Cornell University
Electrical and Computer Engineering

ECE 5210: Theory of Linear Systems

Professor David Forbes Delchamps (dfd1)

Compiled by Nathan Lambert (nol5), ECE '17

December 5, 2017

This is made in appreciation of Professor Delchamps and the positivity he imbues on us through his lectures

Contents

1 Introduction **5**
1.1 State Space Linear System Models Definitions . . . 5
1.1.1 Setup . . . 5
1.1.2 Discrete Time Analysis . . . 6
1.1.3 DT Existence and Uniqueness Theorem . . . 7
1.1.4 Continuous Time Existence and Uniqueness . . . 8

2 Matrix Functions **13**
2.1 Motivation with State Equations . . . 13
2.1.1 State Transition Matrix (STM) and Matrix Exponential . . . 13
2.1.2 MATH 2940: Linear Algebra Review . . . 16
2.1.3 Summary: Properties of e^{tA} The Matrix Exponential . . . 18
2.2 Calculating the Matrix Exponential . . . 18
2.2.1 Eigenvalue Method . . . 18
2.2.2 Laplace and Z Transform Method . . . 23
2.2.3 Cayley-Hamilton Theorem Method . . . 28
2.3 General Calculations of Functions of Matrices . . . 31
2.3.1 Problem Formation . . . 31
2.3.2 Eigenstructure and Minimal Polynomial . . . 32
2.3.3 Functions of a Matrix . . . 34

3 Working With State Systems **37**
3.1 Discrete Time . . . 37
3.1.1 Reachability . . . 37
3.1.2 Controllability . . . 42
3.1.3 Summary . . . 44
3.2 Continuous Time . . . 45
3.2.1 Reachability . . . 45
3.2.2 Controllability . . . 49
3.2.3 Summary . . . 49

3.3 Observability . . . 50
3.3.1 Discrete Time . . . 50
3.3.2 Continuous Time . . . 51
3.3.3 Reachability Observability Duality . . . 52
3.4 Stability . . . 52
3.4.1 Discrete Time Facts . . . 54
3.4.2 Continuous Time Facts . . . 55
3.5 Lyapunov Lemmas . . . 56
3.5.1 Continuous Time . . . 56
3.5.2 Lyapunov Function . . . 57
3.5.3 Discrete Time . . . 58

4 Feedback and Observers **60**
4.1 Introduction . . . 60
4.1.1 Wonham's Theorem and Analysis . . . 60
4.1.2 Observers . . . 62
4.2 I/O Linear System Models . . . 63
4.2.1 Weighting Pattern and Time-Invariance . . . 63
4.2.2 Impulse Response . . . 66
4.3 Realization Theory . . . 68
4.3.1 Realizations and Stability . . . 69
4.3.2 Hankel Matrix . . . 71
4.3.3 Fundamental Theorem of Realizations . . . 72
4.3.4 State Space Isomorphism Theorem . . . 73
4.3.5 Canonical Structure Theorem . . . 74
4.4 Stability for I/O Systems . . . 75
4.4.1 BIBO Stable and Realizations . . . 75
4.4.2 Stabilizable and Detectable . . . 77
4.5 Interconnects . . . 77
4.5.1 Basic Typologies . . . 77
4.5.2 H^∞ Design . . . 79
4.5.3 Observer Construction . . . 81

5 Linear Quadratic Regulator **84**
5.1 Finite {Time Horizon} Problem . . . 84
5.1.1 Introduction . . . 84
5.1.2 Discrete Time Solution . . . 85
5.1.3 Continuous Time Solution . . . 87

5.2 Infinite Time Horizon LQR . . . 91
5.2.1 Continuous Time Solution . . . 91
5.2.2 Discrete Time Notes . . . 95

6 Appendix: 96
6.1 Linear Algebra Foundations . . . 96
6.1.1 Notation, Groundwork, and Abstract Math Concepts . . . 96
6.1.2 Vectors . . . 99
6.1.3 Linear Maps . . . 102
6.1.4 Matrix Representation . . . 104
6.1.5 Matrix Operations . . . 105
6.2 Linear Algebra Operations . . . 107
6.2.1 Norms . . . 107
6.2.2 Inner Products and Orthogonality . . . 110
6.2.3 Singular Value Decomposition . . . 114
6.2.4 Least Squares Optimization . . . 115
6.3 Differential Equations . . . 117
6.3.1 Theorem & Definitions . . . 117
6.3.2 Fundamental Theorem . . . 119
6.4 Alternate State Space Definitions: . . . 124
6.4.1 Dynamical Systems: . . . 124
6.4.2 Time-Invariant Dynamical Systems . . . 126
6.5 Other System Representation Examples . . . 127
6.5.1 Jacobian Linearization . . . 127
6.5.2 Affine Systems . . . 127
6.6 Eigenvalue Placement by State Feedback . . . 128
6.6.1 SISO Systems, alternate derivation . . . 128

Chapter 1

Introduction

1.1 State Space Linear System Models Definitions

1.1.1 Setup

Notation:

- $\mathbb{R}$ = set of all real numbers
- $\mathbb{Z}$ = set of all integers
- $\mathbb{R}^n$ = set of all column n-vectors with real entries (States live here)
- $\mathbb{R}^{m\times n}$ = set of all $m \times n$ matrices with real entries; say $m \times n$ if A ϵ $\mathbb{R}^{m\times n}$, $[A]_{i,j}$ = i,j element of A.
- $x : \mathbb{R} \mapsto \mathbb{R}^n$ = 'x is a mapping from $\mathbb{R}$ to $\mathbb{R}^n$'; or x assigns to each t ϵ $\mathbb{R}$ on x(t) ϵ $\mathbb{R}^n$ or $t \mapsto x(t)$.

Big Picture: real world process.

Model as Follows

$$(I) \quad \begin{aligned} \dot{x}(t) &= A(t)x(t) + B(t)u(t) \\ y(t) &= C(t)x(t) + D(t)u(t) \quad t \in \mathbb{R} \end{aligned}$$

Where...

- $\dot{x}(t)$ = time derivative of x(t)
- $t \mapsto A(t) : (n \times n)$, $t \mapsto B(t) : (n \times m)$, $t \mapsto C(t) : (p \times n)$, $t \mapsto D(t) : (p \times m)$, are real matrix valued functions of $t \in \mathbb{R}$.
- $t \mapsto u(t) \in \mathbb{R}^m$, and $t \mapsto y(t) \in \mathbb{R}^p$ are the input and output functions
- $t \mapsto x(t) \in \mathbb{R}^n$ is the *State* of the system model

Typical Problem: given $t_0 \in \mathbb{R}$ and $x_0 \in \mathbb{R}^n$, find $x(t), t \geq t_0$ and $y(t) t \geq t_0$ given $x(t_0) = x_0$ and that you apply $u(t), t \geq t_0$.

Another: given $x_1 \in \mathbb{R}^n$, can we find some $u : [0, t] \mapsto \mathbb{R}^m$ s.t. $x(t_1) = x_1$ when $x(0) = 0$ and you apply input u from $t = 0$ to $t = t_1$? Does the answer depend on the value of t_1?

Analogous Setup in DT

$$(II) \qquad \begin{aligned} x(k+1) &= A(k)x(k) + B(k)u(k) \\ y(k) &= C(k)x(k) + D(k)u(k) \quad k \in \mathbb{Z} \end{aligned}$$

Where...

- $k \mapsto A(k) : (n \times n)$, $k \mapsto B(k) : (n \times m)$, $k \mapsto C(k) : (p \times n)$, $k \mapsto D(k) : (p \times m)$, are real matrix valued functions of $k \in \mathbb{R}$.
- $k \mapsto u(k) \in \mathbb{R}^m$, and $t \mapsto y(t) \in \mathbb{R}^p$ are the input and output functions
- $k \mapsto x(k) \in \mathbb{R}^n$ is the *State* of the system model

One example of eqn (I)'s model: To be added when learn more diagrams.
Note: Matrix functions A,B,C,D are constant (t-independent), most of our time exists in these situations.

1.1.2 Discrete Time Analysis

First: $x(k+1) = A(k)x(k) + B(k)u(k)$: This is the recipe for next x-value with current x and u.
- Given $k_0 \in \mathbb{Z}$, $x(k_0) = x_0 \in \mathbb{R}^n$, and with u(k) for $k \geq k_0$. find $u(k) for k \geq k_0$.
- See: $x(k_0 + 1) = A(k_0)x_0 + B(k_0)u(k_0)$
$x(k_0 + 2) = A(k_0 + 1)x(k_0 + 1) + B(k_0 + 1)u(k_0)$ etc etc
$= A(k_0 + 1)A(k_0)x_0 + A(k_0 + 1)B(k_0)u(k_0) + B(k_0 + 1)u(k_0)$

Keep this up: for $k \geq k_0$,
$x(k) = \Phi(k, k_0)x_0 + \sum_{l=k_0}^{k-1} \Phi(k, l+1)B(l)u(l)$ and
$y(k) = C(k)[\Phi(k, k_0)x_0 + \sum_{l=k_0}^{k-1} \Phi(k, l+1)B(l)u(l)] + D(k)u(k)$ where

$$\Phi(k, j) = \begin{cases} I_n, k = j \\ A(k-1)A(k-2)..A(j), k > j \\ undefined, k < j \end{cases}$$

Term: The $\mathbb{R}^{n \times n} f(k, j) \mapsto \Phi(k, j)$ when $k \geq j$ is called the *State Transition Matrix of (II)*. It depends only on $k \mapsto A(k)$, s.t m is associated with A.
$\rightarrow$ observe that Φ satisfies linear difference equation of its own for each $j \in \mathbb{Z}$.
$\Phi(k + 1, j) = A(k)\Phi(k, j)$, and $\Phi(j, j) = I_n$.

$\rightarrow \Phi$ satisfies the semi-group property for all $k \geq j \geq l$, $\Phi(k,l) = \Phi(k,j)\Phi(j,l)$.

Idea: in difference equation (II), say we start at time l and apply zero input from $x(l) = x_0$. The state transitions from x_0 at time l to $\Phi(j,k)x_0$ at time j; and from there to $\Phi(k,l)x_0$ at time k. This second transitions happens via $x(k) = \Phi(k,j)x(j) = \Phi(k,j)\Phi(j,l)x_o = \Phi(k,l)x_0$.

Sepcial Case: suppose $k \mapsto A(k), B(k), C(k), D(k)$ *are constant / $k-$independant matrices.*

then

$$\Phi(k,j) = \begin{cases} I_n, k = j \\ A^{k-j}, k > j \\ undefined, k < j \end{cases}$$

Thus with given k_0, x_0, and $x(t_0) = x_0$.

$$x(k) = A^{k-k_0}x_0 + \sum_{l=k_0}^{k-1} A^{k-l-1}B(l)u(l) \qquad k \geq k_0$$

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

1.1.3 DT Existence and Uniqueness Theorem

- Given $x_0 \in \mathbb{R}^n, k_0 \in \mathbb{Z}$ there exists a unique solution $k \mapsto x(k)$, $k \geq k_0$ to be the difference equation in (II) satisfying $x(k_0) = x_0$: $\rightarrow x(k) = \Phi(k,k_0)x_0 + \sum_{l=k_0}^{k-1} \Phi(k,l+1)B(l)u(l) \qquad k \geq k_0$.
- From this, get unique $k \mapsto y(k), k \geq k_0$ using 2^{nd} equation in (II) satisfying the initial conditions.
  - $\rightarrow$ $(k,j) \mapsto \Phi(k,j)$ defined for integers k,j with k≥j is an $n \times n$ matrix valued function.
  - $\rightarrow$ it's the unique solution to the matrix difference equation $\Phi(k+1,j) = A(k)\Phi(k,j)$ for all $k \geq j$ all j. Note that $\Phi(j,j) = I_n$ for all j. Here we are emphasizing the DT result in order to make corresponding CT analysis easier.

  - Given $v \in \mathbb{R}^n$, the Euclidean Norm of V (two norm) is $\|v\|$
    $\|v\| = (v_1^2 + v_2^2 + ... + v_n^2)^{1/2}$.
  - Given $v, w \in \mathbb{R}^n$, the inner product (Euclidean) of v,w is
    $< v, w >= w^T v \in \mathbb{R}$
    Note: $\|v\|^2 =< v, v >$ for all $v \in \mathbb{R}^n$.

  - **Schwarz Inequality**:

    $$| < v, w > | \leq \|v\| \, \|w\| \;\; for\ all\ v, w \in \mathbb{R}^n$$

    ...To see this, for any $\alpha \in \mathbb{R}$

$0 \leq \|v - \alpha w\|^2 =< v - \alpha w, v - \alpha w >$

$=< v, v > -2\alpha < v, w > +\alpha^2 < w, w >$. (Using $< v, w >=< w, v >$).

Schwarz holds trivially when $vorw = 0$, try $\alpha = \frac{<v,w>}{\|w\|^2}$

$0 \leq \|v\|^2 - \frac{2*|<v,w>|^2}{\|w\|^2} + \frac{|<v,w>^2|}{\|w\|^2}$

yields $\|< v, w >\|^2 \leq \|v\|^2 \|w\|^2 - \surd$

- **Euclidean Norm of Matrix**:

Given $A \in \mathbb{R}^{m \times n}$, define the Euclidean Norm as

$$\|A\| = max\{\, \|Av\| : v \in \mathbb{R}^n, \|v\| = 1\}$$

Note: Max argument does not explicitly exists always, turns out $\|A\| = larges\ singular\ value$

Onto a Couple Norm Facts

1. $\|Av\| \leq \|A\| \|v\|$ $for\ all\ v \in \mathbb{R}^n$

   Why?: when $v \neq 0$, $\frac{v}{\|v\|}$ has norm 1, so $\left\|A\frac{v}{\|v\|}\right\| \leq \|A\|$ by definition of norm.
2. $\|AB\| \leq \|A\| \|B\|$ whenever multiplication AB is sensible.

1.1.4 Continuous Time Existence and Uniqueness

Back to (I)... want to show that if given $t_0 \in \mathbb{R}$ and $x_0 \in \mathbb{R}^n, x(t_0) = x_0$, there *exists* an unique solution $t \mapsto x(t)$ for some t's to the differential equation in (I) that satisfies the initial condition.

Consider Continuous Time System Equations

(I) $\dot{x}(t) = A(t)x(t) + B(t)u(t)$ - (Difference Equation DE);

$y(t) = C(t)x(t) + D(t)u(t)$ $\quad t \in \mathbb{R}$ and $x \in \mathbb{R}^n, u \in \mathbb{R}^m, y \in \mathbb{R}^p$

Note: To prove a result similar to Existence and Uniqueness above will require more work than the DT situation.

$\rightarrow$ First will attack *uniqueness*: Suppose we have some interval $[s_1, t_1]$ containing t_0 and two solutions $t \mapsto x(t)$ *and* $t \mapsto \hat{x}(t)$ to the DE and satisfying the initial condition. We will show that $x(t) = \hat{x}(t)$ for all $t \in [s_1, t_1]$.

- let $Z(t) = \|x(t) - \hat{x}(t)\|^2$ $for\ t \in [s_1, t_1]$

  Note: $Z(t) \geq 0$ for all t in interval, from here in derivation suppress t-dependance
- $Z = (x - \hat{x})^T(x - \hat{x})$

  $\dot{Z} = (x - \hat{x})^T(\dot{x} - \dot{\hat{x}}) + (\dot{x} - \dot{\hat{x}})^T(x - \hat{x})$
- Since x and $\hat{x}$ satisfy DE,

  $\dot{x} = Ax + Bu \qquad \dot{\hat{x}} = A\hat{x} + Bu$

  $\rightarrow \dot{x} - \dot{\hat{x}} = A(x - \hat{x})$

$\rightarrow \dot{z} = (x-\hat{x})^T A(x-\hat{x}) + (x-\hat{x})^T A^T(x-\hat{x})$

$|\dot{z}| \quad \leq |(x-\hat{x})^T A(x-\hat{x})| + |(x-\hat{x})^T A^T(x-\hat{x})|$

$\leq \|(x-\hat{x})\| \, \|A(x-\hat{x})\| + \|(x-\hat{x})\| \left\|A^T(x-\hat{x})\right\|$ via Schwarz

$\leq \|A\| \, \|(x-\hat{x})\|^2 + \left\|A^T\right\| \|(x-\hat{x})\|^2$ by $\|Av\| \leq \|A\| \, \|v\|$

$= 2\|A\| \, Z$ by $\left\|A^T\right\| = \|A\|$

$\rightarrow$ says $|\dot{z}| \leq 2\|A\| \, Z \quad = \quad -2\|A\| \leq \dot{z} \leq 2\|A\| \, Z.$

Next will show $z(t) = 0$ for all t in interval $t \in [t_0, t_1]$ and there is a matching argument for the other side of the interval.

- $\dot{z} \leq Z\|A\| \, Z(t) \qquad t \in [t_0, t_1] \qquad z(t_0) = 0$
- let $w(t) = z(t) exp\Big[\int_{t_0}^{t} 2\|A(\tau)\| \, d\tau\Big] \qquad t \in [t_0, t_1]$

  Note $w(t_0) = 0$ *and* $w(t) \geq 0$ *for all* $t \in [t_0, t_1]$

  $\dot{w} = \dot{z}(t) exp\Big[\ \Big] + z(t)(-2\|A(t)\| \, exp\Big[\ \Big])$

  so $\dot{w} \leq 2\|A(t)\| \, Z(t) exp\Big[\ \Big] - 2\|A(t)\| \, exp\Big[\ \Big] Z(t)$ which ≤ 0 for all $t \in [t_0, t_1]$

$\rightarrow$ Thus $\mathbf{w(t) = 0}$ for all $t \in [t_0, t_1]$ since $w(t)$ has been shown to start at 0, be ≥ 0, and have a slope $\leq$0.
Thus it follows, $Z(t) = exp\Big[\int\Big] w(t), \mathbf{z(t) = 0}$, *for* $t \in [t_0, t_1]$.

$\rightarrow \rightarrow$ $Z(t) = 0$ for all $t \in [t_0, t_1]$, so $x(t) = \hat{x}(t)$ for all $t \in [t_0, t_1]$
So, there is at most 1 solution $t \mapsto x(t)$ *to the DE on the interval given satisfying the IC.*

Comment: Argument above assumed some regularity of matrix function $A(t)$; this assumption first occurred when we have $\frac{d}{dt}(-2\int \|A(\tau\|)d\tau$ to be $-2\|A\|$. Turns out, it is sufficient and almost necessary that $t \mapsto x(t)$ be a continuous function of t.

Prove Existence of Solution:
Drawing inspiration from scalar case ($n = m = p = 1$), $\dot{x}(t) = a(t)x(t) + b(t)u(t), \ t \in \mathbb{R}, \ x(t_0) = x_0 \in \mathbb{R}$

- Classic way to solve:

$$define \ : \ \phi(t,s) = exp\Big[\int_s^t a(\tau)d\tau\Big] \qquad all \ t, s \in \mathbb{R}$$

- Turns out this is a solution to the differential equation and initial condition $x(t_0) = x_0$:

$$x(t) = \phi(t, t_0)x_0 + \int_{t_0}^{t} \phi(t,\tau)b(\tau)u(\tau)d\tau \ t \in \mathbb{R}.$$

(background assumption: $t \mapsto a(t), b(t), u(t)$ and they are continuous)

Note $(t,s) \mapsto \phi(t,s)$ satisfies $\frac{d}{dt}\phi(t,s) = a(t)\phi(t,s)$ and $\phi(s,s) = 1 \qquad t,s \in \mathbb{R}$.

- We'll show that for the solution to the general case, the solution to the differential equation with IC $x(t_0) = x_0$ takes the form:

$$x(t) = \Phi(t,t_0)x_0 + \int_t^{t_0} \Phi(t,\tau)B(\tau)u(\tau)d\tau \qquad t \in \mathbb{R}$$

- Unlike the scalar case, there is no nice formula for $\Phi(t,s)$ in particular

$$\Phi(t,s) \neq exp\Big[\int_s^t A(\tau)d\tau\Big] \qquad t,s \in \mathbb{R}$$

- However $\Phi(t,s)$ satisfies what we call the State Transition Matrix

$$(STM)\ \frac{d}{dt}\Phi(t,s) = A(t)\Phi(t,s) \qquad \Phi(s,s) = I_n \qquad t,s \in \mathbb{R}$$

→ First make sure x given above satisfies the differential equation with the initial condition provided Φ satisfies the STM. Specifically plug in $t = t_0 \qquad x(t_0) = \Phi(t_0,t_0)x_0 + \int_{t_0}^{t_0}\ldots^{0}$

giving $x(t_0) = x_0$

- Meanwhile $\frac{d}{dt}x(t) = \frac{d}{dt}\Phi(t,t_0)x_0 + \Phi(t,t)B(t)u(t) + \int_{t_0}^t \frac{d}{dt}\Phi(t,\tau)B(\tau)u(\tau)d\tau$
  $= A(t)\Phi(t,t_0)x_0 + B(t)u(t) + A(t)\int_{t_0}^t \frac{d}{dt}\Phi(t,\tau)B(\tau)u(\tau)d\tau$
  $= A(t)\Big[\Phi(t,t_0)x_0 + \int_{t_0}^t \frac{d}{dt}\Phi(t,\tau)B(\tau)u(\tau)d\tau\Big] + B(t)u(t)$
  $= A(t)x(t) + B(t)u(t).$

  (Assuming $t \mapsto A(t), B(t), u(t)$, is continuous enough to make all derivative-type manipulation legal)

→ Thus, provided we have $(t,s) \mapsto \Phi(t,s)$ satisfying (STM) we have a solution to the differential equation with initial condition $x(t_0) = x_0$.

→ **Observe:** uniqueness theorem applies to (STM), thus if we have a solution, it's the only one.

***Solve* using Picard Iteration**

- Recall (STM) $\frac{d}{dt}\Phi(t,s) = A(t)\Phi(t,s) \qquad \Phi(s,s) = I_n \qquad t,s \in \mathbb{R}.$
- Start by integrating STM from s to t, remember $\int_s^t \frac{d}{dt}f(\tau)d\tau = f(t) - f(s)$

$\Phi(t,s) - \Phi(s,s) = \int_s^t A(\tau)\Phi(\tau,s)d\tau$

$\star \rightarrow \Phi(t,s) = I_n + \int_s^t A(\tau)\Phi(\tau,s)d\tau \star$

$\Phi(t,s) = I_n + \int_s^t A(\mu)\Phi(\mu,s)d\mu \qquad \tau \rightarrow \mu$

$\Phi(\mu,s) = I_n + \int_s^\mu A(\tau)\Phi(\tau,s)d\tau$

Thus, $\Phi(\mu,s) = I_n + \int_s^t A(\mu)\Big[I_n + \int_s^\mu A(\tau)\Phi(\tau,s)d\tau\Big]d\mu$

$$= I_n + \int_s^t A(\mu)d\mu + \int_s^t \int_s^\mu A(\mu)A(\tau)\Phi(\tau,s)d\tau d\mu$$

→ Keep plugging in $\star$ to the integral featuring Φ to ∞.

$$\Phi(t,s) = \sum_{k=0}^{\infty} M_k(t,s) \qquad \textit{the Peano Baker Series}$$

where $M_0(t,s) = I_n, M_1(t,s) = \int_s^t A(\tau_1)d\tau_1$
→ $M_k(t,s) = \int_s^t \int_s^{\tau_1} ... \int_s^{\tau_{k-1}} A(\tau_1)A(\tau_2)...A(\tau_k)d\tau_k...d\tau_1$
Note, because matrices A might not commute, order of factors in integrals matters.

Question: How, if at all, does the **Peano-Baker Series** converge?

Answer: As nicely as you could possible wish for.

= converges for all $t, s \in \mathbb{R}$

= converges uniformly

= thus can take $\frac{d}{dt}$ term by term, etc.

Prove this using **Weierstrass's m-test**:

- Need to bound from above the (Euclidean matrix) norm of each term.

$$\|M_k(t,s)\| \leq \int_s^t \int_s^{\tau_1} ... \int_s^{\tau_{k-1}} \|A(\tau_1)A(\tau_2)...A(\tau_k)\| \, d\tau_k...d\tau_1$$
$$\leq \int_s^t \int_s^{\tau_1} ... \int_s^{\tau_{k-1}} \|A(\tau_1)\| \, \|A(\tau_2)\| \, ... \, \|A(\tau_k)\| \, d\tau_k...d\tau_1$$
$$\frac{1}{k!}\left(\int_s^t \|A(\tau)\| \, d\tau \right)^k$$

- look at $k = 1, 2$

$$k = 1 \qquad \|M_1(t,s)\| \leq \int_s^t \|A(\tau)\| \, d\tau$$
$$k = 2 \qquad \|M_2(t,s)\| \leq \int_s^t \int_s^{\tau_1} \|A(\tau_1)\| \, \|A(\tau_2)\| \, d\tau_2 d\tau_1$$
$$\leq \int_s^t \|A(\tau_1)\| \left(\int_s^t \|A(\tau_2)\| \, d\tau_2 \right) d\tau_1$$

- In any case, series converges for all $t, s \in \mathbb{R}$ and thus defines a $\mathbb{R}^{m \times n}$ valued $(t,s) \mapsto \Phi(t,s)$.
- Check to make sure it satisfies the (STM).

$$\Phi(s,s) = \sum_{k=0}^{\infty} M_k(s,s) \qquad M_0(s,s) = I_n \qquad M_k(s,s) = 0_n \; all \; k > 0$$
so $\Phi(s,s) = I_n$

- Also, by niceness of convergence

$$\frac{d}{dt}\Phi(t,s) = \sum_{k=0}^{\infty} M_k(t,s)$$
$$\frac{d}{dt}M_0(t,s) = \frac{d}{dt}(I_n) =)_n$$
$$\frac{d}{dt}M_1(t,s) = \frac{d}{dt}\int_s^t A(\tau_1)d\tau_1 = A(t)$$
$$\frac{d}{dt}M_2(t,s) = \frac{d}{dt}\int_s^t \int_s^{\tau_1} A(\tau_1)A(\tau_2)d\tau_2 d\tau_1 = \int_s^t A(t)A(\tau_2)d\tau_2 = A(t)M_1(t,s)$$
Similarly, $\frac{d}{dt}M_3(t,s) = A(t)M_2(t,s)$.

→ for all $k > 0 \frac{d}{dt}M_k(t,s) = A(t)M_{k-1}(t,s)$
→ $\frac{d}{dt}\Phi t, s = A(t)\Phi(t,s)$

Term: $(t,s) \mapsto \Phi(t,s)$ is State Transition Matrix Function associated with $t \mapsto A(t)$ Two important properties of Φ :

1. Semi-group property $\Phi(t,s)\Phi(s,\tau) = \Phi(t,\tau)$ all $t, s, \tau \in \mathbb{R}$ (no ordering required as there is with k's and l's in the DT case)
2. $\Phi(t,s)$ is invertible for $t, s \in \mathbb{R}$

   $\Phi(t,s)^{-1} = \Phi(s,t) \rightarrow$ reversing of the state transitions to say

Before specializing to constant A case, think about why we call x (in both continuous and discrete time) the state of the system model.

From these, note *State-Output Equations:*

$$y(t) = C(t)\Phi(t,t_0)x_0 + \int_{t_0}^{t} C(t)\Phi(t,\tau)B(\tau)u(\tau) + D(t)u(t) \qquad all\ t \in \mathbb{R}$$

$$y(k) = C(k)\Phi(k,k_0)x_0 + \sum_{l=k_0}^{k-1} C(k)\Phi(k,l+1)B(l)u(l) + D(k)u(k) \qquad all\ l \geq k_0$$

Chapter 2

Matrix Functions

2.1 Motivation with State Equations

2.1.1 State Transition Matrix (STM) and Matrix Exponential

Question: given $t_0 \in \mathbb{R}$, what do I need to know about *stuff* in the box to figure out y(t) for all $t \geq t_0$ given u(t) for all $t \geq t_0$ (or $k \geq k_0$ in DT)

Answer: Knowing $x, x(t_0) = x$ or $x(k_0)$ suffices!

