system is described by

$$\begin{aligned}\dot{x} &= (A + BK)x + BK\tilde{x} \\ \dot{\tilde{x}} &= (A + LC)\tilde{x}\end{aligned} \tag{17.7}$$

This is an LTI system with "A matrix"

$$\mathcal{A} = \begin{pmatrix} A + BK & BK \\ 0 & A + LC \end{pmatrix}$$

Noting that

$$\det(sI - \mathcal{A}) = \det(sI - A - BK)\det(sI - A - LC) \tag{17.8}$$

it follows that the set of eigenvalues of the closed loop system are simply the union of those of $A + BK$ and those of $A + LC$. So, if both $A + BK$ and $A + LC$ are asymptotically stable, the closed loop system is asymptotically stable. If $(A,\, B)$ is stabilizable, one can choose $K$ so that $A + BK$ is asymptotically stable. If $(C,\, A)$ is detectable, one can choose $L$ so that $A + LC$ is asymptotically stable. Combining these observations with Lemma 21 leads to the following result.

**Theorem 28** *The following statements are equivalent.*

*(a)* $(A, B)$ *is stabilizable and* $(C, A)$ *is detectable.*

*(b) Plant (17.1) is stabilizable via a LTI dynamic output feedback controller.*

*(c) Plant (17.1) is stabilizable via a LTI dynamic output feedback controller whose order is less than or equal to that of the plant.*

**Example 203** The unattached mass with position measurement

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= u \\ y &= x_1 \end{aligned}$$

Observer based controllers are given by

$$\begin{array}{rcllll} \dot{\hat{x}}_1 &=& \hat{x}_2 & + & l_1(\hat{x}_1 - y) \\ \dot{\hat{x}}_2 &=& u & & l_2(\hat{x}_1 - y) \\ u &=& k_1\hat{x}_1 & + & k_2\hat{x}_2 \end{array}$$

which is the same as

$$\begin{array}{rcllllll} \dot{\hat{x}}_1 &=& l_1\hat{x}_1 & + & \hat{x}_2 & - & l_1 y \\ \dot{\hat{x}}_2 &=& (k_1 + l_2)\hat{x}_1 & + & k_2\hat{x}_2 & - & l_2 y \end{array}$$

The closed loop system is described by

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= k_1x_1 + k_2x_2 + k_1\tilde{x}_1 + k_2\tilde{x}_2 \\ \dot{\tilde{x}}_1 &= l_1\tilde{x}_1 + \tilde{x}_2 \\ \dot{\tilde{x}}_1 &= l_2\tilde{x}_1 \end{aligned}$$

where $\tilde{x}_i = \hat{x}_i - x_i$, $i = 1, 2$ are the estimation error state variables. The characteristic polynomial of the closed loop system is

$$p(s) = (s^2 - k_2 s - k_1)(s^2 - l_1 s - l_2)$$

Hence, if

$$k_1, k_2, l_1, l_2 < 0$$

the closed loop system is asymptotically stable.

## 17.4 Discrete-time systems

Basically, everything said above for continuous-time systems also holds for discrete-time systems. The only differences are the way systems are described are the characterization of stability, stabilizability, and detectability in terms of eigenvalues.

*Plant.*

$$\begin{aligned} x(k+1) &= Ax(k)+Bu(k) && (17.9a)\\ y(k) &= Cx(k) && (17.9b) \end{aligned}$$

with state $x(k) \in \mathbb{R}^n$, control input $u(k) \in \mathbb{R}^m$, and measured output $y(k) \in \mathbb{R}^p$.

### 17.4.1 Memoryless output feedback

A *memoryless (or static) linear output feedback controller* is of the form

$$\boxed{u(k) = Ky(k)}$$

where $K$ is a real $m \times p$ matrix, sometimes called a *gain matrix*. This controller results in the following *closed loop system*:

$$x(k+1) = (A+BKC)x(k)$$

### 17.4.2 Dynamic output feedback

In general, a *linear dynamic output feedback controller* is described by

$$\boxed{\begin{aligned} x_c(k+1) &= A_c x_c(k) + B_c y(k)\\ u(k) &= C_c x_c(k) + D_c y(k) \end{aligned}} \qquad (17.10)$$

where $x_c(k) \in \mathbb{R}^{n_c}$ is called the *controller state*; $n_c$ is the *order of the controller.*

*Closed loop system.*

$$\begin{aligned} x(k+1) &= (A+BD_cC)x(k) &+ \quad BC_cx_c(k)\\ x_c(k+1) &= B_cCx(k) &+ \quad A_cx_c(k) \end{aligned} \qquad (17.11)$$

This is an lti system with state $\begin{pmatrix} x \\ x_c \end{pmatrix}$ and "A-matrix"

$$\mathcal{A} = \begin{pmatrix} A+BD_cC & BC_c \\ B_cC & A_c \end{pmatrix}$$

### 17.4.3 Observer based controllers

Consider

$$
\begin{aligned}
x(k+1) &= Ax(k) + Bu(k) && (17.12a)\\
y(k) &= Cx(k) + Du(k) && (17.12b)
\end{aligned}
$$

An observer based controller has the following structure:

$$
\boxed{
\begin{aligned}
\hat{x}(k+1) &= A\hat{x}(k) + Bu + L\left[C\hat{x}(k) + Du(k) - y(k)\right]\\
u(k) &= K\hat{x}(k)
\end{aligned}
}
\tag{17.13}
$$

*Closed loop system.* Let $\tilde{x}$ be the estimation error, that is, $\tilde{x} = \hat{x} - x$. Then the closed loop system is described by

$$
\begin{aligned}
x(k+1) &= (A+BK)x(k) + BK\tilde{x}(k)\\
\tilde{x}(k+1) &= (A+LC)\tilde{x}(k)
\end{aligned}
\tag{17.14}
$$

This is an lti system with "A matrix"

$$
\mathcal{A} = \begin{pmatrix} A+BK & BK \\ 0 & A+LC \end{pmatrix}
$$

**Dead-beat controllers.** Place all eigenvalues of $A+BK$ and $A+LC$ at zero.

# Exercises

**Exercise 138** Consider the system,

$$\begin{aligned}\dot{x}_1 &= x_2 + u\\ \dot{x}_2 &= x_1\\ y &= x_1\end{aligned}$$

where all quantities are scalar. Obtain an output feedback controller which results in an asymptotically stable closed loop system.

**Exercise 139** Consider the system with input $u$, output $y$, and state variables $x_1, x_2, x_3, x_4$ described by

$$\begin{aligned}\dot{x}_1 &= -x_1\\ \dot{x}_2 &= x_1 - 2x_2\\ \dot{x}_3 &= x_1 - 3x_3 + u\\ \dot{x}_4 &= x_1 + x_2 + x_3 + x_4\\ y &= x_3\end{aligned}$$

(a) Is this system observable? (b) Is this system stabilizable using output feedback? Justify your answers.

**Exercise 140** Consider the system,

$$\begin{aligned}\dot{x}_1 &= x_2\\ \dot{x}_2 &= x_1 + u\\ y &= x_1\end{aligned}$$

where all quantities are scalar. Obtain an output feedback controller which results in an asymptotically stable closed loop system.

**Exercise 141** Consider the system

$$\begin{aligned}\dot{x}_1 &= -x_1 + x_3\\ \dot{x}_2 &= u\\ \dot{x}_3 &= x_2\\ y &= x_3\end{aligned}$$

with scalar control input $u$ and scalar measured output $y$.

(a) Obtain an observer-based output feedback controller which results in an asymptotically stable closed loop system.

(b) Can all the eigenvalues of the closed loop system be arbitrarily placed?

**Exercise 142** Consider the system,

$$\begin{aligned}\dot{x}_1 &= x_2 + u_1\\ \dot{x}_2 &= u_2\\ y &= x_1\end{aligned}$$

where all quantities are scalar. Obtain an output feedback controller which results in an asymptotically stable closed loop system.

**Exercise 143** *Stabilization of 2-link manipulator with dynamic output feedback.*

$$\begin{aligned}[m_1 lc_1^2 + m_2 l_1^2 + I_1]\ddot{q}_1 + [m_2 l_1 lc_2 \cos(q_1 - q_2)]\ddot{q}_2 + m_2 l_1 lc_2 \sin(q_1 - q_2)\dot{q}_2^2 \qquad&\\ - [m_1 lc_1 + m_2 l_1] g \sin(q_1) &= u\\[1em] [m_2 l_1 lc_2 \cos(q_1 - q_2)]\ddot{q}_1 + [m_2 lc_2^2 + I_2]\ddot{q}_2 - m_2 l_1 lc_2 \sin(q_1 - q_2)\dot{q}_1^2 - m_2 g lc_2 \sin(q_2) &= 0\\ y &= q_2\end{aligned}$$

Recall the linearizization of this system about

$$q_1 = q_2 = u = 0$$

and recall its state space description.

Use parameter values previously given.

Is the linearization observable? Design a observer based dynamic output feedback controller which asymptotically stabilizes this linearized system. Apply the controller to the nonlinear system and simulate the nonlinear closed loop system for various initial conditions. (Use zero initial conditions for the observer.)

# Chapter 18

# Constant output tracking in the presence of constant disturbances

## 18.1 Zeros

Consider a SISO system described by

$$\dot{x} = Ax + Bu \tag{18.1a}$$
$$y = Cx + Du \tag{18.1b}$$

where $u(t)$ and $y(t)$ are scalars while $x(t)$ is an $n$-vector. The transfer function $\hat{g}$ associated with this system is given by

$$\hat{g}(s) = C(sI - A)^{-1}B + D\,. \tag{18.2}$$

Recall that a scalar $\lambda$ is a zero of $\hat{g}$ if

$$\hat{g}(\lambda) = 0\,.$$

Since every rational scalar transfer function can be written as $\hat{g} = n/d$ where $n$ and $d$ are polynomials with no common zeros, it follows that the zeros of $\hat{g}$ are the zeros of $n$ while the poles of $g$ are the zeros of $d$. Hence, a scalar cannot be simultaneously a pole and a zero of $\hat{g}$. Also, the eigenvalues of $A$ which are neither uncontrollable nor unobservable are the poles of $\hat{g}$.

### 18.1.1 A state space characterization of zeros

For each scalar $\lambda$ we define $T_\lambda$ to be the $(n+1) \times (n+1)$ matrix given by

$$T_\lambda = \begin{pmatrix} A - \lambda I & B \\ C & D \end{pmatrix}. \tag{18.3}$$

This matrix yields the following state space characterization of zeros.

**Theorem 29** *Consider a SISO system described by (18.1) and let $T_\lambda$ be the matrix defined by (18.3). Then $\det T_\lambda = 0$ if and only if at least one of the following conditions hold.*

*(a) $\lambda$ is an uncontrollable eigenvalue of $(A, B)$.*

*(b) $\lambda$ is an unobservable eigenvalue of $(C, A)$.*

*(c) $\lambda$ is a zero of $\hat{g}$.*

PROOF. We first demonstrate that if at least one of conditions (a), (b), or (c) hold, then $\det T_\lambda = 0$. When $\lambda$ is an uncontrollable eigenvalue of $(A, B)$, we have

$$\operatorname{rank} \begin{pmatrix} A - \lambda I & B \end{pmatrix} < n\,.$$

This implies that $\operatorname{rank} T_\lambda < n + 1$; hence $\det T_\lambda = 0$. In a similar fashion, if $\lambda$ is an unobservable eigenvalue of $(C, A)$, then

$$\operatorname{rank} \begin{pmatrix} A - \lambda I \\ C \end{pmatrix} < n\,.$$

From this it also follows that $\det T_\lambda = 0$. Now consider the case in which $\lambda$ is a zero of $\hat{g}$ but is neither an uncontrollable eigenvalue of $(A, B)$ nor an unobservable eigenvalue of $(C, A)$. Since $\lambda$ is a zero of $\hat{g}$ it follows that $\lambda$ is not a pole of $\hat{g}$. Thus $\lambda$ is not an eigenvalue of $A$. Hence $\lambda I - A$ is invertible. Letting $x_0 = (\lambda I - A)^{-1} B$, we obtain that

$$(A - \lambda I)x_0 + B = 0$$

and

$$Cx_0 + D = C(\lambda I - A)^{-1}B + D = \hat{g}(\lambda) = 0\,.$$

This implies that

$$\begin{pmatrix} A - \lambda I & B \\ C & D \end{pmatrix} \begin{pmatrix} x_0 \\ 1 \end{pmatrix} = 0\,.$$

Hence, $\det T_\lambda = 0$.

We now show that if $\det T_\lambda = 0$, then at least one of conditions (a), (b), or (c) must hold. To this end suppose that $\det T_\lambda = 0$ and (a) and (b) do not hold. Then we need to show that (c) holds. Since $\det T_\lambda = 0$, there is an $n$-vector $x_0$ and a scalar $u_0$, such that

$$\begin{pmatrix} A - \lambda I & B \\ C & D \end{pmatrix} \begin{pmatrix} x_0 \\ u_0 \end{pmatrix} = 0\,. \tag{18.4}$$

and $u_0$ and $x_0$ cannot both be zero. We now claim that $u_0$ cannot be zero. If $u_0 = 0$, then $x_0 \neq 0$ and

$$\begin{pmatrix} A - \lambda I \\ C \end{pmatrix} x_0 = 0\,.$$

This contradicts the assumption that $\lambda$ is not an unobservable eigenvalue of $(C, A)$. Hence $u_0$ is non-zero and, by scaling $x_0$, we can consider $u_0 = 1$. It follows from (18.4) that

$$(A - \lambda I)x_0 + B = 0 \tag{18.5a}$$

$$Cx_0 + D = 0 \tag{18.5b}$$

Hence, $B = -(A - \lambda I)x_0$, that is $B$ is a linear combination of the columns of $A - \lambda I$. This implies that

$$\operatorname{rank} \begin{pmatrix} A - \lambda I & B \end{pmatrix} = \operatorname{rank}(A - \lambda I)\,.$$

Using this conclusion and the fact that $\lambda$ is not an uncontrollable eigenvalue of $(A, B)$, we obtain that

$$n = \operatorname{rank} \begin{pmatrix} A - \lambda I & B \end{pmatrix} = \operatorname{rank}(A - \lambda I)\,.$$

Since $\operatorname{rank}(A - \lambda I) = n$, it follows that $\lambda$ is not an eigenvalue of $A$. It now follows from (18.5a) that $x_0 = (\lambda I - A)^{-1}B$; hence, using (18.5a), we obtain

$$0 = Cx_0 + D = C(\lambda I - A)^{-1}B + D = \hat{g}(\lambda)\,,$$

that is, $\lambda$ is a zero of $\hat{g}$. ∎

**Corollary 4** *Suppose $\{A, B, C, D\}$ is a minimal realization of a SISO transfer function $\hat{g}$. Then a scalar $\lambda$ is a zero of $\hat{g}$ if and only if*

$$\det \begin{pmatrix} A - \lambda I & B \\ C & D \end{pmatrix} = 0\,.$$

Proof. If $\{A, B, C, D\}$ is a minimal realization of a SISO transfer function then, $(A, B)$ is controllable and $(C, A)$ is observable. The result now follows from Theorem 29.

**Example 204** Consider the transfer function

$$\hat{g}(s) = \frac{s^2 - 1}{s^2 + 1}\,.$$

The zeros of this transfer function are 1 and $-1$. A minimal realization of this transfer function is given by

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad B = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \qquad C = \begin{pmatrix} -2 & 0 \end{pmatrix}, \qquad D = 1\,.$$

For this realization,

$$\det T_\lambda = \det \begin{pmatrix} A - \lambda I & B \\ C & D \end{pmatrix} = \begin{pmatrix} -\lambda & 1 & 0 \\ -1 & -\lambda & 1 \\ -2 & 0 & 1 \end{pmatrix} = \lambda^2 - 1\,.$$

Clearly, the zeros of $\hat{g}$ are given by $\det T_\lambda = 0$.


Figure 18.1: A block diagram

## 18.2 Tracking and disturbance rejection with state feedback

Consider a system described by

$$\begin{aligned} \dot{x} &= Ax + B_1 w + B_2 u \\ z &= Cx + D_1 w + D_2 u \end{aligned} \tag{18.6}$$

where the $n$-vector $x(t)$ is the state; the $m_1$-vector $w$ is an unknown, constant, disturbance input; the $m_2$-vector $u(t)$ is the control input; and the $p$-vector $z(t)$ is a performance output. Initially, we assume that the state $x$ and the performance output $z$ can be measured.

The control problem we consider here is the following: Suppose $r$ is any constant reference output; we wish to design a feedback controller such that for all initial conditions $x(0) = x_0$, we have

$$\lim_{t\to\infty} z(t) = r$$

and the state $x(t)$ is bounded for all $t \geq 0$.

**Assumptions.** We will assume that the pair $(A, B_2)$ is stabilizable. We will also need the following condition:

$$\boxed{\operatorname{rank}\begin{pmatrix} A & B_2 \\ C & D_2 \end{pmatrix} = n + p} \tag{18.7}$$

that is, the the rank of the above matrix must equal the number of its rows.

For a system with scalar control input $u$ and scalar output $z$, the above condition is equivalent to

$$\det\begin{pmatrix} A & B_2 \\ C & D_2 \end{pmatrix} \neq 0\,.$$

By Theorem 29, this is equivalent to the requirement that 0 is neither an uncontrollable eigenvalue of $(A, B_2)$, an unobservable eigenvalue of $(C, A)$, nor a zero of the transfer function $T_{zu}(s) = C(sI - A)^{-1}B_2 + D_2$.

Why do we need the condition in (18.7)? *If it does not hold then, there is a reference output $r_0$ such that one cannot achieve $\lim_{t\to\infty} z(t) = r_0$ with a bounded state history.* To see this, suppose the above condition does not hold. Then there is a nonzero vector

$$\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

with $v_1 \in R^n$ and $v_2 \in \mathbb{R}^p$ such that

$$\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}^* \begin{pmatrix} A & B_2 \\ C & D_2 \end{pmatrix} = 0\,,$$

that is,

$$v_1^*A + v_2^*C = 0\,, \qquad v_1^*B_2 + v_2^*D_2 = 0$$

Since $(A,\, B_2)$ is stabilizable, we must have $v_2 \neq 0$. (Why?) Consider $w = 0$, premultiply the first and second equations in the system description by $v_1^*$ and $v_2^*$, respectively, and add to obtain

$$v_1^*\dot{x} + v_2^*z = 0\,,$$

that is,

$$\frac{dv_1^*x}{dt} = -v_2^*z$$

Suppose

$$\lim_{t\to\infty} z(t) = r_0 := -v_2$$

Then

$$\frac{dv_1^*x}{dt} = ||v_2||^2 - v_2^*(z - r_0);$$

Since $v_2 \neq 0$ and $\lim_{t\to\infty}(z(t) - r_0) = 0$, it follows that there exists a time $T$ and a positive number $\epsilon$ such that

$$||v_2||^2 - v_2^*(z(t) - r_0) \geq \epsilon \quad \text{for all} \quad t \geq T$$

Hence, for $t \geq T$,

$$v_1^*x(t) \geq v_1^*x(T) + \epsilon(t - T)$$

This implies that $v_1^*x$ is unbounded, So, $x$ is unbounded. ■

**Example 205** Consider

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -x_1 - x_2 + u \\ z &= x_2 \end{aligned}$$

Here

$$\begin{pmatrix} A & B_2 \\ C & D_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ -1 & -1 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

Clearly rank condition (18.7) does not hold. It should be clear from the first differential equation describing this system that, in order to keep $z = x_2$ equal to a nonzero constant, $x_1$ must be unbounded. Note that the transfer function from $u$ to $z$, given by

$$\frac{s}{s^2 + s + 1}\,,$$

has a zero at 0.

• Since the rank of a matrix is always less than or equal to the number of its columns, it follows that if rank condition (18.7) holds then,

$$p \leq m_2 ,$$

that is, the number of performance output variables is less than or equal to the number of control input variables.

### 18.2.1 Nonrobust control

Consider first the question of the existence of steady state values of $x$ and $u$ such that $z$ has a steady state value of $r$. Letting $x(t) \equiv x^e$, $u(t) \equiv u^e$, and $z(t) \equiv r$ in system description (18.6) results in

$$\begin{pmatrix} A & B_2 \\ C & D_2 \end{pmatrix} \begin{pmatrix} x^e \\ u^e \end{pmatrix} = \begin{pmatrix} -B_1 w \\ -D_1 w + r \end{pmatrix}$$

Since the matrix $\begin{pmatrix} A & B_2 \\ C & D_2 \end{pmatrix}$ has full row rank, there exists a solution $(x^e, u^e)$ to the above equation. Consider now the controller

$$u = u^e + K[x - x^e]$$

Letting

$$\delta x := x - x^e$$

the closed loop system is described by

$$\begin{aligned} \delta \dot{x} &= [A + B_2 K]\delta x \\ z &= C\delta x + r \end{aligned}$$

Choosing $K$ so that $A+B_2K$ is asymptotically stable results in $\lim_{t\to\infty} \delta x(t) = 0$ and, hence, $\lim_{t\to\infty} z(t) = r$.

In addition to requiring knowledge of $w$, this approach also requires exact knowledge of the matrices $A, B_2, C, D_2$. Inexact knowledge of these matrices yields inexact values of the required steady values $x_e$ and $u_e$. This results in the steady state value of $z$ being offset from the desired value $r$.

### 18.2.2 Robust controller

**Integration of output error.** Introduce a new state variable $x_I$ given by

$$\boxed{\dot{x}_I = z - r}$$

where $x_I(0) = x_{I0}$ and $x_{I0}$ is arbitrary. Roughly speaking, this state is the integral of the output tracking error, $z - r$; specifically, it is given by

$$x_I(t) = x_{I0} + \int_0^t (z(\tau) - r)\, d\tau \,.$$

Introducing new state, disturbance input and performance output,

$$\tilde{x} := \begin{pmatrix} x \\ x_I \end{pmatrix}, \qquad \tilde{w} := \begin{pmatrix} w \\ r \end{pmatrix}, \qquad \tilde{z} := z - r\,,$$

respectively, we obtain the **augmented plant**,

$$\begin{array}{rcccccc} \dot{\tilde{x}} & = & \tilde{A}\tilde{x} & + & \tilde{B}_1\tilde{w} & + & \tilde{B}_2 u \\ \tilde{z} & = & \tilde{C}\tilde{x} & + & \tilde{D}_1\tilde{w} & + & \tilde{D}_2 u \end{array}$$

where

$$\tilde{A} = \begin{pmatrix} A & 0 \\ C & 0 \end{pmatrix} \qquad \tilde{B}_1 = \begin{pmatrix} B_1 & 0 \\ D_1 & -I \end{pmatrix} \qquad \tilde{B}_2 = \begin{pmatrix} B_2 \\ D_2 \end{pmatrix}$$

$$\tilde{C} = \begin{pmatrix} C & 0 \end{pmatrix} \qquad \tilde{D}_1 = \begin{pmatrix} D_1 & -I \end{pmatrix} \qquad \tilde{D}_2 = D_2$$

We have now the following result.

**Lemma 22** *The pair $(\tilde{A},\, \tilde{B}_2)$ is controllable (stabilizable) if and only if*

*(a) the pair $(A,\, B_2)$ is controllable (stabilizable), and*

*(b) rank condition (18.7) holds.*

Proof. The pair $(\tilde{A},\, \tilde{B}_2)$ is controllable (stabilizable) iff

$$\operatorname{rank}\begin{pmatrix} \tilde{A} - \lambda I & \tilde{B}_2 \end{pmatrix} = n + p$$

for all complex numbers $\lambda$ (for all complex $\lambda$ with $\Re(\lambda) \geq 0$). Also,

$$\begin{pmatrix} \tilde{A} - \lambda I & \tilde{B}_2 \end{pmatrix} = \begin{pmatrix} A - \lambda I & 0 & B_2 \\ C & -\lambda I & D_2 \end{pmatrix}$$

Considering $\lambda = 0$, we obtain

$$\operatorname{rank}\begin{pmatrix} \tilde{A} - \lambda I & \tilde{B}_2 \end{pmatrix} = \operatorname{rank}\begin{pmatrix} A & 0 & B_2 \\ C & 0 & D_2 \end{pmatrix} = \operatorname{rank}\begin{pmatrix} A & B_2 \\ C & D_2 \end{pmatrix}. \tag{18.8}$$

Considering $\lambda \neq 0$, we obtain

$$\begin{aligned} \operatorname{rank}\begin{pmatrix} \tilde{A} - \lambda I & \tilde{B}_2 \end{pmatrix} &= \operatorname{rank}\begin{pmatrix} A - \lambda I & 0 & B_2 \\ C & -\lambda I & D_2 \end{pmatrix} \\ &= \operatorname{rank}\begin{pmatrix} A - \lambda I & B_2 & 0 \\ 0 & 0 & -\lambda I \end{pmatrix} \\ &= \operatorname{rank}\begin{pmatrix} A - \lambda I & B_2 \end{pmatrix} + p \end{aligned} \tag{18.9}$$

Since the pair $(A,\, B_2)$ is controllable (stabilizable) iff $\operatorname{rank}\,[\, A - \lambda I \quad B_2 \,] = n$ for all complex numbers $\lambda$ (for all complex $\lambda$ with $\Re(\lambda) \geq 0$), the desired result now follows from (18.8) and (18.9). ■

**Controller design.** Suppose $(A,\, B_2)$ is stabilizable and rank condition (18.7) holds. Then $(\tilde{A},\, \tilde{B}_2)$ is also stabilizable. Choose *any* matrix $\tilde{K}$ such that $\tilde{A}+\tilde{B}_2\tilde{K}$ is asymptotically stable. Let

$$u = \tilde{K}\tilde{x}$$

If we partition $\tilde{K}$ as

$$\tilde{K} = \begin{pmatrix} K_P & K_I \end{pmatrix}$$

where $K_P$ is given by the first $n$ columns of $\tilde{K}$ and $K_I$ is given by the last $p$ columns of $\tilde{K}$, then the controller is given by

$$\boxed{\begin{array}{rcl} \dot{x}_I &=& z - r \\ u &=& K_P x + K_I x_I \end{array}}$$

Note that this controller can be written as

$$u = K_P x + K_I \int (z - r)$$

This is a generalization of the classical PI (Proportional Integral) Controller.