Thus, for any t_0 or k_0, the state of at the time t_0 or k_0 gives you enough info to figure out:

- State at all $t \geq t_0$ or $k \geq k_0$
- Output at all $t \geq t_0$ or $k \geq k_0$

$\rightarrow$ Given the input for all $t \geq t_0$ or $k \geq k_0$ is "knowing x tells you internal state of system"

State Transition Matrix when A is a constant and $A \in \mathbb{R}^{n \times n}$:

- Show in DT that

$$\Phi(k,j) = \begin{cases} I_n, k = j \\ A^{k-j}, k > j \\ undefined, k < j \end{cases}$$

- In CT have the Peano-Baker Series

$$\Phi(t,s) = \sum_{k=0}^{\infty} M_k(t,s) \qquad t, s \in \mathbb{R}$$

where $M_k(t,s) = \int_s^t \int_s^{\tau_1} ... \int_s^{\tau_{k-1}} A(\tau_1)A(\tau_2)...A(\tau_k)d\tau_k...d\tau_1$, but because A is constant, changes to $M_k(t,s) = \int_s^t \int_s^{\tau_1} ... \int_s^{\tau_{k-1}} A^k d\tau_k...d\tau_1$.

Therefore,

$M_0(t,s) = I_n$

$M_1(t,s) = \int_s^t A d\tau = (t-s)A$

$$M_2(t,s) = \int_s^t \int_s^{\tau_1} A^2 d\tau_2 d\tau_1 = \frac{1}{2}(t-s)A^2 \Big|_{\tau=s}^{\tau=t} = \frac{(t-s)^2}{2}A^2$$

You can easily find with induction that $M_k(t,s) = \frac{(t-s)^k}{k!}A^k$

If you know $M_k(t,s) = \frac{(t-s)^k}{k!}A^k$ then follows that

$$M_{k+1}(t,s) = \int_s^t \int_s^{\tau_1} \ldots \int_s^{\tau_k} A^{k+1} d\tau_{k+1} d\tau_1$$

$$= \int_s^t A M_k(\tau 1, s) d\tau_1$$

$$= \int_s^t \frac{(\tau_1 - s)^k}{k!} A^{k+1} d\tau_1 = \frac{(\tau_1 - s)^{k+1}}{(k+1)!} A^{k+1} \Big|_{\tau=s}^{\tau=t} = \frac{(t-s)^{k-1}}{(k+1)!} A^{k+1}$$

Thus, P-B Series for a constant A (remember zeroth term is I_n)

$$\Phi(t,s) = \sum_{k=0}^{\infty} \frac{(t-s)^k}{k!} A^k \qquad t,s \in \mathbb{R}$$

Remember the Taylor series $e^{-z} = \sum_{k=0}^{\infty} \frac{z^k}{k!}$ for all $z \in \mathbb{R}$. From this we form the **Matrix Exponential**.

$$\Phi(t,s) = e^{(t-s)A} = exp\big[(t-s)A\big] \; for \; all \; t,s \in \mathbb{R}$$

Note above: $e^{(t-s)A} \neq$ elementwise $e^{(t-s) \; element \; of A}$. ie $\left[e^{(t-s)A}\right]_{i,j} \neq e^{[A]_{i,j}(t-s)}$.

Thus, when A,B,C,D are constant in (I), we have integral-ish terms for x, y, for all $t_0 \in \mathbb{R}$, $x_0 \in \mathbb{R}^n$

$$x(t) = e^{(t-t_0)A}x_0 + \int_t^{t_0} e^{(t-t_0)A} B u(\tau) d\tau \qquad t \in \mathbb{R}$$

$$y(t) = Ce^{(t-t_0)A}x_0 + \int_{t_0}^{t} e^{(t-t_0)A}) B u(\tau) + Du(t) \qquad all \; t \in \mathbb{R}$$

How do we compute e^{tA} for $t \in \mathbb{R}$ when $A \in \mathbb{R}^{n \times n}$:

$\rightarrow$ Sometimes (rarely) it is easy

- if $A \in \mathbb{R}^{n \times n}$ is diagonal ie

$$A = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & 0 \\ & & \ddots & \\ & 0 & & \lambda_{n-1} \\ & & & & \lambda_n \end{pmatrix}$$

- Then for $k \geq 0$

$$A^k = \begin{pmatrix} \lambda_1^k & & & \\ & \lambda_2^k & & 0 \\ & & \ddots & \\ & 0 & & \lambda_{n-1}^k \\ & & & & \lambda_n^k \end{pmatrix}$$

so,

$$\frac{t^k}{k!}A^k = \begin{pmatrix} \frac{(\lambda_1 t)^k}{k!} & & & \\ & \frac{(\lambda_2 t)^k}{k!} & & 0 \\ & & \ddots & \\ & 0 & & \frac{(\lambda_{n-1} t)^k}{k!} \\ & & & & \frac{(\lambda_n t)^k}{k!} \end{pmatrix}$$

and finally

$$e^{tA} = \sum_{k=0}^{\infty} \frac{t^k}{k!}A^k = \begin{pmatrix} e^{\lambda_1 t} & & & \\ & e^{\lambda_2 t} & & 0 \\ & & \ddots & \\ & 0 & & e^{\lambda_{n-1} t} \\ & & & & e^{\lambda_n t} \end{pmatrix}$$

- Reality check: because $e^{tA} = \Phi(t,0) \rightarrow e^{tA}$ is invertible for all $t \in \mathbb{R}$ This will always be so, no matter how non-invertible A is. Example, $A = 0, e^{tA} = I_n$ all t.
- if $A \in \mathbb{R}^{n \times n}$ is *nilpotent:* if $A^d = 0_n$ for some $d > 0$

  $\rightarrow$ the ostensibly infinite series for e^{tA} becomes

$$e^{tA} = \sum_{k=0}^{\infty} \frac{t^k}{k!}A^k \qquad all\ t \in \mathbb{R}$$

  $\rightarrow$ in particular, every entry in e^{tA} is a polynomial in t.

A couple of e^{tA} properties. $A \in \mathbb{R}^{n \times n}$:

1. $e^{(t+s)A} = e^{tA}e^{sA} \qquad all\ t, s \in \mathbb{R}$

   to see this, think $\qquad e^{(t+s)A} = e^{(t-(-s))A} = \Phi(t,-s)$

   $= \Phi(t,0)\Phi(0,-s)$ *by semigroup*

   $= e^{tA}e^{sA}$

2. Suppose $A_1, A_2 \in \mathbb{R}^{n\times n}$: $e^{t(A_1+A_2)} \neq e^{tA_1}e^{tA_2}$

necessarily true for all t $\iff$ A_1, A_2 commute ie $A_2A_1 = A_1A_2$

To see this, consider $LHS(t) = e^{t(A_1+A_2)}$ *and* $RHS(t) = e^{tA_1}e^{tA_2}$

$\frac{d}{dt}LHS(t) = (A_1 + A_2)e^{t(A_1+A_2)}$

$\frac{d}{dt}RHS(t) = A_1e^{t(A_1)}e^{t(A_2)} + e^{t(A_1)}A_2e^{t(A_2)}$

$\frac{d^2}{dt^2}LHS(t) = (A_1 + A_2)^2e^{t(A_1+A_2)}$

$\frac{d^2}{dt^2}RHS(t) = A_1^2e^{t(A_1)}e^{t(A_2)} + A_1e^{t(A_1)}A_2e^{t(A_2)} + A_1e^{t(A_1)}A_2e^{t(A_2)} + e^{t(A_1)}A_2^2e^{t(A_2)}$

evaluate both expressions at t = 0 to get

$(A_1 + A_2)^2 \stackrel{?}{=} A_1^2 + 2A_1A_2 + A_2^2$

= when A_1 and A_2 commute!

if $e^{t(A_1+A_2)} = e^{tA_1}e^{tA_2}$ then $A_2A_1 = A_1A_2$.

if $A_2A_1 = A_1A_2$ then

$\frac{d}{dt}LHS(t) = (A_1 + A_2)LHS(t) \qquad L(0) = I_n$

$\frac{d}{dt}RHS(t) = A_1RHS(t) + e^{t(A_1)}A_2e^{t(A_2)} \qquad where\ e^{t(A_1)}A_2 = A_2e^{t(A_1)}$

$= (A_1 + A_2)RHS(t)$

$\rightarrow$ so it is apparent that the LHS and RHS satisfies the same DE with the same initial condition and therefore the LHS = RHS all t

2.1.2 MATH 2940: Linear Algebra Review

- given $A \in \mathbb{R}^{n\times n}$, can view A as defining a linear mapping $v \mapsto Av$ from $\mathbb{C}^n$ *to* $\mathbb{C}^n$.
- say $v_0 \in \mathbb{C}^n$ is an eigenvector of A when $v_0 \neq 0$ and $Av_0 = \lambda_0 v_0$ for some $\lambda \in \mathbb{C}$ (In this case, V_0 is the eigenvector corresponding with the eigenvalue λ_0). Eigenvalue facts:
  - $A \in \mathbb{R}^{n\times n}$ has at most n distinct eigenvectors
  - because A is real and $A \in \mathbb{R}^{n\times n}$, eigenvalues come in complex conjugate pairs.
    $\lambda_0 \leftrightarrow v_0 \iff \bar{\lambda_0} \leftrightarrow \bar{v_0}$ where $^-$ complex conjugate
  - if λ_0 is an eigenvalue of A, define:
    $E(\lambda_0)$ = subspace of $\mathbb{C}^n$ spanned by the eigenvectors of A corresponding with eigenvalue λ_0 (note : every nonzero $v \in E(\lambda_0)$ is an eigenvector of $A \leftrightarrow \lambda_0$)
  - If λ_0 is an eigenvalue of A,
    $G(\lambda_0)$ = set of all $v \in \mathbb{C}^n$ such that $(A - \lambda_0 I_n)^k \cdot v = 0$ some $k > 0$
    $G(\lambda_0)$ = the generalize eigenspace $\leftrightarrow \lambda_0$
    Nonzero $v \in G(\lambda_0)$ are generalized eigenvectors $\leftrightarrow \lambda_0$
  - Observe:
    - $E(\lambda_0) \subset G(\lambda_0)$
    - $E(\lambda_0) = nullspace(A - \lambda_0 I_n)$
    - If $\lambda_1, \cdots, \lambda_s$ are the distinct eigenvalues of A ($s \leq n$), then $G(\lambda_j)\ for\ 1 \leq j \leq s$ are disjoint subspaces of $\mathbb{C}^n$ that 'span' $\mathbb{C}^n$ in the sense that you can write

any $v \in \mathbb{C}^n$ (uniquely) as the sum of vectors in the $G(\lambda_0)$ ie sum of generalize eigenvectors.

- Thus, can always find a basis for $\mathbb{C}^n$ consisting solely of generalized eigenvectors of A. Also, if complex dimension $dim(G(\lambda_0)) = d_j \qquad 1 \leq j \leq s$
  then, $d_1 + d_2 + \cdots + d_s = n$
- Call d_j the algebraic multiplicity of λ_j

- $G(\lambda_0) = nulspace(A - \lambda_j I_n)^{d_j} \qquad 1 \leq j \leq s$
- Follows that $\star \rightarrow (A - \lambda_1 I_n)^{d_1}(A - \lambda_2 I_n)^{d_2} \cdots (A - \lambda_s I_n)^{d_s} \cdot v = 0$ for all $v \in \mathbb{C}^n$
  Why is this true? Think about the $s = 2$ case, with two distinct eigenvalues λ_1, λ_2
  You will have $G(\lambda_1), G(\lambda_2)$ - given $v \in \mathbb{C}^n$, write $v = v_1 + v_2, v_1 \in G(\lambda_1), v_2 \in G(\lambda_2)$
  $(A - \lambda_1 I_n)^{d_1}(A - \lambda_2 I_n)^{d_2} \cdot v =$
  $(A - \lambda_2 I_n)^{d_2}(A - \lambda_1 I_n)^{d_1} v_1 + (A - \lambda_1 I_n)^{d_1}(A - \lambda_2 I_n)^{d_2} v_2$
  which $= 0$ by definition as solution of nullspace, sending $(A - \lambda_1 I_n)^{d_1} v_1 = 0$
- **Term:** The polynomial below is the characteristic polynomial of A

$$(\lambda - \lambda_1)^{d1}(\lambda - \lambda_2)^{d2} \cdots (\lambda - \lambda_s)^{ds}$$

(note: eigenvalues are roots of characteristic polynomial)

- Turns out that the Characteristic Polynomial $= det(\lambda I_n - A)$
- $\star$ above is know as the Cayley - Hamilton Theorem
  " A satisfies its own characteristic polynomial"
  If you plug in A for λ in the characteristic polynomial, you get the 0 matrix.
  Idea is that

$$(A - \lambda_1 I_n)^{d_1}(A - \lambda_2 I_n)^{d_2} \cdots (A - \lambda_s I_n)^{d_s} \cdot v = 0 \; all \; v \in \mathbb{C}^n$$

$$\Updownarrow$$

$$(A - \lambda_1 I_n)^{d_1}(A - \lambda_2 I_n)^{d_2} \cdots (A - \lambda_s I_n)^{d_s} = 0_n$$

- Can write the characteristic polynomial in the unfactored form

$$\lambda^n + q_1\lambda^{n-1} + q_2\lambda_{n-2} + \cdots + q_{n-1}\lambda + q_n \qquad where \; q_1, \cdots, q_n \in \mathbb{R}$$

- Important application of the Cayley-Hamilton gives us a way to express high powers of A as linear combos of $I_n, A, A^2, \cdots A^{n-1}$
  Cay Ham $\rightarrow A^n + q_1 A^{n-1} + \cdots + q_{n-1}A + q_n I_n = 0_n$
  $\rightarrow A^n = -q_1 A^{n-1} - \cdots - q_{n-1}A - q_n I_n$. To get A^{n+1} multiply by A
  $A^{n+1} = -q_1 A^n - \cdots - q_{n-1}A^2 - q_n A$
  $= -q_1(above \; formula) - q_2 A^{n-1} - \cdots - q_{n-1}A^2 - q_n A.$

This result is key later on

- Back to $\lambda_1, \cdots, \lambda_s$ distinct eigenvalues, $E(\lambda_j), G(\lambda_j), 1 \leq j \leq s, d_j = dim(G(\lambda_j)), E(\lambda_j) \subset G(\lambda_j)\ for\ 1 \leq j \leq s$ which implies
  $m_j = dim(E(\lambda_j)) \leq dim(G(\lambda_j)) = d_j$
  call m_j the geometric multiplicity of λ_j for $1 \leq j \leq s$
- when $m_j = d_j$ for all j, $E(\lambda_j) = G(\lambda_j)$ for all j's.
  Meaning every generalized eigenvector of A is a (true) eigenvector of A. In this case, can find a basis for $\mathbb{C}^n$ consisting solely of (true) eigenvectors of A.
- **Term**: when $m_j = d_j$ for all j, we say A is diagonalizable.
  Generalized eigenvectors of A always span $\mathbb{C}^n$ and the (true) eigenvectors span $\mathbb{C}^n \iff$ A is diagonalizable.

2.1.3 Summary: Properties of e^{tA} The Matrix Exponential

Properties:

1. $e^0 = I$
2. $e^{A(t+s)} = e^{At}e^{As}$
3. $e^{(A+B)t} = e^{At}e^{Bt} \quad iff \quad AB = BA$
4. $(e^{At})^{-1} = e^{-At}$
5. $\frac{d}{dt}a^{At} = Ae^{At} = e^{At} \cdot A$
6. let $z(t) \in \mathbb{R}^{n \times n}$. Then the solution to

$$\dot{z}(t) = Az(t) \quad with \quad z(0) = I \quad \rightarrow \quad z(t) = e^{At}$$

2.2 Calculating the Matrix Exponential

2.2.1 Eigenvalue Method

Suppose A is diagonalizable...

Aside on diagonalizability:

- can take any $A \in \mathbb{R}^{n \times n}$ and perturb its elements by an arbitrarily small amount and get another diagonalizable matrix.
  "diagonalizable matrices are dense in $\mathbb{R}^{n \times n}$
  - if $A \in \mathbb{R}^{n \times n}$ has n distinct eigenvalues, A is diagonalizable

· eigenvalues are roots of characteristic polynomial $\lambda^n+q_1\lambda^{n-1}+q_2\lambda_{n-2}+\cdots+q_{n-1}\lambda+q_n$ and you can tweak the q's by an arbitrary small amount to get n distinct roots

· Tweaking $q_j's$ by tweaking elements of A, since $q_j's$ are just polynomials in A's entries

- Given probability density on $\mathbb{R}^{n\times n}$, then probability that a randomly drawn matrix A is diagonalizable is 1.

→ given A diagonalizable, compute e^{tA} as follows, with introduction of "the $\mathbb{X}$ - maneuver":

- find $\lambda_1,\cdots,\lambda_n$ not necessarily distinct eigenvalues of A,(Each distinct eigenvalue appears on list as many times as its algebraic multiplicity) and a set of corresponding linearly independent eigenvectors $v_1,\cdots,v_n$ (remember we can find them since A diagonalizable).
- form $\mathbb{X}\in\mathbb{C}^{n\times n}$ - $\mathbb{X}$ has v_j as its jth column for $1\le j\le s$
- then $A\mathbb{X}=$

$$A\mathbb{X}=A\left[\begin{array}{c|c|c|c} v_1 & v_2 & \cdots & v_n \end{array}\right]=\left[\begin{array}{c|c|c|c} Av_1 & Av_2 & \cdots & Av_n \end{array}\right]$$

$$=\left[\begin{array}{c|c|c|c} \lambda_1v_1 & \lambda_2v_2 & \cdots & \lambda_nv_n \end{array}\right]$$

$$=\left[\begin{array}{c|c|c|c} v_1 & v_2 & \cdots & v_n \end{array}\right]\left[\begin{array}{cccc} \lambda_1 & & & \\ & \lambda_2 & & 0 \\ & & \ddots & \\ & 0 & & \lambda_{n-1} \\ & & & & \lambda_n \end{array}\right]$$

$$=\mathbb{X}\Lambda$$

- $\mathbb{X}$ columns linearly independent → X^{-1} exists

$A=\mathbb{X}\Lambda\mathbb{X}^{-1}$ and $\mathbb{X}^{-1}A\mathbb{X}=\Lambda$

→ Apply to computing e^{tA} (also helps for computing A^k)

$A^2=(\mathbb{X}\Lambda\mathbb{X}^{-1})(\mathbb{X}\Lambda\mathbb{X}^{-1})=\mathbb{X}\Lambda^2\mathbb{X}^{-1}$

$A^3=(\mathbb{X}\Lambda\mathbb{X}^{-1})(\mathbb{X}\Lambda\mathbb{X}^{-1})(\mathbb{X}\Lambda\mathbb{X}^{-1})=\mathbb{X}\Lambda^3\mathbb{X}^{-1}$ which is a telescoping product

for $k>0$ $A^k=\mathbb{X}\Lambda^k\mathbb{X}^{-1}$, and note $k=0\rightarrow A^0=I_n$

- Def:

$$e^{tA}=\sum_{k=0}^{\infty}\frac{t^k}{k!}A^k=sum_{k=0}^{\infty}\frac{t^k}{k!}\mathbb{X}\Lambda^k\mathbb{X}^{-1}$$

$$=\mathbb{X}\Big(\sum_{k=0}^{\infty}\frac{t^k}{k!}\Lambda^k\Big)\mathbb{X}^{-1}$$

$$= \mathbb{X} e^{t\Lambda} \mathbb{X}^{-1}$$

Qualitative Observations: (important for stability) when A is diagonalizable.

· entries in A^k are linear combinations of terms of the form λ_0^k, where λ_0 is an eigenvalue of A, for $k \geq 0$

· entries in e^{tA} are linear combos of terms of the function $e^{\lambda_0 t}$, where λ_0 is an eigenvalue of A,

What if A is not Diagonalizable?

- Suppose $\lambda_1, \lambda_2, \cdots, \lambda_s$ are A's distinct eigenvalues with respective algebraic multiplicities of $d_1, d_2, \cdots, d_s$ thus $d_1 + d_2 + \cdots + d_s = n$
- let $G(\lambda_j)$ = generalized eigenspace $\leftrightarrow \lambda_j \qquad 1 \leq j \leq s$, then $dim(G(\lambda_j)) = d_j$

→ Form $\mathbb{X} \in \mathbb{C}^{n \times n}$ as follows:

· first d_1 columns form a basis for $G(\lambda_1), \cdots$, last d_s columns for a basis for the last generalized eigenspace $G(\lambda_s)$

· Columns of $\mathbb{X}$ together form a basis for $\mathbb{C}^{n \times n}$, thus $\mathbb{X}$ exists

Aside: { each of $G(\lambda_1)$ is a subspace of $\mathbb{C}^n$ invariant under A in a sense that if $v \in G(\lambda_1)$, then $Av \in G(\lambda_j)$ for all j. Because, if $v \in G(\lambda_j)$, then for some k, $(A - \lambda_j I_n)^k v = 0, k \geq 0$, thus $(A - \lambda_j In)^k (Av) = A(A - \lambda_j In)^k = 0$ (by commutivity of A across generalized eigen vector space). This implies that $Av \in G(\lambda_j)$ as well. Also note: powers of A commute in matrix form.}

· Invariance applies that $A\mathbb{X} =$

$$= \mathbb{X} \begin{bmatrix} A_1 & & & \\ & A_2 & & 0 \\ & & \ddots & \\ & 0 & & A_{s-1} \\ & & & & A_s \end{bmatrix}$$

Herein, $A_1, A_2, \cdots, A_s$ are matrices taking up areas along the center of the matrix. A later example will clarify finding them. For now, say $s = 2$. Then

$$AX = A \left[\begin{array}{c|c} \textit{All in } G(\lambda_1) & \textit{All in } G(\lambda_2) \end{array} \right] = \mathbb{X} \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}$$

Where *All in* $G(\lambda_i)$ means that area of the matrix is linear combos of the first d_j columns of A

Discover that, for each j, $(A_j - \lambda_j I_{d_j})^{d_j} = 0_{d_j}$. Thus,

$$\mathbb{X}^{-}1A\mathbb{X} = \begin{bmatrix} A_1 & & & & \\ & A_2 & & 0 & \\ & & \ddots & & \\ & 0 & & A_{s-1} & \\ & & & & A_s \end{bmatrix} = \Lambda + S$$

Where,

$$S = \begin{bmatrix} (A_1 - \lambda_1 I_{d_1}) & & & & \\ & (A_2 - \lambda_2 I_{d_2}) & & 0 & \\ & & \ddots & & \\ & 0 & & (A_{s-1} - \lambda_{s-1} I_{s_1}) & \\ & & & & (A_s - \lambda_s I_{d_s}) \end{bmatrix}$$

$$and\ \Lambda = \begin{bmatrix} \lambda_1 I_{d1} & & & & \\ & \lambda_2 I_{d2} & & 0 & \\ & & \ddots & & \\ & 0 & & \lambda_{s-1} I_{d_{s-1}} & \\ & & & & \lambda_1 I_{d_s} \end{bmatrix}$$

satisfying

$$\Lambda S = S\Lambda \qquad S^d = 0_n\ where\ d = max(d_j)$$

Here is an example of forming the $A_j's$ for the new $\mathbb{X}$ transform. Start with

$$A = \begin{bmatrix} 3 & 1 & 1 \\ 0 & 3 & 1 \\ 0 & 0 & 2 \end{bmatrix} \rightarrow \lambda_1 = 3, d_1 = 2; \lambda_2 = 2, d_2 = 1$$

The basis for $G(\lambda_2)$

$$(A - 2I_n)v = 0 \rightarrow \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},\ yields\ v\ = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$$

The basis for $G(\lambda_1)$

$$(A-3I_n)v=0 \rightarrow \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \; yields \; v \; = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

To find the second eigenvalue for $G(\lambda_1)$, solve

$$(A-3I_n)^2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 1 \end{bmatrix}, (A-3I_n)^2\lambda = 0 \rightarrow v_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

So the basis for $G(\lambda_1)$ is

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

Next, find $\mathbb{X}$ with the eigenvectors from $G(\lambda_1)$ first followed by $G(\lambda_2)$

$$\mathbb{X} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & -1 \end{bmatrix} \; \& \; \mathbb{X}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & -1 \end{bmatrix}$$

and,

$$A\mathbb{X} = \begin{bmatrix} 3 & 1 & 0 \\ 0 & 3 & 2 \\ 0 & 0 & -2 \end{bmatrix} = \mathbb{X} \begin{bmatrix} 3 & 1 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{bmatrix}, \; where \; A_1 = \begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix}, \; A_2 = 2$$

Bottom Line:

$\mathbb{X}^{-1}\mathbb{A}\mathbb{X} = \Lambda + S \qquad S^d = 0_n$

$A = \mathbb{X}(\Lambda + S)\mathbb{X}^{-1}$

$A^k = \mathbb{X}(\Lambda + S)^k\mathbb{X}^{-1} \qquad k \geq 0$

so

$$\rightarrow \; e^{tA} = \sum_{k=0}^{\infty} \frac{t^k}{k!} A^k = \mathbb{X} \sum_{k=0}^{\infty} \frac{t^k}{k!} (\Lambda + S)^k \mathbb{X}^{-1} =$$

$$\mathbb{X}e^{t(\Lambda+S)}\mathbb{X}^{-1}$$

Because $\Lambda S = S\Lambda$, $e^{t(\Lambda+S)} = e^{t\Lambda}e^{tS}$ for all t

Conclude

$$e^{tA} = \mathbb{X}e^{t\Lambda}e^{tS}\mathbb{X}^{-1}$$

Where e^{tA} and e^{tS} are easy to calculate, the later because $S^d = 0_n$ which bounds the sum

So, how does this help with calculating A^k, because $S\Lambda = \Lambda S$!

$$(\Lambda + S)^k = \sum_{l=0}^{k} \binom{n}{k} S^l \Lambda^{k-l}$$

because $S^d = 0$, when $k \geq d$, sum cuts off at $l = d - 1$

$$(\Lambda + S)^k = \sum_{l=0}^{d-1} \binom{n}{k} S^l \Lambda^{k-l} \qquad k \geq d$$

Thus, $A^k = \mathbb{X}(\Lambda + S)^k \mathbb{X}^{-1}$

$$= \begin{cases} \mathbb{X}(\sum_{l=0}^{k} \binom{n}{k} S^l \Lambda^{k-l})\mathbb{X}^{-1} & k \leq d \\ \mathbb{X}(\sum_{l=0}^{d-1} \binom{n}{k} S^l \Lambda^{k-l})\mathbb{X}^{-1} & d \leq k \end{cases}$$

All items at most d for an k, so they are 'easy' to find.

Qualitative Observations

- Every entry in e^{tA} is a linear combination of terms from $(\textit{polynomial in } t) \cdot e^{\lambda_0 t}$ where λ_0 is an eigenvector of A
- Entries in A^k, $k \geq 0$ are linear combinations of terms from $(\textit{polynomial in } t) \cdot \lambda_0^k$ where λ_0 is an eigenvector of A

2.2.2 Laplace and Z Transform Method

Quick Laplace Transform Review

- Given $F : \mathbb{R} \mapsto \mathbb{C}$, the Laplae Transform of F (if it exists) is the function $F(S)$ with a region of convergence (ROC) defined by

$$F(S) = \int_{-\infty}^{\infty} f(t) e^{-st} dt \qquad \textit{on R.O.C. } \sigma_a < \mathbb{R}\{s\} < \sigma_b$$

- The requirements of F for $F(s)$ to exist (in this class)
  - = exists some $\sigma \in \mathbb{R}$ s.t. (converges on F) $\lim_{t \to \pm\infty} e^{-\sigma t} f(t) = 0$
  - = set $\sigma_a = inf(\{\sigma \in \mathbb{R} : e^{-\sigma t} f(t) \to 0 \textit{ as } t \to +\infty\})$
    set $\sigma_b = sup(\{\sigma \in \mathbb{R} : e^{-\sigma t} f(t) \to 0 \textit{ as } t \to -\infty\})$
    ($\sigma_a = -\infty$ *and / or* $\sigma_b = \infty$ are allowed), thus $\sigma_a < \sigma_b$
  - = When there is an integral defining $F(s)$, it exists for all s satisfying $\sigma_a < \mathbb{R}\{s\} < \sigma_b$

- Note, when $f(t) = 0$ all $t < 0$, $\sigma_b = \infty$, so if there is a σ_a that works, ROC is of the form $\sigma_a < \mathbb{R}\{s\} < \infty$
- A non-transformable F is one that grows to fast as $t \to \infty$ for any $e^{-\sigma t}$ no matter how big σ is. Example is:

$$f(t) = \begin{cases} 0 & t < 0 \\ e^{t^2} & t \geq 0 \end{cases}$$

- Here are some classic examples of Laplace Transform Pairs

$$f(t) = \begin{cases} e^{s_0 t} & t \geq 0 \\ 0 & t < 0 \end{cases} \qquad F(s) = \frac{1}{s - s_0} \text{ on ROC } \mathbb{R}\{s_0\} < \mathbb{R}\{s\} < \infty$$

To see this, set $F(s) = \int_{-\infty}^{\infty} f(t) e^{-st} dt = \int_0^{\infty} e^{-(s-s_0)t} dt$
Clearly, integral exists only when $\mathbb{R}\{s - s_0\} > 0$, giving ROC as $\mathbb{R}\{s_0\} < \mathbb{R}\{s\} < \infty$

$$F(s) = \frac{-1}{s - s_0} e^{-(s-s_0)t} \Big|_{t=0}^{t=\infty} = \frac{1}{s - s_0}$$

(evaluation at $t = \infty$ is zero because $\mathbb{R}\{s - s_0\} > 0$)

$$f(t) = \left\{ \frac{t^m}{m!} e^{s_o t} \mathbb{1}(t) \qquad m > 0 \right. \qquad \overset{\mathbb{L}}{\leftrightarrow} \qquad F(s) = \frac{1}{(s - s_0)^{M+1}} \text{ on ROC } \mathbb{R}\{s_0\} < \mathbb{R}\{s\} < \infty$$

- In general, we'll encounter things where the Laplace transform of the form $F(s) = \frac{p(S)}{q(s)}$, where $p(s), q(s)$ are polynomials in S space. Such an $F(s)$ is a rational function of S. $F(s)$ is a proper rational function when $deg(p(s)) \leq deg(q(s))$, and it is a strictly proper rational function when $deg(p(s)) < deg(q(s))$. Note: all prototypical examples have $F(s)$ being strictly proper.
- If we know ahead of time that the $f(t)$ that goes with $F(s)$ satisfies $f(t) = 0$ for $t < 0$, then the ROC comes out to be $\mathbb{R}\{s_0\} < \mathbb{R}\{s\}, \infty$ where the $\mathbb{R}\{s_0\}$ turns out to be the max of real parts of the poles of $F(s)$ = roots of $q(s)$.

Now, we will work through a **prototypical problem** that we will encounter when using the Laplace Transform in this way. Suppose $f(t) = 0$ for all $t < 0$, and we have $F(s) =$ "a strictly proper rational function of s." Find F

Solve by knowing $F(s)$ is strictly proper, so we can expand $F(s)$ with partial fractions, where every term will be of the form

$$\frac{constant}{(s - s_0)^{M+1}}$$

for some $m \geq 0$. Here s_0 is some pole of $F(s)$. Each term gives rise to a term in f(t) of the form

$$(constant) \frac{t^m}{m!} e^{s_0 t} \mathbb{1}(t)$$

Example: take

$$F(s) = \frac{1}{(s-1)(s+2)(s-2)} \qquad poles = 1, -2, 3$$

Partial fractions by zen yields

$$= \frac{\frac{-1}{6}}{s-1} + \frac{\frac{1}{15}}{s+2} + \frac{\frac{1}{10}}{s-3}$$

And Laplace rules give

$$f(t) = \big(-\frac{1}{6}e^{t} + \frac{1}{6}e^{-2t} + \frac{1}{6}e^{3t} \big) \mathbb{1}(t) \qquad t \in \mathbb{R}$$

Next, how do we apply this to finding e^{tA}? Look at an example matrix valued function $g(t) = e^{tA}\mathbb{1}(t), t \in \mathbb{R}$. Last time we same that every element in e^{tA} is a linear combination of terms of the form $(polynomial\ in\ t) \cdot e^{\lambda_0 t}$ where λ_0 is an eigenvalue of A.

- Thus, elements in $e^{tA}\mathbb{1}(t)$ are all linear combinations of the prototypical examples.
- Additionally, all functions of that form have Laplace Transforms and are strictly proper ration functions of s.
- Can compute $e^{tA}, t \geq 0$ element-by-element from partial fractions, etc. We can also think bigger and do the whole matrix at once, as follows:
  1) Note $G(s)$
  $$= \int_{-\infty}^{\infty} g(t)e^{-st}dt = \int_{0}^{\infty} e^{tA}e^{-st}dt$$
  2) Knowing $e^{tA}e^{-st} = e^{-(sI_n - A)t}$. Then when $(sI - n - A)^{-1}$ exists (this is true when s is not an eigenvalue of A) we have, by magic manipulation:
  $$e^{-(sI_n - A)t} = -\frac{d}{dt}(sI_n - A)^{-1}e^{-(sI_n - A)t}$$
  3) Plugging this into 1) yields
  $$G(s) = \int_{0}^{\infty} -\frac{d}{dt}(sI_n - A)^{-1}e^{-(sI_n - A)t}dt$$
  $$= -(sI_n - A)^{-1}e^{-(sI_n - A)t}\Big|_{t=0}^{t=\infty}$$
  · Assuming $\mathbb{R}\{s\} > \mathbb{R}\{\lambda_0\}$ for every λ_0 eigenvalue of A. This shows that $e^{-(sI_n - A)t} \to 0\ as\ t \to \infty$ because all entries are linear combinations of the form $(polynomial\ in\ t) \cdot e^{-(s-\lambda_0)t}$
  4) Thus, every entry in $G(s)$ is a strictly proper rational function of s where poles lie

among the eigenvalues of A.

$$G(s) = (sI_n - A)^{-1}$$

$$g(t) = e^{tA}\mathbb{1}(t) \overset{\mathbb{L}}{\leftrightarrow} G(s) = (sI_n - A)^{-1}$$

Bottom Line: to find $e^{tA}\mathbb{1}(t)$

- Compute $G(s) = (sI_n - A)^{-1}$
- Take element wise inverse Laplace Transform using prototypical examples

Example: Say,

$$A = \begin{bmatrix} 3 & 7 \\ 0 & 2 \end{bmatrix}$$

$$(sI_n - A) = \begin{bmatrix} s-3 & -7 \\ 0 & s-2 \end{bmatrix}, \rightarrow (sI_n - A)^{-1} = \frac{1}{(s-3)(s-2)} \begin{bmatrix} s-3 & -7 \\ 0 & s-2 \end{bmatrix}$$

Partial fractions yields

$$\frac{7}{(s-3)(s-2)} = \frac{7}{(s-3)} + \frac{-7}{(s-2)}$$

Giving

$$e^{tA} = \begin{bmatrix} e^{3t} & (7e^{3t} - 7e^{3t}) \\ 0 & e^{2t} \end{bmatrix}$$

Now, Calculate A^k with the Z transform for $k \geq 0$. We'll start with an intro to the Z-Transform.

- Given $f : \mathbb{Z} \mapsto \mathbb{C}$, the Z-Transform of f, if it exists, is

$$F(z) = \sum_{k=-\infty}^{\infty} f(k)z^{-k} \qquad R.O.C. = R_A < |z| < R_B$$

Where the R.O.C. is the set of z-values (essentially) that makes the series converge

- if $f(k) = 0$ for $k < 0$, ROC is of the form $R.O.C. = R_A < |z| < \infty$
- Note the following about geometric series

$$\sum_{k=0}^{\infty} \gamma^k = \frac{1}{1-\gamma}, \gamma < 1, \qquad and \qquad \sum_{k=0}^{N-1} \gamma^k = \frac{1-\gamma^N}{1-\gamma}, \gamma \neq 1$$

$\rightarrow$ Onto some prototypical examples:

- Say $f(k) = Z_0^k(k)$ where $\mathbb{1}(k)$ is the discrete time unit step which equals one for

nonnegative time. Then,

$$F(z) = \sum_{k=-\infty}^{\infty} f(k) z^{-k} = \sum_{k=0}^{\infty} z_0^k z^{-k} = \sum_{k=0}^{\infty} \frac{z_0}{z}^k$$

Giving $F(z)$

$$= \frac{1}{1 - \frac{z_0}{z}} = \frac{z}{z - z_0} \qquad ROC: \ |\frac{z_0}{z}| < 1 \ or \ |z_0| < |z| < \infty$$

- or if $f(t) = \binom{k}{m} z_0^{k-m} \mathbb{1}(k) \qquad k \in \mathbb{Z}$. Gives

$$F(z) = \frac{z}{(z - z_0)^{M+1}} \qquad ROC: |z_0| < |z| < \infty$$

Problem: Given $F(z)$, knowing its the z-transform of some function f satisfying $f(k) = 0, k < 0$, and $F(z)$ is a proper rational function of z, find f:

Solution: Take the approach that $\frac{F(z)}{z}$ is a strictly proper rational function, so you can expand in partial fractions with terms of the form $\frac{constant}{z}$ pr $\frac{constant}{(z-z_0)^{m+1}}$ where $z_0 =$ pole of the $F(z)$. Multiply through by z to get $F(z) =$ "A sum of prototypical examples plus, perhaps, a constant. If there's a constant C_0, it results in a term of the form $C_0 \delta(k) = C_0 \ if \ k = 0, \ 0 \ if \ k \neq 0$ in $f(k)$.

Example Given

$$F(z) = \frac{z}{(z+1)(z-3)}$$

$$\frac{F(z)}{z} = \frac{1}{(z+1)(z-3)} = \frac{\frac{-1}{4}}{(z+1)} + \frac{\frac{1}{4}}{(z-3)}$$

$$\Rightarrow F(z) = \frac{\frac{-z}{4}}{(z+1)} + \frac{\frac{z}{4}}{(z-3)} \ and \ f(k) = \big(-\frac{1}{4}(-1)^k + \frac{1}{4}(3)^k \big) \mathbb{1}(k)$$

So, how do we use this to find $A^k \ f \geq 0$?: Start with $g(k) = A^k \mathbb{1}(k)$. We saw last time that each element of A^k is a linear combination of terms of the form $(polynomial \ in \ k) \cdot \lambda_0^k$ where λ_0 is a eigenvalue of A. Thus, all the entries have z-transforms, etc. etc, but lets go for the fold. Find $G(z)$

$$G(z) = \sum_{k=-\infty}^{\infty} g(k) z^{-k} = \sum_{k=0}^{\infty} z^{-k} A^k$$

(Note: because of the form of A^k's entries, computes when $|z| > max(\lambda_0)$ where λ_0 is the eigenvalue of A)

But, what does the series converge to? Consider

$$(I_n - z^{-}1A)\sum_{k=0}^{N-1} z^{-k}A^k = I_n - z^{-N}A^N$$

By the canceling of

$$(I_n + z^{-1}A + z^{-2}A^2 + \cdots + z^{-(n-1)}A^{N-1}) - (z^{-1}A - z^{-2}A^2 - \cdots - z^{-N}A^N)$$

E.g.:

$$A = \begin{bmatrix} 3 & 11 \\ 0 & 7 \end{bmatrix}$$

$$zI_2 - A = \begin{bmatrix} z-3 & -11 \\ 0 & z-7 \end{bmatrix} \rightarrow z(zI_2 - A)^{-1} = \begin{bmatrix} \frac{z}{z-3} & \frac{11z}{(z-3)(z-7)} \\ 0 & \frac{z}{z-7} \end{bmatrix}$$

by examples, yields

$$\begin{aligned} \left[A^k\mathbb{1}(k)\right]_{11} &= 3^k\mathbb{1}(k) \\ \left[A^k\mathbb{1}(k)\right]_{22} &= 7^k\mathbb{1}(k) \\ \left[A^k\mathbb{1}(k)\right]_{21} &= 0 \\ \left[A^k\mathbb{1}(k)\right]_{12} &= \rightarrow Expand \end{aligned}$$

$$\left[A^k\mathbb{1}(k)\right]_{12} = \frac{11z}{(z-3)(z-7)} = \frac{\frac{11z}{4}}{(z-7)} + \frac{\frac{-11z}{4}}{(z-3)} = \frac{-11}{4}(3)^k\mathbb{1}(k) + \frac{11}{4}(7)^k\mathbb{1}(k)$$

A final note on the second dubious method of calculation:

$$e^{tA}\mathbb{1}(t) \overset{\mathbb{L}}{\leftrightarrow} (sI_n - A)^{-1} \qquad and \qquad A^k\mathbb{1}(k) \overset{\mathbb{L}}{\leftrightarrow} z(zI_n - A)^{-1}$$

2.2.3 Cayley-Hamilton Theorem Method

- Given $A \in \mathbb{R}^{n\times n}$, let the characteristic polynomial of A =

$$\lambda^n + q_1\lambda^{n-1} + q_2\lambda_{n-2} + \cdots + q_{n-1}\lambda + q_n \qquad where\ q_1, \cdots, q_n \in \mathbb{R}$$

- Noting that the characteristic polynomial $= det(\lambda I_n - A)$ or $(\lambda-\lambda_1)^{r1}(\lambda-\lambda_2)^{r2}\cdots(\lambda-\lambda_s)^{rs}$ where $\lambda_1, \cdots, \lambda_s$ are A's eigenvalues and $r_1, \cdots, r_s$ are their algebraic multiplicities.
- Thus,

$$A^n + q_1A^{n-1} + \cdots + q_{n-1}A + q_nI_n = 0_n$$

- Use this for forumal to find higher powers of A as linear combinations of lower powers of A - specifically $I_n, A, \cdots, A^{n-1}$.

Start by considering the Power Series:

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

- Imagine plugging in the Cayley-Hamilton formula for A^k in terms of $I_n, A, \cdots, A^{k-1}$. Then gathering all the coefficients of I_n, the coefficients of A, $\cdots$, the coefficients of A^{k-1} together. One yields

$$e^{tA} = g_0(t)I_n + g_1(t)A + \cdots + g_{n-1}(t)A^{n-1}$$

  i.e. e^{tA} is a 'time varying linear combination" of $I_n, A, \cdots, A^{k-1}$
- *Caution*: sometimes there's many g's that work here. In essence, the expansion above is not unique. E.g. suppose $A = I_n$, the two below are equal, but with different g's

$$e^{tA} = e^t I_n \leftrightarrow I_n + (e^t - 1)A$$

- Here's how to find one n-tuple set of $g_k's$ that works
  - Start with
$$e^{tA} = g_0(t)I_n + g_1(t)A + \cdots + g_{n-1}(t)A^{n-1}$$
  - Take the time derivative of both sides to yield
$$\frac{d}{dt}e^{tA} = \dot{g}_0(t)I_n + \dot{g}_1(t)A + \cdots + \dot{g}_{n-1}(t)A^{n-1}$$
  - Knowing $\frac{d}{dt}e^{tA} = Ae^{tA}$, get
$$g_0(t)A + g_1(t)A^2 + \cdots + g_{n-1}(t)A^n$$
  - And by Cayley-Hamilton:
$$A^n = -q_1A^{n-1} - q_2A^{n-2} - \cdots - q_{n-1}A - q_nI_n$$
  - Plugging the previous two formulas together to get

$$\begin{aligned} &- q_n g_{n-1}(t)I_n + (-q_{n-1}g_{n-1}(t) + g_0(t))A \\ &\qquad + (-q_{n-2}g_{n-1}(t) + g_1(t))A^2 + \cdots + (-q_1 g_{n-1}(t) + g_{n-2}(t))A^{n-1} \end{aligned} \quad (2.1)$$

· Then, equating the two expressions for $\frac{d}{dt}e^{tA}$ term-by-term to get

$$\begin{aligned}\dot{g}_0(t) &= -q_n g_{n-1}(t)\\ \dot{g}_1(t) &= -q_n g_{n-1}(t) + g_0(t))\\ &\vdots\\ \dot{g}_{n-1}(t) &= -q_1 g_{n-1}(t) + g_{n-2}(t))\end{aligned}$$

· then set

$$g(t) = \begin{bmatrix}\dot{g}_0(t) & \dot{g}_1(t) & \cdots & \dot{g}_{n-1}(t)\end{bmatrix}^T$$

· which turns out to be a companion matrix of A, named A_c where $\dot{g}(t) = A_c g(t)$:

$$\dot{g}(t) = \begin{bmatrix} 0 & 0 & \cdots & 0 & -q_n \\ 1 & 0 & \cdots & 0 & -q_{n-1} \\ 0 & 1 & 0 & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & -q_2 \\ 0 & 0 & \cdots & 1 & -q_1 \end{bmatrix} g(t)$$

· This is called a companion form matrix. Turns out, by an easy induction, it's characteristic polynomial is the same as A's. ie $\lambda^n + q_1\lambda^{n-1} + q_2\lambda_{n-2} + \cdots + q_{n-1}\lambda + q_n$

· Initialize $\dot{g}(t) = A_c g(t)$ at $t = 0$ with

$$g(0) = \begin{bmatrix}1 & 0 & \cdots & 0\end{bmatrix}^T$$

· Then, solve for $g(t), t \in R$, and you get a set of coefficients $g(t)$ such that $e^{tA} = \sum_{k=0}^{N-1} g_k(t)A^k$

→ The above method has essentially reduced computation of e^{tA} to the computation of e^{tA_c} because $g(t) = e^{tA_c}g(0)$ for all t.

- In fact, all we need is the *1st column* of e^{tA_c} to get e^{tA} because of the form of $g(0)$

Next, we will breeze through a similar approach to find A^k with Cayley Hamilton

- Start with similar functions $g_0, \cdots, g_{n-1}$ of k:

$$A^k = g_0(k)I_n + g_1(k)A + \cdots + g_{n-1}(k)A^{n-1}$$

- Again, g's that work are not unique. An easy example is $A = I_n$

$$A^k = I_n^k \leftrightarrow \frac{1}{2}I_n + \frac{1}{2}A$$

- Find 1 ntuple that works as follows.

$$A^{k+1} = g_0(k+1)I_n + g_1(k+1)A + \cdots + g_{n-1}(k+1)A^{n-1}$$

- And meanwhile, by multiplying through by the original

$$A^{k+1} = g_0(k)A + g_1(k)A^2 + \cdots + g_{n-1}(k)A^n$$

- Plug in with $A^n = -q_1A^{n-1} - q_2A^{n-2} - \cdots - q_{n-1}A - q_nI_n$

  and set coefficients in the earlier equation of A^{k+1} in terms of $I_n, A, \cdots A^{n-1}$ to equal the above and get

$$g(k) = \begin{bmatrix} g_0(k) & \cdots & g_{n-1}(k) \end{bmatrix}^T$$

$\rightarrow$ Finding, whereA_c is the same as before,

$$g(k+1) = A_c g(k)|$$

Initialzing with $k = 0$,

$$g(0) = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}^T$$

then results for $g(k), k \geq 0$ works!

2.3 General Calculations of Functions of Matrices

2.3.1 Problem Formation

If $\hat{p}_1(s)$ and $\hat{p}_2(s)$ are two polynomials in s with:

$$\frac{\hat{p}_1(s)}{\hat{x}_A(s)} = \hat{q}_1(s) + \frac{\hat{r}_1(s)}{\hat{x}_A(s)}$$

$$\frac{\hat{p}_2(s)}{\hat{x}_A(s)} = \hat{q}_2(s) + \frac{\hat{r}_2(s)}{\hat{x}_A(s)}$$

Then if $\hat{p}_1(s) \neq \hat{p}_2(s)$, if $\hat{r}_1(s) \equiv \hat{r}_2(s)$, then we will gat that $\hat{p}_1(A) \neq \hat{p}_2(A)$!

Thus, every polynomial function of A can be written as a function of $I, A, A^2, \cdots A^{n-1}$.

How about non-polynomial functions, like $e^{At}, cos(A), log(sinA), or\sqrt{(}A)$. To understand the form of non-polynomial functions of A, we have to study the eigenvalues and eigenvectors of A. The first method comes when you have distinct eigenvalues and linearly independent eigenvectors, the second case is more generalizable.

A-Invariant Subspaces and the Second Representation Theorem

Consider a vector space $(V, \mathbb{R})$ and a linear map $A : V \mapsto V$. A subspace $M \subset V$ is said to be

A-Invariant, or invariant under A, if given $x \in M, Ax \in M$ (often written as $A[M] \subset M$). Below are some examples of A-invariant subspaces:

i. $N(A)$
ii. $R(A)$
iii. $N(A - \lambda_i I)$ where $\lambda_i \in \sigma(A)$
iv. if $P(A) - A^k + \alpha_1 A^{k-1} + \cdots + \alpha_{k-1} A + \alpha_k I$, then $N(p(A))$
v. let the subspaces M_1 and M_2 be A-invariant. Let $M_1 + M_2 := \{x \in V : x_1 + x_2, x_i \in M_i \; for \; i = 1, 2\}$. $M_1 \cap M_2$ and $M_1 + M_2$ are A-invariant.

Direct sum of Subspaces: We say that V is the direct sum of $M_1, M_2, \cdots, M_k$, denotes as $V = M_1 \oplus M_2 \oplus \cdots \oplus M_k$ if $\forall x \in V, \exists! x_i \in M_i, i = 1 \cdots k$, such that $x = x_1 + x_2 + \cdots + x_k$. Direct sum is the generalization of linear independence. Check for example that if $V = M_1 \oplus M_2 \oplus \cdots \oplus M_k$, then $M_i \cap M_j = \{\theta\}$.

Example: Let $A \in \mathbb{R}^{n \times n}$ have n *distinct* eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_n \in \mathbb{C}$. Then:

$$\mathbb{C}^n = N(A - \lambda_1 I) \oplus N(A - \lambda_2 I) \oplus \cdots \oplus N(A - \lambda_n I)$$

2nd Representation Theorem: Let $V = M_1 \oplus M_2$, with $dim(V) = n, dim(M_1) = k, dim(M_2) = n - k$, be a finite dimensional vector space. If M_1 is A-invariant, then A has a representation of the form:

$$A = \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline 0 & A_{22} \end{array}\right]$$

where, $A_{11} \in \mathbb{C}^{k \times k}, A_{12} \in \mathbb{C}^{k \times (n-k)}, A_{22} \in \mathbb{C}^{(n-k) \times (n-k)}$ are block matrices making up A.

2.3.2 Eigenstructure and Minimal Polynomial

We know that:

$$det(sI - A) = \hat{X}_A(s) \qquad (characteristic \; polynomial \; of \; A)$$

We can write:

$$\hat{X}_A(s) = (s - \lambda_1)^{d_1} (s - \lambda_2)^{d_2} \cdots (s - \lambda_\sigma)^{d_\sigma}$$

where $d_1, d_2, \cdots, d_\sigma$ are the multiplicities of $\lambda_1, \lambda_2, \cdots, \lambda_\sigma \in \mathbb{C}$, where $d_1 + d_2 + \cdots + d_\sigma = n$. By the Cayley-Hamilton Theorem, we know that $\hat{X}_A(A) = \theta_{n \times n}$. Let $\hat{\xi}_A(s)$ be the polynomial of least degree such that $\hat{\xi}_A(A) = \theta_{n \times n}$. $\hat{\xi}_A(s)$ divides $\hat{X}_A(s)$, and is called the minimal polynomial of A:

$$\hat{\xi}_A(s) = (s - \lambda_1)^{m_1} \cdots (s - \lambda_\sigma)^{m_\sigma}$$

with $m_1 \leq d_1, m_2 \leq d_2, \cdots m_\sigma \leq d_\sigma$. Below are a couple examples:

$$1)\ A = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{bmatrix} \qquad \begin{aligned} \hat{X}_A(s) &= (s-\lambda_1)^2(s-\lambda_2) \\ \hat{\xi}_A(s) &= (s-\lambda_1)(s-\lambda_2) \end{aligned}$$

$$2)\ A = \begin{bmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_1 \end{bmatrix} \qquad \begin{aligned} \hat{X}_A(s) &= (s-\lambda_1)^3 \\ \hat{\xi}_A(s) &= (s-\lambda_1)^2 \end{aligned}$$

Theorem (proof in textbook)

$$\mathbb{C}^n = N(A-\lambda_1 I)^{m_1} \oplus N(A-\lambda_2 I)^{m_2} \oplus \cdots \oplus N(A-\lambda_\sigma I)^{m_\sigma}$$

Fact: $DimN(A-\lambda_i I)^{m_i} = d_i$

Geometric Structure of Eigenspaces: Consider

$$\begin{aligned} \hat{X}_A(s) &= (s-\lambda_1)^{d_1} \qquad & d_1 &= n \\ \hat{\xi}_A(s) &= (s-\lambda_1)^{m_1} \qquad & m_1 &\geq 1 \end{aligned}$$

for this example, let us say that $m_1 = 3, n = 6$

Let,

$N(A-\lambda I) = Sp\{e_1, e_2, e_3\}$

$N(A-\lambda I)^2 = Sp\{e_1, e_2, e_3, v_1, v_2\} \supset N(A-\lambda I)$

ie $N(A-\lambda I)^2 = N(A-\lambda I) \oplus Sp\{n_1, v_2\}$

Then, we set v_1 and v_2 to be the independent solutions of

$(A-\lambda I)x = e_1$

$(A-\lambda I)x = e_2$

$(A-\lambda I)x = e_3$

Since $R(A-\lambda I) \not\subseteq \mathbb{C}^n$, the equations may not have solutions for all $e_i, i = 1, 2, 3$. Let us say there are solutions for e_1 and e_2; then:

$(A-\lambda I)v_1 = e_1$

$(A-\lambda I)v_2 = e_2$

In summary, the eigenspaces of lower powers of $N(A-\lambda I)^l$ are subsets of higher powers, that all map to the zero vector.

Eigenvector and Generalized Eigenvector Chairs

$$Sp\{e_1, v_1, w_1\} \oplus Sp\{e_2, v_2\} \oplus Sp\{e_3\} = \mathbb{R}^6$$

Thus, the representation of A with respect to $\{e_1, v_1, w_1, e_2, v_2, e_3\}$ is

$$\left[\begin{array}{ccc|cc|c} \lambda & 1 & 0 & & & \\ 0 & \lambda & 1 & & & \\ 0 & 0 & \lambda & & & \\ \hline & & & \lambda & 1 & \\ & & & 0 & \lambda & \\ \hline & & & & & \lambda \end{array}\right] = TAT^{-1}, \qquad T^{-1} = \left[\begin{array}{cccccc} | & | & | & | & | & | \\ e_1 & v_1 & w_1 & e_2 & v_2 & e_3 \\ | & | & | & | & | & | \end{array}\right]$$

And,

$$\begin{array}{lll} Ae_1 = \lambda e_1 & Av_1 = \lambda v_1 + e_1 & Aw_1 = \lambda w_1 + v_1 \\ Ae_2 = \lambda e_2 & Av_2 = \lambda v_2 + e_2 & \\ Ae_3 = \lambda e_3 & & \end{array}$$

2.3.3 Functions of a Matrix

Definition: Let $\hat{f}(s)$ be any function of s analytic on the spectrum of A and $\hat{p}(s)$ be a polynomial such that:

$$\hat{f}^k(\lambda_e) = \hat{p}^k(\lambda_e) \qquad 0 \le k \le m_{e-1} \qquad 1 \le e \le \sigma$$

Then,

$$\hat{f}(A) = \hat{p}(A)$$

Aside: remember that an analytic function is a function that is locally given by a convergent power series, so it is differentiable on the order proportional to the number of eigenvalues with component analysis.