**Closed loop system.** The closed loop system is given by

$$\begin{array}{rcl} \dot{\tilde{x}} &=& [\tilde{A} + \tilde{B}_2\tilde{K}]\tilde{x} \quad + \quad \tilde{B}_1\tilde{w} \\ \tilde{z} &=& [\tilde{C} + \tilde{D}_2\tilde{K}]\tilde{x} \quad + \quad \tilde{D}_1\tilde{w} \end{array}$$

Since the matrix $\tilde{A} + \tilde{B}_2\tilde{K}$ is asymptotically stable and $\tilde{w}$ is constant, all solutions $\tilde{x}$ are bounded for $t \geq 0$. Thus, $x(t)$ is bounded for all $t \geq 0$.

Letting $\eta = \dot{\tilde{x}}$ and noting that $\tilde{w}$ is constant, we obtain that

$$\dot{\eta} = [\tilde{A} + \tilde{B}_2\tilde{K}]\eta\,.$$

Since the matrix $\tilde{A} + \tilde{B}_2\tilde{K}$ is asymptotically stable, we must have

$$\lim_{t\to\infty} \dot{\tilde{x}} = \lim_{t\to\infty} \eta(t) = 0\,.$$

In particular, we have

$$\lim_{t\to\infty} \dot{x}_I = 0\,;$$

hence

$$\lim_{t\to\infty} z(t) = r\,.$$

Thus, the proposed controller achieves the desired behavior.

**Example 206** Consider the constantly disturbed and controlled harmonic oscillator described by

$$\begin{array}{rcl} \dot{x}_1 &=& x_2 \\ \dot{x}_2 &=& -x_1 + w + u \\ z &=& x_1 \end{array}$$

Here

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad B_1 = B_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \qquad C = \begin{pmatrix} 1 & 0 \end{pmatrix}, \qquad D_1 = D_2 = 0$$

The pair $(A,\, B_2)$ is controllable and

$$\text{rank} \begin{pmatrix} A & B_2 \\ C & D_2 \end{pmatrix} \quad = \quad \text{rank} \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} = 3 = n + p$$

Hence

$$\tilde{A} = \begin{pmatrix} A & 0 \\ C & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \qquad \text{and} \qquad \tilde{B}_2 = \begin{pmatrix} B_2 \\ D_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

Letting

$$\tilde{K} = \begin{pmatrix} k_1 & k_2 & k_I \end{pmatrix}$$

we obtain

$$\tilde{A} + \tilde{B}_2\tilde{K} = \begin{pmatrix} 0 & 1 & 0 \\ -1 + k_1 & k_2 & k_I \\ 1 & 0 & 0 \end{pmatrix}$$

The characteristic polynomial $p$ of $\tilde{A} + \tilde{B}_2\tilde{K}$ is given by

$$p(s) = s^3 - k_2 s^2 + (1 - k_1)s - k_I$$

Choosing $k_1, k_2, k_I$ so that this polynomial has roots with negative real parts, a controller yielding constant output tracking, rejecting constant disturbances and yielding bounded states is given by

$$\begin{aligned} \dot{x}_I &= x_1 - r \\ u &= k_1 x_1 + k_2 x_2 + k_I x_I \end{aligned}$$

## 18.3 Measured Output Feedback Control

Consider a system described by

$$\begin{aligned} \dot{x} &= Ax + B_1 w + B_2 u \\ z &= C_1 x + D_{11} w + D_{12} u \\ y &= C_2 x + D_{21} w \end{aligned}$$

where the $n$-vector $x(t)$ is the state; the $m_1$-vector $w$ is an unknown, constant, disturbance input; the $m_2$-vector $u(t)$ is the control input; and the $p_1$-vector $z(t)$ is a performance output. We will assume $z(t)$ can be measured and that instead of $x$ we can only measure the measured output $y(t)$ which is a $p_2$-vector.

Suppose the $p_1$-vector $r$ is any constant reference output and we wish to design a controller such that for any initial condition $x(0) = x_0$,

$$\lim_{t \to \infty} z(t) = r$$

and the state $x(t)$ is bounded for all $t \geq 0$. To achieve this objective, we need the following assumptions.

**Assumptions.**

(a) The pair $(A,\, B_2)$ is stabilizable.

(b) The pair $(C_2,\, A)$ is detectable.

(c)

$$\operatorname{rank} \begin{pmatrix} A & B_2 \\ C_1 & D_{12} \end{pmatrix} = n + p_1$$

As before, we introduce a new state variable $x_I$ given by

$$\dot{x}_I = z - r$$

where $x_I(0)$ is arbitrary. Letting

$$\tilde{x} := \begin{pmatrix} x \\ x_I \end{pmatrix}, \qquad \tilde{w} := \begin{pmatrix} w \\ r \end{pmatrix}, \qquad \tilde{z} := z - r\,, \qquad \tilde{y} := \begin{pmatrix} y \\ x_I \\ r \end{pmatrix}$$

we obtain the augmented plant

$$\begin{array}{rcl} \dot{\tilde{x}} & = & \tilde{A}\tilde{x} + \tilde{B}_1\tilde{w} + \tilde{B}_2 u \\ \tilde{z} & = & \tilde{C}_1\tilde{x} + \tilde{D}_{11}\tilde{w} + \tilde{D}_{12} u \\ \tilde{y} & = & \tilde{C}_2\tilde{x} + \tilde{D}_{21}\tilde{w} \end{array} \tag{18.10}$$

with

$$\tilde{A} = \begin{pmatrix} A & 0 \\ C_1 & 0 \end{pmatrix} \qquad \tilde{B}_1 = \begin{pmatrix} B_1 & 0 \\ D_{11} & -I \end{pmatrix} \qquad \tilde{B}_2 = \begin{pmatrix} B_2 \\ D_{12} \end{pmatrix}$$

$$\tilde{C}_1 = \begin{pmatrix} C_1 & 0 \end{pmatrix} \qquad \tilde{D}_{11} = \begin{pmatrix} D_{11} & -I \end{pmatrix} \qquad \tilde{D}_{12} = D_{12}$$

$$\tilde{C}_2 = \begin{pmatrix} C_2 & 0 \\ 0 & I \\ 0 & 0 \end{pmatrix} \qquad \tilde{D}_{21} = \begin{pmatrix} D_{21} & 0 \\ 0 & 0 \\ 0 & I \end{pmatrix}$$

Consider now any output feedback controller of the form

$$\begin{array}{rcl} \dot{x}_c & = & A_c x_c + B_c \tilde{y} \\ u & = & C_c x_c + D_c \tilde{y} \end{array} \tag{18.11}$$

The resulting closed loop system is described by

$$\begin{pmatrix} \dot{\tilde{x}} \\ \dot{x}_c \end{pmatrix} = \underbrace{\begin{pmatrix} \tilde{A} + \tilde{B}_2 D_c \tilde{C}_2 & \tilde{B}_2 C_c \\ B_c \tilde{C}_2 & A_c \end{pmatrix}}_{A_{cl}} \begin{pmatrix} \tilde{x} \\ x_c \end{pmatrix} + \begin{pmatrix} \tilde{B}_1 + \tilde{B}_2 D_c \tilde{D}_{21} \\ D_c \tilde{D}_{21} \end{pmatrix} \tilde{w} \tag{18.12}$$

From the lemma of the previous section, assumptions (a) and (c) imply that the pair $(\tilde{A},\, \tilde{B}_2)$ is stabilizable. Also, detectability of the pair $(C_2, A)$ is equivalent to detectability of the pair $(\tilde{C}_2,\, \tilde{A})$. Hence, one can always choose matrices $A_c, B_c, C_c, D_c$ so that so that the closed loop system matrix $A_{cl}$ is asymptotically stable.

**Controller design.** Choose any matrices $A_c, B_c, C_c, D_c$ so that the closed loop system matrix $A_{cl}$ in (18.12) is asymptotically stable. Then the controller is given by (18.11).

If we let

$$\begin{aligned} B_c &= \begin{pmatrix} B_{cP} & B_{cI} & B_{cr} \end{pmatrix} \\ D_c &= \begin{pmatrix} B_{cP} & B_{cI} & B_{cr} \end{pmatrix} \end{aligned}$$

where the above partitions correspond to the partitions of $\tilde{y}$, we can express this controller as

$$\begin{aligned} \dot{x}_I &= z - r \\ \dot{x}_c &= A_c x_c + B_{cI} x_I + B_{cP} y + B_{cr} r \\ u &= C_c x_c + D_{cI} x_I + D_{cP} y + D_{cr} r \end{aligned}$$

Note that $A_{cl}$ does not depend on $B_{cr}$ and $D_{cr}$; hence these two matrices can be arbitrarily chosen.

**Closed loop system.** The closed loop system is described by (18.12). Since the closed loop system matrix $A_{cl}$ is asymptotically stable and $\tilde{w}$ is constant, it follows that $\tilde{x}$ and $x_c$ are bounded. Hence $x$ is bounded.

Letting

$$\eta = \begin{pmatrix} \dot{\tilde{x}} \\ \dot{x}_c \end{pmatrix}$$

and noting that $\tilde{w}$ is constant, we obtain that

$$\dot{\eta} = A_{cl} \eta$$

Since the matrix $A_{cl}$ is asymptotically stable, we must have

$$\lim_{t \to \infty} \begin{pmatrix} \dot{\tilde{x}} \\ \dot{x}_c \end{pmatrix} = \lim_{t \to \infty} \eta(t) = 0$$

In particular, we have

$$\lim_{t \to \infty} \dot{x}_I = 0\,;$$

hence

$$\lim_{t \to \infty} z(t) = r$$

### 18.3.1 Observer based controllers

As we have seen in a previous chapter, one approach to the design of stabilizing output feedback controllers is to combine an asymptotic state estimator with a stabilizing state feedback gain matrix. Since there is no need to estimate the state $x_I$ of the augmented plant (18.10), we design the observer to estimate only $x$. This yields the following design.

- Choose gain matrix $\tilde{K}$ so that the $\tilde{A} + \tilde{B}_2 \tilde{K}$ is asymptotically stable and let

$$\tilde{K} = \begin{pmatrix} K_P & K_I \end{pmatrix}$$

where $K_P$ is given by the first $n$ columns of $\tilde{K}$ and $K_I$ is given by the last $p_1$ columns of $\tilde{K}$.

- Choose observer gain matrix $L$ so that $A + LC_2$ is asymptotically stable.
- An observer based controller is given by

$$\boxed{\begin{array}{rcll} \dot{x}_I &=& z - r & \text{integrator} \\ \dot{x}_c &=& Ax_c + B_2u + L(C_2x_c - y) & \text{observer} \\ u &=& K_Px_c + K_Ix_I & \text{control input} \end{array}}$$

where $x_c(0)$ and $x_I(0)$ are arbitrary.

If we let

$$e := x_c - x$$

the closed loop system due to this controller is described by

$$\begin{pmatrix} \dot{\tilde{x}} \\ \dot{e} \end{pmatrix} = \begin{pmatrix} \tilde{A} + \tilde{B}_2\tilde{K} & * \\ 0 & A + LC_2 \end{pmatrix} \begin{pmatrix} \tilde{x} \\ e \end{pmatrix} + \begin{pmatrix} * \\ * \end{pmatrix} \tilde{w}$$

Since the eigenvalues of the closed loop system matrix

$$\begin{pmatrix} \tilde{A} + \tilde{B}_2\tilde{K} & * \\ 0 & A + LC_2 \end{pmatrix}$$

are the union of those of $\tilde{A} + \tilde{B}_2\tilde{K}$ and $A + LC_2$, it follows that this matrix is asympotically stable. Hence, this controller yields the desired behavior.

# Exercises

**Exercise 144** Consider a system described by

$$\hat{z}(s) = T_{zw}(s)\hat{w}(s) + T_{zu}(s)\hat{u}(s)$$

with

$$T_{zw}(s) = \frac{1}{s^2 - 1} \qquad \text{and} \qquad T_{zu}(s) = \frac{s - 2}{s^2 - 1}$$

where $\hat{z}$, $\hat{w}$, and $\hat{u}$ are the Laplace transforms of the performance output $z$, disturbance input $w$, and control input $u$, respectively. Let $r$ be a desired constant output and assume the disturbance input $w$ is an unknown constant. Design an output ($z$) feedback controller which guarantees that for all $r$ and $w$,

$$\lim_{t \to \infty} z(t) = r$$

and all signals are bounded. Illustrate your results with simulations.

**Exercise 145** Consider a system described by

$$\begin{array}{rcccccc} \dot{x} & = & Ax & + & B_1 w & + & B_2 u \\ y & = & Cx & + & D_1 w & + & D_2 u \end{array}$$

where $x(t)$, $u(t)$, and $w$, and $y(t) \in \mathbb{R}^p$ are as above. Suppose that we can only measure the control input $u$ and the measured output $y$. It is desired to obtain an asymptotic observer for $x$ and $w$; that is, we wish to design an observer producing estimates $\hat{x}$ and $\hat{w}$ of $x$ and $w$ with the property that for any initial value $x_0$ of $x$ and any $w$, one has

$$\lim_{t \to \infty} [\hat{x}(t) - x(t)] = 0\,, \qquad \lim_{t \to \infty} [\hat{w}(t) - w] = 0$$

(a) State the least restrictive conditions on the system which allows one to construct an asymptotic estimator.

(b) Give a procedure for constructing an asymptotic estimator.

(c) Illustrate your results with example(s) and simulation results.

# Chapter 19

# Lyapunov revisited*

## 19.1 Stability, observability, and controllability

Recall the damped linear oscillator described by

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -\tfrac{k}{m}x_1 - \tfrac{d}{m}x_2 \end{aligned} \tag{19.1}$$

with $m, d$ and $k$ positive. This system is asymptotically stable. The energy of this system is given by

$$V(x) = \frac{k}{2}x_1^2 + \frac{m}{2}x_2^2 = x^*Px$$

where

$$P = \begin{bmatrix} \frac{k}{2} & 0 \\ 0 & \frac{m}{2} \end{bmatrix}.$$

Along any solution, the rate of change of the energy is given by

$$\frac{dV}{dt} = -dx_2^2 = -x^*C^*Cx$$

where

$$C = \begin{bmatrix} 0 & d^{\frac{1}{2}} \end{bmatrix}$$

From this it should be clear that

$$\begin{aligned} P &> 0 \\ PA + A^*P + C^*C &= 0 \end{aligned}$$

The matrix $Q = C^*C$ is hermitian positive semi-definite, but it is not positive definite. Hence, using this $P$ matrix, our existing Lyapunov results only guarantee that the system is stable; they do not guarantee asymptotic stability. Noting that the pair $(C, A)$ is observable, we present a result in this section which allows us to use the above $P$ matrix to guarantee asymptotic stability of (19.1).

Consider a system described by

$$\dot{x} = Ax \tag{19.2}$$

where $x(t)$ is an $n$-vector and $A$ is a constant $n \times n$ matrix. We have the following result.

**Lemma 23** *Suppose there is a hermitian matrix $P$ satisfying*

$$\boxed{\begin{array}{rcl} P & > & 0 \\ PA + A^*P + C^*C & \leq & 0 \end{array}} \tag{19.3}$$

*where $(C, A)$ is observable. Then system (19.2) is asymptotically stable.*

Proof. Suppose there exists a hermitian matrix $P$ which satisfies inequalities (19.3). Consider any eigenvalue $\lambda$ of $A$; we will show that $\Re(\lambda) < 0$. Let $v \neq 0$ be an eigenvector corresponding to $\lambda$, that is,

$$Av = \lambda v\,. \tag{19.4}$$

Hence,

$$v^*PAv = \lambda v^*Pv\;;$$

and taking the complex conjugate transpose of both sides of this last equation yields

$$v^*A^*Pv = \bar{\lambda} v^*Pv$$

Pre- and post-multiplying the second inequality in (19.3) by $v^*$ and $v$, respectively, yields

$$v^*PAv + v^*APv + v^*C^*Cv \leq 0$$

which implies

$$\lambda v^*Pv + \bar{\lambda} v^*Pv + ||Cv||^2 \leq 0$$

that is,

$$2\Re(\lambda)v^*Pv \leq -||Cv||^2\,.$$

Since $P > 0$, we must have $v^*Pv > 0$.

We now show that we must also have $||Cv|| > 0$. Suppose $||Cv|| = 0$. Then $Cv = 0$. Recalling (19.4) results in

$$\begin{bmatrix} A - \lambda I \\ C \end{bmatrix} v = 0$$

Since $(C, A)$ is observable, the above matrix has rank $n$, and hence $v = 0$; this contradicts $v \neq 0$. So, we must have $||Cv|| > 0$.

Hence,

$$\Re(\lambda) = -||Cv||^2/2v^*Pv < 0$$

Since $\Re(\lambda) < 0$ for every eigenvalue $\lambda$ of $A$, system (19.2) is asymptotically stable. ∎

**Example 207** Returning to the damped linear oscillator, we see that inequalities (19.3) are satisfied by

$$P = \begin{bmatrix} \frac{k}{2} & 0 \\ 0 & \frac{m}{2} \end{bmatrix} \qquad C = \begin{bmatrix} 0 & d^{\frac{1}{2}} \end{bmatrix}$$

Since the oscillator is described by (19.2) with

$$A = \begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{d}{m} \end{bmatrix}$$

we have

$$\begin{bmatrix} C \\ CA \end{bmatrix} = \begin{bmatrix} 0 & d^{\frac{1}{2}} \\ \frac{-d^{\frac{1}{2}}k}{m} & -\frac{d^{\frac{3}{2}}}{m} \end{bmatrix}$$

Since this observability matrix is invertible, $(C, A)$ is observable. Using Lemma 23, one can use the above matrix $P$ to guarantee asymptotic stability of the damped oscillator.

**Exercise 146** Prove the following result. Suppose there is a hermitian matrix $P$ satisfying

$$\begin{aligned} P &\geq 0 \\ PA + A^*P + C^*C &\leq 0 \end{aligned}$$

where $(C, A)$ is detectable. Then the system (19.2) is asymptotically stable.

Recalling that controllability of $(A, B)$ is equivalent to observability of $(B^*, A^*)$, and that asymptotic stability of $A$ and $A^*$ are equivalent, we obtain the following result

**Lemma 24** *Suppose there is a hermitian matrix $S$ satisfying*

$$\boxed{\begin{aligned} S &> 0 \\ AS + SA^* + BB^* &\leq 0 \end{aligned}} \tag{19.5}$$

*where $(A, B)$ is controllable. Then system (19.2) is asymptotically stable*

**Exercise 147** Suppose there is a hermitian matrix $S$ satisfying

$$\begin{aligned} S &\geq 0 \\ AS + SA^* + BB^* &\leq 0 \end{aligned}$$

where $(A, B)$ is stabilizable. Then system (19.2) is asymptotically stable

Suppose system (19.2) is asymptotically stable, Then for any matrix $C$ (of appropriate dimensions) the Lyapunov equation

$$PA + A^*P + C^*C = 0 \tag{19.6}$$

has the unique solution

$$P = \int_0^\infty e^{A^*t}C^*Ce^{At}\, dt$$

Recall that this is the infinite time observability grammian associated with $(C, A)$, that is, $P = W_o$; it is always hermitian and positive semidefinite. In addition, as we have seen earlier, $W_o$ is positive definite if and only if $(C, A)$ is observable. These observations and Lemma 23 lead to the following result.

**Theorem 30** *The following statements are equivalent.*

*(a) The system $\dot{x} = Ax$ is asymptotically stable.*

*(b) There exist a positive definite hermitian matrix $P$ and a matrix $C$ with $(C, A)$ observable which satisfy the Lyapunov equation*

$$\boxed{PA + A^*P + C^*C = 0} \tag{19.7}$$

*(c) For* every *matrix $C$ with $(C, A)$ observable, the Lyapunov equation (19.7) has a unique solution for $P$ and this solution is hermitian positive-definite.*

Recall the original Lyapunov equation

$$PA + A^*P + Q = 0$$

If $Q$ is hermitian positive semi-definite, it can always be expressed as $Q = C^*C$ by appropriate choice of $C$ and one can readily show that $(Q, A)$ is observable iff $(C, A)$ is observable. Also $\operatorname{rank} C = n$ if and only if $Q$ is positive definite. If $C$ has rank $n$, the pair $(C, A)$ is observable for any $A$. So, if $Q$ is positive definite hermitian, the original Lyapunov equation is of the form (19.7) with $(C, A)$ observable.

Recalling that controllability of $(A, B)$ is equivalent to observability of $(B^*, A^*)$, and that asymptotic stability of $A$ and $A^*$ are equivalent, we can obtain the following result.

**Theorem 31** *The following statements are equivalent.*

*(a) The system $\dot{x} = Ax$ is asymptotically stable.*

*(b) There exist a positive definite hermitian matrix $S$ and a matrix $B$ with $(A, B)$ controllable which satisfy the Lyapunov equation*

$$\boxed{AS + SA^* + BB^* = 0} \tag{19.8}$$

*(c) For* every *matrix $B$ with $(A, B)$ controllable, the Lyapunov equation (19.8) has a unique solution for $S$ and this solution is hermitian positive-definite.*

Note that if $A$ is asymptotically stable, the solution $S$ to Lyapunov equation (19.8) is the infinite-time controllability grammian associated with $(A, B)$, that is $S = W_c$. Also,

$$S = \int_0^\infty e^{At} BB^* e^{A^*t}\, dt$$

## 19.2 A simple stabilizing controller

Consider a linear time-invariant system described by

$$\dot{x} = Ax + Bu$$

where $x(t)$ is an $n$-vector and $u(t)$ is an $m$-vector. Consider any positive $T > 0$ and let

$$\boxed{W = W_c(-T) := \int_{-T}^{0} e^{At}BB^*e^{A^*t}\,dt}$$

The above matrix can be regarded as the finite time controllability grammian (over the interval $[-T, 0]$) associated with $(A, B)$. If $(A, B)$ is controllable, $W$ is invertible.

We now show that the controller

$$\boxed{u = -B^*W^{-1}x}$$

yields an asymptotically stable closed loop system.

Using the following properties of $e^{At}$,

$$Ae^{At} = \frac{de^{At}}{dt}\,; \qquad e^{A^*t}A^* = \frac{de^{A^*t}}{dt}$$

we obtain

$$\begin{aligned}
AW + WA^* &= \int_{-T}^{0} \left[Ae^{At}BB^*e^{A^*t} + e^{At}BB^*e^{A^*t}A^*\right]\,dt \\
&= \int_{-T}^{0} \left[\frac{de^{At}}{dt}BB^*e^{A^*t} + e^{At}BB^*\frac{de^{A^*t}}{dt}\right]\,dt \\
&= \int_{-T}^{0} \frac{d\left(e^{At}BB^*e^{A^*t}\right)}{dt}\,dt \\
&= BB^* - e^{-AT}BB^*e^{-A^*T}\,,
\end{aligned}$$

that is,

$$AW + WA^* - BB^* = -e^{-AT}BB^*e^{-A^*T}$$

Considering the closed loop system

$$\dot{x} = (A - BB^*W^{-1})x$$

we have

$$\begin{aligned}
(A - BB^*W^{-1})W + W(A - BB^*W^{-1})^* &= AW + WA^* - 2BB^* \\
&= -BB^* - e^{-AT}BB^*e^{-A^*T} \\
&\leq -BB^*\,,
\end{aligned}$$

that is,

$$(A - BB^*W^{-1})W + W(A - BB^*W^{-1})^* + BB^* \leq 0\,.$$

Controllability of $(A,\ B)$ implies controllability of $(A - BB^*W^{-1},\ B)$. This fact, along with $W > 0$ and the previous inequality imply that the closed-loop system is asymptotically stable. ■

So, the above yields another proof of

$$\boxed{\text{controllability} \qquad \Longrightarrow \qquad \text{stabilizability}}$$

# Chapter 20

# Performance*

## 20.1 The $L_2$ norm

### 20.1.1 The rms value of a signal

Consider any scalar signal $s$ which is defined for $0 \le t < \infty$. One measure of the size of the signal is given by the rms value of the signal; we denote this by $\|s\|_2$ and define it by

$$\|s\|_2 = \left( \int_0^\infty |s(t)|^2 \, dt \right)^{\frac{1}{2}} .$$

This is also called the $L_2$ norm of the signal. One can readily show that the three required properties of a norm are satisfied. This norm is simply the generalization of the usual Euclidean norm for $n$-vectors to signals or functions. Clearly, a signal has a finite $L_2$ norm if and only if it is square integrable.

**Example 208** Consider the exponential signal $s(t) = e^{\lambda t}$ where $\lambda = -\alpha + j\omega$ and $\alpha$ and $\omega$ are both real with $\alpha > 0$. For any $T > 0$,

$$\int_0^T |s(t)|^2 \, dt = \int_0^T e^{-2\alpha t} \, dt = \frac{1}{-2\alpha}(e^{-2\alpha T} - 1) .$$

Thus,

$$\int_0^\infty |s(t)|^2 \, dt = \lim_{T \to \infty} \int_0^T |s(t)|^2 \, dt = \lim_{T \to \infty} \frac{1}{-2\alpha}(e^{-2\alpha T} - 1) = \frac{1}{2\alpha} .$$

Hence $\|s\|_2 = 1/\sqrt{2\alpha}$.

Consider now a signal $s$ where $s(t)$ is an $n$-vector for $0 \le t < \infty$. We define the $L_2$ norm of the signal $s$ by

$$\|s\|_2 = \left( \int_0^\infty |s(t)|^2 \, dt \right)^{\frac{1}{2}}$$

where $\|s(t)\|$ is the usual Euclidean norm of the $n$-vector $s(t)$ as given by

$$\|s(t)\|^2 = |s_1(t)|^2 + |s_2(t)|^2 + \cdots + |s_n(t)|^2 .$$

One can readily show that the three required properties of a norm are satisfied.