In fact, if $m := \sum_{i=1}^{\sigma} m_i$ then,

$$\hat{p}(s) = a_1 s^{m-1} + a_2 s^{m-2} + \cdots + a_m s^0$$

where $a_1, a_2, \cdots, a_n$ are functions of $\hat{f}(\lambda_1), \hat{f}'(\lambda_1), \hat{f}^2(\lambda_1), \cdots, \hat{f}^{m_1}(\lambda_1), \hat{f}(\lambda_2), \cdots$ hence,

$$\hat{f}(A) = a_1 A^{m-1} + \cdots + a_m A^0 = \sum_{l=1}^{\sigma} \sum_{k=0}^{m_{e-1}} p_{kl}(A) \cdot f^k(\lambda_l)$$

Functions of a Matrix (Distinct Eigenvalues)

When a matrix $A \in \mathbb{R}^{n \times n}$ has n distinct eigenvalues, that means that the matrix is diagonalizable

through the known similarity transforms from linearly independent eigenvectors.

$$\Sigma = \begin{bmatrix} \lambda_1 & & & & & \\ & \lambda_2 & & & & \\ & & \ddots & & & \\ & & & \ddots & & \\ & & & & \lambda_{n-1} & \\ & & & & & \lambda_n \end{bmatrix} = TAT^{-1}, \qquad T^{-1} = \begin{bmatrix} | & | & & & | & | \\ e_1 & e_2 & \cdots & \cdots & e_{n-1} & e_n \\ | & | & & & | & | \end{bmatrix}$$

Which gives us a function of a matrix becomes the function below by writing any function as a series of power series and powers of A with Cayley-Hamilton.

$$f(A) = T^{-1} f(\Sigma) T$$

Functions of a Matrix (Repeated Eigenvalues)

To begin, consider

$$J \in \mathbb{R}^{n\times n} = \begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \cdots & \cdots & 0 & \lambda \end{bmatrix}$$

Claim

$$f(J) \in \mathbb{R}^{n\times n} = \begin{bmatrix} f(\lambda) & f'(\lambda) & \cdots & \cdots & \frac{f^{(n-1)}(\lambda)}{(n-1)!} \\ 0 & f(\lambda) & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & f'(\lambda) \\ 0 & \cdots & \cdots & 0 & f(\lambda) \end{bmatrix}$$

Proof:

- Start with the min polynomial $= (s-\lambda)^n$

  Thus, $f(J) = \sum_{l=0}^{n-1} f^{(l)} p_l(J)$

- Choose

$$\begin{cases} f_1(s) = 1 \;\Rightarrow f_1(J) = I = f_1^{(0)}(\lambda) p_0(J) & \Rightarrow p_0(J) = I \\ f_2(s) = s - \lambda \;\Rightarrow f_2(J) = J - \lambda I = f_2^{(1)}(\lambda) p_1(J) & \Rightarrow p_1(J) = J - \lambda I \\ f_3(s) = (s-\lambda)^2 \;\Rightarrow f_3(J) = (J-\lambda I)^2 = f_3^{(2)}(\lambda) p_2(J) & \Rightarrow 2p_2(J) = (J-\lambda I)^2 \end{cases}$$

$\Rightarrow$

$$p_0(J)I, \quad P_1(J) = J - \lambda I, \quad p_2(J) = \frac{1}{2}(J - \lambda I)^2$$

Thus, our claim is satisfied! This gives us the Spectral Mapping Theorem:

$$\sigma\big(f(J)\big) = f\big(\sigma(J)\big)\{f(\lambda_1), \cdots, f(\lambda_n)\}$$

and more generally,

$$\sigma\big(f(A)\big) = f\big(\sigma(A)\big)$$

Chapter 3

Working With State Systems

Next, we will focus on some systems type questions about our state space equations (I) and (II), assuming matrices are constant (time independent).

$$(I) \quad \begin{aligned} \dot{x}(t) &= A(t)x(t) + B(t)u(t) \\ y(t) &= C(t)x(t) + D(t)u(t) \end{aligned} \qquad (II) \quad \begin{aligned} x(k+1) &= A(k)x(k) + B(k)u(k) \\ y(k) &= C(k)x(k) + D(k)u(k) \end{aligned}$$

$$where,\ A \in \mathbb{R}^{n\times n},\ B \in \mathbb{R}^{n\times m},\ C \in \mathbb{R}^{p\times n},\ D \in \mathbb{R}^{p\times m}$$

3.1 Discrete Time

3.1.1 Reachability

First, look about questions about linkage between inputs and states - focus on the differential equation $x(k+1) = A(k)x(k) + B(k)u(k)$.

Definition: say $x_1 \in \mathbb{R}^n$ is a reachable state - or reachable w.r.t. (A, B) - when you can steer the state from zero at time 0 to x in finite time by a suitable choice of inputs. More formally put, x_1 is reachable w.r.t. (A, B) when there exists some $k > 0$ and $u(0), u(1), \cdots, u(k_1 - 1) \in \mathbb{R}^m$ so that $x(k_1) = x_1$ when $x(0) = 0$, and you apply the inputs $u(0), u(1), \cdots, u(k_1-1)$ during $0 \leq k \leq k_1-1$.

Question: Is the set of all states reachable w.r.t (A, B) a subspace of $\mathbb{R}^n$. Turns out this is a tricky question because, say you can reach x_1 in time 7 and you can reach x_2 in time 13, is it obvious that there exists some $k_1 > 0$ so that you can reach $x_1 + x_2$ in time k_1?.

Answer: Yes! It is a subspace!

- First, say you can reach x_1 in time k_1 and you can reach x_2 in time k_2, and $k_2 > k_1$. Suppose,
  - $u(0), u(1), \cdots, u(k_1 - 1)$ gets to x_1 at time k_1,
  - $v(0), v(1), \cdots, v(k_1 - 1)$ gets to x_2 at time k_2,

· To get to $c_1x_1 + c_2x_2$ at time k_2, you proceed as follows:

1. Start at time 0 applying $c_2v(0), c_2v(1), \cdots, c_2v(k_1 - 1)$,
2. Chill for a while about reaching x_1.
3. When you get to time $k_2 - k_1$, start applying in addition $c_1u(0), c_1u(1)$, etc, applying

$$w(k) = \begin{cases} c_2v(k) & 0 \le k \le k_2 - k_1 \\ c_1u(k - k_2 - k_1) + c_2v(k) & k_2 - k_1 \le k \le k_2 - 1 \end{cases}$$

Caution: You can't just apply $c_1u(k) + c_2v(k)$ starting from time 0 through time $k = k_1 - 1$, and then apply $c_2v(k)$ for $k_1 \le k \le k_2 - 1$. Why? To see this, proceed as described, then,

$$x(k) = \sum_{l=0}^{k-1} A^{k-l-1}B\big(c_1u(l)\big) + \sum_{l=0}^{k-1} A^{k-l-1}B\big(c_2v(l)\big) \qquad 0 \le k \le k_1$$

So, at $k = k_1$, first term will be

$$\sum_{l=0}^{k_1-1} A^{k-l-1}B\big(c_1u(l)\big) = c_1x_1$$

Thus,

$$x(k_1) = c_1x_1 + \sum_{l=0}^{k_1-1} A^{k-l-1}B\big(c_2v(l)\big)$$

Apply "c_2v" after time k_1, then

$$x(k_2) = c_1A^{k_2-k_1}x_1 + \sum_{l=0}^{k_2-1} A^{k-l-1}B\big(c_2v(l)\big)$$

$$= c_1A^{k_2-k_1}x_1 + c_2x_1$$

This is clearly off by a factor of $A^{k_2-k_1}$

Check: confirm that the proposed input w works as prescribed:

- Know to start with

$$x_1 = \sum_{l=0}^{k_1-1} A^{k-l-1}Bu(l) \qquad x_2 = \sum_{l=0}^{k_2-1} A^{k-l-1}Bv(l)$$

- Lets apply w starting from $x(0) = 0$, and we see what $x(k_2)$ is:

$$x(k_2) = \sum_{l=0}^{k-1} A^{k-l-1}Bv(l)$$

$$= \sum_{l=0}^{k_2-1} A^{k_2-l-1} B c_2 v(l) + \sum_{l=k_2-K-1}^{k_2-1} A^{k_2-l-1} B c_1 u(l-(k_2-k_1))$$

First term because you're applying $c_2 v$ the whole time

And, the second term because you're applying $c_1 U$ starting at time $(k_2 - k_1)$

- Change index of summation in second term to $m = l - (k_2 - k_1)$, and the second term becomes

$$\sum_{m=0}^{k_1-1} A^{k_1-m-1} B c_1 u(m)$$

Bottom Line: $\{x \in \mathbb{R}^n$ *reachable w.r.t.* $(A, B)\}$ is a subspace of $\mathbb{R}^n$ closed under linear combinations. Remember the two cases for a subspace were closed under linear combinations and containing the zero vector. Recall that a subspace $\mathbb{V}$ of $\mathbb{R}^n$ is <u>invariant under A</u> when $Av \in \mathbb{V}$ for every $v \in \mathbb{V}$ - ie A maps v into V.

Fact: $\{x \in \mathbb{R}^n$ *reachable w.r.t.* $(A, B)\}$ is a subspace of $\mathbb{R}^n$ *invariant under A*

To see the invariance part of this statement, proceed as follows. Suppose you can reach x_1 in finite time, how do you reach Ax_1 in finite time?

- Reach x_1, say in time k_1
- Apply the zero input to the system at time k_1
- Then you have reached Ax_1 at time $k_1 + 1$!.

Next question, how *Long* does it take to reach a state reachable w.r.t. (A, B)?

Fact: When $A \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{n\times m}$, and $x_1 \in \mathbb{R}^n$ is reachable w.r.t. (A, B), turns out you can reach x_1 in time at most n.

Why? To see this:

- Suppose you can reach x_1 in time k_1
- This, you can find $u(0), u(1), \cdots, u(k_1 - 1)$ so that $x_1 = \sum_{l=0}^{k_1-1} A^{k-l-1} B u(l)$

1. If $k_1 \leq n$ we're done - ie we've reached x_1 in time $\leq n$
2. If $K_1 > n$: recall from Cayley Hamilton Theorem that all high powers of A can be written as linear combinations of $I_n, A, A^2, \cdots, A^{n-1}$. Thus can re-write the sum above as below, where $v(l)$ is some linear combination of $u(0), u(1), \cdots, u(k_1 - 1)$ for all $0 \leq l \leq n - 1$

$$x_1 = \sum_{l=0}^{n-1} A^{n-l-1} B v(l)$$

- Hence, applying $v(l)$ from 0 to $l = m - 1$ gets us from 0 to x_1 in time n.

→ One additional <u>refinement</u>: in this context, if x_1 is reachable w.r.t. (A, B) then you can reach x_1 in time exactly n.

- To see why: Second part of argument above shows how to reach x_1 in time exactly n

when you can reach x_1 in time $k_1 > n$.

- If you can reach x_1 in time $k_1 < n$, how do you reach it in exactly n?
  1. Chilling for a while starting at time 0
  2. At time $n - k_1$, start applying the inputs that take you from 0 to x_1 in time k_1 - result will be arrival at x_1 at exactly time n

Point-to-Point Reachability: If $x_1, x_2 \in \mathbb{R}^n$ are reachable w.r.t. (A, B), then you can steer the state from $x(0) = x_1$ to the state x_2 in finite time by suitable choice of u.

Why?

- x_1 reachable $\Rightarrow$ $A^k x_1$ reachable for all $k \geq 0$; in particular $A^N x_1$ reachable because of invariance.
- Because reachable states for a subspace, $x_2 - A^n x_1$ is reachable
- So we can reach this state in time exactly n by choice of $u(0), u(1), \cdots, u(n-1)$, so

$$x_2 - A^n x_1 = \sum_{l=0}^{n-1} A^{n-l-1} B u(l)$$

$$x_2 = A^n x_1 + \sum_{l=0}^{n-1} A^{n-l-1} B u(l)$$

- Above equation says, when $x(0) = x_1$, and you apply those inputs, $x(n) = x_2$.

Let's characterize { Reachable States } algebraically!

- Define this (n x mn) matrix:

$$Q_r(A, B) = \left[\begin{array}{c|c|c|c} B & AB & \cdots & A^{n-1}B \end{array}\right]$$

$Q_r(A, B)$ is called the reachability matrix associated with pair (A,B).

- **Claim**: $\{reachable\ states\} = range(Q_r(A, B))$ (Reminder: range of a matrix is defined as set of all linear combinations of columns and the set of all things that you can right multiply the matrix by and get).
- **Reason**:
  1. First, suppose $x_1 \in range(Q_r)$; then we can find a $v \in \mathbb{R}^{nm}$ such that $x_1 = Q_r(A, B) \cdot v$. Parse v into m dimensional subvectors as follows:

$$v = \left[\begin{array}{c|c|c|c} v_{n-1} & \cdots & v_1 & v_0 \end{array}\right]^T$$

Then,

$$x_1 = Q_r(A,B)\cdot v = \left[\begin{array}{c|c|c|c} B & AB & \cdots & A^{n-1}B \end{array}\right]\left[\begin{array}{c|c|c|c} v_{n-1} & \cdots & v_1 & v_0 \end{array}\right]^T$$

$$x_1 = Bv_{n-1} + ABv_{n-2} + \cdots + A^{n-1}Bv_0$$

Thus, if we set $u(0) = v_0, u(1) = v_1, \cdots, u(n-1) = v_{n-1}$, we get

$$x_1 = \sum_{l=0}^{n-1} A^{n-l-1}Bu(l)$$

Implying x_1 is reachable

Finally, everything in $range(Q_r) \subset \{reachable\ states\}$

2. Next, consider the converse. If $x_1 \in \mathbb{R}^n$ is reachable, we can find $u(0), u(1), \cdots, u(n-1)$ so that

$$x_1 = \sum_{l=0}^{n-1} A^{n-l-1}Bu(l)$$

Then form $v \in \mathbb{R}^{nm}$

$$v = \left[\begin{array}{c|c|c|c} u(n-1) & \cdots & u(1) & u(0) \end{array}\right]^T$$

Then $x_1 = Q_r(A,B)$ and $x_1 \in range(Q_r)$ and $range(Q_r) \supset \{reachable\ states\}$

$\Rightarrow \{reachable\ states\} = range(Q_r(A,B)$

Terminology: say (A,B) is a reachable pair when every $x_1 \in \mathbb{R}^n$ is reachable w.r.t. (A,B), so all of the following are true

- $\{reachable\ states\} = \mathbb{R}^n$
- $range(Q_r(A,B)) = \mathbb{R}^n$
- $Q_r(A,B)$ has full rank n (maximum rank)

→ **Question**: Given n, m and $A \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{n\times m}$, how *generic* is reachability of the pair (A,B)?

- · Note: This is same question as "how generic is it for $rank(Q_r(A,B)) = n$?"
- · Note Also: $Q_r(A,B) \in \mathbb{R}^{n\times n}$ when $m = 1$. $Q_r(A,B)$ has way more columns when m is large.

Answer: for all m, n reachability of (A, B) is generic, both

- Topologically: given a pair (A_0, B_0), you can perturb the entries in (A_0, B_0) by an arbitrarily small amount and obtain a reachable pair. More formally: the set of reachable pairs is *dense* in (A, B) space.
- Probabilisticially: if you have a smooth density function on (A, B) space, then a randomly drawn pair (A, B) will be reachable with probability 1.

3.1.2 Controllability

Controllability: This is a concept related to reachability.

- Say $x_0 \in \mathbb{R}^n$ is controllable w.r.t. pair (A, B) when you can drive the state from $x(0) = x_0$ to zero in finite time by a suitable choice of u.
- More formally: $x_0 \in \mathbb{R}^n$ is controllable w.r.t. (A, B) when there exists $k_1 > 0$ and $u(0), u(1), \cdots, u(k_1 - 1)$ so that

$$0 = A^{k_1} x_0 + \sum_{l=0}^{k_1-1} A^{k_1-l-1} B u(l)$$

- Observe: By point-to-point reachability, every reachable state is also controllable (because it is the case where 0 is the other point, and 0 sis reachable).
- The reverse conclusion is false, ie *controllability* $\not\Longrightarrow$ *reachability*

  Example Given:

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$Q_r(A, B) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

So, only reachable states have 0 in last two positions. But, every $x_0 \in \mathbb{R}^n$ is controllable - because $A^3 = 0_3$. Therefore, can 'control' any $x_0 \in \mathbb{R}^n$ to zero in time 3 just by chilling.

Fact: $\{$*controllable states*$\}$ is a subspace of $\mathbb{R}^n$

This is a tricky argument because say x_0 and x_0' are both controllable, so

- x_0 'controllable to 0' in time k_1
- x_0' 'controllable to 0' in time k_1'

- so, it's not obvious that given c_1, c_2, we can find a single time k_2 and choice of inputs on $0, 1, \cdots, k_2 - 1$ so that $c_1 x_0 + c_2 x_0'$ is controlled to 0 in time k_2
- Resolve this using a chill-for-a-while argument, and this is a 'all back end' version:

To see this, suppose x_0, x_0' are controllable where $u(0), \cdots, u(k_1-1)$ drives x_0 to 0 in time k_1 and $v(0), \cdots, v(k_2-1)$ drives x_0' to 0 in time $k_2 \geq k_1$. Drive $c_1x_0 + c_2x_0'$ to 0 in time k_2 as follows:

- On time interval $0 \leq k < k_1$, apply input $c_1u(k) + c_2v(k)$
- On time interval $k_1 \leq k < k_2$ apply input $c_2v(k)$
- Then it turns out, with $x(0) = c_1x_0 + c_2x_0'$, and inputs above, you get $x(k_2) = 0$

  Idea: Find that

$$x(k_2) = A^{k_2-k_1}\big(A^{k_1}c_1x_0 + c_1\sum_{l=0}^{k_1-1}A^{k_1-l-1}Bu(l)\big) + \big[A^{k_2}c_2x_0' + c_2\sum_{l=0}^{k_2-1}A^{k_2-l-1}Bv(l)\big]$$

  · where $\big(\quad\big) = 0$ because c_1u drives c_1x_0 to 0 in time k_1

  · and $\big[\quad\big] = 0$ because c_2v drives c_1x_0' to 0 in time k_2

→ $c_1x_0 + c_2x_0'$ is also controllable - so any linear combination of controllable states is also controllable, or $\{$*controllable states*$\}$ is a subspace of $\mathbb{R}^n$

Noted Previously

$\{$*reachable states*$\} \subset \{$*controllable states*$\}$, where the reachable states are also invariant under A

- Observe: when x_0 is controllable, A^kx_0 is reachable for some $k_1 > 0$
- Idea: if $u(0), \cdots, u(k_1-1)$ drive x_0 to 0 in time k_1, then

$$0 = A^{k_1}x_0 + \sum_{l=0}^{k_1-1}A^{k_1-l-1}Bu(l) \qquad \rightarrow \qquad A^{k_1}x_0 = \sum_{l=0}^{k_1-1}A^{k_1-l-1}B(-u(l))$$

  Showing us that $A^{k_1}x_0$ is also reachable for some $k_1 > 0$

Lin Alg Fact: If $A \in \mathbb{R}^{n\times n}$ and $\mathbb{W}$ is a subspace of $\mathbb{R}^n$ invariant under A, and $A^{k_1}v \in \mathbb{W}$ for some $k_1 \geq 0$, then $A^nv \in \mathbb{W}$

Idea: let k^* be the smallest k_1 such that $A^{k_1} \in \mathbb{W}$

- Suppose $k^* > n$, we will search for a contradiction to support our claim
- Note: $v, Av, A^2v, \cdots, A^nv$ are linearly dependent, so we can find constants c_k not all zero so that

$$\star \quad c_0v + c_1Av + c_2A^2v + \cdots + c_nA^nv = 0$$

- Multiplying this by A^{k^*-1} and get

$$c_0A^{k^*-1}v + \big(c_1A^{k^*}v + c_2A^{k^*+1}v + \cdots + c_nA^{k^*+n-1}v\big) = 0$$

  Where all of the terms in $\big(\quad\big)$ are in $\mathbb{W}$ because $A^{k^*}v \in \mathbb{W}$ and $\mathbb{W}$ is invariant under A

→ by choice of k^*, a k^*-1 v is not in $\mathbb{W}$, thus $c_0 = 0$. This is because k^*-1 would be less than k_1 which was defined as the smallest.

Since otherwise, could express $A^{k^*-1}v$ as linear combinations of vectors in $\mathbb{W}$. Now go back to the original expression $\star$ between $v, Av, A^2v, \cdots, A^nv$ - multiply by A^{k^*-2} - conclude $c_1 = 0, c_2 = 0, \cdots, c_n = 0$ - which is a contradiction because $c_k's$ are not all 0 by design!

Let's apply this result to our controllability situation! x_0 is controllable w.r.t. (A,B) $\rightarrow$ $A^{k_1}x_0$ is reachable w.r.t. (A,B) for some $k_1 > 0$.

$\rightarrow$ A^nx_0 is reachable w.r.t. (A,B) (so is $-A^nx_0$)

$\rightarrow$ can 'reach' $-A^nx_0$ in time exactly n, by choice of u

$\rightarrow$ can find $u(0), \cdots, u(n-1)$ such that

$$-A^nx_0 = \sum_{l=0}^{n-1} A^{n-l-1}Bu(l) \qquad or \qquad 0 = A^nx_0 + \sum_{l=0}^{n-1} A^{n-l-1}Bu(l)$$

Thus, if x_0 is controllable, we can 'control it to zero' in time exactly n (or $\leq$ n in fact)

Fact: $\{controllable\ states\}$ is invariant under A

Idea: x_0 is controllable leads to

$\rightarrow$ A^nx_0 is reachable

$\rightarrow$ $A^n(Ax_0)$ is also reachable by invariance of $\{reachable\ states\}$

$\rightarrow$ Ax_0 is controllable

3.1.3 Summary

= $\{reachable\ states\}$ and $\{controllable\ states\}$ are both subspaces of $\mathbb{R}^n$ invariant under A

= $\{reachable\ states\} = range(Q_r(A,B))$, where $Q_r(A,B) = \begin{bmatrix} B & AB & \cdots & A^{n-1}B \end{bmatrix} \in \mathbb{R}^{n \times nm}$

= $\{reachable\ states\} \subset \{controllable\ states\}$

= (A,B) is a reachable pair when every $x_1 \in \mathbb{R}^n$ is reachable w.r.t. (A,B) (same as saying the $rank(Q_r(A,B) = n$ or is full rank)

= (A,B) is a controllable pair when every $x_0 \in \mathbb{R}^n$ is controllable w.r.t. (A,B)

= By containment relation, (A,B) reachable $\Rightarrow$ (A,B) controllable; (the converse is not in general true)

= Reachability of a pair (A,B), and hence controllability of (A,B) is generic both topologically and probabilistically.

= x_1 reachable $\Rightarrow$ can 'reach' it in time n

= x_0 controllable $\Rightarrow$ can 'control' it to 0 in time n

3.2 Continuous Time

$$(I) \quad \begin{aligned} \dot{x}(t) &= A(t)x(t) + B(t)u(t) \\ y(t) &= C(t)x(t) + D(t)u(t) \end{aligned} \qquad x \in \mathbb{R}^n, u \in \mathbb{R}^m$$

3.2.1 Reachability

Say $x_1 \in \mathbb{R}^n$ is reachable w.r.t (A,B) when you can drive the state from $x(0) = 0$ to x_1 in finite time by suitable choice of u. More formally, $x_1 \in \mathbb{R}^n$ is reachable w.r.t. (A,B) when there exists $t_1 > 0$ and a choice of input function $u : [0, t_1] \to \mathbb{R}^m$, (Require u to be 'reasonable' enough for the differential equation to work) such that when $x(0) = 0$, and you apply u on $[0, t_1]$, you get x_1.

$$x_1 = \int_0^{t_1} e^{(t_1-\tau)A} B u(\tau) d\tau$$

Fact: $\{reachable\ states\}$ is a subspace of $\mathbb{R}^n$. The subtlety of this statement is the same as in discrete time - what happens to the extra terms for the term that takes less time to reach.

say x_1 is reachable in time t_1, using input $u : [0, t_1] \to \mathbb{R}^m$

say x_2 is reachable in time t_2, using input $v : [0, t_2] \to \mathbb{R}^m$

$\to$ Reach $c_1x_1 + c_2x_2$ in time t_2 as follows: Apply the function

$$= \begin{array}{cc} c_2 v(t) & 0 \le t \le t_2 - t_1 \\ c_1 u(t - (t_2 - t_1)) + c_2 v(t) & t_2 - t_1 \le t \le t_2 \end{array}$$

By plugging this in, you discover

$$x(t_2) = c_2 \int_0^{t_2} e^{(t_2-\tau)A} B u(\tau) d\tau + c_1 \int_{t_2-t_1}^{t_2} e^{(t_2-\tau)A} B u(t - (t_2 - t_1)) d\tau$$

The first term clearly $= c_2x_2$ by definition.

As for the second term, let $\eta = \tau - (t_2 - t_1) \to = c_1 \int_0^{t_1} e^{(t_1-\eta)A} B u(\eta) d\eta$

$\Rightarrow$ $x(t_2) = c_1x_1 + c_2x_2$. So, $c_1x_1 + c_2x_2$ is also reachable.

Aside: Suppose $A \in \mathbb{R}^{n\times n}$ and $W \in \mathbb{R}^n$ is subspace of $\mathbb{R}^n$ that is invariant under A. Thus, W is also invariant under e^{tA} for all $t \in \mathbb{R}$.

- **Idea**: Can write $e^{tA} = \sum_{l=0}^{n-1} g_l(t) \cdot A^l$ for some function $t \to g_l(t)$ by Cayley-Hamilton Theorem. Then, given $v \in W$, $e^{tA}v = \sum_{l=0}^{n-1} g_l(t) \cdot A^l v$. This is because W is invariant under A, all of $Av, A^2v, \cdots, A^lv \in W$ as well. Thus, so is $e^{tA}v$ for any t because W is a subspace.
- In fact, converse is true - ie if $W \in \mathbb{R}^n$ is a subspace invariant under e^{tA} for all $t \in \mathbb{R}$, then W is also invariant under A.
- To see this:
  - Suppose $v \in W$. Then $e^{tA}v \in W$ for all $t \ge 0$. Thus, $e^{tA}v - v \in W$ for all $t \ge 0$.

Thus again, for all $t > 0$, $\frac{1}{t}(e^{tA} - I)v \in W$.

· To end, take the $\lim_{t\to 0}$, giving

$$\lim_{t\to 0} \frac{1}{t}(e^{tA} - I) = \frac{d}{dt}(e^{tA})\Big|_{t=0} = Ae^{tA}\Big|_{t=0} = A$$

- Bottom Line: $lim_{t\to 0}$ gives $Av \in W$

Fact the set $\{reachable\ states\}$ is invariant under A

· **Idea**: Suppose x_1 is reachable. Then we can find t_1 and u such that

$$x_1 = \int_0^{t_1} e^{(t_1-\tau)A} Bu(\tau)d\tau$$

Apply u strategy from 0; reach x_1 at time t_1; then chill for any amount of time $h > 0$; get to $e^{hA}x_1$. Thus for any $h \geq 0, e^{hA}x_1$ is also reachable. So $\{reachables\}$ is invariant under e^{tA} for all $h \geq 0$. Thus by aside, invariant under A itself.

· Next, characterize $\{reachables\ states\}$ exactly

First observe, if we set

$$Q_r(A, B) = \begin{bmatrix} B & AB & \cdots & A^{n-1}B \end{bmatrix} \in \mathbb{R}^{n\times mn}$$

Then find,

$$\{reachables\ states\} \subset range(Q_r(A, B))$$

Idea: x_1 reachable. For some t_1 and u.

$$x_1 = \int_0^{t_1} e^{(t_1-\tau)A} Bu(\tau)d\tau$$

And applying Cayley-Hamilton Theorem $e^{tA} = \sum_{l=0}^{n-1} g_l(t) \cdot A^l$ for some g_l's,

$$x_1 = \int_0^{t_1} \sum_{l=0}^{n-1} g_l(t_1 - \tau) \cdot A^l Bu(\tau)d\tau$$

$$= \sum_{l=0}^{n-1} A^l B \int_0^{t_1} \sum_{l=0}^{n-1} g_l(t_1 - \tau)u(\tau)d\tau$$

$$= Q_r(A, B) \cdot w$$

$$w = \Big[\int_0^{t_1} g_0(t_1 - \tau)u(\tau)d\tau \ \Big|\ \int_0^{t_1} g_1(t_1 - \tau)u(\tau)d\tau \ \Big|\ \cdots \ \Big|\ \int_0^{t_1} g_{n-1}(t_1 - \tau)u(\tau)d\tau \Big]^T$$

$\Rightarrow$ Thus, x_1 reachable $\rightarrow x_1 \in range(Q_r(A,B))$

Now introduce, given $t_1 > 0$, the matrix

$$M_r(t_1; A, B) = \int_0^{t_1} e^{(t_1-\tau)A} BB^T e^{(t_1-tau)A^T} d\tau$$

Note: $M_r \in \mathbb{R}^{n\times n}$ and $M_r^T = M_r$.
$M_r(t_1; A, B)$ is called the t_1-reachability Grammian associated with (A,B)

Claim for every y $t_1 > 0$ $range(M_r(t_1; A, B)) \subset \{reachable\ states\}$
- To see why: suppose $x_1 \in range(M_r)$, then we can find a $v \in \mathbb{R}^n$ such that

$$x_1 = M_r(t_1; A, B) \cdot v$$

$$= \Big(\int_0^{t_1} e^{(t_1-\tau)A} BB^T e^{(t_1-tau)A^T} d\tau \Big) \cdot v$$

$$= \int_0^{t_1} e^{(t_1-\tau)A} B\Big(B^T e^{(t_1-tau)A^T} \cdot v\Big) d\tau$$

$$= \int_0^{t_1} e^{(t_1-\tau)A} Bu(\tau) d\tau$$

when $u(\tau) = B^T e^{(t_1-tau)A^T} \cdot v$ for $\tau \in [0, t_1]$

$\rightarrow$ x_1 is reachable w.r.t. (A,B)

So far, for every $t_1 > 0$ we have

$$range(M_r(t_1; A, B)) \subset \{reachable\ states\} \subset range(Q_r(A,B))$$

Turns Out: all three subspaces are equal for all $t_1 > 0$! How does one see this? Show that $rank(M_r(t_1; A, B))$ is, for every $t_1 > 0$, the same as $rank(Q_r(A,B))$, and since $rank = dim(range)$ and $range(M_r) \subset range(Q_r)$, all three must be equal.

Equality of ranks follows from:

$$v \in nullspace(M_r(t_1; A, B))\ iff\ v \in nullspace(Q_r^T(A,B)) \star$$

Because

$$rank(Q_r(A,B)) = n - dim(nullspace(Q_r^T))$$

$$= n - dim(nullspace(M_r))$$

$$= rank(M_r)$$

To show that $\star$ is true,

- $v \in nullspace(M_r(t_1; A, B))$
  - $\rightarrow$ $M_r(t_1; A, B)v = 0$
  - $\rightarrow$ $v^T M_r(t_1; A, B)v = 0$
  - $\rightarrow$ $\int_0^{t_1} v^T e^{(t_1-\tau)A} BB^T e^{(t_1-\tau)A^T} v d\tau = 0$
  - $\rightarrow$ $\int_0^{t_1} \left\| B^T e^{(t_1-\tau)A^T} v \right\|^2 d\tau = 0$
  - $\rightarrow$ $B^T e^{(t_1-\tau)A} v = 0$ all $\tau \in [0, t_1]$ and finally,
  - $\rightarrow$ $B^T v = 0$ by evaluating at $\tau = t$, and $B^T A^T v = 0$ by taking the derivative with respect to t and evaluating at $\tau = t$, and $\cdots$, and $B^T (A^T)^{n-1} v = 0$ by iterated derivatives.
- Thus, $v \in nullspace(M_r(t_1; A, B)) \Rightarrow v \in nullspace(Q_r^T(A, B))$.
- Conversely, if $v \in nullspace(Q_r^T)$, then, $B^T v = 0, B^T A^T v = 0, \cdots, B^T (A^T)^{n-1} v = 0$. So, by Cayley Hamilton Theorem, $B^T (A^T)^k v = 0$
- Thus, $B^T e^{(t_1-\tau)A^T} v = 0$ for all $\tau \in [0, t_1]$ and therefore,

$$e^{(t_1-\tau)A} BB^T e^{(t_1-\tau)A^T} v = 0 \qquad \rightarrow \qquad \int_0^{t_1} e^{(t_1-\tau)A} BB^T e^{(t_1-\tau)A^T} v d\tau = 0 \; all \; \tau$$

- Which is the same as

$$M_r(t_1; A, B)v = 0$$

thus,

$$v \in nullspace(Q_r(A, B)) \Rightarrow v \in nullspace(M_r(t_1; A, B)) \qquad all \; t_1 > 0$$

- And taken together,

$$\{reachables\} = range(Q_r(A, B)) = range(M_r(t_1; A, B)) \qquad all \; t_1 > 0$$

Fallout: $range(M_r(t_1; A, B))$ is the same for all $t_1 > 0$. Saw earlier that if $x_1 \in range(M_r(t_1; A, B))$, you can reach x_1 in time t_1. Because $range(M_r(t_1; A, B))$ is the same for all t_1, and equals $\{reachables\}$, then you can reach x_1 in as short a time as you want!

Caution: Reaching a reachable x_1 really fast will require lots of input energy as

$$x_1 = \int_0^{t_1} e^{(t_1-\tau)A} Bu(\tau) d\tau$$

So, if $t = 10^{-59} \rightarrow u \approx 10^{59}$ to get $x_1 \approx 3$

Term: Say (A,B) is a reachable pair when every $x_1 \in \mathbb{R}^n$ is reachable w.r.t. (A,B)

$$\begin{aligned} (A, B) \; is \; reachable \; pair \quad &\Leftrightarrow \quad Q_r(A, B) \; has \; (full) \; rank = n \\ &\Leftrightarrow \quad M_r(t_1; A, B) \; is \; invertible \; for \; some \; t_1 > 0 \end{aligned}$$

3.2.2 Controllability

We'll see controllability in continuous time is way easier than in discrete time. Say $x_0 \in \mathbb{R}^n$ is controllable w.r.t. (A,B) when you can start at $x(0) = x_0$ and drive $x(t)$ to zero in finite time by suitable choice of input. More formally, x_0 is controllable w.r.t. (A,B) when there exists $t_1 > 0$ and 'nice' input $u : [0, t_1] \to \mathbb{R}^m$ such that

$$0 = e^{t_1 A} x_0 + \int_0^{t_1} e^{(t_1-\tau)A} B u(\tau) d\tau$$

Note: If x_0 is reachable then it's controllable, x_0 reachable $\Rightarrow e^{tA}x_0$ is reachable for all $t_1 \geq 0$ by invariance under A of $\{reachable\ states\}$

- Can pick any $t_1 > 0$ and choose $u : [0, t_1] \to \mathbb{R}^m$ to reach $e^{t_1 A}x_0$ in time exactly t
- $-e^{t_1 A}x_0 = \int_0^{t_1} e^{(t_1-\tau)A} B u(\tau) d\tau$ for that u choice
- $0 = e^{t_1 A}x_0 + \int_0^{t_1} e^{(t_1-\tau)A} B u(\tau) d\tau$ so x_0 also controllable

Moreover, if x_1 is controllable, then x_1 is also reachable. Why?

$\to$ $-e^{t_1 A}x_1$ reachable for some $t_1 > 0$

$\to$ $e^{-t_1 A}(-e^{t_1 A}x_1)$ reachable by invariance of $\{reachables\}$

$\to$ x_1 is also reachable

· Note in DT since A^k is not always invertible, this doesn't work in general

Bottom Line: In continuous time

$$\{reachable\ states\} = \{controllable\ states\}$$

Say (A,B) is a controllable pair when every $x_0 \in \mathbb{R}^n$ is controllable w.r.t. (A,B). In continuous time,

$$(A, B)\ controllable \qquad \leftrightarrow \qquad (A, B)\ reachable$$

3.2.3 Summary

1. If $x_0 \in \mathbb{R}^n$ is controllable w.r.t. (A,B) you can "control x_0 to 0" in a arbitrarily small amount of time by suitable u-choice. (Follows from corresponding reachability result applied to reaching $-e^{t_1 A}x_0$ in time exactly t_1)
2. Point-to-point reachability: If x_1 and x_2 are reachable w.r.t. (A,B), then for any $t_1 > 0$ there exists $u : [0, t_1] \to \mathbb{R}^m$ such that $x_2 = e^{t_1 A}x_0 + \int_0^{t_1} e^{(t_1-\tau)A} B u(\tau) d\tau$. (So "You can drive the state from $x(0) = x_1$ to $x(t_1) = x_2$ in any positive amount of time by suitable u-choice)
3. "Time-varying versions" of these (controllability and reachability) results exists - see book by Rugh

- What makes things hard - dependence of everything on initial times - concepts include "uniform reachability", etc
- Time-varying notions are important when dealing with linearization of nonlinear systems about "non-equilibrium trajectories"

3.3 Observability

Observability is the next major point of discussion. It is about the connection between states and outputs. Reachability and controllability is for the connection between states and inputs.

Question: What can you tell about the state by observing only the output (and manipulating the input)?

3.3.1 Discrete Time

$$(II) \qquad \begin{aligned} x(k+1) &= A(k)x(k) + B(k)u(k) \\ y(k) &= C(k)x(k) + D(k)u(k) \end{aligned} \qquad x \in \mathbb{R}^n, u \in \mathbb{R}^m, y \in \mathbb{R}^p, A \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{n\times m}, \textit{ etc.}$$

Say x_0 is unobservable w.r.t. (A,C) when there's no way to choose inputs $u(k), k \geq 0$ to distinguish whether $x(0) = 0$ or $x(0) = x_0$ by observing outputs $y(k), k \geq 0$ - ie $x(0) = x_0$ gives rise to some output sequence $y(k), k \geq 0$, no matter what u you use. More formally, $x_0 \in \mathbb{R}^n$ is unobservable when for every $k \mapsto u(k), k \geq 0$ we have

$$CA^k x_0 + \sum_{l=0}^{k-1} CA^{k-l-1}Bu(l) + Du(k)$$

$$= 0 + \sum_{l=0}^{k-1} CA^{k-l-1}Bu(l) + Du(k)$$

Where the first equation is $y(k)$ when you start at $x(0) = x_0$ for all $k \geq 0$, and the second is $y(k)$ when start at $x(0) = 0$.

It follows that x_0 is unobservable w.r.t. (A,C) if and only iff

$$CA^k x_0 = 0 \qquad for\ all\ k \geq 0$$

Define: Observability Matrix: $Q_O(A, C)$

$$Q_O(A,C) = \begin{bmatrix} C & | & CA & | & CA^2 & | & \cdots & | & CA^{n-1} \end{bmatrix}^T \in \mathbb{R}^{np\times n}$$

$$\begin{aligned} x_0 \text{ unobservable} &\Rightarrow Q_O(A,C)x_0 = 0 (np \text{ dim vector of } 0s) \\ &\Rightarrow \quad x_0 \in nullspace(Q_O(A,C)) \end{aligned}$$

In fact, if $x_0 \in nullspace(Q_O(A,C))$ then, $Cx_0 = 0, CAx_0 = 0, \cdots CA^{n-1}x_0 = 0$

By Cayley-Hamilton, you can write any power of A^k as a linear combination of $I_n, A, A^2, \cdots, A^{n-1} \rightarrow$ so $CA^k x_0 = 0$ for all $k \geq 0$ when $x_0 \in nullspace(Q_O(A,C))$

Bottom Line:

$$\{unobservable\ states\} = nullspace(Q_O(A,C))$$

Conclude That:

$=$ $\{unobservable\ states\}$ is a subspace of $\mathbb{R}^n$

$=$ $\{unobservable\ states\}$ is invariant under A $(CA^k x_0 \rightarrow CA^k(Ax_0) = 0)$

Say (A,C) is a observable pair when the only unobservable $x_0 \in \mathbb{R}^n$ is $x_0 = 0$. Thus,

$$\begin{aligned} (A,C) \text{ observable pair} &\Leftrightarrow \quad nullspace(Q_O(A,C)) = \{0\} \\ &\Leftrightarrow \quad Q_O(A,C) \text{ has } (full) \text{ rank} = n \end{aligned}$$

3.3.2 Continuous Time

$$(I) \qquad \begin{aligned} \dot{x}(t) &= A(t)x(t) + B(t)u(t) \\ y(t) &= C(t)x(t) + D(t)u(t) \end{aligned}$$

Say $x_0 \in \mathbb{R}^n$ is unobservable w.r.t. (A,C) when there's no way to distinguish whether you started at $x(0) = 0$ or $x(0) = x_0$ by choosing $u(t), t \geq 0$ and observing the resulting $y(t), t \geq 0$. Formally, x_0 is unobservable when for all $u : [0, \infty] \rightarrow \mathbb{R}^m$, you have

$$Ce^{tA}x_0 + \int_0^t Ce^{(t-\tau)A}Bu(\tau)d\tau + Du(t)$$

$$= 0 + \int_0^t Ce^{(t-\tau)A}Bu(\tau)d\tau + Du(t) \qquad for\ all\ t \geq 0$$

Accordingly, x_0 unobservable if and only if

$$Ce^{tA}x_0 = 0 \qquad for\ all\ t \geq 0$$

Fact $Ce^{tA}x_0 = 0$ for all $t \geq 0 \Leftrightarrow Q_O(A,C)x_0 = 0$

Observability Grammian: is a positive semi-definite matrix:

$$W_o[t_0, t_1] = \int_{t_0}^{t_1} \Phi^*(\tau, t_0) C^*(\tau) C(\tau) \Phi(\tau, t_0) d\tau \in \mathbb{R}^{n \times n}$$

$$completely\ observable \Leftrightarrow rank(W_o) = n$$

3.3.3 Reachability Observability Duality

One final detail: **Duality** between reachability and observability

· **Fact**:

$$\begin{array}{rcl} (A, B)\ reachable\ pair & \Leftrightarrow & (A^T, B^T)\ observable\ pair \\ & and & \\ (A, C)\ observable\ pair & \Leftrightarrow & (A^T, C^T)\ reachable\ pair \end{array}$$

· To see this:

- $(A, B)\ reachable \Leftrightarrow rank(Q_r(A, B)) = n \Leftrightarrow rank(Q_r^T(A, B)) = n$, because the $rank(A) = rank(A^T)$.
- But,

$$Q_r^T = \begin{bmatrix} B & AB & A^2B & \cdots & A^{n-1}B \end{bmatrix}^T$$

$$= \begin{bmatrix} B^T \\ B^T A^T \\ B^T (A^T)^2 \\ \vdots \\ B^T (A^T)^{n-1} \end{bmatrix} = Q_O(A^T, B^T)$$

- Then,

$$\begin{array}{rcl} rank(Q_r(A, B)) = n & \Leftrightarrow & rank(Q_O(A^T, B^T)) = n \\ & \Leftrightarrow & (A^T, B^T)\ is\ observable \end{array}$$

· A similar argument for other direction - start with $Q_O(A, C)$...

3.4 Stability

Next topic: **Stability** of state space linear system models for constant matrices - for time varying matrices see Rugh's book. There are various ways to approach the concept of stability, but they all end up being "all about the eigenstructure of A".

General Definition: system (I) or (II) is stable in the sense of Lyapunov when: For any two initial conditions x_0, x_0' at time 0, and any single input function u for time > 0, the difference

$\|x_0 - x_0'\|$, where x is starting from x_0 and applying u, x' is starting from x_0' and applying u, is bounded for all time > 0. Ie there's a single $R > 0$ such that $\|x - x'\| \leq R$ for all times > 0.

Say the system is asymptotically stable (in the sense of Lyapunov), when not only is it stable, but in the notation above

$$\|x - x'\| \to 0 \; as \; t \to \infty$$

In equations:

Discrete Time: Stable when any $x_0, x_0' \in \mathbb{R}^n$, and any choice $k \mapsto u(k) \in \mathbb{R}^m, k \geq 0$, we have

$$\|x(k) - x'(k)\| \leq R \; for \; all \; k \geq 0, \qquad when$$

$$x(k) = A^k x_0 + \sum_{l=0}^{k-1} A^{k-l-1} B u(l) \qquad k \geq 0$$

$$x'(k) = A^k x_0' + \sum_{l=0}^{k-1} A^{k-l-1} B u(l) \qquad k \geq 0$$

And asymptotically stable when stable, and for all x_0, x_0', u have

$$\lim_{k \to \infty} \|x(k) - x'(k)\| = 0$$

Continuous Time: Stable when for every $x_0, x_0' \in \mathbb{R}^n$, and $u : [0, \infty) \to \mathbb{R}^n$, there exists $R > 0$ such that

$$\|x(t) - x'(t)\| \leq R \; for \; all \; t \geq 0, \qquad when$$

$$x(t) = e^{tA} x_0 + \int_0^t B u(\tau) d\tau \qquad t \geq 0$$

$$x'(t) = e^{tA} x_0' + \int_0^t B u(\tau) d\tau \qquad t \geq 0$$

and asymptotically stable when stable and

$$\lim_{t \to \infty} \|x(t) - x'(t)\| = 0$$

- Note: in either case, the u-terms disappear when you enter the expressions into the definition. Find that,

  CT Stable when for every $x_0, x_0' \in \mathbb{R}^n$, there exists $R > 0$ such that $\left\|e^{tA}(x_0 - x_0')\right\| \leq R$, and asymptotically stable when $\left\|e^{tA}(x_0 - x_0')\right\| \to 0 \; as \; t \to \infty$. Which is the same as saying, stable when for all $v \in \mathbb{R}^n$, $\left\|e^{tA} v\right\|$ is bounded on $[0, \infty)$ and asymptotically stable when for all $v \in \mathbb{R}^n$, $\left\|e^{tA} v\right\| \to 0 \; as \; t \to \infty$.

  DT Stable when for all $v \in \mathbb{R}^n$, $\left\|A^k v\right\|$ is bounded on $0 \geq k < \infty$, asymptotically stable when for all v, $\left\|A^k v\right\| \to 0 \; as \; \to \infty$

- **Turns out**: behavior (boundedness or unboundedness) of e^{tA} or A^k for $t \in [0, \infty)$ or $0 \leq k < \infty$ - and it's asymptotics as $k \to \infty, t \to \infty$ are all about the eigenstructures of A.

3.4.1 Discrete Time Facts

1) If A has a single eigenvalue λ_0 with $|\lambda_0| > 1$, then not stable
   Idea:
   - if $v \in \mathbb{C}^n$, eigenvector for λ_0, then $\left\|A^k v\right\| = |\lambda_0|^k \|v\|$ for all $k \geq 0$ which is unbounded in k or $0 \leq k < \infty$
   - if $v \in \mathbb{R}^n$ and $\lambda_0 \in \mathbb{R}$, this contradicts stability
   - if $v \in \mathbb{C}^n$ and $\lambda_0 \in \mathbb{C}$, either $\left\|A^k Re\{v\}\right\|$ or $\left\|A^k Im\{v\}\right\|$ must be unbounded on $0 \leq k < \infty$ (or both - also contradicts stability)
2) If all eigenvalues of A have magnitudes ≤ 1, and at least one λ_0 has magnitude $= 1$, then not asymptotically stable (make stable).
   Reason: if v is eigenvector $\leftrightarrow \lambda_0$, then $\left\|A^k v\right\| = |\lambda_0|^k \|v\| = \|v\|$, which does not $\to 0$ as $k \to \infty$ (complex / real thing as in 1)
$\star$ 3) Asymptotically stable $\Leftrightarrow$ all A's eigenvalues have magnitudes < 1
   Term: say $A \in \mathbb{R}^{n \times n}$ is Schur when all A's eigenvalues have magnitudes < 1
   - This is easiest to see when A diagonalizable... in that case, can find $\mathbb{X}$ such that

$$A = \mathbb{X}\Lambda\mathbb{X}^{-1} \qquad so\ A^k = \mathbb{X}\Lambda^k\mathbb{X}^{-1}\ all\ k > 0$$

   - Given $v \in \mathbb{C}^n$

$$\left\|A^k v\right\| = \left\|\mathbb{X}\Lambda^k\mathbb{X}^{-1}v\right\| \leq \|\mathbb{X}\| \cdot \left\|\mathbb{X}^{-1}\right\| \cdot \left\|\Lambda^k\right\| \cdot \|v\| \quad k > 0$$

     where $\|\Lambda\| = max\{|\lambda_0| : \lambda_0\ an\ eigenvalue\ of\ A\} < 1$

   - Thus,

$$\left\|A^k v\right\| \to 0\ as\ k \to \infty$$

   $\Rightarrow$ If, A is Schur, then asymptotically stable (converse follows from 1) and 2) above)
   - For non-diagonalizable case, refer to end of Matrix Handout - show, when A is Schur, and $\rho = max\{\lambda_0's\ magnitudes\}$, then for any $\epsilon > 0$ there exists some $M > 0$ such that

$$\left\|A^k v\right\| \leq M(\rho + \epsilon)^k \|v\| \qquad all\ v \in \mathbb{R}^n$$

     Thus, where A is Schur, so $\rho < 1$, can choose ϵ small enough so $\rho + \epsilon < 1$, showing $\left\|A^k v\right\| \to 0\ as\ k \to \infty$ for all $v \in \mathbb{R}^n$
4) If A has eigenvalues, all of which have magnitudes ≤ 1, and at least 1 with magnitude $= 1$, then:

Stable $\Leftrightarrow$ (for every λ_0 with $|\lambda_0| = 1$, the algebraic multiplicity and geometric multiplicity of λ_0 are the same - ie every generalized eigenvector $\leftrightarrow \lambda_0$ is a true eigenvector $\mapsto \lambda_0$ (see handout...))

3.4.2 Continuous Time Facts

1) If A has at least one eigenvalue λ_0 with $Re\{\lambda_0\} > 0$, then not stable. The reasoning is below:
   say $v_0 \in \mathbb{C}^n$ eigenvector $\leftrightarrow \lambda_0$, then
   $e^{tA}v_0 = e^{\lambda_0 t}v_0 \Rightarrow \left\|e^{tA}v_0\right\| = |e^{\lambda_0 t}| \, \|v_0\|$
   $|e^{\lambda_0 t}| = |e^{(\sigma_0 + jw_0)t}| = e^{\sigma_0 t}$ *with* $|e^{jw_0 t}| = 1$
   And because $\sigma_0 > 0$, $e^{\sigma_0 t}$ grows as $t \to \infty$, and $\left\|e^{tA}v_0\right\|$ is not bounded on $[0, \infty) \Rightarrow$ *not stable*
2) If A has all eigenvalues with ≤ 0 real parts, and it has at least one eigenvalue λ_0 with $Re\{\lambda_0\} = 0$, then not asymptotically stable (maybe stable). Reason:
   Say $\lambda_0 = jw_0$ & $v_0 \leftrightarrow \lambda_0$. Then,
   $\left\|e^{tA}v_0\right\| = |e^{tA}| \, \|v_0\| = \|v_0\| \;\; allt \geq 0$
   so $\left\|e^{tA}v_0\right\| \not\to 0 \; as \; t \to \infty \Rightarrow$ *not asymptotically stable*
⋆ 3) Asymptotically Stable $\Leftrightarrow$ all of A's eigenvalues have strictly negative real parts. All eigenvalues in left half plane.
   Term: in this case, call A a Hurwitz Matrix.
   - Easiest way to see when A Diagonalizable, so $A = \mathbb{X}\Lambda\mathbb{X}^{-1}, e^{tA} = \mathbb{X}e^{t\Lambda}\mathbb{X}^{-1}$
     Thus,
     $\left\|e^{tA}v\right\| \leq \|\mathbb{X}\| \cdot \|\mathbb{X}^{-1}\| \cdot \left\|e^{\Lambda t}\right\| \cdot \|v\| \;\; all \; t \geq 0$
   - When $Re\{eigenvalues\} < 0$, diagonal elements in $e^{t\Lambda}$ decay exponentially as $t \to \infty$, so $\left\|e^{t\Lambda}\right\| \to 0 \; as \; t \to \infty$, so $\left\|e^{tA}v\right\| \to 0 \; as \; t \to \infty$, so *asymptotically stable* because v is arbitrary.
   - When A is not diagonalizable, can show that (see handout), for every $t > 0$, there exists $M > 0$ such that (where $\hat{\sigma_0} = max\{Re\{eigenvalues \; of \; A\}\}$)

$$\left\|e^{tA}v\right\| \leq Me^{\hat{\sigma_o}+\epsilon)t} \|v\| \qquad all \; t \; \geq 0 \; and \; v \in \mathbb{R}^n$$

   Thus, if A Hurtwitz, so $\hat{\sigma_o} < 0$, can pick ϵ - say $\epsilon = -\frac{\hat{\sigma_o}}{\pi}$ - so $\hat{\sigma_o} + \epsilon < 0$, implying that $\left\|e^{tA}v\right\| \to 0 \; as \; t \to \infty$, so asymptotically stable
   - The converse that asymptotitcally stable $\Rightarrow$ A is Hurwitz follows an argument from points 1) and 2).
4) If all eigenvalues of A have ≤ 0 real parts, and at least one has $Re\{\} = 0$. Then, stable $\Leftrightarrow$ for every eigenvalue λ_0 with $Re\{\lambda\} = 0$, the algebraic multiplicity of that eigenvalue is equal to the geometric multiplicity. Ie the generalized eigenvectors are true eigenvectors of A.

3.5 Lyapunov Lemmas

Clearly, items 3) in continuous and discrete time are the most important of them in some sense. Basically in each case have:

DT asymptotically stable $\Leftrightarrow$ A is Schur

CT asymptotically stable $\Leftrightarrow$ A is Hurwitz

There must be some neat way to answer the yes-no questions: Hurwitz? Schur? about A without finding all the eigenvalues of A. The Answer is by means of the Lyapunov Lemmas.