### 20.1.2 The $L_2$ norm of a time-varying matrix

Consider now a function $G$ defined on $[0, \infty)$ where $G(t)$ is a matrix for $0 \le t < \infty$. We define the $L_2$ norm of $G$ by

$$\|G\|_2 = \left( \int_0^\infty \mathrm{trace}\,(G^*(t)G(t))\, dt \right)^{\frac{1}{2}} .$$

Suppose $G(t)$ is $m \times n$ and the $m$-vectors $g_1(t), g_2(t), \cdots, g_n(t)$ are the columns of $G(t)$. Then

$$\begin{aligned} \mathrm{trace}\,(G(t)^*G(t)) &= g_1(t)^*g_1(t) + g_2(t)^*g_2(t) + \cdots + g_n(t)^*g_n(t) \\ &= \|g_1(t)\|^2 + \|g_2(t)\|^2 + \cdots + \|g_n(t)\|^2 . \end{aligned}$$

So we conclude that

$$\|G\|_2^2 = \sum_{i=1}^n \int_0^\infty \|g_i(t)^2\|^2\, dt = \sum_{i=1}^n \|g_i\|_2^2 \tag{20.1}$$

where the $g_i$ is the $i$-th column of $G$.

## 20.2 The $H_2$ norm of an LTI system

Consider an input-output system with disturbance input $w$ and performance output $z$ described by

$$\begin{aligned} \dot{x} &= Ax + Bw \\ z &= Cx \end{aligned} \tag{20.2}$$

and suppose $A$ is asymptotically stable. Recall that zero initial condition input-output behavior of this system can be completely characterized by its impulse response

$$G(t) = Ce^{At}B . \tag{20.3}$$

Consider first a SISO system with impulse response $g$. When $A$ is asymptotically stable, it follows that $g$ is square integrable and the $L_2$ norm $\|g\|_2$ of $g$ is finite where

$$\|g\|_2 = \left( \int_0^\infty |g(t)|^2\, dt \right)^{\frac{1}{2}} . \tag{20.4}$$

We refer to $\|g\|_2$ as the $H_2$ norm of the system. Thus the $H_2$ norm of an asymptotically stable SISO system is the rms value of its impulse response. This norm is a a performance measure for the system; it can be regarded as a measure of the ability of the system to mitigate the effect of the disturbance $w$ on the performance output $z$.

Consider now a general linear time-invariant system described by (20.2) with inpulse response given by (20.3). When $A$ is asymptotically stable, if follows that each element of $G$ is square integrable, and we define the $L_2$ norm of the impulse response by

$$\|G\|_2 = \left( \int_0^\infty \mathrm{trace}\,(G^*(t)G(t))\, dt \right)^{\frac{1}{2}} . \tag{20.5}$$

We refer to $\|G\|_2$ as the $H_2$ norm of the system.

We now obtain a frequency domain interpretation of $\|G\|_2$. Recall that the transfer function of this system is given by $\hat{G}(s) = C(sI - A)^{-1}B$. When $A$ is asymptotically stable, we define the 2-norm of the transfer function by

$$\|\hat{G}\|_2 = \left( \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{trace}\,(G(\jmath\omega)^* G(\jmath\omega))\, d\omega \right)^{\frac{1}{2}} . \tag{20.6}$$

For a SISO system with transfer function $\hat{g}$, we have

$$\|\hat{g}\|_2^2 = \frac{1}{2\pi} \int_{-\infty}^{\infty} |\hat{g}(\jmath\omega)|^2\, d\omega \,.$$

Since $\hat{G}$ is the Laplace transform of $G$, it follows from Parseval's Theorem that

$$\|\hat{G}\|_2 = \|G\|_2 \,, \tag{20.7}$$

that is, the 2-norm of the transfer function equals the $L_2$ norm of the impulse response.

## 20.3 Computation of the $H_2$ norm

We now show how the $H_2$-norm can be computed using Lyapunov equations. Suppose $A$ is asymptotically stable and let $W_o$ be the observability grammian associated with $(C, A)$, that is, $W_o = \int_0^\infty e^{A^*t} C^* C e^{At}\, dt$. Recall that $W_o$ is the unique solution to the Lyapunov equation

$$W_o A + A^* W_o + C^* C = 0 \,.$$

We now obtain that

$$\int_0^\infty G^*(t)G(t)\, dt = \int_0^\infty B^* e^{A^*t} C^* C e^{At} B\, dt = B^* \int_0^\infty e^{A^*t} C^* C e^{At}\, dt\, B = B^* W_o B \,,$$

that is,

$$\int_0^\infty G^*(t)G(t)\, dt = B^* W_o B \,. \tag{20.8}$$

Using the linearity of the trace operation, it now follows that

$$\|G(\cdot)\|_2^2 = \int_0^\infty \mathrm{trace}\,(G^*(t)G(t))\, dt = \mathrm{trace}\left( \int_0^\infty G^*(t)G(t)\, dt \right) = \mathrm{trace}\,(B^* W_o B) \,;$$

hence

$$\boxed{\|G\|_2^2 = \mathrm{trace}\,(B^* W_o B)} \,. \tag{20.9}$$

We can also compute the $H_2$-norm using the controllability grammian. Since

$$\mathrm{trace}\,(G^*(t)G(t)) = \mathrm{trace}\,(G(t)G^*(t)) \,;$$

it follows that the $H_2$-norm is also given by

$$\|G\|_2 = \left( \int_0^\infty \operatorname{trace}\left( G(t)G^*(t) \right) dt \right)^{\frac{1}{2}} . \tag{20.10}$$

Let $W_c$ be the controllability grammian associated with $(A, B)$, that is, $W_o = \int_0^\infty e^{At}BB^*e^{A^*t}\,dt$. Recall that $W_c$ is the unique solution to the Lyapunov equation

$$AW_c + W_cA^* + BB^* = 0 .$$

Hence,

$$\int_0^\infty G(t)G(t)^*\,dt = \int_0^\infty Ce^{At}BB^*e^{A^*t}C^*\,dt = C\int_0^\infty e^{At}BB^*e^{A^*t}\,dt\,C^* = CW_cC^* ,$$

that is,

$$\int_0^\infty G(t)G^*(t)\,dt = CW_cC^* . \tag{20.11}$$

Using the linearity of the trace operation, it now follows that

$$\|G(\cdot)\|_2^2 = \int_0^\infty \operatorname{trace}\left( G(t)G^*(t) \right) dt = \operatorname{trace}\left( \int_0^\infty G(t)G^*(t)\,dt \right) = \operatorname{trace}\left( CW_cC^* \right) ;$$

hence

$$\boxed{\|G\|_2^2 = \operatorname{trace}\left( CW_cC^* \right)} . \tag{20.12}$$

**Example 209** Consider the scalar system described by

$$\begin{aligned} \dot{x} &= -\alpha x + u \\ y &= x \end{aligned}$$

with $\alpha > 0$. This system has transfer function $\hat{g}(s) = \frac{1}{s+\alpha}$. Hence,

$$\begin{aligned} \|\hat{g}\|_2^2 &= \frac{1}{2\pi}\int_{-\infty}^\infty |\hat{g}|^2\,d\omega = \frac{1}{2\pi}\int_{-\infty}^\infty \frac{1}{\alpha^2+\omega^2}\,d\omega = \frac{1}{2\pi\alpha}\int_{-\infty}^\infty \frac{1}{1+\eta^2}\,d\eta \qquad (\eta = \omega/\alpha) \\ &= \frac{1}{2\pi\alpha}\,\Big|_{-\infty}^{\infty} \tan^{-1}(\eta) \\ &= \frac{1}{2\alpha} . \end{aligned}$$

The impulse response for this system is given by $g(t) = e^{-\alpha t}$. Hence

$$\begin{aligned} \|g\|_2^2 &= \int_0^\infty g(t)^2\,dt = \int_0^\infty e^{-2\alpha t}\,dt = \Big|_0^\infty \frac{1}{-2\alpha}e^{-2\alpha t} \\ &= \frac{1}{2\alpha} . \end{aligned}$$

This clearly illustrates the fact that $\|\hat{g}\|_2 = \|g\|_2$ and the $H_2$ norm of this system is $(1/2\alpha)^{\frac{1}{2}}$.

The observability grammian $W_o$ is given by the solution to the Lyapunov equation

$$-W_o\alpha - \alpha W_o + 1 = 0\,.$$

This yields $W_o = 1/2\alpha$. Since $B = 1$, we readily see that the $H_2$ norm is given by trace $(B^* W_o B)^{\frac{1}{2}}$. The controllability grammian $W_c$ is given by the solution to the Lyapunov equation

$$-\alpha W_c - W_c\alpha + 1 = 0\,.$$

This yields $W_c = 1/2\alpha$. Since $C = 1$, we readily see that the $H_2$ norm is given by trace $(C W_c C^*)^{\frac{1}{2}}$.

**MATLAB**

```
>> help normh2

 NORMH2 Continuous H2 norm.

  [NMH2] = NORMH2(A,B,C,D) or
  [NMH2] = NORMH2(SS_) computes the H2 norm of the given state-space
     realization. If the system not strictly proper, INF is returned.
     Otherwise, the H2 norm is computed as

             NMH2 = trace[CPC']^0.5 = trace[B'QB]^0.5

     where P is the controllability grammian, and Q is the observability
     grammian.


>> help norm2

  nh2 = norm2(sys)

  Computes the H2 norm of a stable and strictly proper LTI
  system
                                 -1
                SYS = C (sE-A)  B

  This norm is given by
                 --------------
                /
               V  Trace(C*P*C')

  where P solves the Lyapunov equation:  A*P+P*A'+B*B' = 0.

  See also  LTISYS.
```

## Exercises

**Exercise 148** Consider the system described by

$$\begin{aligned}\dot{x}_1 &= x_2\\ \dot{x}_2 &= -2x_1 - 3x_2 + w\\ y &= 4x_1 + 2x_2\end{aligned}$$

Obtain the $H_2$ norm of this system using the following methods.
(a) The definition in (20.4).
(b) Equation (20.9).
(c) Equation (20.11).
(d) MATLAB.

**Exercise 149** Without carrying out any integration, obtain the $H_2$ norm of the system whose impulse response $g$ is given by

$$g(t) = e^{-t} + e^{-2t}\,.$$

# Chapter 21

# Linear quadratic regulators (LQR)

## 21.1 The linear quadratic optimal control problem

In this chapter we learn another method of designing stabilizing state feedback controllers. These controllers result from solving the linear quadratic (LQ) optimal control problem and are called linear quadratic regulators (LQR). This problem is characterized by a linear system and a quadratic cost.

Consider a linear time-invariant system (plant)

$$\boxed{\dot{x} = Ax + Bu} \tag{21.1}$$

with initial condition

$$x(0) = x_0 \tag{21.2}$$

where the real scalar $t$ is the time variable, the $n$-vector $x(t)$ is the plant state, and the $m$-vector $u(t)$ is the control input. When we refer to a control history, control signal, or open loop control we will mean a continuous control input function $u(\cdot)$ which is defined for $0 \leq t < \infty$.

Suppose $Q$ and $R$ are hermitian matrices with

$$R > 0 \quad \text{and} \quad Q \geq 0$$

We call them weighting matrices. For each initial state $x_0$ and each control history $u(\cdot)$, we define the corresponding quadratic cost by

$$\boxed{J(x_0, u(\cdot)) := \int_0^\infty x(t)^* Q x(t) + u(t)^* R u(t)\, dt}$$

where $x(\cdot)$ is the unique solution of (21.1) with initial condition (21.2). We wish to minimize this cost by appropriate choice of the control history. Although the above cost can be infinite, it follows from the assumptions on $Q$ and $R$ that it is always non-negative, that is,

$$0 \leq J(x_0, u(\cdot)) \leq \infty\,.$$

In addition to minimizing the above cost, we will demand that the control input drives the state to zero, that is, we require that

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

So, for a fixed initial state $x_0$, we say that a control history $u(\cdot)$ is admissible at $x_0$ if it results in $\lim_{t\to\infty} x(t) = 0$. We say that a control history $u^{opt}(\cdot)$ is optimal at $x_0$ if it is admissible at $x_0$ and

$$J(x_0, u^{opt}(\cdot)) \le J(x_0, u(\cdot)) \tag{21.3}$$

for every admissible control history $u(\cdot)$. We refer to

$$J^{opt}(x_0) := J(x_0, u^{opt}(\cdot))$$

as the optimal cost at $x_0$.

In general, the optimal control history and the optimal cost depend on the initial state. Also, for a fixed initial state, the optimal cost is unique whereas the optimal control history need not be. In some cases an optimal control history might not exist. Since the cost is bounded below by zero, there is always an infimal cost $J^{inf}(x_0)$. That is, one cannot achieve better than this cost but can come arbitrarily close to it; more precisely, for every admissible control history $u(\cdot)$, one has

$$J^{inf}(x_0) \le J(x_0, u(\cdot))$$

and for every $\epsilon > 0$ there is an admissible control history $u^\epsilon(\cdot)$ such that

$$J(x_0, u^\epsilon(\cdot)) \le J^{inf}(x_0) + \epsilon$$

When an optimal control history exists, we must have $J^{inf}(x_0) = J^{opt}(x_0)$. If there exists at least one admissible control history yielding a bounded cost, then $J^{inf}(x_0) < \infty$.

**Example 210** Consider the scalar integrator plant described by

$$\dot{x} = u$$

with initial condition $x(0) = 1$. Suppose we wish to drive the state of this plant to zero while minimizing the cost

$$\int_0^\infty u(t)^2\, dt\,.$$

Consider any positive $\delta > 0$ and let

$$u(t) = -\delta e^{-\delta t}\,.$$

Then, the resulting state history, $x(t) = e^{-\delta t}$, has the desired behavior and the corresponding cost is

$$\int_0^\infty \delta^2 e^{-2\delta t}\, dt = |_0^\infty \frac{\delta}{-2} e^{-2\delta t} = \frac{\delta}{2}\,.$$

By choosing $\delta > 0$ arbitrarily small, we can make the cost arbitrarily close to zero. Since the cost is bounded below by zero, it follows that the infimal cost is zero. However, we cannot make the cost zero. If the cost is zero, that is $\int_0^\infty u(t)^2\, dt = 0$, then we must have $u(t) \equiv 0$; hence $x(t) \equiv 1$. This control history does not drive the state to zero and, hence, is not admissible. So there is no optimal control history for this problem. Note that in this problem, $(Q, A) = (0, 0)$ has an imaginary axis unobservable eigenvalue.

**Optimal feedback controllers.** Suppose one has a feedback controller $\mathcal{C}$ with the property that for any initial state $x_0$ the control histories generated by the controller are admissible. We say that $\mathcal{C}$ is an optimal feedback controller if for each initial state $x_0$ the corresponding control histories generated by $\mathcal{C}$ are optimal at $x_0$. We will shortly demonstrate the following fact for linear quadratic optimal control problems:

*If there is an optimal control history for every initial state then, for every initial state, the optimal control history is unique and is given by a linear memoryless state feedback controller.*

Since we are going to look controllers which minimize the cost and result in

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

we need to make the following assumption for the rest of the chapter:

- The pair $(A, B)$ is stabilizable.

## 21.2 The algebraic Riccati equation (ARE)

We will see that, when an optimal feedback controller exists, its determination involves the following matrix equation which is known as the Algebraic Riccati Equation (ARE):

$$\boxed{PA + A^*P - PBR^{-1}B^*P + Q = 0} \tag{21.4}$$

The matrices $A, B, Q, R$ are determined by the LQR problem data. Solving the LQR problem involves solving the ARE for a hermitian matrix $P$. Since the ARE depends quadratically on $P$, it is a nonlinear matrix equation and, in general, has more than one solution.

A solution $P$ is said to be a stabilizing solution if the matrix $A - BR^{-1}B^*P$ is Hurwitz, that is, all its eigenvalues have negative real parts. This is the solution we will be looking for. The significance of this property is that the linear state feedback controller

$$u = -R^{-1}B^*Px$$

results in the asymptotically stable closed loop system:

$$\dot{x} = (A - BR^{-1}B^*P)x$$

It shown in Section 21.4 that *if there is an optimal control history for every initial state then, the ARE must have a stabilizing solution.* Hence, if the ARE does not have a stabilizing solution then, there is not an optimal control history for every initial state.

It is also shown in Section 21.4 that *if the pair $(Q, A)$ has no imaginary unobservable eigenvalues then, the ARE must have a stabilizing solution.*

Section 21.4 also shows that *a stabilizing solution is* maximal in the sense that if $P$ is a stabilizing and $\tilde{P}$ is any other solution then, $P \geq \tilde{P}$. Hence there can only be one stabilizing solution.

## 21.3 Linear state feedback controllers

Suppose

$$u(t) = Kx(t)$$

where $K$ is a constant state feedback gain matrix. The resulting closed loop system is described by

$$\dot{x} = (A + BK)x$$

and for each initial state $x_0$ the controller results in the cost

$$J(x_0, u(\cdot)) = \hat{J}(x_0, K) := \int_0^\infty x(t)^*[K^*RK + Q]x(t)\, dt$$

This feedback controller is an optimal feedback controller if and only if it is a stabilizing controller (that is, $A + BK$ is asymptotically stable) and for each initial state $x_0$ we have

$$\hat{J}(x_0, K) \leq J(x_0, u(\cdot))$$

for every admissible control history $u(\cdot)$.

If $K$ is stabilizing then, recalling a result from Lyapunov theory,

$$\hat{J}(x_0, K) = x_0^* P x_0 \tag{21.5}$$

where $P$ is the unique solution to the following Lyapunov equation

$$P(A + BK) + (A + BK)^* P + K^*RK + Q = 0 \tag{21.6}$$

Since $A + BK$ is asymptotically stable and $K^*RK + Q$ is hermitian positive semi-definite, we have $P^* = P \geq 0$.

**Example 211** Consider the simple real scalar system

$$\dot{x} = \alpha x + u$$

with cost

$$\int_0^\infty u(t)^2\, dt$$

If we do not require $\lim_{t\to\infty} x(t) = 0$ then regardless of initial state $x_0$, this cost is uniquely minimized by

$$u(t) \equiv 0$$

and the resulting controlled system is described by

$$\dot{x} = \alpha x$$

This system is not asymptotically stable for $\alpha \geq 0$. So, in looking for an optimal control we restrict consideration to control inputs that result in $\lim_{t\to\infty} x(t) = 0$

Now consider any linear controller $u = kx$. This generates a closed loop system

$$\dot{x} = (\alpha + k)x$$

which is asymptotically stable if and only if $\alpha + k < 0$. Utilizing (21.5), (21.6), the cost corresponding to a stabilizing linear controller is $Px_0^2$ where

$$P = \frac{k^2}{-2(\alpha + k)}$$

If $\alpha = 0$, $P = -k/2$. Since we require asymptotic stability of the closed loop system, this cost can be made arbitrarily small but nonzero. So for $\alpha = 0$, we have an *infimum cost* (that is, zero) but no minimum cost. The infimum cannot be achieved by controls which drive the state to zero. Note that in this case, $(Q, A) = (0, 0)$ has an imaginary axis unobservable eigenvalue.

## 21.4 Infimal cost*

### 21.4.1 Reduced cost stabilizing controllers.

Suppose one has a stabilizing gain matrix $K$ and wishes to obtain another stabilizing gain matrix which results in reduced cost. We will show that this is achieved by the gain matrix

$$\tilde{K} = -R^{-1}B^*P \tag{21.7}$$

where $P$ solves the Lyapunov equation (21.6).

**Fact 13** *Suppose $A + BK$ is asymptotically stable and $\tilde{K}$ is given by (21.7) and (21.6). Then,*

*(i) the matrix $A + B\tilde{K}$ is asymptotically stable;*

*(ii) for every initial state $x_0$,*

$$\hat{J}(x_0, \tilde{K}) \leq \hat{J}(x_0, K)$$

Proof. The Appendix contains a proof. ■

### 21.4.2 Cost reducing algorithm

The above result suggests the following algorithm to iteratively reduce the cost.

**Jack Benny algorithm.**

(0) Choose any gain matrix $K_0$ so that $A + BK_0$ is asymptotically stable.

(1) For $k = 0, 1, 2, \cdots$, let $P_k$ be the unique solution to the Lyapunov equation

$$P_k(A + BK_k) + (A + BK_k)^*P_k + K_k^*RK_k + Q = 0 \tag{21.8}$$

and let

$$K_{k+1} = -R^{-1}B^*P_k \tag{21.9}$$

**Fact 14** *Suppose $(A, B)$ is stabilizable, $R^* = R > 0$ and $Q^* = Q \geq 0$. Consider any sequence of matrices $\{P_k\}$ generated by the Jack Benny algorithm. Then this sequence has a limit $\bar{P}$, that is,*

$$\lim_{k\to\infty} P_k = \bar{P}.$$

*Furthermore, $\bar{P}$ has the following properties.*

*(a)*

$$\bar{P}^* = \bar{P} \geq 0$$

*(b) $P = \bar{P}$ solves*

$$\boxed{PA + A^*P - PBR^{-1}B^*P + Q = 0} \tag{21.10}$$

*(c) Every eigenvalue of the matrix $A - BR^{-1}B^*\bar{P}$ has real part less than or equal to zero.*

*(d) If $(Q, A)$ has no imaginary axis unobservable eigenvalues, then every eigenvalue of the matrix $A - BR^{-1}B^*\bar{P}$ has negative real part.*

Proof. (a) Since $P_k$ solves (21.8) we have $P_k \geq 0$. From consequence (ii) of fact 13, we have that $P_{k+1} \leq P_k$. Hence,

$$0 \leq P_{k+1} \leq P_k$$

From this one can show that there is a matrix $\bar{P}$ such that

$$\bar{P} := \lim_{k\to\infty} P_k$$

and $\bar{P} \geq 0$.

(b) In the limit, and using (21.9),

$$\bar{K} := \lim_{k\to\infty} K_k = \lim_{k\to\infty} K_{k+1} = -R^{-1}B^*\bar{P}$$

hence, it follows from (21.8) that

$$\bar{P}(A + B\bar{K}) + (A + B\bar{K})^*\bar{P} + \bar{K}^*R\bar{K} + Q = 0 \tag{21.11}$$

Substituting

$$\bar{K} = -R^{-1}B^*\bar{P}$$

results in (21.10).

(c) For each $k$, $\Re(\lambda) < 0$ for all eigenvalues $\lambda$ of $A + B\bar{K}$. Since $\bar{K} = \lim_{k\to\infty} K_k$ and the eigenvalues of a matrix depend continuously on its coefficients, we must have $\Re(\lambda) \leq 0$ for all eigenvalues $\lambda$ of $A + B\bar{K}$.

(d) Now suppose $A + B\bar{K}$ has an eigenvalue $\lambda$ with $\Re(\lambda) = 0$. Letting $v$ be an eigenvector 0 to $\lambda$ and pre- and post-multiplying (21.11) by $v^*$ and $v$ yields

$$v^*\bar{K}^*R\bar{K}v + v^*Qv = 0$$

Hence,

$$\bar{K}v = 0 \qquad Qv = 0$$

and

$$\lambda v = (A + B\bar{K})v = Av$$

So,

$$\begin{pmatrix} A - \lambda I \\ Q \end{pmatrix} v = 0$$

that is, $\lambda$ is an unobservable eigenvalue of $(Q, A)$. So, if $(Q, A)$ has no imaginary axis unobservable eigenvalues, $A + B\bar{K}$ has no imaginary axis eigenvalues. So all the eigenvalues of $A + B\bar{K}$ have negative real parts. ■

**Example 212** Recall first example. Here the algorithm yields:

$$\begin{aligned} P_k &= \frac{K_k^2}{-2(\alpha + K_k)} \\ K_{k+1} &= -P_k \end{aligned}$$

Hence,

$$P_{k+1} = \frac{P_k^2}{2(-\alpha + P_k)}$$

1. $(\alpha < 0)$: Here $\bar{P} = \lim_{k\to\infty} P_k = 0$ and $A - BR^{-1}B^*\bar{P} = \alpha < 0$.
2. $(\alpha > 0)$: Here $\bar{P} = \lim_{k\to\infty} P_k = 2\alpha$ and $A - BR^{-1}B^*\bar{P} = -\alpha < 0$.
3. $(\alpha = 0)$: Here $P_{k+1} = P_k/2$ for all $k$. Hence, $\bar{P} = \lim_{k\to\infty} P_k = 0$ and $A - BR^{-1}B^*\bar{P} = 0$.

   Note that this is the only case in which $(Q, A)$ has an imaginary axis unobservable eigenvalue.

**Remark 11** The matrix equation (21.10) is called an algebraic Riccati equation (ARE). Since it depends quadratically on $P$, it is a nonlinear matrix equation and, in general, has more than one solution.

**Remark 12** We say that a hermitian matrix $P$ is a stabilizing solution to ARE (21.10) if all the eigenvalues of the matrix $A - BR^{-1}B^*P$ have negative real parts. The significance of this is that the controller

$$u = -R^{-1}B^*Px$$

results in the asymptotically stable closed loop system:

$$\dot{x} = (A - BR^{-1}B^*P)x$$

### 21.4.3 Infimal cost

The following result provides a lower bound on the infimal cost $J^{inf}(x_0)$ in terms of a specific solution to the ARE.

**Fact 15** (Completion of squares lemma.) *Suppose $R^* = R > 0$ and $P$ is any hermitian solution to ARE (21.10). Then for every initial condition $x_0$ and every continuous control input history $u(\cdot)$ which guarantees*

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

*we have*

$$x_0^* P x_0 \leq J(x_0, u(\cdot))\,;$$

*hence*

$$x_0^* P x_0 \leq J^{inf}(x_0)\,.$$

Proof. Suppose $P$ is *any* hermitian solution $P$ to the ARE (21.10), let $x_0$ be any initial state, and consider any control history $u(\cdot)$ which results in $\lim_{t\to\infty} x(t) = 0$. Consider the function $V(x(\cdot))$ given by

$$V(x(t)) = x^*(t)Px(t)$$

Its derivative is given by

$$\begin{aligned}\frac{dV(x(t))}{dt} &= 2x^*(t)P\dot{x}(t)\\ &= 2x^*(t)PAx + 2x^*(t)PBu(t)\end{aligned}$$

Utilizing the ARE (21.10) we obtain

$$\begin{aligned}2x^*PAx &= x^*(PA + A^*P)x\\ &= x^*PBR^{-1}B^*Px - x^*Qx\end{aligned}$$

Hence (omitting the argument $t$ for clarity),

$$\begin{aligned}\frac{dV(x(t))}{dt} &= x^*PBR^{-1}B^*Px + 2x^*PBu - x^*Qx\\ &= x^*PBR^{-1}B^*Px + 2x^*PBu + u^*Ru - x^*Qx - u^*Ru\\ &= (u + R^{-1}B^*Px)^*R(u + R^{-1}B^*Px) - x^*Qx - u^*Ru\end{aligned}$$

Hence

$$x^*Qx + u^*Ru = -\frac{dV(x)}{dt} + (u + R^{-1}B^*Px)^*R(u + R^{-1}B^*Px)$$

Considering any $t \geq 0$ and integrating over the interval $[0, t]$ yields

$$\int_0^t (x^*Qx + u^*Ru)\,dt = -V(x(t)) \;+\; V(x(0)) \;+\; \int_0^t (u + R^{-1}B^*Px)^*R(u + R^{-1}B^*Px)\,dt$$

Recall $V(x) = x^*Px$, $x(0) = x_0$, and $\lim_{t\to\infty} x(t) = 0$; so, considering the limit as $t \to \infty$ and recalling the definition of the cost, we have

$$\boxed{J(x_0, u(\cdot)) = x_0^*Px_0 + \int_0^\infty (u + R^{-1}B^*Px)^*R(u + R^{-1}B^*Px)\, dt} \tag{21.12}$$

Since $R > 0$, we immediately have

$$J(x_0, u(\cdot)) \geq x_0^*Px_0 \tag{21.13}$$

■

The next result states that for any initial state $x_0$, the infimal cost is $x_0^*\bar{P}x_0$ where $\bar{P}$ is the limit of any sequence $\{P_k\}$ generated by Jack Benny algorithm.

**Lemma 25** (Infimal cost lemma.) *Suppose $(A, B)$ is stabilizable, $R^* = R > 0$ and $Q^* = Q \geq 0$. Consider any sequence of matrices $\{P_k\}_{k=0}^\infty$ generated by the Jack Benny algorithm. Then this sequence has a limit $\bar{P}$, that is,*

$$\lim_{k\to\infty} P_k = \bar{P}$$

*and $\bar{P}$ has the following properties.*

*(a) If $P$ is any hermitian solution to ARE (21.10), then*

$$\bar{P} \geq P$$

*(b) For every initial state $x_0$,*

$$\boxed{J^{inf}(x_0) = x_0^*\bar{P}x_0}$$

*(c) If there exists an optimal control history for every initial state $x_0$ then $\bar{P}$ is a stabilizing solution to ARE (21.10).*

PROOF. Consider any sequence of matrices $\{P_k\}$ generated by the Jack Benny algorithm. Then, from fact (14), this sequence has a limit $\bar{P}$, that is,

$$\lim_{k\to\infty} P_k = \bar{P}$$

and it satisfies ARE (21.10).

Consider any initial state $x_0$ and any $k = 1, 2, \ldots$. The control history given by

$$u_k(t) = -R^{-1}B^*P_{k-1}x(t)$$

results in the cost

$$J(x_0, u_k(\cdot)) = x_0^*P_kx_0$$

Hence

$$J^{\text{inf}}(x_0) \leq x_0^*P_kx_0$$

Since $\bar{P}$ is a solution to the ARE, it follows from the completion of squares lemma that

$$x_0^*\bar{P}x_0 \leq J^{\inf}(x_0)$$

Hence

$$x_0^*\bar{P}x_0 \leq J^{\inf}(x_0) \leq x_0^*P_kx_0$$

Considering the limit as $k \to \infty$ we must have

$$x_0^*\bar{P}x_0 \leq J^{\inf}(x_0) \leq x_0^*\bar{P}x_0\,;$$

so

$$J^{inf}(x_0) = x_0^*\bar{P}x_0$$

If $P$ is any hermitian solution to ARE (21.10), we must have

$$x_0^*Px_0 \leq J(x_0, u_k(\cdot)) = x_0^*P_kx_0$$

Taking the limit as $k \to \infty$, we obtain

$$x_0^*Px_0 \leq x_0^*\bar{P}x_0$$

Since this is true for all $x_0$ we have $P \leq \bar{P}$.

Suppose an optimal control history $u^{opt}$ exists for every $x_0$. Then

$$J(x_0, u^{opt}) = J^{opt}(x_0) = J^{\inf}(x_0) = x_0^*\bar{P}x_0$$

Recalling the proof of the completion of squares lemma, we have

$$J(x_0, u^{opt}(\cdot)) \;=\; x_0^*\bar{P}x_0 + \int_0^\infty (u^{opt} + R^{-1}B^*\bar{P}x)^*R(u^{opt} + R^{-1}B^*\bar{P}x)\,dt$$

Hence

$$\int_0^\infty (u^{opt} + R^{-1}B^*\bar{P}x)^*R(u^{opt} + R^{-1}B^*\bar{P}x)\,dt = 0$$

Since $R > 0$ we must have

$$u^{opt} = -R^{-1}B^*\bar{P}x$$

Since $u^{opt}(\cdot)$ is admissable, it results in $\lim_{t\to\infty}$. Since the above is true for any initial state $x_0$, it follows that the system

$$\dot{x} = (A - BR^{-1}B^*)x$$

must be asymptotically stable. Hence $\bar{P}$ is a stabilizing solution to ARE. ■

**Remark 13** A solution to the ARE which satisfies condition (a) above is called the maximal solution to the ARE. Clearly, it must be unique.

## 21.5 Optimal stabilizing controllers

The solution to the LQR problem is based on the following matrix equation which is known as the Algebraic Riccati Equation:

$$\boxed{PA + A^*P - PBR^{-1}B^*P + Q = 0} \tag{21.14}$$

Since this is a nonlinear equation in $P$ it can have more that one solution. We will be interested in a stabilizing solution which is defined as a hermitian matrix $P$ with the property that all the eigenvalues of the matrix $A - BR^{-1}B^*P$ have negative real parts.

**Theorem 32** *Suppose $R^* = R > 0$ and the ARE (21.14) has a stabilizing solution $P$. Then for every initial state $x_0$, there is a unique optimal control history $u^{opt}$; it is given by*

$$\boxed{u^{opt}(t) = -R^{-1}B^*Px(t)}$$

*and the optimal cost is*

$$J^{opt}(x_0) = x_0^*Px_0\,.$$

PROOF. Let $x_0$ be any initial state, and consider any control history $u(\cdot)$ which results in $\lim_{t\to\infty} x(t) = 0$. Define the function $V$ given by

$$V(t) = x^*(t)Px(t)$$

The derivative of $V$ is given by

$$\begin{aligned}\frac{dV}{dt} &= 2x^*(t)P\dot{x}(t)\\ &= 2x^*(t)PAx(t) + 2x^*(t)PBu(t)\end{aligned}$$

Utilizing the ARE (21.14) we obtain

$$\begin{aligned}2x^*PAx &= x^*(PA + A^*P)x\\ &= x^*PBR^{-1}B^*Px - x^*Qx\end{aligned}$$

Hence (omitting the argument $t$ for clarity),

$$\begin{aligned}\frac{dV}{dt} &= x^*PBR^{-1}B^*Px + 2x^*PBu - x^*Qx\\ &= x^*PBR^{-1}B^*Px + 2x^*PBu + u^*Ru - x^*Qx - u^*Ru\\ &= (u + R^{-1}B^*Px)^*R(u + R^{-1}B^*Px) - x^*Qx - u^*Ru\end{aligned}$$

Hence

$$x^*Qx + u^*Ru = -\frac{dV}{dt} + (u + R^{-1}B^*Px)^*R(u + R^{-1}B^*Px)$$

Considering any $t \geq 0$ and integrating over the interval $[0, t]$ yields

$$\int_0^t (x^*Qx + u^*Ru)\,dt = -V(t) \;+\; V(0) \;+\; \int_0^t (u + R^{-1}B^*Px)^*R(u + R^{-1}B^*Px)\,dt$$

Recall $V(t) = x^*(t)Px(t)$, $x(0) = x_0$, and $\lim_{t\to\infty} x(t) = 0$; so, considering the limit as $t \to \infty$ and recalling the definition of the cost, we have

$$\boxed{J(x_0, u(\cdot)) = x_0^*Px_0 + \int_0^\infty (u + R^{-1}B^*Px)^*R(u + R^{-1}B^*Px)\,dt} \tag{21.15}$$

Since $P$ is a stabilizing solution to ARE (21.10) the control history given by $u(t) = \hat{u}(t) := -R^{-1}B^*Px(t)$ results in $\lim_{t\to\infty} x(t) = 0$; hence

$$J(x_0, \hat{u}(\cdot)) = x_0^*Px_0$$

Consider any other admissible control history $u(\cdot)$. Since $R > 0$, we have

$$J(x_0, u(\cdot)) \geq x_0^*Px_0$$

This implies that $\hat{u}$ is optimal at $x_0$ and the optimal cost is $x_0^*Px_0$.

Also, (21.15) implies that if any other control $\tilde{u}$ yields the optimal cost $x_0^*Px_0$, then the $\tilde{u}(t) = -R^{-1}B^*Px(t)$; hence $\tilde{u}(t) = \hat{u}(t)$. So, $\hat{u}$ is a unique optimal control history. ■

**Example 213** Returning to Example 211 we have $A = \alpha$, $B = 1$, $Q = 0$, and $R = 1$. Hence the ARE is

$$2\alpha P - P^2 = 0$$

This has two solutions

$$P = P_1 := 0 \qquad P = P_2 := 2\alpha$$

Hence

$$A - BR^{-1}B^*P_1 = \alpha \qquad A - BR^{-1}B^*P_2 = -\alpha$$

We have three cases.
Case 1. $(\alpha < 0)$: $P_1 = 0$ is the stabilizing solution and optimal controller is $u = 0$.
Case 2. $(\alpha > 0)$: $P_1 = 2\alpha$ is the stabilizing solution and optimal controller is $u = -2\alpha x$.
Case 3. $(\alpha = 0)$: No stabilizing solution.

**Remark 14** If $P$ is a stabilizing solution to ARE (21.10) then, it can be shown that $P$ must be the maximal solution $\bar{P}$. This follows from

$$x_0^*Px_0 = J^{opt}(x_0) = J^{\inf}(x_0) = x_0^*\bar{P}x_0\,.$$

Hence there can be at most one stabilizing solution.

### 21.5.1 Existence of stabilizing solutions to the ARE

**Lemma 26** *Suppose $(A, B)$ is stabilizable, $R^* = R > 0$ and $Q^* = Q$. Then ARE (21.10) has a stabilizing solution if and only if $(Q, A)$ has no imaginary axis unobservable eigenvalues.*

PROOF. From Fact 14 it follows that if $(Q, A)$ has no imaginary unobservable eigenvalues then the ARE has a stabilizing solution.

We now prove the converse by contradiction. Suppose $P$ is a stabilizing solution to ARE and $\lambda$ is an imaginary unobservable eigenvalue of $(Q, A)$. Then, there is a nonzero vector $v$ such that

$$Av = \lambda v \qquad \text{and} \qquad Qv = 0$$

Pre- and postmultiplying the ARE by $v^*$ and $v$ yields

$$2\Re(\lambda)v^*Pv - v^*PBR^{-1}B^*Pv = 0$$

Hence, $R^{-1}B^*Pv = 0$ and

$$\begin{aligned}(A - BR^{-1}B^*P)v &= Av - BR^{-1}B^*Pv \\ &= \lambda v\end{aligned}$$

that is, $\lambda$ is an eigenvalue of the matrix $A - BR^{-1}B^*P$. This contradicts the asymptotic stability of this matrix. ■

**Lemma 27** *Suppose $(A, B)$ is stabilizable, $R^* = R > 0$ and $Q^* = Q \geq 0$. If $(Q, A)$ is detectable, then*

*(a) The ARE has a stabilizing solution.*

*(b) $P$ is a stabilizing solution if and only if $P \geq 0$*

PROOF. Part (a) follows from previous lemma.

(b) We have already seen that if $P$ is stabilizing then $P \geq 0$. So, all we have to show is that if $P$ solves the ARE and $P \geq 0$ then $P$ must be stabilizing. To prove this we write the ARE as

$$P(A + BK) + (A + BK)^*P + K^*RK + Q = 0$$

where $K = -R^{-1}B^*P$. Hence,

$$P(A + BK) + (A + BK)^*P + Q \leq 0$$

Since $P \geq 0$ and $(Q, A)$ is detectable, it follows from a previous Lyapunov result that $A+BK$ is asymptotically stable; hence $P$ is stabilizing. ■

## 21.6 Summary

The main results of this chapter are summarized in the following theorem.

**Theorem 33** *Suppose $(A, B)$ is stabilizable, $R^* = R > 0$ and $Q^* = Q \geq 0$. Then the following statements are equivalent.*

*(a) For every initial state $x_0$, there is an optimal control history.*

*(b) $(Q, A)$ has no imaginary axis unobservable eigenvalues.*

*(c) ARE (21.10) has a stabilizing solution.*

*(d) ARE (21.10) has a unique stabilizing solution $P$; also $P$ is positive semi-definite and is the maximal solution.*

*Suppose any of conditions (a)-(d) hold. Then for every initial state $x_0$, the optimal control history is unique and is generated by the linear state feedback controller*

$$\boxed{u(t) = -R^{-1}B^*Px(t)}$$

*where $P$ is the stabilizing solution to ARE (21.10). Also, the optimal cost is given by*

$$\boxed{J^{opt}(x_0) = x_0^*Px_0}$$

**Example 214** Consider the problem of finding an optimal stabilizing controller in the problem of minimizing the cost,

$$\int_0^\infty x_2^2 + u^2 \, dt$$

for the system,

$$\begin{aligned} \dot{x}_1 &= \alpha x_1 + x_2 \\ \dot{x}_2 &= u \end{aligned}$$

This is an LQR (linear quadratic regulator) problem with

$$A = \begin{pmatrix} \alpha & 1 \\ 0 & 0 \end{pmatrix} \qquad B = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \qquad Q = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \qquad R = 1$$

The pair $(Q, A)$ has a single unobservable eigenvalue $\alpha$. Hence an optimal stabilizing controller exists if and only if $\alpha \neq 0$.

Considering

$$P = \begin{pmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{pmatrix}$$

(we are looking for a symmetric $P$) the ARE yields the following three scalar equations

$$\begin{array}{rclcrcl} 2\alpha p_{11} & - & p_{12}^2 & & & = & 0 \\ p_{11}+\alpha p_{12} & - & p_{12}p_{22} & & & = & 0 \\ 2p_{12} & - & p_{22}^2 & + & 1 & = & 0 \end{array} \tag{21.16}$$

We consider three cases:

Case 1 ($\alpha < 0$.) The first equation in (21.16) yields $p_{11} = p_{12}^2/\alpha$. Since we require $p_{11} \geq 0$ we must have

$$p_{11} = p_{12} = 0$$

The second equation in (21.16) is now automatically satisfied and (requiring $p_{22} \geq 0$) the third equation yields

$$p_{22} = 1$$

Since

$$P = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

is the only positive semi-definite solution to the ARE; it must be the stabilizing one. The fact that it is stabilizing is evident from

$$A_{cl} = A - BR^{-1}B^*P = \begin{pmatrix} \alpha & 1 \\ 0 & -1 \end{pmatrix}$$

So, the optimal stabilizing controller for this case is given by

$$u = -x_2$$

Case 2 ($\alpha > 0$.) We obtain three positive semi-definite solutions:

$$P_i \;=\; \begin{pmatrix} \beta_i^2/2\alpha & \beta_i \\ \beta_i & (1+2\beta_i)^{\frac{1}{2}} \end{pmatrix} \qquad \beta_1 = 2\alpha(\alpha+1), \quad \beta_2 = 2\alpha(\alpha-1), \quad \beta_3 = 0$$

Since $(P_1)_{11} \geq (P_2)_{11}, (P_3)_{11}$, and the stabilizing solution is maximal, the matrix $P_1$ must be the stabilizing solution. Let us check that $P_1$ is indeed a stabilizing solution. The corresponding closed loop matrix,

$$A_{cl} = A - BR^{-1}B^*P_1 = \begin{pmatrix} \alpha & 1 \\ -\beta_1 & -(1+2\beta_1)^{\frac{1}{2}} \end{pmatrix}$$

has characteristic polynomial

$$q(s) = s^2 + a_1 s + a_0$$

with

$$\begin{array}{rclcll} a_1 & = & (1+2\beta_1)^{\frac{1}{2}} - \alpha & = & (4\alpha^2+4\alpha+1)^{\frac{1}{2}} - \alpha & > \; 0 \\ a_0 & = & \beta_1 - \alpha(1+2\beta_1)^{\frac{1}{2}} & = & 2\alpha(\alpha+1) - \alpha(4\alpha^2+4\alpha+1)^{\frac{1}{2}} & > \; 0 \end{array}$$

Since all coefficients of this second order polynomial are positive, all roots of this polynomial have negative real part; hence $A_{cl}$ is asymptotically stable. So, the optimal stabilizing controller for this case is given by

$$u = -2\alpha(\alpha+1)x_1 - (4\alpha^2 + 4\alpha + 1)^{\frac{1}{2}}x_2$$

Case 3 ($\alpha = 0$.) We already know that there is no optimal stabilizing controller for this case. However, to see this explicitly note that the only positive semi-definite solution to the ARE is

$$P = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

Since

$$A_{cl} = A - BR^{-1}B^*P = \begin{pmatrix} 0 & 1 \\ 0 & -1 \end{pmatrix},$$

this solution is not stabilizing; hence there there is no stabilizing solution and no optimal stabilizing controller.

## 21.7 MATLAB

```
>> help lqr

 LQR    Linear quadratic regulator design for continuous systems.
        [K,S,E] = LQR(A,B,Q,R)  calculates the optimal feedback gain
        matrix K such that the feedback law  u = -Kx  minimizes the cost
        function:
                J = Integral {x'Qx + u'Ru} dt

        subject to the constraint equation:
                .
                x = Ax + Bu

        Also returned is S, the steady-state solution to the associated
        algebraic Riccati equation and the closed loop eigenvalues E:
                                     -1
                0 = SA + A'S - SBR  B'S + Q     E = EIG(A-B*K)

        [K,S,E] = LQR(A,B,Q,R,N) includes the cross-term N that relates
        u to x in the cost function.

                J = Integral {x'Qx + u'Ru + 2*x'Nu}

        The controller can be formed with REG.

        See also: LQRY, LQR2, and REG.
>> help lqr2

 LQR2   Linear-quadratic regulator design for continuous-time systems.
        [K,S] = LQR2(A,B,Q,R)  calculates the optimal feedback gain matrix
        K such that the feedback law  u = -Kx  minimizes the cost function

              J = Integral {x'Qx + u'Ru} dt
                                                .
        subject to the constraint equation:   x = Ax + Bu

        Also returned is S, the steady-state solution to the associated
        algebraic Riccati equation:
                                     -1
                0 = SA + A'S - SBR  B'S + Q

        [K,S] = LQR2(A,B,Q,R,N) includes the cross-term 2x'Nu that
        relates u to x in the cost functional.
```

```
        The controller can be formed with REG.

        LQR2 uses the SCHUR algorithm of [1] and is more numerically
        reliable than LQR, which uses eigenvector decomposition.

        See also: ARE, LQR, and LQE2.
>> help are

 ARE    Algebraic Riccati Equation solution.
        X = ARE(A, B, C) returns the stablizing solution (if it
        exists) to the continuous-time Riccati equation:

                A'*X + X*A - X*B*X + C = 0

        assuming B is symmetric and nonnegative definite and C is
        symmetric.

        See also: RIC.
```

## 21.8 Minimum energy controllers

Suppose we are interested in stabilizing a system with the minimum amount of control effort. Then, we might consider the problem of minimizing

$$\int_0^\infty \|u(t)\|^2 \, dt$$

subject to

$$\dot{x} = Ax + Bu \qquad \text{and} \qquad x(0) = x_0 \qquad \text{and} \qquad \lim_{t\to\infty} x(t) = 0\,. \tag{21.17}$$

This is an LQR problem with $Q = 0$ and $R = I$. The corresponding ARE is

$$PA + A^*P - PBR^{-1}B^*P = 0\,. \tag{21.18}$$

Since $Q = 0$, it follows from our main LQR result that that an optimal stabilizing controller exists if and only if $A$ has no imaginary eigenvalues.

The optimal controllers for this problem result in a closed loop system having the following interesting property.

*If $\lambda$ is an eigenvalue of $\tilde{A} := A + BK = A - BR^{-1}B^*P$, then either $\lambda$ or $-\bar{\lambda}$ is an eigenvalue of $A$.*

To demonstrate the above property, rewrite ARE as

$$P\tilde{A} + A^*P = 0\,.$$

Suppose $\lambda$ is an eigenvalue of $\tilde{A}$. Then there is a nonzero vector $v$ such that $\tilde{A}v = \lambda v$. Post-multiplying the ARE by $v$ yields

$$\lambda Pv + A^* Pv = 0$$

If $Pv \neq 0$, then $Pv$ is an eigenvector of $A^*$ with eigenvalue $-\lambda$. Hence , $-\bar{\lambda}$ is an eigenvalue of $A$. If $Pv = 0$, then $\tilde{A}v = Av - BB^*Pv = Av$; hence $Av = \lambda v$ and $\lambda$ is an eigenvalue of $A$.

**Example 215** Suppose we wish to obtain a stabilizing state feedback controller which always minimizes $\int_0^\infty u(t)^2\, dt$ for the scalar system

$$\dot{x} = ax + u\,.$$

Here

$$A = a\,, \quad B = 1\,, \quad R = 1\,, \quad Q = 0\,.$$

So the problem has a solution provided $a$ is not on the imaginary axis. The ARE corresponding to this problem,

$$(a + \bar{a})P - P^2 = 0\,,$$

has two solutions

$$P = 0 \quad \text{and} \quad P = a + \bar{a}\,.$$

The two corresponding values of $\tilde{A} := A - BR^{-1}B^*P$ are

$$a \quad \text{and} \quad -\bar{a}$$

respectively.

From this we can explicitly that, when $a$ is on the imaginary axis, there is no stabilizing $P$.

When $a$ has negative real part, $P = 0$ is stabilizing, the optimal controller is zero and $\tilde{A} = A$.

When $a$ has positive real part, $P = a + \bar{a}$ is stabilizing, the optimal controller is $u = -(a + \bar{a})x$ and $\tilde{A} = -\bar{a}$.

## 21.9 $H_2$ optimal controllers*

Consider a system described by

$$\begin{aligned} \dot{x} &= Ax + B_1 w + B_2 u \\ z &= Cx + Du \end{aligned}$$

where the $m_1$-vector $w(t)$ is a disturbance input; the $m_2$-vector $u(t)$ is a control input; and the $p$-vector $z$ performance output. We assume that

(i) $\text{rank} D = m_2$.

(ii) $D^*C = 0$.

With the above assumptions, the energy $\|z\|_2$ associated with the output is given by

$$\|z\|_2^2 = \int_0^\infty \|z(t)\|^2\, dt = \int_0^\infty \|Cx(t) + Du(t)\|^2\, dt = \int_0^\infty (x(t)^*Qx(t) + u(t)^*Ru(t))\, dt$$

with

$$Q = C^*C \qquad \text{and} \qquad R = D^*D\,.$$

Since $D$ has full column rank, it follows that $R$ is positive definite.

A linear state feedback controller

$$u = Kx$$

results in the closed loop system

$$\dot{x} = (A + B_2K)x + B_1w \tag{21.19a}$$

$$z = (C + DK)x \tag{21.19b}$$

Suppose that $A + B_2K$ is asymptotically stable and recall the definition of the $H_2$-norm of a linear system. If $G_K$ is the impulse response matrix from $w$ to $z$ then the $H_2$ norm of the above system is given by $\|G_K\|_2 = \left(\text{trace}\left(\int_0^\infty G_K^*(t)G_K(t)\, dt\right)\right)^{\frac{1}{2}}$. Noting that

$$(C + DK)^*(C + DK) = C^*C + K^*D^*DK$$

it follows that

$$\|G_K\|_2^2 = \text{trace}\, B_1^* P_K B_1$$

where $P_K$ is the unique solution to

$$P_K(A + B_2K) + (A + B_2K)^*P_K + K^*D^*DK + C^*C = 0\,.$$

From this one can readily show that if one considers the LQ problem with

$$B = B_2\,, \qquad R = D^*D\,, \qquad Q = C^*C$$

then an optimal LQ controller minimizes the $H_2$ norm of the closed loop system.

The observability condition on $(Q, A)$ is equivalent to:

• The pair $(C, A)$ has no imaginary axis unobservable eigenvalues.

So provided this condition holds the controller which minimizes $\|G_K\|_2$ is given by

$$u = -(D^*D)^{-1}B_2^*P$$

where $P$ is the stabilizing solution to

$$PA + A^*P - PB_1^*(D^*D)^{-1}B_1^*P + C^*C = 0$$

• If $A + B_2K$ is asymptotically stable, it follows from an identity in Lyapunov revisited that

$$\|G_K\|_2^2 = \operatorname{trace}(C + DK)^*X(C + DK)$$

where $X$ is the unique solution to

$$(A + B_2K)X + X(A + B_2K)^* + B_1^*B_1 = 0$$

The hermitian positive semi-definite matrix $X$ is sometimes referred to as the *state covariance matrix* associated with the above system.

## 21.10 Exercises

**Exercise 150** Consider the system

$$\dot{x} = 2x + u\,,$$

where $x$ and $u$ are real scalars. Obtain an optimal stabilizing feedback controller which minimizes

$$\int_0^\infty u(t)^2\,dt\,.$$

**Exercise 151** Consider the system

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

where $x_1, x_2$ and $u$ are real scalars. Obtain an optimal stabilizing feedback controller which minimizes

$$\int_0^\infty u(t)^2\,dt\,.$$

**Exercise 152** Consider the system

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

where all quantities are scalar. Obtain a optimal stabilizing feedback controller which minimizes

$$\int_0^\infty u(t)^2\,dt\,.$$

**Exercise 153** For the system

$$\begin{aligned}\dot{x}_1 &= \alpha x_1 + x_2\\ \dot{x}_2 &= -x_2 + u\end{aligned}$$

Obtain a optimal feedback controller which minimizes

$$\int_0^\infty u(t)^2\,dt$$

You may have to distinguish between different cases of $\alpha$.