3.5.1 Continuous Time

Recall that $Q \in \mathbb{R}^{n\times n}$ is positive definite when Q is symmetric and all the eigenvalues of Q are positive (same as saying Q is symmetric and $v^TQv > 0$ for all nonzer $v \in \mathbb{R}^n$). **Lyapunov Lemma Part 1**: If $A \in \mathbb{R}^{n\times n}$ is Hurwitz, then for every symmetric $Q \in \mathbb{R}^{n\times n}$ there exists a unique symmetric $P \in \mathbb{R}^{n\times n}$ satisfying the Lyapunov Equation:

$$PA + A^TP = -Q$$

Furthermore - if Q is positive definite, so is P

- To see this: think of the mapping $P \mapsto PA + A^TP$ as a linear mapping from the space of symmetric matrices to itself - we'll exhibit a P satisfying $f(P) = Q$ for a given Q, making f surjective, hence injective, giving uniqueness. Given Q,

$$P = \int_0^\infty e^{tA^T}Qe^{tA}dt$$

- The integral converges because A Hurwitz and entries in e^{tA} and e^{tA^T} take the form of linear combinations of terms of the form $(\textit{polynomial in } t)\cdot e^{\lambda_0 t}$ where λ_0 is an eigenvalue of A. By Hurwitzness of A, these all integrate nicely from 0 to ∞. Check that $f(P) = Q$

$$\begin{aligned} PA + A^TP &= \int_0^\infty \left(e^{tA^T}Qe^{tA}A + A^Te^{tA^T}Qe^{tA}\right)dt \\ &= \int_0^\infty \tfrac{d}{dt}\left(e^{tA^T}Qe^{tA}\right)dt \\ &= e^{tA^T}Qe^{tA}\Big|_{t=0}^{t=\infty} \\ &= 0 - Q \end{aligned}$$

- $\Rightarrow$ f is bijective, meaning we can always solve for P given Q and that a solution P is unique given$Q$. To see why P is positive definite when Q is, consider $v^TPv = \int_0^\infty (v^Te^{tA^T}Qe^{tA}v)dt$ when $v \neq 0, e^{tA}v \neq 0$, so the integrand > 0 because Q is positive definite, so the integral is > 0. $\Rightarrow$ Q is positive definite implies P is positive definite.

Problem: Say you have a Hurwitz matrix A, and someone brings you a symmetric matrix Q and asks you to find P satisfying the Lyapunov Equation.

- How not to find P: find e^{tA}, and compute $P = \int_0^\infty e^{tA^T} Q e^{tA} dt$
- How to solve for P: recognize that the Lyapunov Equation is a set of $\frac{n(n+1)}{2}$ linear equations in the $\frac{n(n+1)}{2}$ unknown entries in P. Can find P exactly by Gauss elimination or whatever.

Example! Take A, and someone gives you Q wanting a symmetric P satisfying

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

$$A = \begin{bmatrix} -1 & 2 \\ 0 & -3 \end{bmatrix} \ (is\ Hurwitz),\ and\ Q = \begin{bmatrix} 7 & 0 \\ 0 & 2 \end{bmatrix},\ with\ P = \begin{bmatrix} \alpha & \beta \\ \beta & \gamma \end{bmatrix}$$

$$\begin{bmatrix} \alpha & \beta \\ \beta & \gamma \end{bmatrix}\begin{bmatrix} -1 & 2 \\ 0 & -3 \end{bmatrix} + \begin{bmatrix} -1 & 2 \\ 0 & -3 \end{bmatrix}\begin{bmatrix} \alpha & \beta \\ \beta & \gamma \end{bmatrix} = \begin{bmatrix} -7 & 0 \\ 0 & -2 \end{bmatrix}$$

$$-\alpha - \gamma = -7, \qquad 2\alpha - 3\beta - \beta = 0, \qquad 2(2\beta - 3\gamma) = -2$$

Lyapunov Lemma Part 2: (more than a converse) If for some positive definite matrix $Q \in \mathbb{R}^{n\times n}$, there exists a positive definite P satisfying the Lyapunov Equation, then A is Hurwitz.

Sketch of an argument for this, as a rigorous proof is lengthy:

- Suppose P, Q are positive definite and satisfy Lyapunov's Equation; want to show A is Hurwitz
- Consider the function $f : \mathbb{R}^n \mapsto \mathbb{R}$ defined by $f(v) = v^T P v$ for all $v \in \mathbb{R}^n$
- Note that $f(v) > 0$ for all $v \neq 0$
- Pick an initial condition $x_) \neq 0$ for $\dot{x} = Ax$; then $x(t) = e^{tA}x_0$ for all $t \geq 0$
- Note: $e^{tA}x_0 \neq 0$ all $t \geq 0$ because e^{tA} is always invertible
- Consider $t \mapsto f(e^{tA}) = x_0^T e^{tA^T} P e^{tA} x_0 = g(t),\ t \geq 0$

$$\begin{aligned} \tfrac{d}{dt}g(t) &= x_0^T(e^{tA^T}A^T P e^{tA} + e^{tA^T} P A e^{tA})x_0 \\ &= x_0^T e^{tA^T}(A^T P + PA)e^{tA}x_0 \\ &= -x_0^T e^{tA^T} Q e^{tA} x_0,\ t \geq 0 \\ &< 0 \text{ for all } t > 0 \text{ because Q positive definite and the initial condition} \end{aligned}$$

- Since $g(t) > 0$ for all $t \geq 0$ and is strictly decreasing and continuous in t, $g(t)$ must approach a limit $\hat{g}$ as $t \to \infty$
- So, $\frac{d}{dt}g(t) \to 0$ as $t \to \infty$, which means $e^{tA}x_0 \to 0$ as $t \to \infty$ because Q positive definite
- Thus, have shown that $e^{tA}x_0 \to 0$ as $t \to \infty$ for all $x_0 \in \mathbb{R}^n \Rightarrow A$ is Hurwitz!

3.5.2 Lyapunov Function

Terminology: The function $v \mapsto v^T P v$, is a Lyapunov Function for the differential equation $\dot{x} = Ax$

Question: How does Lyapunov Lemma Part 2 save computations? How does it make answer the H question algorithmic?Think about if someone says: here's an A, is it Hurwitz? Proceed as follows:

- Pick your favorite positive definite Q (ex is identity) and try to solve for P $\in \mathbb{R}^{n\times n}$, symmetric in

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

  - · if no solution exists, A is not Hurwitz
  - · if a solution exists that is not positive definite, A is not Hurwitz
  - · if a positive definite solution exists, A is Hurwitz
- Note, can test for positive definiteness of P without finding P's eigenvalues. A quick couple things to look for in positive definiteness space are that diagonal elements are strictly positive and diagonal elements dominate in the way that elements decay in magnitude from the diagonal.

  → P is positive definite *iff* all principal, minor determinants of P a strictly positive.

3.5.3 Discrete Time

Lyapunov Lemma Part 1: If $A \in \mathbb{R}^{n\times n}$ is Schur, then for every symmetric $Q \in \mathbb{R}^{n\times n}$ there exists a unique symmetric $P \in \mathbb{R}^{n\times n}$ satisfying the Lyapunov DT Equation (And if Q is positive definite, so is P):

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

- To see why, note that given a Q, the following P satisfies the equation

$$P = \sum_{k=0}^{\infty} (A^T)^k Q A^k$$

- Note that the sum converges because A being Schur shows that all entries in A^k or $(A^T)^k = (A^k)^T$ are linear combinations of terms like $(\textit{polynomial in } k) \cdot \lambda_0^k$ where $|\lambda_0| < 1$, so they sum up from $k = 0$ *to* ∞.
- Uniqueness of P given Q follows same vector space and dimension argument as in continuous time.

Lyapunov Lemma Part 2: (more than a converse) If for some positive definite matrix $Q \in \mathbb{R}^{n\times n}$, there exists a positive definite P satisfying the Lyapunov DT Equation, then A is Schur.

- Show by setting $f(v) = v^T Pv,\ v \in \mathbb{R}^n$, pick $x_0 \in \mathbb{R}^n$, and look at $g(k) = g(A^k x_0) = x_0^T (A^T)^k P A^k x_0,\ k \geq 0$ (which should look similar to the CT case)
- Assume $A^k x_0 \neq 0$ for any $k > 0$ - consider x_0's violating this later...

- Then $g(k) > 0$ for all $k \geq 0$, and $g(k+1) - g(k) = x_0^T (A^T)^{k+1} P A^{k+1} x_0 - x_0^T (A^T)^k P A^k x_0$ which reduces to $-x_0^T (A^T)^k Q A^k x_0 < 0$ because Q is positive definite.
- Thus $g(k) > 0$ and is strictly decreasing in k, so it approaches a limit as $k \to \infty$, so $g(k+1) - g(k) \to 0$ as $k \to \infty$, so $A^k x_0 \to 0$ as $k \to \infty$
- This applies to all x_0 such that $A^k x_0 \neq 0$ for any $k > 0$, but if $A^k x_0 = 0$ for any $k > 0$, $A^k x_0 \to 0$ trivially.

$\Rightarrow$ $A^k x_0 \to 0$ as $k \to \infty$ for all $x_0 \in \mathbb{R}^n$, thus A is Schur

Chapter 4

Feedback and Observers

4.1 Introduction

Feedback and **Observers** in the state space linear system models of the form (I) and (II) with constant matrices (See Rugh's book for time-varying case). Given that you understand why using feedback in general can be useful, we define a constant-gain state feedback control law for a system (I) or (II) as a choice of u given by

$$u = -Kx + v$$

where $K \in \mathbb{R}^{m \times n}$ and v is some exogenous input.

Implementing such control laws lead to a closed-loop state equation:

$$\dot{x} = (A - BK)x + Bv \qquad x(k+1) = (A - BK)x(k) + Bv$$

Thus, $(A - BK)$ is approximately the A matrix of the closed loop system

4.1.1 Wonham's Theorem and Analysis

Question: How can we design a closed loop A matrix $(A - BK)$ by means of choosing K? Ex. can we choose K such that A-BK is Hurwitz or Schur? Can we find a stabilizing feedback control law?

W. Murray Wonham (1968) proved an amazing theorem: if (A,B) reachable, then you can assign the eigenvalues of A-BK any way you want by choice of $K \in \mathbb{R}^{m \times n}$

Technically: Given any n complex numbers, not necessarily distinct, that could be the eigenvalues (counted with multiplicities) of a real (n x n) matrix. When (A,B) is reachable, you can find K$\in \mathbb{R}^{m \times n}$ such that (A-BK) has those eigenvalues. The handout proves the general case, where we will go over the $m = 1$ case in class. The handout shows how to reduce the general m case to $m = 1$ based on Heymann's Lemma

Strategy: Discuss the case for m=1; refer details to the handout; note how the $m > 1$ case reduces to m=1 case via Heymann's Lemma

- Given (A,B) reachable with $B \in \mathbb{R}^{n\times 1}$, can show that (handout) there exists some $X \in \mathbb{R}^{n\times n}$ such that

$$\mathbb{X}^{-1}A\mathbb{X} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ 0 & 0 & 0 & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & 1 \\ -q_n & -q_{n-1} & \cdots & -q_2 & -q_1 \end{bmatrix}$$

$$\mathbb{X}^{-1}B = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}^T$$

- $\mathbb{X}^{-1}A\mathbb{X}$ is in companion form, so the characteristic polynomial of $\mathbb{X}^{-1}A\mathbb{X}$, and hence that of A, is

$$\lambda^n + q_1\lambda^{n-1} + \cdots + q_{n-1}\lambda + q_n$$

- Suppose we want $A - BK$ to have a certain set of eigenvalues that are roots of

$$\lambda^n + p_1\lambda^{n-1} + \cdots + p_{n-1}\lambda + p_n$$

- If we set $\hat{K}$,

$$\hat{K} = \begin{bmatrix} (p_n - q_n) & (p_{n-1} - q_{n-1}) & \cdots & (p_1 - q_1) \end{bmatrix} \in \mathbb{R}^{1\times n}$$

Giving,

$$\mathbb{X}^{-1}A\mathbb{X} - \mathbb{X}^{-1}B\hat{K} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ 0 & 0 & 0 & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & 1 \\ -p_n & -p_{n-1} & \cdots & -p_2 & -p_1 \end{bmatrix}$$

$$= \mathbb{X}^{-1}(A - B\hat{K}\mathbb{X}^{-1})\mathbb{X}$$

- Thus, by setting $K = \hat{K}\mathbb{X}$, the eigenvalues of $\mathbb{X}^{-1}(A - BK)\mathbb{X}$ - same eigenvalues of A - are roots of

$$\lambda^n + p_1\lambda^{n-1} + \cdots + p_{n-1}\lambda + p_n$$

→ Reduce $m > 1$ case to m=1 by using Heymann's Lemma:

  · If (A,B) reachable and j such that jth column - call it B_j - of B is nonzero, there exists some $K_j \in \mathbb{R}^{m\times 1}$ such that $(A - BK_j, B_j)$ is reachable on the m=1 level.

  · From there, move eigenvalues of $(A - BK_j - B_jK')$ by choice of $K' \in \mathbb{R}^{1\times n}$; then

make $K \in \mathbb{R}^{m \times n}$ via

$$K = K_j - \begin{bmatrix} & 0 & \\ \cdots & K' & \cdots \\ & 0 & \end{bmatrix}$$

· Where K' is on the jth row. Then, $A - BK = A - BK_j - B_jK'$ has the eigenvalues you want

4.1.2 Observers

One other feedback-related topic: Observers for state space linear system models with constant matrices. For a change, we'll pay attention to the $y = Cx + Du$ part of the models.

Problem: would like to implement a constant-gain state feedback control law

$$u = -Kx + v$$

but, you don't have access to x, only to y

- A possible workaround (stay with continuous time for the moment) is to build a secondary system where you still know z, u, y':

$$\textit{system}: \quad \begin{aligned} \dot{x}(t) &= A(t)x(t) + B(t)u(t) \\ y(t) &= C(t)x(t) + D(t)u(t) \end{aligned}$$

$$\textit{Auxiliary system} \quad \begin{aligned} \dot{z}(t) &= A(t)z(t) + B(t)u(t) \\ y'(t) &= C(t)z(t) + D(t)u(t) \end{aligned}$$

- If we could initialize the auxiliary system such that $z(0) = x(0$, then for any choice of u, we'll have $z(t) = x(t)$ for all $t > 0$ and $y'(t) = y(t)$ for all $t > 0$. Thus, if we set $u = -Kz + v$ we'd implement the state feedback law $u = -Kx + v$ in the original system.
- Sadly, we cannot do this as we do not have access to $x(0)$

→ Workaround:

· make z track x asymptotically

· use $u = -Kz + v$ in original system

- This will at least asymptotically (over the long haul), make the original system behave 'as if' we had been able to set $u = -Kx + v$

→ Proceed as follows

· set $u = -Kz + v$

· Add on a correlation term to the $\dot{z}$ equation $L(Y' - Y), L \in \mathbb{R}^{n \times p}$, giving

$$\dot{z} = Az + B(-Kz + v) + LC(z - x)$$

· In original system, set $u = -Kz + v$, giving

$$\dot{x} = Ax - BKz + Bv$$

- **Claim**: if we choose $L \in \mathbb{R}^{n \times p}$ correctly, z will track x asymptotically assuming (A,C) is observable.

  $\frac{d}{dt}(z - x) = \dot{z} - \dot{x}$
  $= Az + LC(z - x) - Ax$
  $= (A + LC)(z - x)$

  $\rightarrow$ z tracks x asymptotically means $z(t) - x(t) \rightarrow 0$ as $t \rightarrow \infty$, which means/ requires (A+LC) is Hurwitz (all eigenvalues are in open left half plane).

Question: Why do we know, given (A,C) observable, that we can find $L \in \mathbb{R}^{n \times p}$ such that (A+LC) is Hurwitz?

Answer: Duality and Wonham!

$$\begin{aligned} (A,C) observable \quad &\Rightarrow \ (via\ duality) \quad && (A^T, C^T)\ reachable \\ &\Rightarrow \ (via\ Wonham) && A^T - C^T K\ Hurwitz\ for\ some\ K \in \mathbb{R}^{p \times n} \\ &\Rightarrow && A^T + C^T L^T\ Hurwitz\ when\ L = -K^T \\ &\Rightarrow && A + LC\ Hurwitz\ for\ some\ L - choice \end{aligned}$$

Auxiliary system so constructed - the z-system, where z tracks x asymptotically - called an observer or state-simulator for original system. Idea due originally to Dr. Luenberger. Main user choice is L; Wonham implies that you can choose L so $x - z \rightarrow 0$ as fast as you want, but might not in some cases, want super fast tracking.

Note: Whole construction works the same in discrete time, expect choose $z(k+1) - x(k+1) = (A + LC)(z(k) - x(k))$ and the problem boils down to choosing $L \in \mathbb{R}^{n \times p}$ such that (A+LC) is Schur - can do that in fact, by Wonham, place eigenvalues of (A+LC) anywhere when (A,B) observable.

4.2 I/O Linear System Models

Shift to talking about I/O linear system models.

4.2.1 Weighting Pattern and Time-Invariance

Consider first a state space model (I) with possibly time varying matrices of the usual sizes. Given some initial time $t_0 \in \mathbb{R}$, perform the following experiment. Set $x(t_o) = 0$; apply an input $u : \mathbb{R} \rightarrow$

$\mathbb{R}^m$ starting at time t_0; observe $y(t)$ for $t \geq t_0$. Discover that

$$y(t) = \int_{t_0}^{t} C(t)\Phi(t,\tau)B(\tau)u(\tau)d\tau + D(t)u(t) \qquad all\ t \geq t_0$$

can think of the state space system model as defining a family of linear mappings S_{t_0} - one such mapping for each $t_0 \in \mathbb{R}$ - where S_{t_0} maps the set of all 'reasonable' $\mathbb{R}^m$ valued functions - call that set U - to the set Y_{t_0} of all reasonable $\mathbb{R}^p$ valued functions defined on $[t_0, \infty)$

Prompts this abstract definition for an m-input p-output I/O linear system model: Such a thing consists of

- a set of $\mathbb{R}^m$-valued input functions U defined on all $t \in \mathbb{R}$
- for each $t_0 \in \mathbb{R}$, a linear mapping $S_{t_0} : U \mapsto Y_{t_0}$, where Y_{t_0} is the set of all $\mathbb{R}^p$-valued output functions defined on $[t_0, \infty)$
- plus a few regularity conditions to make everything work (skip these...)

Note: any I/O system model that arrises as above forms a state space model in such such a thing, eg could set $U =$ all continuous $u : \mathbb{R} \mapsto \mathbb{R}^m$ and S_{t_0} defined as follows: $S_{t_0}(u)$ has specification $S_{t_0}(u)(t) = RHS\ of\ above\ eqn.$

We are going to restrict our attention to I/O system models that arise as follows:

- There exists a $(p \times m)$-matrix valued function defined only for $t \geq \tau$

$$(t, \tau) \mapsto W_0(t, \tau)$$

- Such that for an $u \in U$ and $t_0 \in \mathbb{R}$ we have for some $D : \mathbb{R} \mapsto \mathbb{R}^(p \times m)$

$$S_{t_0}(u)(t) = \int_{t_0}^{t} W_0(t,\tau)u(\tau)d(\tau) + D(t)u(t) \qquad t \geq \tau$$

Term: The Weighting Pattern of such an I/O linear system model is

$$W(t,\tau) = W_0(t,\tau) + D(t)\delta(t-\tau) \qquad t \geq \tau$$

Thus,

$$S_{t_0}(u)(t) = \int_{t_0}^{t} W(t,\tau)u(\tau)d\tau \qquad t \geq t_0,\ u \in U$$

So, W describes how the system weights values of u on the interval $[t_0, t]$ to yield value of S_{t_0} at time t.

Summarize so far:

· A continuous time m-input p-output I/O linear system model has an input function space U and I/O mappings $S_{t_0} : U \mapsto Y_{t_0}$ (for every $t_0 \in \mathbb{R}$) given by

$$S_{t_0}(u)(t) = \int_{t_0}^{t} W(t,\tau)u(\tau)d\tau \qquad t \geq t_0,\ u \in U$$

. Where W_0 is nice

$$W(t,\tau) = W_0(t,\tau) + D(t)\delta(t-\tau) \qquad t \geq \tau$$

· Every state-space linear system model (I) gives rise to one of the

$$W(t,\tau) = C(t)\Phi(t,\tau)B(\tau) + D(t)\delta(t-\tau) \quad t \geq \tau$$

· Turns out not every I/O linear system model arise from a state space model (I), but we have not proved this

Now for discrete time systems: one can go through the same drill for these systems - can define I/O models abstractly, then restrict systems where I/O mappings arise from weighting patterns, etc. Arrive at summary so for in discrete time:

· A discrete time m-input, p-output I/O linear system model has an input function space U and I/O mappings $S_{k_0} : U \mapsto Y_{k_0}$ (every $k_0 \in \mathbb{Z}$) given by (Note no need for the $W = W_0 + \delta$ thin in DT)

$$S_{k_0}(u)(k) = \sum_{l=k_0}^{k} W(k,l)u(l) \qquad k \geq k_0,\ u \in U$$

· Every state space model (II) gives rise to one of these -

$$W(k,l) = C(k)\Phi(k,l+1)B(l)\mathbb{1}(k-l-1) + D(k)\delta(k-l) \qquad k \geq l$$

· Note every I/O model arises from a state space model.

Let's do a quick reality check on the discrete time W formula: Start with (II); pick $k_0 \in \mathbb{Z}$, set $x(k_0 = 0$; apply u for times $\geq k_0$; find

$$S_{k_0}(u)(k) = y(k) = \sum_{l=k_0}^{k-1} C(k)\Phi(k,l+1)B(l)u(l) + D(k)u(k) \qquad k \geq k_0$$

To absorb the $D(k)u(k)$ into the sum, factor out a $u(k)$ and the inside is the $W(k,l)$

Next: Define time-invariance for I/O system models.

Start with new notation for Shift Notation:

CT: if $s \in \mathbb{R}$ and $f : \mathbb{R} \mapsto something$, $Shift_s(f)$ has specification

$$Shift_s(f)(t) = f(t-s) \qquad all\ t$$

DT: if $j \in \mathbb{Z}$ and $f : \mathbb{R} \mapsto something$, , $Shift_j(f)$ has specification

$$Shift_j(f)(k) = f(k-j) \qquad all\ k$$

From there, we get a definition of time invariance as follows:

CT: for every $t_0 \in \mathbb{R}, s \in \mathbb{R}$, and $u \in U$, we have

$$S_{t_0+s}(Shift_s(u))(t+s) = S_{t_0}(u)(t) \qquad t \geq t_0$$

DT: for all k_0, u, j,

$$S_{k_0+j}(Shift_j(u))(t+j) = S_{k_0}(u)(k) \qquad k \geq k_0$$

And, time-invariance of weighting patterns:

CT: $sys\ time-invariant \Leftrightarrow W(t+s, \tau+s) = W(t,\tau) \qquad t \geq \tau, s \in \mathbb{R}$

DT: $sys\ time-invariant \Leftrightarrow W(k+j, l+j) = W(k,l) \qquad k \geq l, j \in \mathbb{Z}$

{To see why this condition on W implies time-invariance, look into formulas. In continuous time with given parameters above, start with $S_{t_0+s}(Shift_s(u))(t+s)$ and execute a change of variables $\tau = \rho - s$ to remove the difference in time dependance while the differential elements change one to one.}

$\Rightarrow$ People often say a system is time-invariant if and only if the value of its W depends only on the time difference between W's two arguments

4.2.2 Impulse Response

Say you have a time-invariant system with weighting pattern W, define the systems Impulse Response H as the $(p \times m)$-matrix valued function with specification

$$\textbf{CT}: \qquad H(t) = \begin{cases} 0 & t < 0 \\ W(t,0) & t \geq 0 \end{cases} \qquad \textbf{DT}: \qquad H(k) = \begin{cases} 0 & k < 0 \\ W(k,0) & k \geq 0 \end{cases}$$

Describe system's I/O behavior in terms of impulse responses:

$$\textbf{CT}: S_{t_0}(u)(t) = \begin{array}{ll} \int_{t_0}^{t} W(t,\tau)u(\tau)d(\tau) & t \geq t_0 \\ \int_{t_0}^{t} W(t-\tau,0)u(\tau)d(\tau) & t \geq t_0 \\ \int_{t_0}^{t} H(t-\tau)u(\tau)d(\tau) & t \geq t_0 \end{array}$$

$$\textbf{DT}: S_{k_0}(u)(k) = \begin{array}{ll} \sum_{l=k_0}^{k} W(k,l)u(l) & k \geq k_0 \\ \sum_{l=k_0}^{k} W(k-l,0)u(l) & k \geq k_0 \\ \sum_{l=k_0}^{k} H(k-l)u(l) & k \geq k_0 \end{array}$$

Next, because $H(t-\tau) = 0$ for $t < \tau$ and $H(k-l) = 0$ for $k < l$, can expand expressions as

$$S_{t_0}(u)(t) = \int_{t_0}^{\infty} H(t-\tau)u(\tau)d(\tau) \qquad t \geq t_0$$

$$S_{k_0}(u)(k) = \sum_{l=k_0}^{\infty} H(k-l)u(l) \qquad k \geq k_0$$

Terminology: given input function u, the <u>all-time response</u> (if it exists) of an I/O system to an input u is function $S(u)$ with specification

$$S(u)(t) = \lim_{t_0 \to -\infty} S_{t_0}(u)(t) \qquad all\ t \in \mathbb{R}$$

$$S(u)(k) = \lim_{k_0 \to -\infty} S_{k_0}(u)(k) \qquad all\ k \in \mathbb{Z}$$

All-time response to u exists iff limits above exist for all t or k

Thus, when all-time response $S(u)$ to input u exists, we have

$$S(u)(t) = \int_{-\infty}^{\infty} H(t-\tau)u(\tau)d(\tau) \qquad allt \in \mathbb{R}$$

$$S(u)(k) = \sum_{l=-\infty}^{\infty} H(k-l)u(l) \qquad allk \in \mathbb{Z}$$

Which are convolutions!

Suppose we have a state-space model (I) or (II) with constant matrices. Weighting pattern of the I/O system model to which state space systems gives rise is:

$$W(t,\tau) = Ce^{(t-\tau)A}B + D\delta(t-\tau) \qquad t \geq \tau$$

$$W(k,l) = CA^{k-l-1}B\mathbb{1}(k-l-1) + D\delta(k-l) \qquad k \geq l$$

Thus, the I/O system model is time invariant in the I/O sense, and the impulse response is:

$$H(t) = Ce^{tA}B\mathbb{1}(t) + D\delta(t) \qquad t \in \mathbb{R}$$

$$H(k) = CA^{k-1}B\mathbb{1}(k-1) + D\delta(k) \qquad k \in \mathbb{Z}$$

4.3 Realization Theory

Focus on time-invariant case: Can two different state space models with respective (A, B, C, D) and $(\hat{A}, \hat{B}, \hat{C}, \hat{D})$ give rise to the same I/O system model?

Continuous Time: suppose two matrix quadruples are related by an $\mathbb{X}$-maneuver, ie $\hat{A} = \mathbb{X} - A\mathbb{X}$, $\hat{B} = \mathbb{X}^{-1}B$, $\hat{C} = C\mathbb{X}$, $\hat{D} = D$ for some invertible $\mathbb{X} \in \mathbb{R}^{n\times n}$. Then,

$$e^{t\hat{A}} = e^{t(\mathbb{X}-A\mathbb{X})} = \mathbb{X}^{-1}e^{tA}\mathbb{X}$$

$$\hat{C}e^{t\hat{A}}\hat{B}\mathbb{1}(t) + \hat{D}\delta(t) = C(\mathbb{X})(\mathbb{X}^{-1}e^{tA}\mathbb{X})(\mathbb{X}^{-1}B) + D\delta(t)$$

$$= Ce^{tA}B + D\delta(t) \qquad all\ t \in \mathbb{R}$$

Essentially, $\mathbb{X}$-maneuver changing the basis of the state-space.

Here's another kind of nonuniqueness. Given (A, B, C, D), let

$$\hat{A} = \begin{bmatrix} & & & 0 \\ & A & & 0 \\ & & & \vdots \\ & & & 0 \\ 0 & 0 & \cdots & 0 \quad 10^{59} \end{bmatrix} \qquad \hat{B} = \begin{bmatrix} & B & \\ & & \\ 0 & 0 & 0 \quad 0 \quad 0 \end{bmatrix} \qquad \hat{C} = \begin{bmatrix} & & 0 \\ & C & 0 \\ & & 0 \\ & & 0 \\ & & 0 \end{bmatrix} \qquad \hat{D} = D$$

Note: $\hat{A} \in \mathbb{R}^{(n+1)\times(n+1)}$ etc - more states in the hatted model.