**Exercise 154** *2-link manipulator* Recall the two link manipulator with measured output $y = q_2$. Recall the linearization of this system about

$$q_1 = q_2 = u = 0$$

and recall its state space description.

Obtain an LQG controller for this linearized system. Apply the controller to the nonlinear system and simulate the nonlinear closed loop system for various initial conditions. (Use zero initial conditions for the observer.)

## 21.11 Appendix

### 21.11.1 Proof of fact 1

*Asymptotic Stability*

$$\begin{aligned} P(A+B\tilde{K})+(A+B\tilde{K})^*P &= P(A-BR^{-1}B^*P)+(A-BR^{-1}B^*P)^*P \\ &= PA+A^*P-2PBR^{-1}B^*P \\ &= -PBK-K^*B^*P-K^*RK-2PBR^{-1}B^*P-Q \\ &= -PBR^{-1}B^*P-Q-(K-\tilde{K})^*R(K-\tilde{K}) \end{aligned}$$

Hence

$$P(A+B\tilde{K})+(A+B\tilde{K})^*P+PBR^{-1}B^*P+Q+(K-\tilde{K})^*R(K-\tilde{K})=0 \qquad (21.20)$$

Consider any eigenvalue $\lambda$ of $A+B\tilde{K}$ and let $v \neq 0$ be a corresponding eigenvector. Then

$$2\Re(\lambda)v^*Pv+v^*(K-\tilde{K})^*R(K-\tilde{K})v \leq 0$$

First note that if

$$v^*(K-\tilde{K})^*R(K-\tilde{K})v=0$$

we must have

$$\tilde{K}v=Kv$$

and hence

$$(A+B\tilde{K})v=(A+BK)v$$

that is, $\lambda$ is an eigenvalue of $A+BK$. Since $K$ is stabilizing, $\Re(\lambda)<0$. So suppose $v^*(K-\tilde{K})R(K-\tilde{K})v>0$. Then

$$2\Re(\lambda)v^*Pv<0$$

Since $v^*Pv \geq 0$ we must have $\Re(\lambda)<0$. Since $\Re(\lambda)<0$ for every eigenvalue $\lambda$ of $A+B\tilde{K}$, we have asymptotic stability of $A+B\tilde{K}$.

*Reduced cost:* The cost associated with $\tilde{K}$ is given by

$$\hat{J}(x_0,\tilde{K})=x_0^*\tilde{P}x_0$$

where $\tilde{P}$ is the unique solution to

$$\tilde{P}(A+B\tilde{K})+(A+B\tilde{K})^T\tilde{P}+\tilde{K}^TR\tilde{K}+Q=0$$

Substition for $\tilde{K}$ yields

$$\tilde{P}(A+B\tilde{K})+(A+B\tilde{K})^*\tilde{P}+PBR^{-1}B^*P+Q=0 \qquad (21.21)$$

Combining (21.21) and (21.20) yields

$$(P-\tilde{P})(A+B\tilde{K})+(A+B\tilde{K})^*(P-\tilde{P}) \leq 0$$

In other words, for some hermitian positive semi-definite matrix $\Delta Q$,

$$(P - \tilde{P})(A + B\tilde{K}) + (A + B\tilde{K})^*(P - \tilde{P}) + \Delta Q = 0$$

Since $A + B\tilde{K}$ is asymptotically stable,

$$P - \tilde{P} = \int_0^\infty e^{(A+B\tilde{K})^* t} \Delta Q e^{(A+B\tilde{K})t} \, dt$$

Hence $P - \tilde{P} \geq 0$, or

$$\tilde{P} \leq P$$

∎

# Chapter 22

# LQG controllers

In the previous chapter, we saw how to obtain stabilizing state feedback gain matrices $K$ by solving a Ricatti equation. Anytime one has a technique for obtaining stabilizing state feedback gain matrices, one also has a technique for obtaining stabilizing observer gain matrices $L$. This simply follows from the fact the $A + LC$ is asymptotically stable if and only if the same is true for $A^* + C^*K$ with $L = K^*$. In this chapter, we use the results of the last chapter to obtain an observer design method which is based on a Ricatti equation which we call the observer Ricatti equation. We then combine these results with the results of the previous chapter to obtain a methodolology for the design of stabilizing output feedback controllers via the solution of two uncoupled Ricatti equations.

## 22.1 LQ observer

• Plant

$$\begin{aligned}\dot{x} &= Ax + Bu \\ y &= Cx\end{aligned}$$

where $t$ is the time variable, the $n$-vector $x(t)$ is the plant state, the $m$-vector $u(t)$ is the control input and the $p$-vector $y(t)$ is the measured output.

• Observer

$$\dot{x_c} = Ax_c + Bu + L(Cx_c - y)$$

• Error dynamics : $\tilde{x} = x_c - x$

$$\dot{\tilde{x}} = (A + LC)\tilde{x}$$

In what follows $V$ is a $p \times p$ matrix and $W$ is a $n \times n$ matrix.

**Lemma 28** *Suppose $(C, A)$ is detectable, $V^* = V > 0$, and $W^* = W \geq 0$. Then the following are equivalent.*

*(i) The pair $(A, W)$ has no imaginary axis uncontrollable eigenvalues.*

*(ii) The Riccati equation*

$$\boxed{AS + SA^* - SC^*V^{-1}CS + W = 0} \tag{22.1}$$

*has a hermitian solution $S$ with $A - SC^*V^{-1}C$ asymptotically stable.*

Proof. By duality. ■

• We will refer to ARE (22.1) as an observer Riccati equation associated with the plant.

• We say that $S$ is a stabilizing solution to (22.1) if $A - SC^*V^{-1}C$ is asymptotically stable. The significance of this is that if one chooses the observer gain $L$ as

$$\boxed{L = -SC^*V^{-1}} \tag{22.2}$$

then the error dynamics are described by the asymptotically stable system

$$\dot{\tilde{x}} = (A - SC^*V^{-1}C)\tilde{x}$$

## 22.2 LQG controllers

Plant

$$\boxed{\begin{array}{rcl} \dot{x} & = & Ax + Bu \\ y & = & Cx \end{array}}$$

Basically we combine an LQR regulator with an LQ observer.

So suppose the the plant satisfies the following minimal requirements for stabilization via output feedback:

(i) $(A, B)$ is stabilizable.

(ii) $(C, A)$ is detectable.

Choose weighting matrices

$$Q \quad R \quad W \quad V$$

(of appropriate dimensions) which satisfy the following conditions:

(i) $R^* = R > 0$ and $Q^* = Q \geq 0$, and $(Q, A)$ has no imaginary axis unobservable eigenvalues.

(ii) $V^* = V > 0$ and $W^* = W \geq 0$, and $(A, W)$ has no imaginary axis uncontrollable eigenvalues.

Let $P$ and $S$ be the stabilizing solutions to the Riccati equations

$$\boxed{\begin{array}{rcl} PA + A^*P - PBR^{-1}B^*P + Q & = & 0 \\ \\ AS + SA^* - SC^*V^{-1}CS + W & = & 0 \end{array}}$$

**LQG Controller** The corresponding LQG controller is given by

$$
\begin{aligned}
\dot{x}_c &= (A - BR^{-1}B^*P)x_c - SC^*V^{-1}(Cx_c - y) \\
u &= -R^{-1}B^*Px_c
\end{aligned}
$$

Note that this can be written as

$$
\boxed{
\begin{aligned}
\dot{x}_c &= (A - BR^{-1}B^*P - SC^*V^{-1}C)x_c + SC^*V^{-1}y \\
\\
u &= -R^{-1}B^*Px_c
\end{aligned}
}
$$

Letting $\dot{\tilde{x}} = x_c - x$, the closed loop system is described by

$$
\begin{aligned}
\dot{x} &= (A - BR^{-1}B^*P)x \quad - \quad BR^{-1}B^*P\tilde{x} \\
\dot{\tilde{x}} &= (A - SC^*V^{-1}C)\tilde{x}
\end{aligned}
$$

As we have seen before the eigenvalues of this closed loop system are simply the union of those of the asymptotically matrices $A - BR^{-1}B^*P$ and $A - SC^*V^{-1}C$. Hence, the closed loop system is asymptotically stable.

## 22.3 $H_2$ Optimal controllers*

### 22.3.1 LQ observers

Consider

$$
\begin{aligned}
\dot{x} &= Ax + B_1w + B_2u \\
y &= Cx + Dw
\end{aligned}
$$

where the measurement vector $y(t)$ is a $p$-vector. We assume that

(i) $\text{rank}D = p$

(ii) $DB_1^* = 0$

With an observer

$$\dot{x}_c = Ax_c + B_2u + L(Cx_c - y)\,,$$

the corresponding observer error dynamics are described by

$$\dot{\tilde{x}} = (A + LC)\tilde{x} + (B_1 - LD)w\,.$$

Note that

$$(B_1 - LD)(B_1 - LD)^* = B_1B_1^* + LDD^*L^*$$

Consider any "performance variable"

$$\tilde{z} = M\tilde{x}$$

Suppose $A + LC$ is asymptotically stable and $G_L$ is the impulse response matrix from $w$ to $\tilde{z}$. Then, the $H_2$ norm of $G_L$ is given by

$$\|G_L\|_2^2 = \int_0^\infty \text{trace } (G^*(t)G(t)) \ dt = MS_LM^*$$

where $S_L$ uniquely solves

$$(A + LC)S_L + S_L(A + LC)^* + LDD^*L^* + B_1B_1^* = 0$$

With

$$V = DD^* \qquad W = B_1B_1^*$$

the corresponding LQ observer minimizes the $H_2$ norm of $G_L$.

The controllability condition on $(A, W)$ is equivalent to:

- The pair $(A, B_1)$ has no imaginary axis uncontrollable eigenvalues.

So provided this condition holds the observer gain which minimizes the $H_2$ norm of $G_L$ is given by

$$L = -SC^*(DD^*)^{-1}$$

where $S$ is the stabilizing solution to

$$AS + SA^* - SC^*(DD^*)^{-1}CS + B_1B_1^* = 0$$

### 22.3.2 LQG controllers

Consider

$$\begin{array}{rcccccc} \dot{x} &=& Ax &+& B_1w &+& B_2u \\ z &=& C_1x & & &+& D_{12}u \\ y &=& C_2x &+& D_{21}w & & \end{array}$$

# Chapter 23

# Systems with inputs and outputs: II*

Recall

$$\longleftarrow \boxed{\begin{array}{rcl} \dot{x} &=& Ax + Bu \\ y &=& Cx + Du \end{array}} \longleftarrow \tag{23.1}$$

## 23.1 Minimal realizations

Recall that $(A, B, C, D)$ is a realization of a transfer matrix $\hat{G}$ if

$$\hat{G}(s) = C(sI - A)^{-1}B + D.$$

• A realization is minimal if the dimension of its state-space is less than or equal to the state space dimension of any other realization.

**SISO systems.** Consider a SISO system described by

$$\hat{G}(s) = \frac{n(s)}{d(s)}$$

where $n$ and $d$ are polynomials. If $n$ and $d$ are coprime (i.e., they have no common zeros), then the order of a minimal realization is simply the degree of $d$.

**Equivalent systems.** System

$$\begin{array}{rcl} \dot{x} &=& Ax + Bu \\ y &=& Cx + Du \end{array}$$

is equivalent to system

$$\begin{array}{rcl} \dot{\tilde{x}} &=& \tilde{A}\tilde{x} + \tilde{B}u \\ y &=& \tilde{C}\tilde{x} + \tilde{D}u \end{array}$$

if there is a nonsingular matrix $T$ such that

$$\begin{array}{rclcrcl} \tilde{A} &=& T^{-1}AT & \qquad & \tilde{B} &=& T^{-1}B \\ \tilde{C} &=& CT & & \tilde{D} &=& D \end{array}$$

**Theorem 34**

*(i) $(A, B, C, D)$ is a minimal realization of $\hat{G}(s)$ iff $(A, B)$ is controllable and $(C, A)$ is observable.*

*(ii) All minimal realizations are equivalent.*

## 23.2 Exercises

**Exercise 155** Obtain a state space realization of the transfer function,

$$\hat{G}(s) = \begin{bmatrix} \dfrac{1}{s-1} & \dfrac{1}{s-1} \\ \\ \dfrac{1}{s+1} & -\dfrac{1}{s+1} \end{bmatrix}.$$

Is your realization minimal?

**Exercise 156** Obtain a state space realization of the transfer function,

$$\hat{G}(s) = \begin{bmatrix} \dfrac{s}{s-1} & \dfrac{1}{s+1} \end{bmatrix}.$$

Is your realization minimal?

**Problem 2** Obtain a minimal state space realization of the transfer function

$$\hat{G}(s) = \frac{s^2 - 3s + 2}{s^2 + s - 2}.$$

# Chapter 24

# Time-varying systems*

## 24.1 Linear time-varying systems (LTVs)

Here we are concerned with linear time-varying systems (LTVs) systems described by

$$\boxed{\begin{array}{rcl} \dot{x}(t) & = & A(t)x(t) + B(t)u(t) \\ y(t) & = & C(t)x(t) + D(t)u(t) \end{array}} \tag{24.1}$$

where $t$ is the time variable, the state $x(t)$ is an $n$-vector, the input $u(t)$ is an $m$-vector while the output $y(t)$ is a $p$-vector. We suppose that the time-varying matrices $A(t), B(t), C(t), D(t)$ depend continuously on $t$.

### 24.1.1 Linearization about time-varying trajectories

Consider a nonlinear system described by

$$\dot{x} = F(x, u) \tag{24.2a}$$

$$y = H(x, u)\,. \tag{24.2b}$$

Suppose $\bar{u}(\cdot)$ is a specific time-varying input history to the nonlinear system and $\bar{x}(\cdot)$ is any corresponding state trajectory. Then

$$\dot{\bar{x}}(t) = F(\bar{x}(t), \bar{u}(t))\,.$$

Let $\bar{y}(\cdot)$ be the corresponding output history, that is,

$$\bar{y}(t) = F(\bar{y}(t), \bar{u}(t))\,.$$

Then, introducing the perturbed variables

$$\delta x = x - \bar{x} \quad \text{and} \quad \delta u = u - \bar{u} \quad \text{and} \quad \delta y = y - \bar{y}\,,$$

the linearization of system (24.2) about $(\bar{x}, \bar{u})$ is defined by

$$\delta\dot{x} = A(t)\delta x + B(t)\delta u \tag{24.3a}$$

$$\delta y = C(t)\delta x + D(t)\delta u \tag{24.3b}$$

where

$$
\begin{aligned}
A(t) &= \frac{\partial F}{\partial x}(\bar{x}(t), \bar{u}(t)) \quad \text{and} \quad B(t) = \frac{\partial F}{\partial x}(\bar{u}(t), \bar{u}(t)) \\
C(t) &= \frac{\partial H}{\partial x}(\bar{x}(t), \bar{u}(t)) \quad \text{and} \quad D(t) = \frac{\partial H}{\partial u}(\bar{x}(t), \bar{u}(t)) .
\end{aligned}
$$

Note that, although the original nonlinear system is time-invariant, its linearization about a time-varying trajectory can be time-varying.

## 24.2 The state transition matrix

Consider a linear time-varying system described by

$$\boxed{\dot{x} = A(t)x} \tag{24.4}$$

where $t$ is the time variable, $x(t)$ is an $n$-vector and the time-varying matrix $A(t)$ depends continuously on $t$.

**Fact 16** *We have existence and uniqueness of solutions, that is, for each initial condition,*

$$x(t_0) = x_0 \,, \tag{24.5}$$

*there is a unique solution to the differential equation (24.4) which satisfies the initial condition. Also, this solution is defined for all time.*

**Fact 17** *Solutions depend linearly on initial state. More specifically, for each pair of times $t, t_0$, there is a matrix $\Phi(t, t_0)$ such that for every initial state $x_0$, the solution $x(t)$ at time $t$ due to initial condition (24.5) is given by*

$$\boxed{x(t) = \Phi(t, t_0)x_0}$$

The matrix $\Phi(t, t_0)$ is called the state transition matrix associated with system (24.4). Note that it is uniquely defined by

$$
\begin{aligned}
\Phi(t_0, t_0) &= I \\
\frac{\partial \Phi}{\partial t}(t, t_0) &= A(t)\Phi(t, t_0)
\end{aligned}
$$

**LTI systems**

$$\dot{x} = Ax$$

$$\Phi(t, t_0) = e^{A(t-t_0)}$$

**Scalar systems.** Consider any scalar system described by

$$\dot{x} = a(t)x$$

where $x(t)$ and $a(t)$ are scalars. The solution to this system corresponding to initial condition $x(t_0) = x_0$ is given by

$$x(t) = e^{\int_{t_0}^{t} a(\tau)\, d\tau} x_0 \,.$$

Hence,

$$\Phi(t, t_0) = e^{\int_{t_0}^{t} a(\tau)\, d\tau}$$

**Some properties of the state transition matrix**

(a)

$$\Phi(t_2, t_1)\Phi(t_1, t_0) = \Phi(t_2, t_0)$$

(b) $\Phi(t, t_0)$ is invertible and

$$\Phi(t_0, t) = \Phi(t, t_0)^{-1}$$

(c)

$$\frac{\partial \Phi}{\partial t_0}(t, t_0) = -\Phi(t, t_0)A(t_0)$$

(d)

$$\det[\Phi(t, t_0)] = \exp \int_{t_0}^{t} \text{trace}\,[A(\tau)]\, d\tau$$

**Computation of the state transition matrix.** Let $X$ be the solution of

$$\dot{X}(t) = A(t)X(t) \tag{24.6a}$$

$$X(0) = I \tag{24.6b}$$

where $I$ is the $n \times n$ identity matrix. Then $X(t) = \Phi(t, 0)$. Hence, using property (b) above, $X(t_0)^{-1} = \Phi(0, t_0)$ Now use property (a) to obtain

$$\Phi(t, t_0) = X(t)X(t_0)^{-1} \tag{24.7}$$

## 24.3 Stability

Recall that when $A$ is a constant matrix, system (24.4) is asymptotically stable if and only if all the eigenvalues of $A$ have negative real parts. This leads to the following natural conjecture: system (24.4) is asymptotically stable if the eigenvalues of $A(t)$ have negative real parts for all time $t$. This conjecture is, in general, false. This is demonstrated by the following example. Actually, in this example, the eigenvalues of $A(t)$ are the same for all $t$ and have negative real parts. However, the system has unbounded solutions and is unstable.

**Example 216** (Markus and Yamabe, 1960.) This is an example of an unstable second order system whose time-varying $A$ matrix has constant eigenvalues with negative real parts. Consider

$$A(t) = \begin{bmatrix} -1 + a\cos^2 t & 1 - a\sin t\cos t \\ -1 - a\sin t\cos t & -1 + a\sin^2 t \end{bmatrix}$$

with $1 < a < 2$. Here,

$$\Phi(t,0) = \begin{bmatrix} e^{(a-1)t}\cos t & e^{-t}\sin t \\ -e^{(a-1)t}\sin t & e^{-t}\cos t \end{bmatrix}$$

Since $a > 1$, the system corresponding to this $A(t)$ matrix has unbounded solutions. However, for all $t$, the characteristic polynomial of $A(t)$ is given by

$$p(s) = s^2 + (2-a)s + 2$$

Since $2 - a > 0$, the eigenvalues of $A(t)$ have negative real parts.

## 24.4 Systems with inputs and outputs

$$\begin{aligned} \dot{x} &= A(t)x + B(t)u \qquad x(t_0) = x_0 && (24.8) \\ y &= C(t)x + D(t)u && (24.9) \end{aligned}$$

• The solution to (24.8) is uniquely given by

$$\boxed{x(t) = \Phi(t,t_0)x_0 + \int_{t_0}^{t} \Phi(t,\tau)B(\tau)u(\tau)\,d\tau}$$

**Impulse response matrix**

$$y(t) = C(t)\Phi(t,t_0)x_0 + \int_{t_0}^{t} C(t)\Phi(t,\tau)B(\tau)u(\tau)\,dt + D(t)u(t)$$

Hence, for $x(t_0) = 0$,

$$y(t) = \int_{t_0}^{t} G(t,\tau)u(\tau)\,d\tau$$

where

$$G(t,\tau) := C(t)\Phi(t,\tau)B(\tau) + D(\tau)\delta(\tau)$$

## 24.5 DT

$$x(k) = \underbrace{\Phi(k,k_0)}_{\text{state transition matrix}} x_0$$

**Fact 18** *(a)*

$$\Phi(k_0, k_0) = I$$

*(b)*

$$\Phi(k_2, k_1)\Phi(k_1, k_0) = \Phi(k_2, k_0)$$

*(c)*

$$\Phi(k+1, k_0) = A(k)\Phi(k, k_0)$$

LTI

$$\Phi(k) = A^k$$

# Chapter 25

# Appendix A: Complex stuff

---

## 25.1 Complex numbers

Think about ordered pairs $(\alpha, \beta)$ of real numbers $\alpha, \beta$ and suppose we define addition operation and a multiplication operation by

$$(\alpha, \beta) + (\gamma, \delta) = (\alpha + \gamma, \beta + \delta)$$

$$(\alpha, \beta)(\gamma, \delta) = (\alpha\gamma - \beta\delta, \alpha\delta + \beta\gamma)$$

respectively. It can readily be shown that these two operations satisfy the 'usual' rules of addition and multiplication for real numbers.

Consider now any two ordered pairs of the form $(\alpha, 0)$ and $(\gamma, 0)$. Using the above definitions of addition and multiplication we obtain

$$(\alpha, 0) + (\gamma, 0) = (\alpha + \gamma, 0)$$

and

$$(\alpha, 0)(\beta, 0) = (\alpha\beta, 0)$$

(check it out) So we can identify each ordered pair of the form $(\alpha, 0)$ with the real number $\alpha$ and the above definitions of addition and multiplication reduce to the usual ones for real numbers.

Now consider $\jmath := (0, 1)$. Using the above definition of multiplication, we obtain

$$j^2 = (0, 1)(0, 1) = (-1, 0)$$

In other words (identifying $(-1, 0)$ with the real number $-1$),

$$j^2 = -1$$

We also note that for any real number $\beta$,

$$j\beta = (0, 1)(\beta, 0) = (0, \beta)$$

hence,

$$(\alpha, \beta) = (\alpha, 0) + (0, \beta) = \alpha + j\beta$$

If we regard the set $\mathbb{C}$ of complex numbers as the set of ordered pairs of real numbers for which an addition and multiplication are as defined above, we arrive at the following representation of complex nubers.

- A complex number $\lambda$ can be written as

$$\lambda = \alpha + \jmath\beta$$

where $\alpha$, $\beta$ are real numbers and $j^2 = -1$. The real numbers $\alpha$ and $\beta$ are called the *real part* and *imaginary part* of $\lambda$ and are denoted by

$$\alpha := \Re(\lambda) \qquad \beta := \Im(\lambda)$$

respectively.

Note

$$(\alpha + j\beta) \; + \; (\gamma + j\delta) = \alpha + \gamma + j(\beta + \delta)$$

and

$$\begin{aligned} (\alpha + j\beta)(\gamma + j\delta) &= \alpha\gamma + j\alpha\delta + j\beta\gamma + j^2\beta\delta \\ &= (\alpha\gamma - \beta\delta) + j(\alpha\delta + \beta\gamma) \end{aligned}$$

In other words, we recovered our original definitions of complex addition and multiplication.

- If $\lambda = \alpha + \jmath\beta$, its *complex conjugate* $\bar{\lambda}$ is defined by

$$\bar{\lambda} := \alpha - \jmath\beta$$

Some properties:

$$\begin{aligned} \Re(\lambda) &= \frac{1}{2}(\lambda + \bar{\lambda}) \\ \Im(\lambda) &= \frac{1}{2j}(\lambda - \bar{\lambda}) \end{aligned}$$

$$\begin{aligned} \overline{\lambda + \eta} &= \bar{\lambda} + \overline{\eta} \\ \overline{\lambda\eta} &= \bar{\lambda}\overline{\eta} \end{aligned}$$

- *The absolute value of $\lambda$:*

$$\begin{aligned} |\lambda| &= \sqrt{\bar{\lambda}\lambda} \\ &= \sqrt{\alpha^2 + \beta^2} \end{aligned}$$

• Suppose $p$ is an *n-th order polynomial*, i.e.,

$$p(s) = a_0 + a_1 s + \ldots + a_n s^n$$

where each coefficient $a_i$ is a complex number and $a_n \neq 0$. A complex number $\lambda$ is a *root* or *zero* of $p$ if

$$p(\lambda) = 0$$

Each $p$ has at least one root and at most $n$ distinct roots. Suppose $p$ has $l$ *distinct* roots, $\lambda_1, \lambda_2, \ldots, \lambda_l$; then

$$p(s) = a_n \prod_{i=i}^{l} (s - \lambda_i)^{m_i}$$

The integer $m_i$ is called the *algebraic multiplicity* of $\lambda_i$.

Suppose

$$p(s) = a_0 + a_1 s + \ldots + a_n s^n$$

and define

$$\bar{p}(s) = \bar{a_0} + \bar{a_1} s + \ldots + \bar{a_n} s^n$$

Then $\lambda$ is a root of $p$ iff $\bar{\lambda}$ is a root of $\bar{p}$. To see this note that $\overline{p(\lambda)} = \bar{p}(\bar{\lambda})$, hence

$$p(\lambda) = 0 \qquad \text{iff} \qquad \bar{p}(\bar{\lambda}) = 0$$

*Real coefficients.* If the coefficients $a_i$ of $p$ are real, then the roots occur in complex conjugate pairs, i.e., if $\lambda$ is a root then so is $\bar{\lambda}$.

### 25.1.1 Complex functions

The *exponential function*

$$\boxed{e^\lambda = 1 + \lambda + \tfrac{\lambda^2}{2!} + \tfrac{\lambda^3}{3!} + \ldots + \tfrac{\lambda^n}{n!} + \ldots}$$

It satisfies:

$$\frac{de^z}{dz} = e^z$$

and

$$e^{(\lambda+\eta)} = e^\lambda e^\eta$$

Can define $\cos(\lambda)$ and $\sin(\lambda)$. They satisfy

$$\cos(\lambda)^2 + \sin(\lambda)^2 = 1$$

Also, for any complex $\lambda$

$$\boxed{e^{\jmath\lambda} = \cos(\lambda) + \jmath \sin(\lambda)}$$

Hence, for any real number $\beta$,

$$e^{\jmath\beta} = \cos(\beta) + \jmath \sin(\beta)$$

and

$$|e^{\jmath\beta}| \quad = \quad \sqrt{\cos(\beta)^2 + \sin(\beta)^2} = 1$$

Also,

$$e^{(\alpha+\jmath\beta)} = e^{\alpha}e^{\jmath\beta} \ = \ e^{\alpha}\cos\beta + \jmath e^{\alpha}\sin\beta$$

Hence,

$$|e^{(\alpha+\jmath\beta)}| = |e^{\alpha}|\,|e^{\jmath\beta)}| = e^{\alpha}$$

## 25.2 Complex vectors and $\mathbb{C}^n$

Read all the material on real vectors and replace the word real with the word complex and $\mathbb{R}$ with $\mathbb{C}$.

If $x \in \mathbb{C}^n$, the *complex conjugate of $x$:*

$$\bar{x} := \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \vdots \\ \bar{x}_n \end{bmatrix}$$

Some properties:

$$\begin{aligned} \overline{x+y} &= \bar{x} + \bar{y} \\ \overline{\lambda x} &= \bar{\lambda}\bar{x} \end{aligned}$$

$$\begin{aligned} \Re(x) &= \frac{1}{2}(x+\bar{x}) \\ \Im(x) &= \frac{1}{2j}(x-\bar{x}) \end{aligned}$$

## 25.3 Complex matrices and $\mathbb{C}^{m\times n}$

Read all the material on real matrices and replace the word real with the word complex and $\mathbb{R}$ with $\mathbb{C}$

If $A \in \mathbb{C}^{m\times n}$, the *complex conjugate of $A$:*

$$\bar{A} := \begin{bmatrix} \bar{a}_{11} & \bar{a}_{12} & \dots & \bar{a}_{1n} \\ \bar{a}_{21} & \bar{a}_{22} & \dots & \bar{a}_{2n} \\ \vdots & \vdots & & \vdots \\ \bar{a}_{m1} & \bar{a}_{m2} & \dots & \bar{a}_{mn} \end{bmatrix}$$

Some properties:

$$\begin{aligned}\overline{A+B} &= \overline{A}+\overline{B}\\ \overline{\lambda A} &= \bar{\lambda}\bar{A}\\ \overline{AB} &= \bar{A}\bar{B}\end{aligned}$$

$$\Re(A) = \frac{1}{2}(A+\bar{A})$$
$$\Im(A) = \frac{1}{2j}(A-\bar{A})$$

The *complex conjugate transpose* of $A$:

$$\begin{aligned}A^* &:= \bar{A}^T\\ &= \begin{bmatrix} \bar{a}_{11} & \bar{a}_{21} & \dots & \bar{a}_{m1}\\ \bar{a}_{12} & \bar{a}_{22} & \dots & \bar{a}_{m2}\\ \vdots & \vdots & & \vdots\\ \bar{a}_{1n} & \bar{a}_{2n} & \dots & \bar{a}_{mn}\end{bmatrix}\end{aligned}$$

Some properties:

$$\begin{aligned}(A+B)^* &= A^*+B^*\\ (\lambda A)^* &= \bar{\lambda}A^*\\ (AB)^* &= B^*A^*\end{aligned}$$

**Defn.** $A$ is *hermitian* if

$$A^* = A$$

# Chapter 26

# Appendix B: Norms

## 26.1 The Euclidean norm

What is the size or magnitude of a vector? Meet *norm*.

- Consider any complex $n$ vector $x$.

  The *Euclidean norm* or 2-*norm* of $x$ is the nonegative real number given by

$$\boxed{||x|| := (|x_1|^2 + \ldots + |x_n|^2)^{\frac{1}{2}}}$$

  Note that

$$\begin{aligned} ||x|| &= (\bar{x}_1 x_1 + \ldots + \bar{x}_n x_n)^{\frac{1}{2}} \\ &= (x^* x)^{\frac{1}{2}} \end{aligned}$$

  If $x$ is *real*, these expressions become

$$\begin{aligned} ||x|| &= (x_1^2 + \ldots + x_n^2)^{\frac{1}{2}} \\ &= (x^T x)^{\frac{1}{2}} \end{aligned}$$

```
>>  norm([3; 4])

ans = 5

>>  norm([1; j])

ans = 1.4142
```

Note that in the last example $x^T x = 0$, but $x^* x = 2$.

**Properties of** $|| \cdot ||$. The Euclidean norm has the following properties.

(i) For every vector $x$,

$$||x|| \geq 0$$

and

$$||x|| = 0 \qquad \text{iff} \qquad x = 0$$

(ii) (*Triangle Inequality.*) For every pair of fectors $x$, $y$,

$$||x + y|| \leq ||x|| + ||y||$$

(iii) For every vector $x$ and every scalar $\lambda$.

$$||\lambda x|| = |\lambda| ||x||$$

Any real valued function on a vector space which has the above three properties is defined to be a norm. Two other commonly encountered norms on the space of $n$-vectors are the 1-norm:

$$\|x\|_1 := |x_1| + \ldots + |x_n|$$

and the $\infty$-norm:

$$\|x\|_\infty := \max_i |x_i|$$

As a example of a norm on the set of $m \times n$ matrices $A$, consider

$$\|A\| = \mathrm{Tr}(A'A)^{\frac{1}{2}} = \left( \sum_{i=1}^{m} \sum_{j=1}^{n} |A_{ij}|^2 \right)^{\frac{1}{2}}$$

where $\mathrm{Tr}(M)$ denotes the trace of a square matrix $M$, that is the sum of its diagonal elements. This norm is called the Frobenius norm and is the matrix analog of the Euclidean norm for $n$-vectors. We will meet other matrix norms later.

# Chapter 27

# Appendix C: Some linear algebra

All the definitions and results of this section are stated for real scalars. However, they also hold for complex scalars; to get the results for complex scalars, simply replace 'real' with 'complex' and $\mathbb{R}$ with $\mathbb{C}$.

## 27.1 Linear equations

Many problems in linear algebra can be reduced to that of solving a bunch of linear equations in a bunch of unknowns. Consider the following $m$ linear scalar equations in $n$ scalar unknowns $x_1, x_2, \ldots, x_n$:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned} \tag{27.1}$$

Given the scalars $a_{ij}$ and $b_i$, a basic problem is that of determining whether or not the above system of linear equations has a solution for $x_1, x_2, \cdots, x_n$ and, if a solution exists, determining all solutions.

One approach to solving the above problem is to carry out a sequence of elementary operations on the above equations to reduce them to a much simpler system of equations which have the same solutions as the original system. There are three elementary operations:

*1) Interchanging two equations.*
*2) Multiplication of an equation by a nonzero scalar.*
*3) Addition of a multiple of one equation to another different equation.*

Each of the above operations are reversible and do not change the set of solutions.

The above approach is usually implemented by expressing the above equations in matrix form and carrying out the above elementary operations via elementary matrix row operations.

Introducing

$$b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}, \qquad A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \qquad \text{and} \qquad x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},$$

the above system of scalar equations can written in matrix form as:

$$Ax = b \tag{27.2}$$

where $x$ is an $n$-vector, $b$ is an $m$-vector and $A$ is an $m \times n$ matrix.

**Elementary row operations.** The above elementary operations on the original scalar equations are respectively equivalent to the following elementary row operations on the augmented matrix:

$$\begin{pmatrix} A & b \end{pmatrix}$$

associated with the scalar equations.

*1) Interchanging two rows.*
*2) Multiplication of a row by a nonzero scalar.*
*3) Addition of a multiple of one row to another different row.*

**The reduced row echelon form of a matrix.** Consider any matrix $M$. For each nonzero row, the leading element of that row is the first non-zero element of the row. The matrix $M$ is said to be in reduced row echelon form if its structure satisfies the following four conditions.

*1) All zero rows come after the nonzero rows.*
*2) All leading elements are equal to one.*
*3) The leading elements have a staircase pattern.*
*4) If a column contains a leading element of some row, then all other elements of the column are zero.*

This is illustrated with the following matrix.

$$M = \begin{pmatrix} 0 & \cdots & 0 & \boxed{1} & * & \cdots & 0 & * & \cdots & 0 & \cdots & * \\ 0 & \cdots & 0 & 0 & 0 & \cdots & \boxed{1} & * & \cdots & 0 & \cdots & * \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & \boxed{1} & \cdots & * \\ \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & 0 \end{pmatrix} \tag{27.3}$$

So, the matrices

$$\begin{pmatrix} 0 & \boxed{1} & 0 & 0 & 2 \\ 0 & 0 & \boxed{1} & 0 & 3 \\ 0 & 0 & 0 & \boxed{1} & 4 \end{pmatrix} \qquad \begin{pmatrix} \boxed{1} & 2 & 0 & 0 & 2 \\ 0 & 0 & \boxed{1} & 0 & 3 \\ 0 & 0 & 0 & \boxed{1} & 4 \end{pmatrix} \qquad \begin{pmatrix} \boxed{1} & 0 & 0 & 2 & 0 \\ 0 & 0 & \boxed{1} & 3 & 0 \\ 0 & 0 & 0 & 0 & \boxed{1} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

are in reduced row echelon form, whereas the following matrices are not in reduced row echelon form.

$$\begin{pmatrix} \boxed{1} & 2 & 0 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \boxed{1} & 4 \end{pmatrix} \qquad \begin{pmatrix} \boxed{1} & 0 & 0 & 3 \\ 0 & \boxed{2} & 0 & 0 \\ 0 & 0 & \boxed{1} & 4 \end{pmatrix} \qquad \begin{pmatrix} 0 & \boxed{1} & 2 & 3 \\ \boxed{1} & 0 & 5 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad \begin{pmatrix} \boxed{1} & 2 & 0 & 0 \\ 0 & \boxed{1} & 2 & 0 \\ 0 & 0 & 0 & \boxed{1} \end{pmatrix}$$

The above matrices respectively violate conditions 1, 2, 3, 4 above.

**Fact 19** *Every matrix can be transformed into reduced row echelon form by a sequence of elementary row operations. Although the sequence of row operations is not unique, the resulting row echelon form matrix is unique.*

We will denote the row echelon form of a matrix $M$ by rref $(M)$. The following example illustrates how to obtain the reduced row echelon form of a matrix.

**Example 217**

$$M = \begin{pmatrix} 0 & 2 & 2 & 4 & -2 \\ 1 & 2 & 3 & 4 & -1 \\ 2 & 0 & 2 & 0 & 2 \end{pmatrix}$$

We reduce $M$ to its reduced row echelon form as follows:

$$\begin{pmatrix} 0 & 2 & 2 & 4 & -2 \\ 1 & 2 & 3 & 4 & -1 \\ 2 & 0 & 2 & 0 & 2 \end{pmatrix}$$

Interchange Rows 1 and 2.

$$\begin{pmatrix} 1 & 2 & 3 & 4 & -1 \\ 0 & 2 & 2 & 4 & -2 \\ 2 & 0 & 2 & 0 & 2 \end{pmatrix}$$

Add multiples of row 1 to all suceeding rows.

$$\begin{pmatrix} 1 & 2 & 3 & 4 & -1 \\ 0 & 2 & 2 & 4 & -2 \\ 0 & -4 & -4 & -8 & 4 \end{pmatrix}$$

Multiply row 2 by 1/2.

$$\begin{pmatrix} 1 & 2 & 3 & 4 & -1 \\ 0 & 1 & 1 & 2 & -1 \\ 0 & -4 & -4 & -8 & 4 \end{pmatrix}$$

Add multiples of row 2 to all succeeding rows.

$$\begin{pmatrix} 1 & 2 & 3 & 4 & -1 \\ 0 & 1 & 1 & 2 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Add multiples of row 2 to all preceeding rows.

$$\begin{pmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 2 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Sometimes the leading elements of a reduced row echelon matrix are called pivots.

**Gauss-Jordan procedure for solving systems of linear equations.** Recall the problem of solving the system of scalar linear equations (27.1) or their vector/matrix counterpart (27.2). The Gauss-Jordan procedure is a systematic method for solving such equations. One simply uses elementary row operations to transform the augmented matrix $[A\,b]$ into its reduced row echelon form which we denote by $[\hat{A}\ \hat{b}]$. Since $[\hat{A}\ \hat{b}]$ is obtained from $[A\ b]$ by elementary row operations, the set of vectors $x$ which satisfy $Ax = b$ is exactly the same as the set which satisfy $\hat{A}x = \hat{b}$. However, since $[\hat{A}\ \hat{b}]$ is in reduced row echelon form, it can be simply determined whether or not the equation $\hat{A}x = \hat{b}$ has a solution for $x$, and, if a solution exists, to determine all solutions.

**Example 218** Determine whether or not the following system of linear equations has a solution. If a solution exists, determine whether or not it is unique; if not unique, obtain an expression for all solutions and give two solutions.

$$\begin{array}{rcr} y-2z &=& 4 \\ x-y &=& 1 \\ 3x-3y+4z &=& -9 \\ -2y+10z &=& -26 \end{array}$$

Here the augmented matrix is given by

$$\begin{pmatrix} A & b \end{pmatrix} = \begin{pmatrix} 0 & 1 & -2 & 4 \\ 1 & -1 & 0 & 1 \\ 3 & -3 & 4 & -9 \\ 0 & -2 & 10 & -26 \end{pmatrix}$$

Carrying our elementary row operations, we obtain

$$\begin{pmatrix} 0 & 1 & -2 & 4 \\ 1 & -1 & 0 & 1 \\ 3 & -3 & 4 & -9 \\ 0 & -2 & 10 & -26 \end{pmatrix} \longrightarrow \begin{pmatrix} \boxed{1} & -1 & 0 & 1 \\ 0 & 1 & -2 & 4 \\ 3 & -3 & 4 & -9 \\ 0 & -2 & 10 & -26 \end{pmatrix} \longrightarrow \begin{pmatrix} \boxed{1} & -1 & 0 & 1 \\ 0 & \boxed{1} & -2 & 4 \\ 0 & 0 & 4 & -12 \\ 0 & -2 & 10 & -26 \end{pmatrix} \longrightarrow \begin{pmatrix} \boxed{1} & -1 & 0 & 1 \\ 0 & \boxed{1} & -2 & 4 \\ 0 & 0 & 4 & -12 \\ 0 & 0 & 6 & -18 \end{pmatrix}$$

$$\longrightarrow \begin{pmatrix} \boxed{1} & -1 & 0 & 1 \\ 0 & \boxed{1} & -2 & 4 \\ 0 & 0 & \boxed{1} & -3 \\ 0 & 0 & 6 & -18 \end{pmatrix} \longrightarrow \begin{pmatrix} \boxed{1} & -1 & 0 & 1 \\ 0 & \boxed{1} & -2 & 4 \\ 0 & 0 & \boxed{1} & -3 \\ 0 & 0 & 0 & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} \boxed{1} & -1 & 0 & 1 \\ 0 & \boxed{1} & 0 & -2 \\ 0 & 0 & \boxed{1} & -3 \\ 0 & 0 & 0 & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} \boxed{1} & 0 & 0 & -1 \\ 0 & \boxed{1} & 0 & -2 \\ 0 & 0 & \boxed{1} & -3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

The reduced row echelon form of $[A\ b]$ (the last matrix above) tells us that the original equations are equivalent to:

$$\begin{array}{rcl} x &=& -1 \\ y &=& -2 \\ z &=& -3 \end{array}$$

Thus, the original equations have a unique solution given by $x = -1$, $y = -2$ and $z = -3$.

From the above procedure we see that there are three possibilities when it comes to solving linear equations:

1) A unique solution.

2) Infinite number of solutions.

3) No solution.

**MATLAB.**

```
help rref

RREF   Reduced row echelon form.
   R = RREF(A) produces the reduced row echelon form of A.

   [R,jb] = RREF(A) also returns a vector, jb, so that:
       r = length(jb) is this algorithm's idea of the rank of A,
       x(jb) are the bound variables in a linear system, Ax = b,
       A(:,jb) is a basis for the range of A,
       R(1:r,jb) is the r-by-r identity matrix.

   [R,jb] = RREF(A,TOL) uses the given tolerance in the rank tests.

   Roundoff errors may cause this algorithm to compute a different
   value for the rank than RANK, ORTH and NULL.

   See also RREFMOVIE, RANK, ORTH, NULL, QR, SVD.

Overloaded methods
   help sym/rref.m
```

**Exercises**

**Exercise 157** Determine whether or not the following system of linear equations has a solution. If a solution exists, determine whether or not it is unique, and if not unique, obtain an expression for all solutions.

$$\begin{array}{rcrcrcrcr} x_1 & - & x_2 & + & 2x_3 & + & x_4 & = & 5 \\ x_1 & - & x_2 & + & x_3 & & & = & 3 \\ -2x_1 & + & 2x_2 & & & + & 2x_4 & = & -2 \\ 2x_1 & - & 2x_2 & - & x_3 & - & 3x_4 & = & 0 \end{array}$$

**Exercise 158** Determine whether or not the following system of linear equations has a solution. If a solution exists, determine whether or not it is unique; if not unique, obtain an expression for all solutions and given two solutions.

$$\begin{array}{rcrcrcr} x_1 & - & 2x_2 & - & 4x_3 & = & -1 \\ 2x_1 & - & 3x_2 & - & 5x_3 & = & -1 \\ x_1 & & & + & 2x_3 & = & 1 \\ x_1 & + & x_2 & + & 5x_3 & = & 2 \end{array}$$

**Exercise 159** Determine whether or not the following system of linear equations has a solution. If a solution exists, determine whether or not it is unique; if not unique, obtain an expression for all solutions and give two solutions.

$$\begin{array}{rcrcrcrcr} x_1 & - & 2x_2 & & x_3 & - & 2x_4 & = & -1 \\ -3x_1 & + & 7x_2 & - & 3x_3 & + & 7x_4 & = & 4 \\ 2x_1 & - & x_2 & + & 2x_3 & - & x_4 & = & 2 \end{array}$$

**Exercise 160** Determine whether or not the following system of linear equations has a solution. If a solution exists, determine whether or not it is unique; if not unique, obtain an expression for all solutions and give two solutions.

$$\begin{array}{rcrcrcrcr} & & x_2 & & & + & x_4 & = & 2 \\ 2x_1 & + & 3x_2 & + & 2x_3 & + & 3x_4 & = & 8 \\ x_1 & + & x_2 & + & x_3 & + & x_4 & = & 3 \end{array}$$

**Exercise 161** Determine whether or not the following system of linear equations has a solution. If a solution exists, determine whether or not it is unique; if not unique, obtain an expression for all solutions and given two solutions.

$$\begin{array}{rcrcrcr} x_1 & - & 2x_2 & - & 4x_3 & = & -1 \\ 2x_1 & - & 3x_2 & - & 5x_3 & = & -1 \\ x_1 & & & + & 2x_3 & = & 1 \\ x_1 & + & x_2 & + & 5x_3 & = & 2 \end{array}$$

**Exercise 162** Determine whether or not the following system of linear equations has a solution. If a solution exists, determine whether or not it is unique; if not unique, obtain an expression for all solutions and given two solutions.

$$\begin{aligned} y - 2z &= 4 \\ x - y &= 1 \\ 3x - 3y + 4z &= -9 \\ -2y + 10z &= -26 \end{aligned}$$

**Exercise 163** We consider here systems of linear equations of the form

$$Ax = b$$

$$\text{with} \qquad x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

For each of the following cases, we present the reduced row echelon form of the augmented matrix $[A\,b]$. Determine whether or not the corresponding system of linear equations has a solution. If a solution exists, determine whether or not it is unique; if not unique, obtain an expression for all solutions and given two solutions.

$$\text{(a)} \begin{pmatrix} 1 & 5 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{(b)} \begin{pmatrix} 1 & -2 & 0 & 6 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{(c)} \begin{pmatrix} 1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{(d)} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

## 27.2 Subspaces

Recall that a vector space is a set equipped with two operations, vector addition and multiplication by a scalar, and these operations must satisfy certain properties. The following concept generalizes the notion of a line or a plane which passes through the origin.

*A non-empty subset $\mathcal{S}$ of a vector space is a* subspace *if for every pair of elements $x, y$ in $\mathcal{S}$ and every pair of scalars $\alpha, \beta$, the vector $\alpha x + \beta y$ is contained in $\mathcal{S}$.*

In other words, a subset of a vector space is a subspace if it is closed under the operations of addition and multiplication by a scalar. Note that a subspace can be considered a vector space (Why?). An immediate consequence of the above definition is that a subspace must contain the zero vector; to see this consider $\alpha$ and $\beta$ to be zero.

**Example 219** Consider the set of real 2-vectors $x$ which satisfy

$$x_1 - x_2 = 0$$

This set is a subspace of $\mathbb{R}^2$. Note that the set of vectors satisfying

Figure 27.1: A simple subspace

$$x_1 - x_2 - 1 = 0$$

is not a subspace.

**Example 220** Consider any real $m \times n$ matrix $A$ and consider the set of real $n$-vectors $x$ which satisfy

$$Ax = 0$$

This set is a subspace of $\mathbb{R}^n$.

**Example 221** The set of real matrices of the form

$$\begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix}$$

is a subspace of the space of real $2 \times 2$ matrices.

**Example 222** As a more abstract example of a subspace space, let $\mathcal{V}$ be the set of continuous real-valued functions which are defined on the interval $[0, \infty)$and let addition and scalar multiplication be as defined earlier Suppose $\mathcal{W}$ be the subset of $\mathcal{V}$ which consists of the differentiable functions $\mathcal{V}$. Then $\mathcal{W}$ is a subspace of $\mathcal{V}$.

## 27.3 Basis

### 27.3.1 Span (not Spam)

Consider a bunch $\mathcal{X}$ of vectors. A linear combination of these vectors is a vector which can be expressed as

$$\alpha_1 x^1 + \alpha_2 x^2 + \ldots + \alpha_m x^m$$

where each $x^i$ is in the bunch and each $\alpha_i$ is a scalar.

*The* span *of a bunch $\mathcal{X}$ of vectors is the set of all vectors of the form*

$$\alpha_1 x^1 + \alpha_2 x^2 + \ldots + \alpha_m x^m$$

*where each $x^i$ is in the bunch and each $\alpha_i$ is a scalar; it is denoted by span $\mathcal{X}$.*

In other words, the span of a bunch of vectors is the set of all linear combinations of vectors from the bunch.

**Exercise 164** Show that span $\mathcal{X}$ is a subspace.

**Example 223**

$$\operatorname{span}\left\{\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\right\} = \mathbb{R}^2$$

To see this, note that any element $x$ of $\mathbb{R}^2$ can be expressed as

$$\begin{pmatrix}x_1\\x_2\end{pmatrix} = x_1\begin{pmatrix}1\\0\end{pmatrix} + x_2\begin{pmatrix}0\\1\end{pmatrix}$$

**Example 224**

$$\operatorname{span}\left\{\begin{pmatrix}1&0\\0&0\end{pmatrix}, \begin{pmatrix}0&1\\0&0\end{pmatrix}, \begin{pmatrix}0&0\\1&0\end{pmatrix}, \begin{pmatrix}0&0\\0&1\end{pmatrix}\right\} = \mathbb{R}^{2\times 2}$$

To see this, note that any element $A$ of $\mathbb{R}^{2\times 2}$ can be expressed as:

$$\begin{pmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{pmatrix} = a_{11}\begin{pmatrix}1&0\\0&0\end{pmatrix} + a_{12}\begin{pmatrix}0&1\\0&0\end{pmatrix} + a_{21}\begin{pmatrix}0&0\\1&0\end{pmatrix} + a_{22}\begin{pmatrix}0&0\\0&1\end{pmatrix}$$

**Fact 20** *Consider a bunch $\mathcal{X}$ of vectors. The following 'experiences' do not affect span.*

*(a) Multiplication of one of the vectors by a nonzero scalar.*

*(b) Replacing a vector by the sum of itself and a scalar multiple of another (different) vector in the bunch.*

*(c) Removal of a vector which is a linear combination of other vectors in the bunch.*

*(d) Inclusion of a vector which is in the span of the original bunch.*

**Span and rref.** One can use rref to solve the following problem. *Given a bunch of $n$-vectors $x^1, x^2, \cdots, x^m$ and another $n$-vector $b$, determine whether or not $b$ is in the span of the bunch.* To solve this problem, consider any linear combination $\alpha_1 x^1 + \alpha_2 x^2 + \ldots + \alpha_m x^m$ of the vectors $x^1, x^2, \cdots, x^m$. Let $A$ be the matrix whose columns are the vectors $x^1, x^2, \cdots, x^m$, that is,

$$A = \begin{pmatrix} x^1 & x^2 & \cdots & x^m \end{pmatrix}$$

and let $\alpha$ be the $m$-vector whose components are $\alpha_1, \cdots, \alpha_m$, that is,

$$\alpha = [\ \alpha_1 \quad \alpha_2 \quad \cdots \quad \alpha_m\ ]^T .$$

Then the above linear combination can be expressed as

$$\alpha_1 x^1 + \alpha_2 x^2 + \ldots + \alpha_m x^m = \begin{pmatrix} x^1 & x^2 & \cdots & x^m \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{pmatrix} = A\alpha .$$

Thus, $b$ is in the span of the bunch if and only if the equation

$$A\alpha = b$$

has a solution for $\alpha$. One can determine whether or not a solution exists by considering $\hat{M} = \text{rref}(M)$ where

$$M = [A\ b] = \begin{pmatrix} x^1 & x^2 & \cdots & x^m & b \end{pmatrix} .$$

**Example 225** Suppose we wish to determine whether or not the vector $(1, 1, -1)$ is a linear combination of the vectors

$$\begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}, \qquad \begin{pmatrix} 0 \\ 3 \\ 1 \end{pmatrix}, \qquad \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}.$$

Since

$$\text{rref}\left( \begin{pmatrix} 2 & 0 & -1 & 1 \\ 1 & 3 & 1 & 1 \\ 1 & 1 & 0 & -2 \end{pmatrix} \right) = \begin{pmatrix} 1 & 0 & -0.5 & 0 \\ 0 & 1 & 0.5 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

it follows that the first vector is not a linear combination of the remaining three.