Thus,

$$\hat{C}e^{t\hat{A}}\hat{B} = \begin{bmatrix} & & 0 \\ & C & 0 \\ & & 0 \\ & & 0 \\ & & 0 \end{bmatrix} \begin{bmatrix} & & 0 \\ & e^{tA} & 0 \\ & & \vdots \\ & & 0 \\ 0 & 0 & \cdots \quad 0 \quad 10^{59} \end{bmatrix} \begin{bmatrix} & & \\ & B & \\ & & \\ 0 & 0 & 0 \quad 0 \quad 0 \end{bmatrix} = Ce^{tA}B \qquad all\ t \in \mathbb{R}$$

Note: same two things works in discrete time.

Terminology: Given impulse response H of a time-invariant I/O system model, we say that (A, B, C, D) is a realization of the system (or of H) when:

$$H(t) = Ce^{tA}B\mathbb{1}(t) + D\delta(t) \qquad t \in \mathbb{R}$$

$$H(k) = CA^{k-1}B\mathbb{1}(k-1) + D\delta(k) \qquad k \in \mathbb{Z}$$

Say a time-invariant system is realizable when it has a realization.

So far have seen:

- any realizable system has infinitely many realizations
- any realizable system has realizations with an arbitrarily large n = number of state variables

Turns out: for any realizable I/O system, there's a lower bound on the number of state variables in any realization.

Terminology: (A, B, C, D) is a minimal realization of an I/O system when $size(A)$ is smallest possible amount all realizations of the system. Think of a non-minimal realization as having extraneous states - ie states that don't effect the system's I/O behavior

4.3.1 Realizations and Stability

Question we haven't addressed: When is a time-invariant I/O system realizable? Specifically, what properties of H corresponds to realizability?

Continuous time: say H is realizable - $H(T) = Ce^{tA}B\mathbb{1}(t) + D\delta(t)$. Take the Laplace Transform of H - call it $G(s)$, and the dubious methods shows that:

$$e^{tA}\mathbb{1}(t) \overset{L}{\leftrightarrow} (sI_n - A)^{-1}$$

$$H(t) \overset{L}{\leftrightarrow} C(sI_n - A)^{-1}B + D$$

Discrete time: $H(k) = CA^{k-1}B\mathbb{1}(k-1) + D\delta(k)$

$$A^{k-1}\mathbb{1}(k-1) \overset{z}{\leftrightarrow} (zI_n - A)^{-1}$$

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

Answer / Terminology: in each case G is the Transfer Function of the system (which is a $(p \times m)$ matrix valued function).

Observe: entries in $G(s)$ and $G(z)$ are proper rational functions of s and z. So, $F(s) = \frac{p(s)}{q(s)}$ and $degree(p(s)) \leq degree(q(s))$, for strict proper rational function, $\leq \to <$.

Note: Entries in $(sI_n - A)^{-1}$ and $(zI_n - A)^{-1}$ are strictly proper rational functions. Adding the D term at worst makes these entries proper, but still rational.

$\Rightarrow$ Thus, if system is realizable, then the transfer function has proper rational entries (in continuous and discrete time). The converse is also true!

- The converse is also true: If transfer function entries are proper rational, then realizable
  - Look first at scalar ($m = p = 1$) case and look at continuous time because the discrete time algebra works out exactly the same.
  - We have:

$$G(s) = \frac{p(s)}{q(s)} \qquad degree(p(s)) \leq degree(q(s))$$

· Let $D = \lim_{s\to\infty} G(s)$, (note $D = 0$ when $degree(p) < degree(q)$)

Then,

$$G_0(s) = G(s) - D = \frac{r(s)}{q(s)} \qquad deg(r) < deg(q)$$

$$= \frac{r_1 s^{n-1} + r_2 s^{n-2} + \cdots + r_{n-1}s + r_n}{s^n + q_1 s^{n-1} + \cdots + q_{n-1}s + q_n}$$

Let,

$$A = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ 0 & 0 & 0 & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & 1 \\ -q_n & -q_{n-1} & \cdots & -q_2 & -q_1 \end{bmatrix}, \; B = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}, \; C = \begin{bmatrix} r_n & r_{n-1} & \cdots & r_2 & r_1 \end{bmatrix}$$

· Easy to show,

$$G_0(s) = C(sI_n - A)^{-1}B$$

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

showing that (A, B, C, D) realizes the system

- Example, n=2 case

$$sI_n - A = \begin{bmatrix} s & -1 \\ q_2 & s + q_1 \end{bmatrix}, \qquad (sI_n - A)^{-1} = \frac{1}{s^2 + q_1 s + q_2}\begin{bmatrix} s + q_1 & 1 \\ -q_2 & s \end{bmatrix}$$

$$(sI_n - A)^{-1}B = \frac{1}{s^2 + q_1 s + q_2}\begin{bmatrix} 1 \\ s \end{bmatrix}, \qquad \Rightarrow \frac{r_1 s + r_2}{s^2 + q_1 s + q_2}$$

General Case ($m > 1$ and or $p > 1$):

- let $D = \lim_{s\to\infty} G(s)$, giving strictly proper rational function

$$G_0(s) = G(s) - D$$

- Then let $q(s) = s^n + q_1 s^{n-1} + \cdots + q_{n-1}s + q_n$ be the lowest common denominator of all the entries in $G_0(s)$
- Then, $R(s) = q(s)G_0(s)$ is a $(p \times m)$ matrix of polynomials of $degree \leq n - 1$
- So,

$$R(s) = R_1 s^{n-1} + R_2 s^{n-2} + \cdots + R_{n-1}s + R_n$$

Where each R_j is a $p \times m$ contant matrix

- See that the following A, B with $(m \times m)$ blocks and C with $(p \times m)$ blocks emerge

$$A = \begin{bmatrix} 0_m & I_m & 0_m & \cdots & 0_m \\ 0_m & 0_m & I_m & \cdots & 0_m \\ 0_m & 0_m & 0_m & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & I_m \\ -q_n I_m & -q_{n-1} I_m & \cdots & -q_2 I_m & -q_1 I_m \end{bmatrix}, \; B = \begin{bmatrix} 0_m \\ 0_m \\ \vdots \\ 0_m \\ I_m \end{bmatrix}, \; C = \begin{bmatrix} R_n & R_{n-1} & \cdots & R_2 & R_1 \end{bmatrix}$$

- Easy to show,

$$G_0(s) = C(sI_{nm} - A)^{-1}B$$

$$G(s) = C(sI_{nm} - A)^{-1}B + D$$

showing that (A, B, C, D) is a realization of the system

$\Rightarrow$ System is realizable *if and only if* the transfer function is proper rational

Next, want to talk about minimal realizations of realizable I/O systems - realizations with the smallest possible size A matrix (fewest number of states).

Observation: (A, B, C, D) and $(\hat{A}, \hat{B}, \hat{C}, \hat{D})$ realize the same system *if and only if* $\hat{D} = D$, $CA^{k-1}B = \hat{C}\hat{A}^{k-1}\hat{B}$ for all $k \geq 1$, both in continuous and discrete time. To see why this is an iff statement, take the derivative with respect to t of each side and evaluate at $t = 0$ without the delta functions terms.

$$\mathbf{DT}: CA^{k-1}B\mathbb{1}(k-1) + D\delta(k) = \hat{C}\hat{A}^{k-1}\hat{B}\mathbb{1}(k-1) + \hat{D}\delta(k)$$

$$\mathbf{CT}: Ce^{tA}B\mathbb{1}(t) + D\delta(t) = \hat{C}e^{t\hat{A}}\hat{B}\mathbb{1}(t) + \hat{D}\delta(t)$$

4.3.2 Hankel Matrix

Terminology: set of $CA^{k-1}B$, $k \geq 1$ called the Markov Parameter of the I/O system. Given A, B, C of sizes $(n \times n), (n \times m), (p \times m)$, define the Hankel Matrix associated with these of degree n as a $(p \times m)$ block matrix of form

$$H_n = \begin{bmatrix} CB & CAB & CA^2B & \cdots & CA^{n-1}B \\ CAB & CA^2B & & & \\ CA^2B & & & & \vdots \\ \vdots & & & \ddots & \\ CA^{n-1}B & & \cdots & & CA^{2(n-1)}B \end{bmatrix}$$

$$H_n = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} \begin{bmatrix} B & AB & \cdots & A^{n-1}B \end{bmatrix} = Q_0(A,C) \cdot Q_r(A,B)$$

If (A,B) reachable *and* (A,C) observable, then $rank(H_n) = n$. This is because $rank(Q_r) = n = rank(Q_0)$, and multiplying matrixes only maintains or reduces rank. Alternatively, because $rank(Q_r) = n$, so $range(Q_r)$ has dimension n, so $null(Q_0) = 0$, so $Q_0 \times (range(Q_r))$ has dimension n in the space $\mathbb{R}^{np}$.

4.3.3 Fundamental Theorem of Realizations

Fundamental Theorem of I/O Realization Theory:
(A, B, C, D) is a minimal realization of the associated I/O system *if and only if* (A, B) reachable and (A, C) observable.

First show: If reachable and observable, then minimal
- If reachable and observable, and $A \in \mathbb{R}^{n \times n}$, then Hankel matrix has rank n
- Suppose had $(\hat{A}, \hat{B}, \hat{C}, \hat{D})$ beign another realization with a smaller $\hat{A} \in \mathbb{R}^{r \times r},\ r < n$.
- Since $\hat{C}\hat{A}^{k-1}\hat{B} = CA^{k-1}B$ all $k \geq 1$,

$$H_n = \begin{bmatrix} \hat{C} \\ \hat{C}\hat{A} \\ \vdots \\ \hat{C}\hat{A}^{n-1} \end{bmatrix} \begin{bmatrix} \hat{B} & \hat{A}\hat{B} & \cdots & \hat{A}^{n-1}\hat{B} \end{bmatrix}$$

- The above Hankel matrix is the product of an $(i \times r)$ matrix and a $(r \times j)$ matrix, it will have $rank(H_n) \leq r < n$.

→ Thus, because (A,B) reachable and (A,C) observable, $rank(H_n) = n$ by definition and there is a contradiction!

Converse Proof: If minimal, then reachable and observable
- Suppose (A,B,C,D) is minimal, say $A \in \mathbb{R}^{n \times n}$
- Recall that for any invertible $\mathbb{X} \in \mathbb{R}^{n \times n}$, $(\mathbb{X}^{-1}A\mathbb{X}, \mathbb{X}^{-1}B, C\mathbb{X}, D)$ is another realization
- Suppose (A,B) not realizable, then $\{reachable\ states\}$ is proper subset of $\mathbb{R}^n$ - say $dim = r < n$
- Let $\mathbb{X}$ be an invertible $(n \times n)$ matrix where first r columns form a basis for $\{reachable\ states\}$

- let $\hat{B} = \mathbb{X}^{-1}B$; note then $B = \mathbb{X}\hat{B}$. Then, because $range(B) \subset \{reachable\ states\}$,

$$\hat{B} = \begin{bmatrix} B_1 \\ 0 \end{bmatrix}$$

Where the first r rows of $\hat{B}$ are the original matrix B, and the last $n - r$ are 0.

- Because $\{reachable\ states\}$ invariant under A, form a block matrix $A\mathbb{X}$ where $A_1 \in \mathbb{R}^{r\times r}$ and the rest of the zeros and don't cares fill $A\mathbb{X} \in \mathbb{R}^{n\times n}$

$$A\mathbb{X} = \mathbb{X}\begin{bmatrix} A_1 & /// \\ O & /// \end{bmatrix} \rightarrow \mathbb{X}^{-1}A\mathbb{X} = \begin{bmatrix} A_1 & /// \\ O & /// \end{bmatrix}$$

- Partition $C\mathbb{X}$ as follows

$$C\mathbb{X} = \begin{bmatrix} C_1 & /// \end{bmatrix}$$

- We know $(\mathbb{X}^{-1}A\mathbb{X}, \mathbb{X}^{-1}B, C\mathbb{X}, D)$ $(\hat{A}, \hat{B}, \hat{C}, \hat{D})$ also realizes the system, thus

$$\hat{C}\hat{A}^{k-1}\hat{B} = CA^{k-1}B \qquad k \geq 1$$

- Because of hatted matrices special forms,

$$\hat{C}\hat{A}^{k-1}\hat{B} = C_1A_1^{k-1}B_1 \qquad k \geq 1$$

$\Rightarrow$ (A_1, B_1, C_1, D) also realizes the system $\rightarrow$ contradiction! because we assumed (A, B, C, D) minimal and A_1 is smaller than A

- Similarly, assuming (A,C) not observable leads via an $\mathbb{X}$-maneuver to a smaller A realization... contradiction

$\Rightarrow$ if minimal, then reachable and observable

4.3.4 State Space Isomorphism Theorem

State Space Isomorphism Theorem: Any two minimal realizations of the same time invariant I/O system are related by an $\mathbb{X}$-maneuver. Ie if (A, B, C, D) and $(\hat{A}, \hat{B}, \hat{C}, \hat{D})$ realize (minimally) to same systems, then there exists an invertible $\mathbb{X} \in \mathbb{R}^{n\times n}$ such that

$$\hat{A} = \mathbb{X}^{-1}A\mathbb{X}, \qquad \hat{B} = \mathbb{X}^{-1}B, \qquad \hat{C} = C\mathbb{X}, \qquad \hat{D} = D$$

- Start with, because hatted and unhatted things realize the same I/O system, they give rise

to the same hatted matrix H_n

$$H_n = \begin{matrix} Q_0(A,C) \cdot Q_r(A,B) = Q_0 Q_r \\ Q_0(\hat{A},\hat{C}) \cdot Q_r(\hat{A},\hat{B}) = \hat{Q}_0 \hat{Q}_r \end{matrix}$$

- Because minimal, $Q_r Q_r^T$ invertible, $\hat{Q}_r \hat{Q}_r{}^T$ also, $Q_0 Q_0^T$ also, $\hat{Q}_0 \hat{Q}_0{}^T$ also (all of size n by n), giving

$$Q_0 Q_r = \hat{Q}_0 \hat{Q}_r$$

$$Q_0 Q_r \hat{Q}_r^T = \hat{Q}_0 \hat{Q}_r \hat{Q}_r^T$$

$$\rightarrow \hat{Q}_0 = Q_0 \underline{Q_r \hat{Q}_r^T (\hat{Q}_r \hat{Q}_r^T)^{-1}}$$

- And

$$Q_0 Q_r = \hat{Q}_0 \hat{Q}_r$$

$$Q_0^T Q_0 Q_r = Q_0^T \hat{Q}_0 \hat{Q}_r$$

$$\rightarrow Q_r = \underline{(Q_0^T Q_0)^{-1} Q_0^T \hat{Q}_0} \hat{Q}_r$$

- Where both of the underline portions are equal to $\mathbb{X}$, and we see

$$Q_r = \mathbb{X}\hat{Q}_r \rightarrow \hat{Q}_r = \mathbb{X}^{-1} Q_r \rightarrow \hat{B} = \mathbb{X}^{-1} B$$

$$\hat{Q}_0 = Q_0 \mathbb{X} \rightarrow \hat{C} = C\mathbb{X}$$

- Remains to verify that $\hat{A} = \mathbb{X}^{-1} A \mathbb{X}$. One does this by verifying $\mathbb{X}\hat{A} = A\mathbb{X}$, which is the same as $\hat{Q}_0 \hat{A} \hat{Q}_r = Q_0 A Q_r$. Details of this in the handout.

4.3.5 Canonical Structure Theorem

Canonical Structure Theorem: Given any (A,B,C,D), need not be reachable or observable, there's an $\mathbb{X}$-maneuver that turns (A,B,C,D) into

$$\hat{A} = \mathbb{X}^{-1} A \mathbb{X} = \begin{bmatrix} A_{11} & A_{12} & A_{13} & A_{14} \\ 0 & A_{22} & 0 & A_{24} \\ 0 & 0 & \underline{A_{33}} & A_{34} \\ 0 & 0 & 0 & A_{44} \end{bmatrix}, \; \hat{B} = \mathbb{X}^{-1} B = \begin{bmatrix} B_1 \\ 0 \\ \underline{B_3} \\ 0 \end{bmatrix}, \; \hat{C} = C\mathbb{X} = \begin{bmatrix} 0 & 0 & \underline{C_3} & C_4 \end{bmatrix}, \; \hat{D} = D$$

and (A_{33}, B_3, C_3, D) is a minimal realization of the I/O system realized by (A, B, C, D). Ie,

$$C(sI_n - A)^{-1} B + D = C_3 (sI - A_{33})^{-1} B_3 + D$$

Basic idea, form $\mathbb{X}$ as follows:

- First few columns of $\mathbb{X}$ are basis for $\{reachable\ states\} \cap \{unobservable\ states\}$
- Next few complete first few to form a basis for $\{unobservable\ states\}$
- Next again, together with first few, make a basis for $\{reachable\ states\}$
- Final few, together with first few, make a basis for $\mathbb{R}^n$

Because...

- set $\{reachable\ states\}$ contains the range of B and is invariant under A

  AND
- set $\{unobservable\ states\} \subset nullspace(C)$ and is invariant under A, the matrices $\hat{A}, \hat{B}, \hat{C}, \hat{D}$ take prescribed forms

Note: From this, you can build an algorithm to take an arbitrary realization to a minimal realization.

Idea: Gauss elimination gives you $nullspace(Q_0(A, C))$, $range(Q_r(A, B))$, *etc.* Turns out this is inefficient, and better ways exists such as the Kalman-Ho Algorithm, Silverman's Algorithm, etc.

4.4 Stability for I/O Systems

Restrict attention to time-invariant I/O systems characterized by impulse response H

4.4.1 BIBO Stable and Realizations

→ H system is Bounded Input Bounded Output (BIBO) Stable when every bounded input signal u leads to a well defined all-time response y that's also a bounded function. ie,

CT BIBO Stable when for every $u : \mathbb{R} \mapsto \mathbb{R}^m$, satisfying $\|u(t)\| \leq R_1$ for all $t \in \mathbb{R}$ for some $R_1 > 0$, y defined by all time response equation

$$y(t) = \int_{-\infty}^{\infty} H(t-\tau)u(\tau)d\tau \qquad t \in \mathbb{R}$$

is well defined and satisfies $\|y(t)\| \leq R_2$ for some $R_2 > 0$

DT BIBO Stable when for every $u : \mathbb{Z} \mapsto \mathbb{R}^m$, satisfying $\|u(k)\| \leq R_1$ for all $k \in \mathbb{Z}$ for some $R_1 > 0$, y defined by all time response equation

$$y(k) = \sum_{l=-\infty}^{\infty} H(k-l)u(k) \qquad k \in \mathbb{Z}$$

is well defined and satisfies $\|y(k)\| \leq R_2$ for some $R_2 > 0$

Fact: System with impulse response H is BIBO Stable *if and only if*

CT $\int_0^\infty \|H(t)\|\, dt$ exists (where $\|\ \|$ is standard matrix Euclidean Norm)

DT $\sum_{k=0}^{\infty} \|H(k)\|$ exists

Easy to show why these conditions imply BIBO Stability, but the converse is more difficult

- Suppose $\int_0^\infty \|H(t)\|\, dt$ exists, let $u : \mathbb{R} \mapsto \mathbb{R}^m$ satisfy $\|u(t)\| \leq R_1$ for all $t \in \mathbb{R}$
- Compute

$$y(t) = \int_{-\infty}^{\infty} H(t-\tau)u(\tau)d\tau$$

$$\|y(t)\| \leq \int_{-\infty}^{\infty} \|H(t-\tau)\|\, \|u(\tau)\|\, d\tau$$

$$\|y(t)\| \leq \Big(\int_{-\infty}^{\infty} \|H(t-\tau)\|\, d\tau \Big) \cdot R_1$$

$$\|y(t)\| \leq \Big(\int_{0}^{\infty} \|H(\rho)\|\, d\rho \Big) \cdot R_1 = R_2$$

- Can view $\int_0^\infty \|H(\rho)\|\, d\rho$ as system gain so to speak.
- A similar argument works in discrete time

Question: What if an I/O system is realizable? Ie, it has a proper rational transfer function $G(s)$ or $G(z)$

Fact: System is BIBO Stable *if and only if*

CT Every pole of every entry in G(s) has strictly negative real part

DT Every pole of every entry in G(z) has magnitude strictly less than 1

idea:

$$H(t) \overset{L}{\longleftrightarrow} G(s) \qquad\qquad H(k) \overset{z}{\longleftrightarrow} G(z)$$

- Meaning entries in

  H(t): $(\textit{polynomial in } t) \cdot e^{\lambda_0 t}$ $\quad$ λ_0 pole of $G(s)$

  H(k): $(\textit{polynomial in } k) \cdot \lambda_0^k$ $\quad$ λ_0 pole of $G(z)$
- and for those to absolutely integrable on $[0, \infty)$ or summable on $0 \leq k < \infty$, which is necessary for BIBO Stability, need $Re\{\lambda_0\} < 0$ or $|\lambda_0| < 1$

Relationship between BIBO Stability of I/O Systems and Lyapunov Stability for state space Realizations (all time invariant).

Fact:

1) if (A, B, C, D) realizes an I/O system, then

   CT: A Hurwitz $\rightarrow$ I/O system BIBO Stable

   DT: A Schur $\rightarrow$ I/O system BIBO Stable

   Reason: $\{poles\ of\ transfer\ function\} \subset \{eigenvalues\ of\ A\}$
2) if I/O system BIBO Stable and (A, B, C, D) realizes it minimally, then

   CT: A Hurwitz

   DT: A Schur

4.4.2 Stabilizable and Detectable

Stabilizable: The LTI system (A, B) is said to be stabilizable if all uncontrollable modes are stable, that is all eigenvalues in the right half plane still correspond to a full rank matrix below:

$$rank[sI - A \ I \ B] = n \qquad \forall s \in \sigma(A) \cap \mathbb{C}_+^o$$

or,

$$\forall \lambda_i \in \mathbb{C}_+^o, \qquad e_i \in sp\{C\} \qquad C: \ Controllability \ Matrix$$

Detectable: The LTI system (A, C) is said to be detectable if all unobservable modes are stable, that is all eigenvalues in the right half plane still correspond to a full rank matrix below:

$$rank\begin{bmatrix} sI - A \\ \hline C \end{bmatrix} = n \qquad \forall s \in \sigma(A) \cap \mathbb{C}_+^o$$

or,

$$\forall \lambda_i \in \mathbb{C}_+^o, \qquad e_i \notin N\{O\} \qquad O: \ Observability \ Matrix$$

4.5 Interconnects

4.5.1 Basic Typologies

Given two m-input p-output time invariant I/O systems with respective impulse responses H_1 and H_2 and transfer functions $G_1(s), G_2(s)$, consider the Parallel Connection:

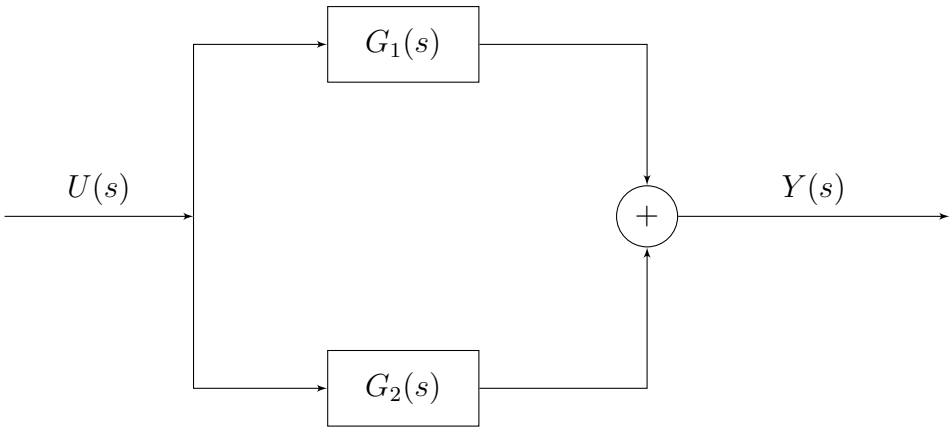


Figure 4.1: Parallel Connection

Overall transfer function $u \mapsto y$ is

$$G(s) = G_1(s) + G_2(s)$$

And overall impulse response

$$H = H_1 + H_2$$

$H_1(G_1)$ is impulse response (transfer function) or an m-input, q-output system and $H_2(G_2)\cdots$ q-input, p-output system can form the Cascade Connection:

Figure 4.2: Cascade Connection

Overall transfer function $u \mapsto y$ is, where G_2 is $p \times q$ and G_1 is $q \times m$, giving and $p \times m$ function. Order of multiplication matters with matrix functions.

$$G(s) = G_2(s)G_1(s)$$

And overall impulse response

$$H = H_1 * H_2 \qquad H(t) = \int_{-\infty}^{\infty} H_2(\tau)H_1(t-\tau)d\tau \qquad t \in \mathbb{R}$$

In either case, if both systems are BIBO stable, so is the interconnection. In general think about L^2-ness of the H's; for rational G's, think about the poles of G's.

Third, interconnections involving feedback are more interesting. Below, $u \in \mathbb{R}^m, y \in \mathbb{R}^p, G_1 : p \times m, G_2 : m \times p$

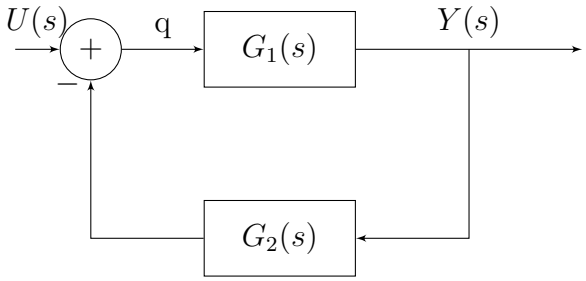


Figure 4.3: Feedback Interconnect

Overall transfer function $u \mapsto y$ is

$$Q(s) = U(s) - G_2(s)Y(s)$$

$$Y(s) = G_1(s)Q(s) = G_1(s)U(s) - G_1(s)G_2(s)Y(s)$$

$$(I_p + G_1(s)G_2(s))Y(s) = G_1(s)U(s)$$

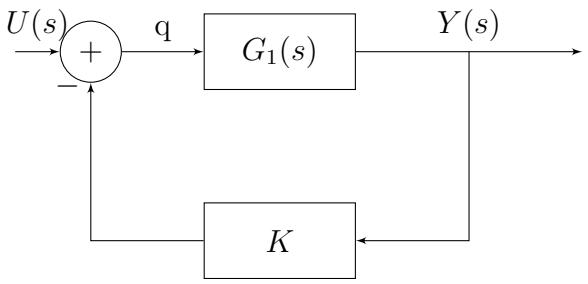
$$Y(s) = (I_p + G_1(s)G_2(s))^{-1}G_1(s)U(s)$$

$$\Rightarrow G(s) = (I_p + G_1(s)G_2(s))^{-1}G_1(s)$$

Caution: even when $G_1(s)$ and $G_2(s)$ transfer functions of BIBO Stable, feedback interconnection above might not be stable!

Example! $m = p = 1$

$$G_1(s) = \frac{s}{s^2 + 3s + 2} \qquad G_2(s) = K = constant\ gain$$



$$G(s) = \frac{G_1(s)}{1 + KG_1(s)} = \frac{s}{s^2 + (K+3)s + 2}$$

- For some k-values, eq $k = 1$, $G(s)$ has poles in the $Re\{s\} < 0 \rightarrow$ overall system BIBO Stable.
- For $k = -3, \pm j\sqrt{2}$ are poles of $G(s)$ - not BIBO Stable.
- Poles both complex with positive real part when $K < -3$

$\rightarrow$ In any event, can pursue this further, but not for us...

4.5.2 H^∞ Design

We'll talk briefly about H^∞ control system design.