## 27.3.2 Linear independence

*A bunch of vectors, $x^1, x^2, \ldots, x^m$ is* linearly dependent *if there exist scalars $\alpha_1, \alpha_2, \ldots, \alpha_m$, not all zero, such that*

$$\alpha_1 x^1 + \alpha_2 x^2 + \ldots + \alpha_m x^m = 0 \tag{27.4}$$

*The bunch is* linearly independent *if it is not linearly dependent, that is, the bunch is linearly independent if and only if (27.4) implies that*

$$\alpha_1, \alpha_2, \ldots, \alpha_m = 0$$

If the bunch consists of a single vector, then it is linearly independent if and only if that vector is not the zero vector. If a bunch consists of more than one vector, then it is linearly dependent if and only if one of the vectors can be expressed as a linear combination of the preceding vectors in the bunch, that is for some $j$,

$$x^j = \beta_1 x^1 + \beta_2 x^2 + \ldots + \beta_{j-1} x^{j-1}$$

for some scalars $\beta_i$.

**Example 226** The following bunches of vectors are linearly dependent.

(i)

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

(ii)

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \end{pmatrix}$$

(iii)

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 7 \\ 6 \\ 5 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}$$

Why?

(iv)

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

**Example 227** The following bunches of vectors are linearly independent.

(a)

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

(b)

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

(c)

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 7 \\ 6 \\ 5 \end{pmatrix}$$

(d)

$$\begin{pmatrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 494 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 0 \\ 5 \\ 6 \\ 4 \\ 74 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 203 \\ 11 \end{pmatrix}$$

To see this:

$$\alpha_1 \begin{pmatrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 494 \end{pmatrix} + \alpha_2 \begin{pmatrix} 0 \\ 0 \\ 5 \\ 6 \\ 4 \\ 74 \end{pmatrix} + \alpha_3 \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 203 \\ 11 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

implies

$$\begin{array}{ccccccc} \alpha_1 & & & & & = & 0 \\ 2\alpha_1 & + & 5\alpha_2 & & & = & 0 \\ 4\alpha_1 & + & 4\alpha_2 & + & 203\alpha_3 & = & 0 \end{array}$$

The first equation implies $\alpha_1 = 0$; combining this with the second equation yields $\alpha_2 = 0$; combining these results with the last equation produces $\alpha_3 = 0$. So, no nontrivial linear combination of this bunch equals zero; hence the bunch is linearly independent.

(e)

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

**Exercise 165** Useful generalization of example 227 (d)

Consider a bunch of $m$ vectors in $\mathbb{R}^n$ ($m \le n$) which have the following structure.

$$\begin{pmatrix} 0 \\ \vdots \\ 0 \\ x^1_{n_1} \\ \vdots \\ \vdots \\ \vdots \end{pmatrix}, \quad \begin{pmatrix} 0 \\ \vdots \\ \vdots \\ 0 \\ x^2_{n_2} \\ \vdots \\ \vdots \end{pmatrix}, \quad \ldots, \quad \begin{pmatrix} 0 \\ \vdots \\ \vdots \\ \vdots \\ 0 \\ x^m_{n_m} \\ \vdots \end{pmatrix}$$

$$x^1_{n_1}, x^2_{n_2}, \ldots, x^m_{n_m} \neq 0$$

In words, the first $n_i - 1$ components of the $i$-th vector are zero; its $n_i$ component is nonzero; and $n_1 < n_2 < \ldots < n_{n_m}$.

Prove that this bunch is linearly independent.

• Consider a bunch

$$x^1, x^2, \ldots, x^m$$

of vectors. It can readily be shown that:

**Fact 21** *The following 'experiences' do not affect linear independence status.*

*(a) Multiplication of one of the vectors by a nonzero scalar.*

*(b) Replacing a vector by the sum of itself and a scalar multiple of another (different) vector in the bunch.*

*(c) Removal of a vector from the bunch.*

*(d) Inclusion of a vector which is* not *in the span of the original bunch.*

**Using rref to determine linear independence.** Suppose we want to determine whether or not a bunch of $n$-vectors, $x^1, x^2, \cdots, x^m$, are linearly independent. We can achieve this using rref. To see this, consider any linear combination $\alpha_1 x^1 + \alpha_2 x^2 + \ldots + \alpha_m x^m$ of the vectors $x^1, x^2, \cdots, x^m$. Let $A$ be the $n \times m$ matrix whose columns are the vectors $x^1, x^2, \cdots, x^m$, that is,

$$A = \begin{pmatrix} x^1 & x^2 & \cdots & x^m \end{pmatrix},$$

and let $\alpha$ be the $m$-vector whose components are $\alpha_1, \cdots, \alpha_m$, that is,

$$\alpha = [\ \alpha_1 \quad \alpha_2 \quad \cdots \quad \alpha_m\ ]^T.$$

Then the above linear combination of vectors can be expressed as

$$\alpha_1 x^1 + \alpha_2 x^2 + \ldots + \alpha_m x^m = \begin{pmatrix} x^1 & x^2 & \cdots & x^m \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{pmatrix} = A\alpha.$$

Thus the linear combination is zero if and only if $A\alpha = 0$. Letting $\hat{A}$ be the reduced row echelon form of $A$, it follows that $\alpha$ satisfies $A\alpha = 0$ if and only if $\hat{A}\alpha = 0$. So, the columns of $A$ are linearly independent if and only if the columns of $\hat{A}$ are linearly independent. Moreover, a column of $A$ is a linear combination of the preceding columns of $A$ if and only if the corresponding column of $\hat{A}$ is the same linear combination of the preceding columns of $\hat{A}$. From the above observations, we can obtain the following results where $\hat{A}$ is the reduced row echelon form of $A$.

(a) *The vectors are linear independent if and only if every column of $\hat{A}$ contains a leading element, that is, the number of leading elements of $\hat{A}$ is the same as the number of vectors.*

(b) If $\hat{A}$ has a zero column, then the corresponding $x$-vector is zero; hence the vectors are linearly dependent.

(c) If a non-zero column of $\hat{A}$ does not contain a leading element, then the corresponding vector can be written as a linear combination of the preceding vectors, hence the vectors are linearly dependent.

**Example 228** Show that the vectors,

$$\begin{pmatrix} 1 \\ 1 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 3 \\ -2 \end{pmatrix},$$

are linear dependent.

*Solution:*

$$\text{rref}\left(\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & -1 & 3 \\ -1 & 0 & -2 \end{pmatrix}\right) = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

The vectors are linearly dependent. The third vector is the twice the first vector minus the second vector.

**Example 229** Are the vectors

$$\begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix},$$

linearly independent?

*Solution:*

$$\text{rref}\left(\begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 2 & 0 & -1 \end{pmatrix}\right) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The vectors are linearly independent.

### 27.3.3 Basis

*A* basis *for a vector space is a bunch of linearly independent vectors whose span is the space.*

A vector space is called finite dimensional if it has a finite basis; otherwise it is infinite dimensional.

For a finite dimensional vector space $\mathcal{V}$, it can be shown that the number of vectors in every basis is the same; this number is called the dimension of the space and is denoted dim $\mathcal{V}$. We refer to a vector space with dimension $n$ as an $n$-dimensional space. In an $n$-dimensional space, every bunch of more than $n$ vectors space must be linearly dependent, and every bunch of $n$ linearly independent vectors must be a basis.

**Example 230** The following vectors constitute a basis for $\mathbb{R}^2$ (Why?) :

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Hence $\mathbb{R}^2$ is a 2-dimensional space (Big surprise!).

**Example 231** The following vectors constitute a basis for $\mathbb{R}^{2\times 2}$ (Why?) :

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

Hence $\mathbb{R}^{2\times 2}$ is a 4-dimensional space.

**Example 232** The following vectors constitute a basis for the subspace of example 221.

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

**Exercise 166** Give an example of a basis for $\mathbb{R}^3$ and for $\mathbb{R}^{2\times 3}$. What are the dimensions of these spaces?

**Coordinates.** If $e^1, e^2, \ldots, e^n$ is a basis for an $n$-dimensional vector space, then every vector $x$ in the space has a *unique* representation of the form:

$$x = \xi_1 e^1 + \xi_2 e^2 + \ldots + \xi_n e^n$$

The scalars $\xi_1, \xi_2, \cdots, \xi_n$ are called the coordinates of $x$ wrt the basis $e^1, e^2, \ldots, e^n$. Thus there is a one-to-one correspondence between any real $n$-dimensional vector space and $\mathbb{R}^n$.

$$x \quad \overrightarrow{\longleftarrow} \quad \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{pmatrix}$$

Any linear map from a real $n$-dimensional $\mathcal{V}$ space to a real $m$-dimensional space $\mathcal{W}$ can be 'uniquely' represented by a real $m \times n$ matrix; that is, there is a one-to-one correspondence between the set of linear maps from $\mathcal{V}$ to $\mathcal{W}$ and $\mathbb{R}^{m\times n}$. This is achieved by choosing bases for $\mathcal{V}$ and $\mathcal{W}$.

## 27.4 Range and null space of a matrix

### 27.4.1 Null space

Consider any real $m \times n$ matrix $A$.

*The* null space of *$A$, denoted by $\mathcal{N}(A)$, is the set of real $n$-vectors $x$ which satisfy*

$$Ax = 0$$

*that is,*

$$\mathcal{N}(A) := \{x \in \mathbb{R}^n : Ax = 0\}$$

- $\mathcal{N}(A)$ is a subspace of $\mathbb{R}^n$. (**Exercise**)

- The nullity of $A$ is the dimension of its null space, that is,

$$\text{nullity}\, A := \dim \mathcal{N}(A)$$

Determining the null space of a matrix $A$ simply amounts to solving the homogeneous equation $Ax = 0$. This we can do by obtaining rref $(A)$.

**Example 233** A null space calculation

Let's compute a basis for the null space of the following matrix.

$$A = \begin{pmatrix} 0 & 2 & 2 & 4 & -2 \\ 1 & 2 & 3 & 4 & -1 \\ 2 & 0 & 2 & 0 & 2 \end{pmatrix}$$

We saw in Example 217 that the reduced row echelon form of $A$ is given by

$$\hat{A} = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 2 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Hence, the vectors in the null space of $A$ are those vectors $x$ which satisfy $\hat{A}x = 0$, that is,

$$\begin{aligned} x_1 + x_3 + x_5 &= 0 \\ x_2 + x_3 + 2x_4 - x_5 &= 0 \end{aligned}$$

these equations are equivalent to

$$\begin{aligned} x_1 &= -x_3 - x_5 \\ x_2 &= -x_3 - 2x_4 + x_5 \end{aligned}$$

Hence, we can consider $x_3, x_4, x_5$ arbitrary and for any given $x_3, x_4, x_5$, the remaining components $x_1$ and $x_2$ are uniquely determined; hence, all vectors in the null space must be of the form

$$\begin{pmatrix} -x_3 - x_5 \\ -x_3 - 2x_4 + x_5 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = x_3 \begin{pmatrix} -1 \\ -1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} 0 \\ -2 \\ 0 \\ 1 \\ 0 \end{pmatrix} + x_5 \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

where $x_3, x_4, x_5$ are arbitrary. Thus the null space of $A$ is the subspace spanned by the three vectors:

$$\begin{pmatrix} -1 \\ -1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \qquad \begin{pmatrix} 0 \\ -2 \\ 0 \\ 1 \\ 0 \end{pmatrix} \qquad \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

Since the three vectors are linearly independent, they form a basis for the null space of $A$; hence the nullity of $A$ is 3.

**Example 234** Consider any non-zero (angular velocity) matrix of the form:

$$A = \begin{pmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{pmatrix}$$

The null space of $A$ is the set of vectors $x$ which satisfy:

$$\begin{array}{rcrcrcl} & & -\omega_3 x_2 & + & \omega_2 x_3 & = & 0 \\ \omega_3 x_1 & & & - & \omega_1 x_3 & = & 0 \\ -\omega_2 x_1 & + & \omega_1 x_2 & & & = & 0 \end{array}$$

Every solution to these equations is a scalar multiple of

$$x = \begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix}$$

Hence, the null space of $A$ is the 1-dimensional subspace of $\mathbb{R}^3$ spanned by this vector.

- *The set of equilibrium states of a continuous LTI system $\dot{x} = Ax$ is the null space of $A$.*

- *The set of equilibrium states of a discrete LTI system $x(k+1) = Ax(k)$ is the null space of $A - I$.*

### 27.4.2 Range

Consider a real $m \times n$ matrix $A$.

*The* range of $A$*, denoted by* $\mathcal{R}(A)$*, is the set of* $m$*-vectors of the form* $Ax$ *where* $x$ *is an* $n$*-vector, that is,*

$$\mathcal{R}(A) := \{Ax : x \in \mathbb{R}^n\}$$

Thus, the range of $A$ is the set of vectors $b$ for which the equation $Ax = b$ has a least one solution for $x$.

- $\mathcal{R}(A)$ is a subspace of $\mathbb{R}^m$. (**Exercise**)

*The* rank *of* $A$ *is the dimension of its range space, that is,*

$$\mathit{rank}\, A := \mathit{dim}\; \mathcal{R}(A)$$

Since

$$A = \begin{pmatrix} a_1 & a_2 & \dots & a_n \end{pmatrix}$$

where $a_1, a_2, \dots, a_n$ are the columns of $A$, it should be clear that

$$Ax = \begin{pmatrix} a_1 & a_2 & \dots & a_n \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_1\, a_1 + x_2\, a_2 + \dots + x_n\, a_n$$

Hence, the range of a matrix is the set of all linear combinations of its columns. Thus the rank of $A$ equals the maximum number of linearly independent columns of $A$.

From Fact 20 it should be clear that if a matrix $\tilde{A}$ is obtained from a matrix $A$ by a sequence of the following operations then $A$ and $\tilde{A}$ have the same range.

(a) Interchanging two columns.

(b) Mutiplication of one of the columns by a nonzero scalar.

(c) Replacing a column by the sum of itself and a scalar multiple of another (different) column in the bunch.

(d) Removal of a column which is a linear combination of other columns in the bunch.

(e) Inclusion of a column which is a linear combination of other columns in the bunch.

#### Calculation of range basis and rank with rref

Here, we present two ways to calculate a basis for the range of a matrix using rref.

**Method one: rref**$(A)$ A basis for the range of $A$ is given by the columns of $A$ corresponding to the columns of rref$(A)$ which contain leading elements. Thus the rank of $A$ is the number of leading elements in rref$(A)$. This method yields a basis whose members are columns from $A$.

**Method two: rref**$(A^T)$ Let $\hat{B}$ be the reduced row echelon form of $A^T$, that is

$$\hat{B} = \mathrm{rref}(A^T)\,.$$

Then the nonzero columns of $\hat{B}^T$ form a basis for the range of $A$. Thus the rank of $A$ is the number of leading elements in rref$(A^T)$.

From the above considerations, it follows that, for any matrix $A$,

$$\mathrm{rank}\, A^T = \mathrm{rank}\, A$$

hence,

$$\begin{aligned}\mathrm{rank}\, A &= \text{maximum number of linearly independent columns of } A\\ &= \text{maximum number of linearly independent rows of } A\end{aligned}$$

**Example 235 (A range calculation)** Let's compute a basis for the range of the matrix,

$$A = \begin{pmatrix} 0 & 1 & 1 & 2 \\ 1 & 2 & 3 & 4 \\ 2 & 0 & 2 & 0 \end{pmatrix}$$

Using Method one, we compute

$$\mathrm{rref}(A) = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Since the leading elements of rref$(A)$ occur in columns 1 and 2, a basis for the range of $A$ is given by columns 1 and 2 of $A$, that is

$$\begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}.$$

Also, the rank of $A$ is two.

Using Method two, we compute

$$\mathrm{rref}(A^T) = \begin{pmatrix} 1 & 0 & -4 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Hence,

$$\text{rref}(A^T)^T = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -4 & 2 & 0 & 0 \end{pmatrix}$$

Since columns 1 and 2 are the non-zero columns of $\text{rref}(A^T)^T$ , a basis for the range of $A$ is given by columns 1 and 2 of $\text{rref}(A^T)^T$, that is

$$\begin{pmatrix} 1 \\ 0 \\ -4 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}.$$

Again we obtain that the rank of $A$ is two.

**Fact 22** *If $A$ is an $m \times n$ matrix,*

$$\boxed{\mathit{dim}\ \mathcal{R}(A) + \mathit{dim}\ \mathcal{N}(A) = n}$$

**Fact 23** *If $A$ and $B$ are two matrices which have the same numbers of columns, then*

$$\mathcal{N}(A) = \mathcal{N}(B) \qquad \textit{if and only if} \qquad \mathcal{R}(A^T) = \mathcal{R}(B^T)$$

**Exercise 167** Picturize the range and the null space of the following matrices.

$$\begin{pmatrix} 2 & 1 \\ 2 & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}$$

**Exercise 168**

(a) Find a basis for the null space of the matrix,

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \end{pmatrix}.$$

(b) What is the nullity of A?

Check your answer using MATLAB.

**Exercise 169**

(a) Find a basis for the range of the matrix,

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \end{pmatrix}.$$

(b) What is the rank of $A$?

Check your answer using MATLAB.

**Exercise 170** Obtain a basis for the range of each of the following matrices. Also, what is the rank and nullity of each matrix?

$$\begin{pmatrix} 4 & 1 & -1 \\ 3 & 2 & -3 \\ 1 & 3 & 0 \end{pmatrix} \qquad \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 4 & 1 & 2 \\ 3 & 4 & 5 & 0 & 0 \end{pmatrix}$$

Check your answers using MATLAB.

**Exercise 171** Obtain a basis for the the null space of each of the matrices in exercise 170. Check your answers using MATLAB

### 27.4.3 Solving linear equations revisited

Recall the problem of solving the following $m$ scalar equations for $n$ scalar unknowns $x_1, x_2, \dots, x_n$:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n &= b_m \end{aligned}$$

These equations can be written in matrix form as:

$$Ax = b$$

where $x$ is $n \times 1$, $b$ is $m \times 1$ and $A$ is $m \times n$.

• This matrix equation has at least one solution for $x$ if and only if $b$ is in the range of $A$. This is equivalent to

$$\text{rank } (A\ b) = \text{rank } A$$

The matrix equation has at least one solution for *every* $b$ if and only if the range of $A$ is $R^m$, that is, the rank of $A$ is $m$, or equivalently, the nullity of $A$ is $n - m$; hence $n \geq m$.

• If $x^*$ is any particular solution of $Ax = b$, then all other solutions have the form

$$x^* + e$$

where $e$ is in the null space of $A$.

To see this:

Hence, a solution is *unique* if and only if the nullity of $A$ is zero or equivalently, the rank of $A$ is $n$; hence $n \leq m$.

• It follows from the above that $A$ is invertible if and only if its rank is $m$ and its nullity is 0; since $\text{rank}\, A + \text{nullity}\, A = n$, we must have $m = n$, that is $A$ must be square.

**MATLAB** A solution (if one exists) to the equation $Ax = b$ can be solved by

```
x= A\b
```

```
>> help orth

 ORTH   Orthogonalization.
        Q = orth(A) is an orthonormal basis for the range of A.
        Q'*Q = I, the columns of Q span the same space as the columns
        of A and the number of columns of Q is the rank of A.

        See also SVD, ORTH, RANK.

>> help rank

 RANK   Number of linearly independent rows or columns.
        K = RANK(X) is the number of singular values of X
        that are larger than MAX(SIZE(X)) * NORM(X) * EPS.
        K = RANK(X,tol) is the number of singular values of X that
        are larger than tol.

>> help null

 NULL   Null space.
        Z = null(A) is an orthonormal basis for the null space of A.
        Z'*Z = I, A*Z has negligible elements, and the number of columns
        of Z is the nullity of A.

        See also SVD, ORTH, RANK.
```

**Exercise 172** Using MATLAB, compute the rank of the matrix

$$\begin{pmatrix} 1 & 0 \\ 0 & \delta \end{pmatrix}$$

for

$$\delta = 1,\ 10^{-3},\ 10^{-10},\ 10^{-15},\ 10^{-16}$$

Any comments?

### Range and null space of linear maps

Suppose $\mathcal{A}$ is a linear map from an $n$-dimensional space $\mathcal{V}$ to an $m$-dimensional space $\mathcal{W}$. Then the above definitions and results for matrices can be generalized as follows.

*Range of $\mathcal{A}$*

$$\mathcal{R}(\mathcal{A}) := \{\mathcal{A}x : x \in \mathcal{V}\}$$

• $\mathcal{R}(\mathcal{A})$ is a subspace of $\mathcal{W}$.

*rank* $\mathcal{A}$ *:= dim* $\mathcal{R}(\mathcal{A})$

*Null space of* $\mathcal{A}$

$$\mathcal{N}(\mathcal{A}) := \{x \in \mathcal{V} : \mathcal{A}x = 0\}$$

• $\mathcal{N}(\mathcal{A})$ is a subspace of $\mathcal{V}$.

*nullity of* $\mathcal{A}$ *:= dim* $\mathcal{N}(\mathcal{A})$

**Fact 24**

$$\mathit{dim}\, \mathcal{R}(\mathcal{A}) + \mathit{dim}\, \mathcal{N}(\mathcal{A}) = n$$

• Suppose $\mathcal{W} = \mathcal{V}$. Then $\mathcal{A}$ is invertible if and only if rank $\mathcal{A} = n$.

Suppose $b$ is given vector in $\mathcal{W}$ and consider the equation.

$$\mathcal{A}x = b$$

This equation has a solution for $x$ if and only if $b$ is in the range of $\mathcal{A}$. If $x^*$ is any particular solution of $\mathcal{A}x = b$, then all other solutions have the form

$$x^* + e$$

where $e$ is in the null space of $\mathcal{A}$.........

## 27.5 Coordinate transformations

Suppose we have $n$ scalar variables $x_1, x_2, \ldots, x_n$ and we implicitly define new scalar variables $\xi_1, \xi_2, \ldots, \xi_n$ by

$$\boxed{x = U\xi}$$

where $U$ is an $n \times n$ invertible matrix. Then, $\xi$ is explicitly given by

$$\xi = U^{-1}x$$

We can obtain a geometric interpretation of this change of variables as follows. First, observe that

$$\underbrace{\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}}_{x} = x_1 \underbrace{\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}}_{e^1} + x_2 \underbrace{\begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}}_{e^2} + \ldots + x_n \underbrace{\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}}_{e^n}$$

that is the scalars $x_1, x_2, \ldots, x_n$ are the coordinates of the vector $x$ wrt the *standard basis* $e^1, e^2, \ldots, e^n$, or, $x$ is the coordinate vector of itself wrt the standard basis. Suppose

$$U = \begin{pmatrix} u^1 & \ldots & u^j & \ldots & u^n \end{pmatrix}$$

that is, $u^j$ is the $j$-th column of $U$. Since $U$ is invertible, its columns $u^1, u^2, \ldots, u^n$ are linearly independent; hence they form a basis for $\mathbb{R}^n$. Since $x = U\xi$ can be written as

$$x = \xi_1 u^1 + \xi_2 u^2 + \ldots + \xi_n u^n$$

we see that $\xi_1, \xi_2, \ldots, \xi_n$ are the coordinates of $x$ wrt to the new basis $u^1, u^2, \cdots, u^n$ and the vector $\xi$ is the coordinate vector of the vector $x$ wrt this new basis. So $x = U\xi$ defines a *coordinate transformation.*

### 27.5.1 Coordinate transformations and linear maps

Suppose we have a linear map from $\mathbb{R}^n$ into $\mathbb{R}^m$ defined by

$$y = Ax$$

where $A$ is an $m \times n$ matrix. Suppose we introduce coordinate transformations

$$x = U\xi\,, \qquad y = V\eta$$

where $U$ and $V$ are invertible matrices. Then the map is described by

$$\eta = \Lambda\xi$$

where

$$\Lambda = V^{-1}AU$$

Recalling that the columns $u^1, u^2, \ldots, u^n$ of $U$ form a basis for $\mathbb{R}^n$, and the columns $v^1, v^2, \ldots, v^m$ of $V$ form a basis for $\mathbb{R}^m$, we have now the following very useful result.

• **Useful result:** Suppose

$$Au^j = \alpha_{1j}v^1 + \alpha_{2j}v^2 + \ldots + \alpha_{mj}v^m$$

Then the matrix $\Lambda = V^{-1}AU$ is uniquely given by

$$\Lambda_{ij} = \alpha_{ij}$$

that is $\Lambda_{ij}$ is the $i$-th coordinate of the vector $Au^j$ wrt the basis $v^1, v^2, \ldots, v^m$, or, the $j$-th column of $\Lambda$ is the coordinate vector of $Au^j$ wrt the basis $v^1, v^2, \ldots, v^m$.

**Example 236** Suppose $A$ is a $2 \times 2$ matrix and there are two linearly independent vectors $u^1$ and $u^2$ which satisfy:

$$Au^1 = 2u^1 \qquad \text{and} \qquad Au^2 = u^1 + 2u^2$$

Letting $U = \begin{pmatrix} u^1 & u^2 \end{pmatrix}$, we can use the above useful result to obtain that

$$U^{-1}AU = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$$

**Example 237** Suppose $A$ is a $3 \times 2$ matrix while $U = \begin{pmatrix} u^1 & u^2 \end{pmatrix}$ and $V = \begin{pmatrix} v^1 & v^2 & v^3 \end{pmatrix}$ are invertible $3 \times 3$ and $2 \times 2$ matrices, respectively, which satisfy

$$Au^1 = v_1 + 2v_3 \qquad Au_2 = 2v_2 + v_3$$

Then, applying the above useful result, we obtain

$$V^{-1}AU = \begin{pmatrix} 1 & 0 \\ 0 & 2 \\ 2 & 1 \end{pmatrix}$$

*Proof of useful result.* Premultiply $\Lambda = V^{-1}AU$ by $V$ to get:

$$V\Lambda = AU$$

On the right we have

$$AU \;=\; A\begin{pmatrix} u^1 & \ldots & u^j & \ldots & u^n \end{pmatrix} = \begin{pmatrix} Au^1 & \ldots & Au^j & \ldots & Au^n \end{pmatrix}$$

And on the left,

$$V\Lambda = \begin{pmatrix} v^1 & v^2 & \dots & \dots & v^m \end{pmatrix} \begin{pmatrix} \Lambda_{11} & \dots & \Lambda_{1j} & \dots & \Lambda_{1n} \\ \Lambda_{21} & \dots & \Lambda_{2j} & \dots & \Lambda_{2n} \\ \vdots & & \vdots & & \vdots \\ \vdots & & \vdots & & \vdots \\ \Lambda_{m1} & \dots & \Lambda_{mj} & \dots & \Lambda_{mn} \end{pmatrix}$$

$$= \left( \Lambda_{11}v^1 + \dots + \Lambda_{m1}v^m \vdots \quad \dots\dots \quad \vdots \underbrace{\Lambda_{1j}v^1 + \ldots + \Lambda_{mj}v^m}_{j\text{-th column}} \vdots \quad \dots\dots \quad \vdots \Lambda_{1n}v^1 + \ldots + \Lambda_{mn}v^m \right)$$

Comparing the expressions for the $j$-th columns of $AU$ and $V\Lambda$ yields

$$\Lambda_{1j}v^1 + \ldots + \Lambda_{mj}v^m = Au^j = \alpha_{1j}v^1 + \ldots + \alpha_{mj}v^m$$

Since the $\{v^1, \ldots, v^m\}$ is a basis we must have $\Lambda_{ij} = \alpha_{ij}$. ■

### 27.5.2 The structure of a linear map

Consider again a linear map from $\mathbb{R}^n$ into $\mathbb{R}^m$ defined by

$$y = Ax$$

where $A$ is an $m \times n$ matrix. Suppose the rank of $A$ is $r$. Then the dimension of the null space of $A$ is $n - r$. Choose any basis

$$\{v^1, \ldots, v^r\}$$

for the range of $A$ and extend it to a basis

$$\{v^1, v^2, \ldots, v^m\}$$

for $\mathbb{R}^m$. Choose any basis

$$\{u^{r+1}, \ldots, u^n\}$$

for the null space of $A$ and extend it to a basis

$$\{u^1, u^2, \ldots, u^n\}$$

for $\mathbb{R}^n$.

If we introduce the coordinate transformations

$$y = V\eta\,, \qquad x = U\xi$$

where

$$V = [v^1 \ldots v^m]\,, \quad U = [u^1 \ldots u^n]$$

the original linear map is described by

$$\eta = \Lambda\xi$$

where

$$\Lambda = V^{-1}AU \ .$$

Recall the useful result on coordinate transformations. Since the vectors $u^j, j = r+1, \ldots, n$ are in the null space of $A$ we have

$$Au^j = 0 \qquad \text{for} \quad j = r+1, \ldots, n$$

Hence, the last $r$ columns of $\Lambda$ are zero. For $j = 1, \ldots, r$ the vector $Au^j$ is in the range of $A$. Hence it can be expressed as a linear combination of the vectors $v^1, \ldots, v^r$. Consequently, the last $m-r$ rows of $\Lambda$ are zero and $\Lambda$ has the following structure:

$$\Lambda = \begin{pmatrix} \Lambda_{11} & 0 \\ 0 & 0 \end{pmatrix}$$

where $\Lambda_{11} \in \mathbb{R}^{r \times r}$. Since the rank of $A$ equals rank $\Lambda$ which equals rank $\Lambda_{11}$, the matrix $\Lambda_{11}$ must be invertible.

If we decompose $\xi$ and $\eta$ as

$$\xi = \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} , \qquad \eta = \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix}$$

where $\xi_1 \in \mathbb{R}^r$, $\xi_2 \in \mathbb{R}^{n-r}$, $\eta_1 \in \mathbb{R}^r$, $\eta_2 \in \mathbb{R}^{m-r}$, then the linear map is described by

$$\begin{aligned} \eta_1 &= \Lambda_{11}\xi_1 \\ \eta_2 &= 0 \end{aligned}$$

where $\Lambda_{11}$ is invertible.

Also, we have shown that any matrix any $m \times n$ matrix $A$ can be decomposed as

$$\boxed{A = V \begin{pmatrix} \Lambda_{11} & 0 \\ 0 & 0 \end{pmatrix} U^{-1}}$$

where $U$, $V$, and $\Lambda$ are square invertible matrices of dimensions $n \times n$, $m \times m$, and $r \times r$, respectively, and $r$ is the rank of $A$.

# Chapter 28

# Appendix D: Inner products

## 28.1 The usual scalar product on $\mathbb{R}^n$ and $\mathbb{C}^n$

Consider any pair $x, y$ of real $n$-vectors. The scalar product (inner product) of $x$ and $y$ is the real number given by

$$\langle y, x \rangle \quad = \quad y_1 x_1 + \ldots + y_n x_n \,.$$

Regarding $x$ and $y$ as real column matrices, we have

$$\langle y, x \rangle = y^T x \,.$$

Suppose $x$ and $y$ are complex $n$-vectors. If we use the above definition as the definition of inner product, we won't have the property that $\langle x, x \rangle \geq 0$. To see this, note that $x = \jmath$ results in $x^2 = -1$. So we define the scalar product (inner product) of $x$ and $y$ to be the complex number given by

$$\boxed{\langle y, x \rangle := \bar{y}_1 x_1 + \ldots + \bar{y}_n x_n}$$

When $x$ and $y$ are real this definition yields the usual scalar product on $\mathbb{R}^n$. Regarding $x$ and $y$ as column matrices, we have

$$\langle y, x \rangle = y^* x$$

where $y^*$ is the complex conjugate transpose of $y$.

We can obtain the usual Euclidean norm from the usual scalar product by noting that

$$\begin{aligned} \langle x, x \rangle &= \bar{x}_1 x_1 + \ldots + \bar{x}_n x_n \\ &= |x_1|^2 + |x_2|^2 + \ldots + |x_n|^2 \\ &= \|x\|^2 \,. \end{aligned}$$

Hence

$$\|x\| = \langle x, x \rangle^{\frac{1}{2}} \,.$$

**Properties of $\langle \cdot, \cdot \rangle$.** The usual scalar product on $\mathbb{R}^n$ or $\mathbb{C}^n$ has the following properties.

(i) For every pair of vectors $x$ and $y$,

$$\langle x, y \rangle = \overline{\langle y, x \rangle}$$

(ii) For every vector $y$, every pair of vectors $x^1$, $x^2$, and every pair of scalars $\alpha_1, \alpha_2$,

$$\langle y, \alpha_1 x^1 + \alpha_2 x^2 \rangle \; = \; \alpha_1 \langle y, x^1 \rangle + \alpha_2 \langle y, x^2 \rangle$$

(iii) For every vector $x$,

$$\langle x, x \rangle \geq 0$$

and

$$\langle x, x \rangle = 0 \qquad \text{if and only if} \qquad x = 0$$

For real vectors, property (i) reduces to $\langle x, y \rangle = \langle y, x \rangle$. By utilizing properties (i) and (ii), we obtain that for every vector $x$, every pair of vectors $y^1$, $y^2$, and every pair of scalars $\alpha_1, \alpha_2$,

$$\langle \alpha_1 y^1 + \alpha_2 y^2, x \rangle \; = \; \overline{\alpha_1} \langle y^1, x \rangle + \overline{\alpha_2} \langle y^2, x \rangle$$

Any scalar valued function on a real (or complex) vector space which has the above three properties is defined to be an inner product. A vector space equipped with an inner product is called an inner product space. For example, suppose $\mathcal{V}$ is the vector space of all real valued continuous functions defined on the interval $[0, 1]$. For any two functions $f$ and $g$ in $\mathcal{V}$, let

$$\langle f, g \rangle = \int_0^1 f(t) g(t) \, dt \, .$$

One may readily show that properties (i)-(iii) above hold.

## 28.2 Orthogonality

Suppose $\mathcal{V}$ is a real (or complex) vector space which is equipped with an inner product $\langle \cdot, \cdot \rangle$.

*A vector $y$ is said to be* orthogonal *to a vector $x$ if*

$$\langle y, x \rangle = 0$$

For example, in $\mathbb{R}^2$ the vector $(1, 1)$ is orthogonal to vector $(1, -1)$.

If $y$ is orthogonal to $x$, then (using property (ii) above) $x$ is orthogonal to $y$; so, we usually just say $x$ and $y$ are orthogonal.

Figure 28.1: Orthogonality

### 28.2.1 Orthonormal basis

Suppose $\mathcal{V}$ is finite dimensional and

$$b^1, b^2, \dots b^n$$

is a basis for $\mathcal{V}$. Then this basis is an orthonormal basis if for every $i, j$,

$$\begin{aligned} \langle b^i, b^i \rangle &= 1 \\ \langle b^i, b^j \rangle &= 0 \qquad \text{when} \quad i \neq j \end{aligned}$$

that is each basis vector has magnitude 1 and is orthogonal to all the other basis vectors.

#### Unitary coordinate transformations

Suppose $b^1, b^2, \dots, b^n$ is an orthonormal basis for $\mathbb{C}^n$ and consider the transformation matrix

$$T = \begin{pmatrix} b^1 & b^2 & \dots & b^n \end{pmatrix}$$

Then

$$\begin{aligned} T^*T &= \begin{pmatrix} b^{1*} \\ b^{2*} \\ \vdots \\ b^{n*} \end{pmatrix} \begin{pmatrix} b^1 & b^2 & \dots & b^n \end{pmatrix} \\ &= \begin{pmatrix} b^{1*}b^1 & b^{1*}b^2 & \dots & b^{1*}b^n \\ b^{2*}b^1 & b^{2*}b^2 & \dots & b^{2*}b^n \\ \vdots & \vdots & \vdots & \\ b^{n*}b^1 & b^{n*}b^2 & \dots & b^{n*}b^n \end{pmatrix} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}, \end{aligned}$$

that is, $T^*T = I$.

*A square complex matrix $T$ is said to be* unitary *if*

$$T^*T = I\,.$$

If $T$ is unitary, then

$$T^{-1} = T^*\,.$$

**Example 238** The following matrix is unitary for any $\theta$.

$$T = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

Consider a unitary coordinate transformation

$$x = T\xi$$

with $T$ unitary. Then for any two $n$ vectors $\xi^1$ and $\xi^2$,

$$\begin{aligned} \langle T\xi^1, T\xi^2 \rangle &= (T\xi^1)^* T\xi^2 \\ &= {\xi^1}^* T^* T\xi^2 \\ &= {\xi^1}^* \xi^2 \\ &= \langle \xi^1, \xi^2 \rangle \end{aligned}$$

hence, $\langle T\xi^1, T\xi^2 \rangle = \langle \xi^1, \xi^2 \rangle$; that is, unitary transformations preserve the scalar product and, hence, the Euclidean norm.

## 28.3 Hermitian matrices

In this section, the square $(n \times n)$ complex matrix $P$ is hermitian, that is

$$\boxed{P^* = P}$$

If $P$ is real, then hermitian is equivalent to symmetric, that is, $P^T = P$.

• *For any complex $n$-vector $x$, the scalar $x^*Px$ is real:*

$$\begin{aligned}\overline{x^*Px} &= (x^*Px)^* \\ &= x^*P^*x \\ &= x^*Px\end{aligned}$$

Since $\overline{x^*Px} = x^*Px$, the scalar $x^*Px$ is real. ■

• *Every eigenvalue $\lambda$ of $P$ real.* Since $\lambda$ is an eigenvalue of $P$, there is a nonzero vector $v$ such that

$$Pv = \lambda v$$

Hence

$$v^*Pv = \lambda v^*v = \lambda ||v||^2$$

Since $v^*Pv$ is real and $||v||^2$ is nonzero real, $\lambda$ must be real. ■

• *Suppose $v_1$ and $v_2$ are eigenvectors corresponding to different eigenvalues of $P$. Then $v_1$ and $v_2$ are orthogonal:*

$$Pv_1 = \lambda_1 v_1$$

implies

$$v_2{}^*Pv_1 = \lambda_1 v_2{}^*v_1$$

Similarly,

$$v_1{}^*Pv_2 = \lambda_2 v_1{}^*v_2$$

Taking the complex conjugate transpose of both sides of this last expression and using $P^* = P$ and $\overline{\lambda_2} = \lambda_2$ yields

$$v_2{}^*Pv_1 = \lambda_2 v_2{}^*v_1$$

Equating the two expressions for $v_2{}^*Pv_1$ yields

$$\lambda_1 v_2{}^*v_1 = \lambda_2 v_2{}^*v_1$$

Since $\lambda_1 \neq \lambda_2$, we must have $v_2{}^*v_1 = 0$, that is, $v_2$ is orthogonal to $v_1$. ■

Actually, the following more general result is true.

**Fact 25** *Every hermitian matrix has an orthonormal basis of eigenvectors.*

It now follows that every hermitian matrix $P$ can be expressed as

$$\boxed{P = T\Lambda T^*}$$

where $T$ is a unitary matrix and $\Lambda$ is a real diagonal matrix whose diagonal elements are the eigenvalues of $P$ and the number of times an eigenvalue appears equals its algebraic multiplicity.

## 28.4 Positive and negative (semi)definite matrices

### 28.4.1 Quadratic forms

For any square $n \times n$ matrix $P$ we can define an associated quadratic form:

$$x^*Px = \sum_{i=1}^{n}\sum_{j=1}^{n} p_{ij}x_i^*x_j \,.$$

Quadratic forms arise naturally is mechanics. For example, suppose one considers a rigid body and chooses an orthogonal coordinate system with origin at the body mass center. If $\omega$ is a real 3-vector consisting of the components of the angular velocity vector of the body and $I$ is the $3 \times 3$ symmetric inertia matrix of the body, then the rotational kinetic energy of the body is given by $\frac{1}{2}\omega^T I\omega$.

If $P$ is hermitian ($P^* = P$) then, as we have already shown, the scalar $x^*Px$ is real for all $x \in \mathbb{C}^n$. The following result provides useful upper and lower bounds on $x^*Px$ .

- *If $P$ is a hermitian $n \times n$ matrix, then for all $x \in \mathbb{C}^n$*

$$\boxed{\lambda_{min}(P)||x||^2 \leq x^*Px \leq \lambda_{max}(P)||x||^2}$$

*where $\lambda_{\min}(P)$ and $\lambda_{\max}(P)$ denote the minimum and maximum eigenvalues of $P$ respectively.*

Proof. Letting $\xi = T^*x$, we have

$$\begin{aligned} x^*Px &= x^*T\Lambda T^*x \\ &= \xi^*\Lambda\xi \\ &= \lambda_1\xi_1^*\xi_1 + \ldots + \lambda_n\xi_n^*\xi_n \\ &= \lambda_1|\xi_1|^2 + \ldots + \lambda_n|\xi_n|^2 \end{aligned}$$

Since, $\lambda_i \leq \lambda_{\max}(P)$ for $i = 1, \ldots, n$ we have $\lambda_i|\xi_i|^2 \leq \lambda_{\max}(P)|\xi_i|^2$; hence

$$x^*Px \leq \lambda_{\max}(P)\left(|\xi_1|^2 + \ldots + |\xi_n|^2\right) = \lambda_{max}(P)||\xi||^2 \,;$$

hence $x^*Px \leq \lambda_{max}(P)||\xi||^2$. Also $||\xi|| = ||x||$; hence,

$$x^*Px \leq \lambda_{max}(P)||x||^2$$

In an analogous fashion one can show that

$$x^*Px \geq \lambda_{min}(P)||x||^2$$

■

### 28.4.2 Definite matrices

**DEFN.** A hermitian matrix $P$ is positive definite (pd) if

$$x^*Px > 0 \qquad \text{for all nonzero } x \text{ in } \mathbb{C}^n$$

We denote this by $P > 0$.

The matrix $P$ is negative definite (nd) if $-P$ is positive definite; we denote this by $P < 0$.

**Example 239** For

$$P = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$$

we have (note the completion of squares trick)

$$\begin{aligned} x^*Px &= x_1^*x_1 - x_1^*x_2 - x_2^*x_1 + 2x_2^*x_2 \\ &= (x_1 - x_2)^*(x_1 - x_2) + x_2^*x_2 \\ &= |x_1 - x_2|^2 + |x_2|^2 \end{aligned}$$

Clearly, $x^*Px \geq 0$ for all $x$. If $x^*Px = 0$, then $x_1 - x_2 = 0$ and $x_2 = 0$; hence $x = 0$. So, $P > 0$.

**Fact 26** *The following statements are equivalent for any hermitian matrix $P$.*

*(a) $P$ is positive definite.*

*(b) All the eigenvalues of $P$ are positive.*

*(c) All the leading principal minors of $P$ are positive; that is,*

$$\begin{aligned} p_{11} &> 0 \\ \det\begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix} &> 0 \\ &\vdots \\ \det(P) &> 0 \end{aligned}$$

**Example 240** Consider

$$P = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$$

Since $p_{11} = 1 > 0$ and $\det(P) = 1 > 0$, we must have $P > 0$.

```
>> eig(P)
ans =
    0.3820
    2.6180
```

Note the positive eigenvalues.

### 28.4.3 Semi-definite matrices

**DEFN.** A hermitian matrix $P$ is positive semi-definite (psd) if

$$x^*Px \geq 0 \qquad \text{for all } x \text{ in } \mathbb{C}^n$$

We denote this by $P \geq 0$

The matrix $P$ is negative semi-definite (nsd) if $-P$ is positive semi-definite; we denote this by $P \leq 0$

**Fact 27** *The following statements are equivalent for any hermitian matrix $P$.*

*(a) $P$ is positive semi-definite.*

*(b) All the eigenvalues of $P$ are non-negative.*

*(c) All the principal minors of $P$ are non-negative.*

**Example 241** This example illustrates that non-negativity of only the leading principal minors of $P$ is not sufficient for $P \geq 0$. One needs non-negativity of all the principal minors.

$$P = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}$$

We have $p_{11} = 0$ and $\det(P) = 0$. However,

$$x^*Px = -|x_2|^2$$

hence, $P$ is not psd. Actually, $P$ is nsd.

**Fact 28** *Consider any $m \times n$ complex matrix $M$ and let $P = M^*M$. Then*

*(a) $P$ is hermitian and $P \geq 0$*

*(b) $P > 0$ if and only if rank $M = n$.*

**Example 242**

$$P = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

Since

$$P = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \end{pmatrix}$$

and

$$\text{rank} \begin{pmatrix} 1 & 1 \end{pmatrix} = 1$$

$P \geq 0$ but $P$ is not pd.

**Exercise 173** (Optional)

Suppose $P$ is hermitian and $T$ is invertible. Show that $P > 0$ if and only if $T^*PT > 0$.

**Exercise 174** (Optional)

Suppose $P$ and $Q$ are two hermitian matrices with $P > 0$. Show that $P + \epsilon Q > 0$ for all real $\epsilon$ sufficiently small; that is, there exists $\bar{\epsilon} > 0$ such that whenever $|\epsilon| < \bar{\epsilon}$, one has $P + \epsilon Q > 0$.

## 28.5 Singular value decomposition

In this section we introduce a matrix decomposition which is very useful for *reliable* numerical computations in linear algebra. First, some notation.

The diagonal elements of an $m \times n$ matrix $A$ are

$$A_{11}, A_{22}, \ldots, A_{pp} \qquad \text{where} \quad p = \min\{m, n\}$$

We say that $A$ is diagonal if all its nondiagonal elements are zero, that is, $A_{ij} = 0$ for $i \neq j$. Examples are:

$$\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 0 \\ 0 & 2 \\ 0 & 0 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \qquad \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

**Theorem 35 (Singular value decomposition)** *If $A$ is a complex $m \times n$ matrix, then there exist unitary matrices $U \in \mathbb{C}^{m\times m}$ and $V \in \mathbb{C}^{n\times n}$ such that*

$$\boxed{A = U\Sigma V^*}$$

*where $\Sigma$ is a diagonal matrix of the same dimension as $A$ and with real nonnegative diagonal elements:*

$$\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_p \geq 0 \qquad p = \min\{m, n\}$$

*If $A$ is real, then $U$ and $V$ are real.*

So, if $A$ is square ($m = n$), then

$$\Sigma = \begin{pmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_n \end{pmatrix}.$$

When $A$ is "fat" ($m < n$), we have

$$\Sigma = \begin{pmatrix} \sigma_1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & \ddots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \sigma_m & 0 & \cdots & 0 \end{pmatrix}.$$

Finally, a "tall" matrix ($m > n$) yields

$$\Sigma = \begin{pmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_n \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}.$$

The scalar $\sigma_i$ is called the $i$-th singular value of $A$. Let $\sigma_r$ be the smallest nonzero singular value of $A$. Then $r$ is smaller than both $m$ and $n$. Moreover, $\Sigma$ can be expressed as

$$\Sigma = \begin{pmatrix} \Sigma_{11} & 0 \\ 0 & 0 \end{pmatrix}$$

where $\Sigma_{11}$ is the square diagonal invertible matrix given by

$$\Sigma_{11} = \begin{pmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & & 0 \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_r \end{pmatrix}$$

If

$$U \begin{pmatrix} \Sigma_{11} & 0 \\ 0 & 0 \end{pmatrix} V^*$$

is a singular value decomposition of $A$, then

$$A^*A = V \begin{pmatrix} \Sigma_{11}^2 & 0 \\ 0 & 0 \end{pmatrix} V^*$$

Hence, the nonzero singular values of $A$ are uniquely given by

$$\sigma_i = \sqrt{\lambda_i}$$

where $\lambda_1, \dots, \lambda_r$ are the nonzero eigenvalues of the hermitian positive semidefinite matrix $A^*A$.

**Example 243** The matrix

$$A = \begin{pmatrix} 1.2 & -0.64 & -0.48 \\ 1.6 & 0.48 & 0.36 \end{pmatrix}$$

has the following singular value decomposition

$$A = \underbrace{\begin{pmatrix} 0.6 & 0.8 \\ 0.8 & -0.6 \end{pmatrix}}_{U} \underbrace{\begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}}_{\Sigma} \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & -0.8 & 0.6 \\ 0 & -0.6 & -0.8 \end{pmatrix}^*}_{V^*}$$

**Useful facts**

• The maximum (or largest) singular value of a matrix satisfies the properties of a norm. Thus, if we consider the set of $m \times n$ matrices $A$ and define

$$||A|| = \sigma_{\max}(A)$$

where $\sigma_{\max}(A)$ denotes the largest singular of $A$, then the following properties hold:

$$\begin{array}{ll} \text{(a)} & ||A|| \geq 0 \quad \text{and} \quad \|A\| = 0 \quad \text{if and only if} \quad A = 0 \\ \text{(b)} & ||\lambda A|| = |\lambda|\|A\| \\ \text{(c)} & \|A + B\| \leq \|A\| + \|B\| \,. \end{array}$$

In addition to the above usual norm properties, this norm also satisfies

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

where the above vector norms are the usual euclidean vector norms. Thus,

$$\|Ax\| \leq \|A\|\|x\|$$

for all $x$.

• Using the above norm, we can define the distance between two matrices $A$ and $B$ of the same dimensions as $\|A - B\|$.

• The rank of $A$ equals the number of its nonzero singular values.

• Suppose $k$ is a nonegative integer strictly less than the rank $r$ of $A$. Then in a certain sense (which can be made precise), a matrix of rank $k$ which is "closest" to $A$ is given by

$$A_k = U\Sigma_k V^*$$

where $\Sigma_k$ is the diagonal matrix with diagonal elements

$$\sigma_1,\ \sigma_2,\ \ldots,\ \sigma_k,\ \underbrace{0,\ \ldots,\ 0}_{p-k \text{ zeros}}$$

that is, the first $k$ diagonal elements of $\Sigma_k$ are the same as those of $\Sigma$ and the remaining diagonal elements are zero. Also, the "distance" between $A_k$ and $A$ is $\sigma_{k+1}$. So, the *smallest nonzero singular value* $\sigma_r$ of $A$ is a measure of the distance between $A$ and the closest matrix of rank $r-1$.

• Suppose $A$ has rank $r$ and

$$U = \begin{pmatrix} u^1 & u^2 & \ldots & u^m \end{pmatrix}, \qquad V = \begin{pmatrix} v^1 & v^2 & \ldots & v^n \end{pmatrix}$$

Then the vectors

$$u^1, u^2, \ldots, u^r$$

form an orthonormal basis for the range of $A$ and the vectors

$$v^{r+1}, v^{r+2}, \ldots, v^n$$

form an orthonormal basis for the null space of $A$

**MATLAB**

```
>> help svd

 SVD    Singular value decomposition.
        [U,S,V] = SVD(X) produces a diagonal matrix S, of the same
        dimension as X and with nonnegative diagonal elements in
        decreasing order, and unitary matrices U and V so that
        X = U*S*V'.

        By itself, SVD(X) returns a vector containing diag(S).

        [U,S,V] = SVD(X,0) produces the "economy size"
        decomposition. If X is m-by-n with m > n, then only the
        first n columns of U are computed and S is n-by-n.
```

**Example 244**

```
A =

     1     1     1
     1    -1     0
     1     1     1
     1     1     1

 [u s v] = svd(A)

u =
    0.5774   -0.0000   -0.8165   -0.0000
   -0.0000   -1.0000   -0.0000   -0.0000
    0.5774   -0.0000    0.4082   -0.7071
    0.5774   -0.0000    0.4082    0.7071

s =
    3.0000         0         0
         0    1.4142         0
         0         0         0
         0         0         0

v =
    0.5774   -0.7071    0.4082
    0.5774    0.7071    0.4082
    0.5774    0.0000   -0.8165
```

So $A$ has rank 2 and a basis for its range is given by

$$\begin{pmatrix} 0.5774 \\ 0 \\ 0.5774 \\ 0.5774 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 0 \\ 0 \end{pmatrix}.$$

A basis for the null space of $A$ is given by

$$\begin{pmatrix} 0.4082 \\ 0.4082 \\ -0.8165 \end{pmatrix}$$

# Bibliography


[1] Chen, C.T., *Linear System Theory and Design, 3rd Ed.,* Oxford University Press, 1998.

[2] T. Kailath, T., *Linear Systems,* Prentice-Hall Information and Systems Science Series, 1979.

[3] Khalil, H., *Nonlinear Systems, 3rd Ed.,* Prentice-Hall, 2002.

[4] Vidasagar, M., *Nonlinear Systems Analysis, 2nd Ed.*, Prentice-Hall, 1993.