First, given BIBO stable system with transfer function $G(s)$ or $G(z)$, the Frequency Response (matrix) of the system is the $p \times m$ matrix valued function:

CT: Transfer function evaluated on the imaginary axis:

$$\omega \mapsto G(j\omega) \qquad \omega \in \mathbb{R}$$

DT: Transfer function evaluated on the unit circle:

$$\omega \mapsto G(e^{j\omega}) \qquad \omega \in \mathbb{R}$$

Terminology: $\underline{H^\infty\text{-norm}}$ of a BIBO Stable system is

CT: $\sup_{\omega\in\mathbb{R}} \sigma_{max}(G(j\omega))$

DT: $\sup_{\omega\in\mathbb{R}} \sigma_{max}(G(e^{j\omega}))$

Where,

$$\sigma_{max}(matrix\ A) = largest\ singular\ value\ of\ A = \sqrt{Largest\ eigenvalue\ of\ A^T A}$$

$$\rightarrow H^\infty - norm \approx max\ possible\ system\ gain \approx \max_{\omega\in\mathbb{R}}\ of\ 2-norm\ of\ G(j\omega)\ or\ G(e^{j\omega})$$

Turns out, this is indeed a norm on $\{BIBO\ Stable\ Systems\}$

Paradigmatic H^∞ control setup:

$d =$ exogenous inputs plus disturbances

$e =$ error signal that you want to be small

- (In general, exogenous signal is something you want to track. Ie, can formulate lots of control problems as keeping some signal e close to zero)

$P =$ the plant - can't do anything with this

$K =$ the controller - a user choice

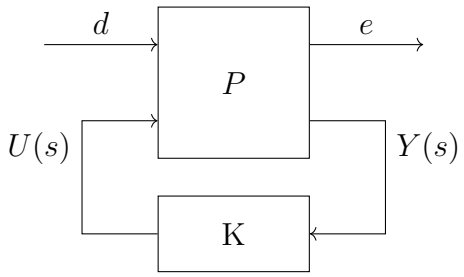


Figure 4.4: Paradigmatic Example

Overarching Goal:

1. K BIBO Stable
2. e stays small

H^∞ Formulation:

· First model P as

$$P = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} - transfer\ function\ block\ matrices$$

ie, get a matrix equation below where P_{11} is open loop transfer function $d \mapsto e$

$$\begin{bmatrix} E(s) \\ Y(s) \end{bmatrix} = \begin{bmatrix} P_{11}(s) & P_{12}(s) \\ P_{21}(s) & P_{22}(s) \end{bmatrix} \begin{bmatrix} D(s) \\ U(s) \end{bmatrix}$$

· But, we're choosing to set $U(s) = K(s)Y(s)$. Plug that in for $U(s)$ - find overall $d \mapsto e$. Closed loop transfer function calculated via:

$$E = P_{11}D + P_{12}U = P_{11}D + P_{12}KY$$

$$Y = P_{21}D + P_{21}U = P_{21} + P_{22}KY$$

$$\downarrow$$

$$Y = (I - P_{22}K)^{-1}P_{21}D$$

$$E = (P_{11} + P_{12}K(I - P_{22}K)^{-1}P_{21})D$$

$$\Rightarrow G = (P_{11} + P_{12}K(I - P_{22}K)^{-1}P_{21})$$

· Take Overarching Goal above and formulate it as a optimization problem
  - Choose K so K and G above are BIBO Stable *AND* H^∞-norm of G is as small as possible.
  - Constructed minimization problem - could be difficult - but people have, over the years, cleverly $K-parameterized$ things to turn this into a convex optimization problem

4.5.3 Observer Construction

One other feedback related item: Call the observer construction, given

$$\dot{x} = Ax + Bu, \qquad y = Cx$$

Want to implement a state feedback $u = -Kx + v$, where v is an exogenous signal, but don't have x. Build an observer

$$\dot{w} = Aw + Bu - L(Cw - y)$$

Find if $(A - LC)$ is Hurwitze, w tracks x asymptotically when you use $u = -Kw + v$ in both systems.

Look at what's happening from an I/O Standpoint
  - Have

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

- Wish to get system with transfer function

$$G(s) = C(sI_n - (A - BK))^{-1}B$$

$v \mapsto y$ transfer function if we could implement state feedback $u = -Kx$ exactly

Question: What is the $v \mapsto y$ transfer function?

- Re-write system equations to have a $2n$-dimensional state $[x \quad w]^T$, output y, and input v

$$\dot{x} = Ax - BKw + BV \qquad Y = Cx$$

$$\dot{w} = Aw - BKw + Bv - LCw + LCx$$

$$\frac{d}{dt}\begin{bmatrix} \dot{x} \\ \dot{w} \end{bmatrix} = \begin{bmatrix} A & -BK \\ LC & (A - LC - BK) \end{bmatrix}\begin{bmatrix} x \\ w \end{bmatrix} + \begin{bmatrix} B \\ B \end{bmatrix} v; \qquad Y = \begin{bmatrix} C & 0 \end{bmatrix}\begin{bmatrix} x \\ w \end{bmatrix}$$

- The matrices above are the A', B', and C' of the $2n$-dimensional state $[x \quad w]^T$
- Perform $\mathbb{X}$-maneuver on system with

$$\mathbb{X} = \begin{bmatrix} I_n & I_n \\ 0 & I_n \end{bmatrix} \qquad \mathbb{X}^{-1} = \begin{bmatrix} I_n & -I_n \\ 0 & I_n \end{bmatrix}$$

$$C'\mathbb{X} = \begin{bmatrix} L & C \end{bmatrix} \qquad \mathbb{X}^{-1}B' = \begin{bmatrix} 0 \\ B \end{bmatrix}$$

$$A'\mathbb{X} = \begin{bmatrix} A & A - BK \\ LC & A - BK \end{bmatrix} \qquad \mathbb{X}^{-1}A'\mathbb{X} = \begin{bmatrix} A - LC & 0 \\ LC & A - BK \end{bmatrix}$$

- $v \mapsto y$ transfer function

$$C'(sI_{2n} - A')^{-1}B' = C\mathbb{X}(sI_{2n} - \mathbb{X}^{-1}A'\mathbb{X})^{-1}(\mathbb{X}^{-1}B)$$

$$sI_{2n} - \mathbb{X}^{-1}A'\mathbb{X} = \begin{bmatrix} sI_n - (A - LC) & 0 \\ -LC & sI_n - (A - BK) \end{bmatrix}$$

$$(sI_{2n} - \mathbb{X}^{-1}A'\mathbb{X})^{-1} = \begin{bmatrix} //////// & 0 \\ //////// & (sI_n - (A - BK))^{-1} \end{bmatrix}$$

$$(sI_{2n} - \mathbb{X}^{-1}A'\mathbb{X})^{-1}(\mathbb{X}^{-1}B') = \begin{bmatrix} 0_n \\ (sI_n - (A - BK))^{-1}B \end{bmatrix}$$

$$C'\mathbb{X}(sI_{2n} - \mathbb{X}^{-1}A'\mathbb{X})^{-1}(\mathbb{X}^{-1}B') = C(sI_n - (A - BK))^{-1}B$$

$\rightarrow$ $v \mapsto y$ transfer function is $C(sI_n - (A - BK))^{-1}B$, which is the same as the $v \mapsto y$ transfer

function of the system:

$$\dot{x} = Ax + B(-kx + v)$$

$$y = Cx$$

which is the system with series observer interpretation of the state feedback control law!

The Story! Start with $\dot{x} = Ax + Bu$; have exogenous input v, and you want to adjust $v \mapsto y$ transfer function by using that feedback, $u = -kx + v$; don't have access to x; use an observer to implement control law approximately; *voila*, you have the exact $v \mapsto y$ transfer function you wanted. The block diagram for the overall system is easy ish to derive.

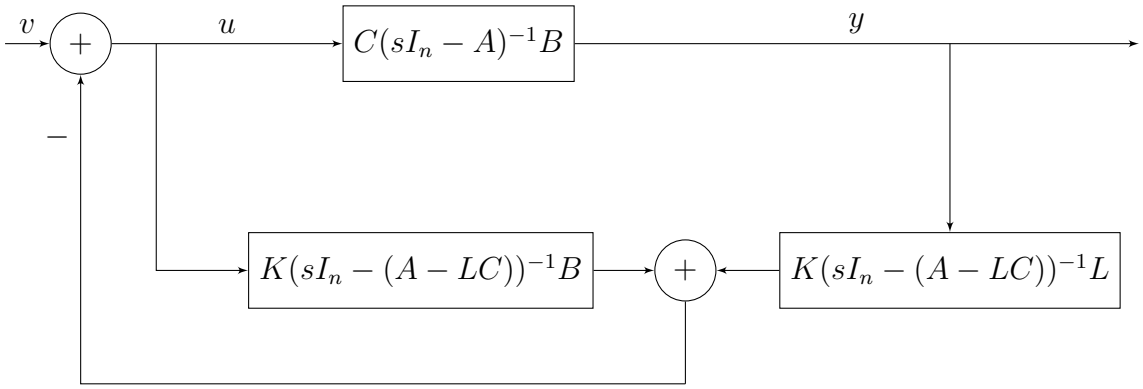


Figure 4.5: Controller-Observer Implementation

- Note, since $(A - LC)$ Hurwitz, feedback loop systems are all BIBO Stable
- Known as (one version of) Controller-Observer Implementation of

$$\dot{x} = Ax + Bu, \qquad u = -kx + v$$

$$\dot{w} = Aw + Bu + L(Cx - Cw), \qquad y = Cx$$

- See:

$$W(s) = (sI_n - (A - LC))^{-1}(BU(s) + LCX(s))$$

$$U(s) = -KW(s) + V(s)$$

$$U(s) \overset{\checkmark}{=} -K(sI_n - (A - LC))^{-1}BU(s) - K(sI_n - (A - LC))^{-1}LY(s) + V(s)$$

Chapter 5

Linear Quadratic Regulator

→ THE paradigmatic solved problem in optimal control, **terminology**: Optimal Control: Choosing control function to minimize some cost function; all this associated with some system connecting input to cost. Look first at discrete time version - simplest form as follows - note that the matrices (A, B, C) need not be constant for now (later they will be).

5.1 Finite {Time Horizon} Problem

5.1.1 Introduction

- Given

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

$$Y(k) = Cx(k), \qquad x(0) + x_0$$

along with $k_1 > 0$ and a nonnegative definite matrix $Q \in \mathbb{R}^{n \times n}$

- Choose $u(0), u(1), \cdots, u(k_1 - 1)$ to minimize this cost function:

$$\sum_{k=0}^{k_1-1} \left(\|u(k)\|^2 + \|y(k)\|^2 \right) + x^T(k_1) Q x(k_1)$$

- Essentially, choose u to keep y small - want u to be relatively small as well - $x^T(k_1)Qx(k_1) \geq 0$ penalizes final state
- Could conceivably rewrite $x(k_1)$ and $y(0), \cdots, y(k_1 - 1)$ in terms of x_0 and u-values - get big function of u-values - take derivative and set equal to 0, etc etc. This is not efficient nor revealing of cool structure.

5.1.2 Discrete Time Solution

- Approach akin to *dynamic programming*, but different:
  - Suppose $u(0), u(1), \cdots, u(k_1-1)$ is an optimizing choice of u; using this u lands you at state $x(k_1-1)$ at time k_1-1
  - Last value of u - ie $u(k_1-1)$ - must be the best possible value of $u(k_1-1)$ given that $x(k_1-1)$ value.
  - Sitting at time k_1-1; you're looking at choosing $u(k_1-1)$ to minimize

$$\|u(k_1-1)\|^2 + \|y(k_1-1)\|^2 + x^T(k_1)Qx(k_1)$$

    These are the only terms in the cost that you could conceivably affect by choice of $u(k_1-1)$ - in fact, can't effect $\|y(k_1-1)\|^2 = \|Cx(k_1-1)\|^2$
  - Thus, at time k_1 want $u(k_1-1)$ to minimize

$$\|u(k_1-1)\|^2 + x^T(k_1)Qx(k_1)$$

$$= u^T(k_1-1)u(k_1-1) + [Ax(k_1-1)+Bu(k_1-1)]^TQ[Ax(k_1-1)+Bu(k_1-1)]$$

  - Take derivative and set equal to 0:

$$u^T(k_1-1) + [Ax(k_1-1)+Bu(k_1-1)]^TB = 0$$

$$\rightarrow u(k_1-1) + B^T[Ax(k_1-1)+Bu(k_1-1)] = 0$$

$$(I_m + B^TQB)u(k_1-1) = -B^TQAx(k_1-1)$$

$$u(k_1-1) = -(I_m + B^TQB)^{-1}B^TQAx(k_1-1)$$

    noting that $(I_m + B^TQB)$ is positive definite and therefore inveritble
  - This is the choice you must make at time k_1-1 given you've learned that $x(k_1-1)$
  - (Can re-work this as $u(k_1-1) = -F(k_1-1)x(k_1-1)$)
  - Now consider: what $u(k_1-2)$ do you use given that you've landed at stable $x(k_1-2)$ at time k_1-2? We know that we'll choose $u(k_1-1)$ as above. The terms we can affect in the cost function by choice of $u(k_1-2)$ are

$$\|u(k_1-2)\|^2 + \big(\|u(k_1-1)\|^2 + \|y(k_1-1)\|^2 + x^T(k_1)Qx(k_1)\big)$$

$$u(k_1-1) = -F(k_1-1)x(k_1-1), \qquad y(k_1-1) = Cx(k_1-1)$$

$$x(k_1) = (A - BF(k_1-1))x(k_1-1)$$

· Thus, the parenthesized term is

$$x^T(k_1-1)\Big[F^T(k_1-1)F(k_1-1)+C^TC+(A-BF(k_1-1))^TQ(A-BF(k_1-1))\Big]x(k_1-1)$$

$$= x(k_1-1)^TP(k_1-1)x(k_1-1)$$

where P is nonnegative definite again...

· Thus, we want $u(k_1-2)$ to minimize

$$\|u(k_1-2)\|^2 + x^T(k_1-1)P(k_1-1)x(k_1-1)$$

which looks a lot like the $u(k_1-1)$ optimized problem, and we know the answer:

$$u(k_1-2) = -F(k_1-2)x(k_1-2)$$

$$= -(I_m + B^TP(k_1-1)B)^{-1}B^TP(k_1-1)Ax(k_1-2)$$

→ For each time $k_1-1, k_1-2, \cdots$ you find that you're solving the same problem: Choose u to minimize $\|u\| + (next\ x)^T(some\ P)(next\ x)$

- Keep this up: if you define $k \mapsto P(k)$ as the solution to the difference equation

$$P(k-1) = F^T(k-1)F(k-1) + C^TC + (A-BF(k-1))^TP(k)(A-BF(k-1))$$

$$P(k_1) = Q\ (final\ condition)$$

and for each $k \le k_1$

$$F(k-1) = (I_m + B^TP(k)B)^{-1}B^TP(k)A$$

- (Can solve this difference equation offline - get $k \mapsto P(k)$ and $k \mapsto F(k)$)
- Then, the optimal control, in feedback form,

$$u(k) = -F(k)x(k) \qquad 0 \le k < k_1$$

a time-varying state feedback control law

- You can write it in *open-loop form* as follows:

→ When you apply the optimal control, x evolves according to

$$x(k+1) = (A-BF(k))x(k), \qquad k \ge 0, \qquad x(0) = x_0$$

$$\mathcal{A}(k) = (A-BF(k))$$

so,

$$x(k) = \Phi_{\mathcal{A}}(k,0)x_0 \qquad 0 \le k < k_1$$

$$u(k) = -F(k)\Phi_{\mathcal{A}}(k,0)x_0 \qquad 0 \le k < k_1$$

- Difference Equation (not it's *nonlinear*) for $k \mapsto P(k)$ is the discrete time Riccoti Equation associated with the LQR Problem
- Again, this solution works even when A, B, C are time varying!

5.1.3 Continuous Time Solution

Given:

$$\dot{x}(t) = Ax(t) + Bu(t), \qquad x(0) = x_0$$

$$y(t) = Cx(t)$$

Choose $u : [0, t_1] \mapsto \mathbb{R}^m$ to minimize

$$\int_0^{t_1} (\|u(t)\|^2 + \|y(t)\|^2)dt + x^T(t_1)Qx(t_1)$$

- Strategy:
  - *motivate* the solution by taking a limit of the problem - this yields a proposed solution
  - check that the proposed solution is the solutions
- Discretize as follows: let h be small (we'll let it go to 0 later), $t_1 = k_1 h$; give discretized system dynamics

$$\frac{x\big((k+1)h\big) - x(kh)}{h} = Ax(kh) + Bu(kh)$$

$$x\big((k+1)h\big) = (I_n + hA)x(kh) + hBu(kh), \qquad x(0h) = x_0$$

$$= \bar{A}x(kh) + \bar{B}u(kh), \qquad Y(kh) = Cx(kh)$$

and divested cost

$$\sum_{k=0}^{k_1-1} h\Big(\|u(kh)\|^2 + \|y(kh)\|^2 \Big) + x^T(k_1 h)Qx(k_1 h)$$

- Analyze this like the discrete time problem with matrices $\bar{A}, \bar{B}, \bar{C}$ and discover optimal control is:

$$hu(kh) = -(I_m + \bar{B}^T\bar{P}\big((k+1)h\big)\bar{B})^T\bar{B}^T\bar{P}\big((k+1)h\big)\bar{A}x(kh)$$

and we want a definition for $\bar{P}\big((k+1)h\big)$. Use $\bar{B} = hB$,

$$u(kh) = -(I_m + \bar{B}^T\bar{P}\big((k+1)h\big)\bar{B})^T B^T\bar{P}\big((k+1)h\big)\bar{A}x(kh)$$

$$= -\bar{F}(kh)x(kh) \qquad 0 \le k < k_1$$

P's arise from difference equation with final condition $\bar{P}(k_1 h) = Q$

$$\bar{P}(kh) = \bar{F}^T(kh)\bar{F}(kh) + C^T C + (\bar{A} - \bar{B}\bar{F}(kh))^T \bar{P}\big((k+1)h\big)(\bar{A} - \bar{B}\bar{F}(kh))$$

- Now, take the limit as h goes to 0; first look at $\bar{F}(kh)$

$$(I_m + \bar{B}^T \bar{P}\big((k+1)h\big)\bar{B})^{-1} = (I_m + h^2 B^T \bar{P} B)^{-1} \to I_m$$

$$B^T \to B^T, \bar{P}\big((k+1)h\big) \to P(t)$$

$$\bar{A} = I_n + hA \to I_n$$

Thus,

$$\bar{F}(kh) \to B^T P(t)$$

which models the form (for the open loop control state feedback)

$$u(t) = -F(t)x(t) = -B^T P(t)x(t)$$

- Taking the limit of $\bar{P}(kh)$ equation, keeping only the 0^{th} and 1^{st} order terms in h

$$\bar{F}^T(kh)\bar{F}(kh) \to P\big((k+1)h\big)BB^T P\big((k+1)h\big), \qquad C^T C \to C^T C$$

Last term,

$$(I_n + hA - hBB^T P\big((k+1)h\big))^T P\big((k+1)h\big)(I_n + hA - hBB^T P\big((k+1)h\big))$$

$$\downarrow$$

$$\bar{P}\big((k+1)h\big) + h\bar{P}\big((k+1)h\big)A - hA^T \bar{P}\big((k+1)h\big)$$

$$-hP\big((k+1)h\big)BB^T P\big((k+1)h\big) - h\bar{P}\big((k+1)h\big)BB^T \bar{P}\big((k+1)h\big)$$

$$\Rightarrow \bar{P}(kh) - \bar{P}\big((k+1)h\big) \approx h\bar{P}\big((k+1)h\big) + hA^T \bar{P}\big((k+1)h\big) - h\bar{P}\big((k+1)h\big)BB^T \bar{P}\big((k+1)h\big) + hC^T C$$

divide by h, and take limit to 0

$$\dot{P}(t) = -\Big(P(t)A + A^T P(t) - P(t)BB^T P(t) + C^T C\Big), \qquad P(t_1) = Q$$

Riccati Differential Equation (RIC)

Bottom line so far is that we have the propose solution. Let $t \mapsto P(t)$, $t \in [0, t_1]$ satisfying

(RIC), then the optimal control in feedback form is given as

$$u(t) = -B^T P(t)x(t) \qquad t \in [0, t_1]$$

Caution: all predicated on existence of required $t \mapsto P(t)$ satisfying (RIC)... deal with later.

Now, verify it's actually the solution. Proceed as follows:

- use input $u(t) = -B^T P(t)x(t) + v(t)$ for $t \in [0, t_1]$
- discover that the best choice of $v = 0$
- given choice of u:

$$\|u\|^2 = \left\|-B^T Px + v\right\|^2 = X^T PBB^T Px + \|v\|^2 - x^T PBv - v^T BPx$$

$$\dot{x} = Ax + Bu = Ax - BB^T Px + Bv \Rightarrow Bv = -\dot{x} - Ax + BB^T Px$$

Thus,

$$\|u\|^2 = X^T PBB^T Px + \|v\|^2 - x^T P\big(\dot{x} - Ax + BB^T Px\big) - \big(\dot{x} - Ax + BB^T\big)^T Px$$

$$= -X^T PBB^T Px - x^T P\dot{x} - \dot{x}^T Px + x^T\big(PA + A^T P\big)x + \|v\|^2$$

$$\|y\|^2 = x^T C^T Cx$$

So,

$$\|u\|^2 + \|y\|^2 = -x^T P\dot{x} - \dot{x}^T Px - x^T\big(- PA - A^T P + PBB^T P - C^T C\big)x + \|v\|^2$$

$$= -\frac{d}{dt}(x^T Px) + \|v\|^2 \qquad when\ \dot{P} = -PA - A^T P + PBB^T P - C^T C$$

- Looking back at the cost,

$$Cost = \int_0^{t_1} \Big(- \frac{d}{dt}(x^T Px) + \|v\|^2 \Big)dt + x^T(t_1)Qx(t_1)$$

$$= -x^T Px|_0^{t_1} + \int_0^{t_1} \big(\|v\|^2 \big)dt + x^T(t_1)Qx(t_1)$$

$$= x(0)^T P(0)x(0) + \int_0^{t_1} \big(\|v\|^2 \big)dt \qquad via P(t_1) = Q$$

- Hence,
  1. $v \equiv 0$ minimizes cost
  2. optimal cost is $x_0^T P(0)x_0$ because $x(0) = x_0$

≈ Again, this applies even when A, B, and or C are time-varying; optimal control is time-varying; state feedback control; optimal cost $= x_0^T P(0) x_0$ is quadratic in the initial state.

- As in discrete time, can write the optimal control in open loop form as follows:

$$\mathcal{A}(t) = A - BB^T P(t) \qquad t \in [0, t_1]$$

then,

$$u(t) = \Phi(t, 0) x_0 \qquad \Phi(STM) \leftrightarrow \mathcal{A}(t)$$

- But wait, need to make sure the required $t \mapsto P(t), t \in [0, t_1]$ exists. From ODE / dynamical system theory, you can show:
  a) you know $t \mapsto P(t), P(t_1) = Q$, exists at least on some small interval $(\hat{t}, t_1]$
  b) you can't extend the solution to times $< \hat{t}$ *if and only if* $v^T P(t) v$ blows up as t goes down towards $\hat{t}$ for some $v \in \mathbb{R}^n$

  Thus, if we can show $v^T P(t) v$ for all $v \in \mathbb{R}^n$ is bounded from above by a common bound for all t for which $P(t)$ is defined, can conclude that $t \mapsto P(t), t \in [0, t_1]$ exists.
- Can bound $v^T P(t) v$ as follows:

$$v^T P(t) v = min \int_0^{t_1} (\|u(t)\|^2 + \|y(t)\|^2) dt + x^T(t_1) Q x(t_1), \ from\ x(t) = v$$

Thus,

$$v^T P(t) v \leq \int_0^{t_1} (\|u(t)\|^2 + \|y(t)\|^2) dt + x^T(t_1) Q x(t_1), \ all\ u$$

Then,

$$v^T P(t) v \leq \int_0^{t_1} (\|y(t)\|^2) dt + v^T \Phi^T(t_1, t) Q \Phi(t_1, t) v$$

$$= v^T \Big(\int_\tau^{t_1} \Phi^T(t_1, \tau) C^T C \Phi(t_1, t) d\tau + Phi^T(t_1, t) Q \Phi(t_1, t) \Big) v$$

Can see inside parentheses is a nice continuous nonnegative definite matrix-valued function on $[0, t_1]$ - hence has a max norm M on that interval. So,

$$v^T P(t) v \leq M \|v\|^2 \qquad all\ t \in [0, t_1], v \in \mathbb{R}^m$$

→ Finished continuous time finite time horizon LQR Problem! Emphasize these above results hold even when A, B, C are time-varying.

5.2 Infinite Time Horizon LQR

Now look at **Infinite Time Horizon** Problem with Constant Matrices A, B, C:

DT: given,

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

$$y(k) = Cx(k), \qquad x(0) = x_0$$

find $k \mapsto u(k), k \geq 0$ to minimize

$$\sum_{k=0}^{\infty} \left(\|u(k)\|^2 + \|y(k)\|^2 \right)$$

CT: given,

$$\dot{x}(t) = Ax(t) + Bu(t), \qquad x(0) = x_0$$

$$y(t) = Cx(t)$$

find $u : [t, \infty] \mapsto \mathbb{R}^m$ to minimize

$$\int_0^{\infty} \left(\|u(t)\|^2 + \|y(t)\|^2 \right) dt$$

Observe: without further assumptions, may not even have a choice of u that makes the cost finite!

$$e.g. \quad \dot{x} = x + (0) \cdot u, y = x, \qquad \rightarrow \textit{no u can control system}, \; y(t) = e^t x(0)$$

→ If we assume (in both discrete and continuous time) that (A,B) is controllable, then we know that there exists an input function that drives the state $x(t)$ from x_0 at time 0 to zero in finite time - so use that input until the state is 0, then $u = 0$ after that makes cost integral or sum finite.

⇒ Assume (A,B) controllable henceforth.

In this case, it turns out that we can motivate the solution by taking a suitable limit of the solution to a finite-horizon problem. Along the way observability of (A,C) will come into play

5.2.1 Continuous Time Solution

Solution follows by taking a limit of finite time horizon LQR Problem. Consider, to start, minimizing the finite time problem cost

$$\int_0^{t_1} \left(\|u\|^2 + \|y\|^2 \right) dt + x^T(t_1) 0_n x(t_1)$$

- Start with

$$u(t) = -B^T P(t)x(t),\ t \in [0, t_1]\ where\ t \mapsto P(t)$$

solves the Riccati Equation

$$\dot{P}(t) = -\Big(P(t)A + A^T P(t) - P(t)BB^T P(t) + C^T C\Big), \qquad P(t_1) = 0_n$$

and the optimal control is

$$x_0^T P(0) x_0,\ P(0) = P_{t_1}(0)$$

- Assume (A,C) observable, then $P_{t_1}(0)$ is positive definite for all $t_1 > 0$

Goal: take $\lim_{t_1 \to \infty} P_{t_1}(0)$

Fact: For every $x_0 \in \mathbb{R}^n,\ t_1 \mapsto x_0^T P_{t_1}(0) x_0$ is <u>increasing</u> in t_1, also <u>bounded</u> from above. Hence it approaches a limit as $t_1 \to \infty$

To See: suppose $t_1 \le t_2$. Then for any u, we have

$$\int_0^{t_1} \big(\|u\|^2 + \|y\|^2 \big) dt \le \int_0^{t_2} \big(\|u\|^2 + \|y\|^2 \big) dt$$

Thus, if u^{**} minimizes the RHS, we have

$$\int_0^{t_1} \big(\|u^{**}\|^2 + \|y^{**}\|^2 \big) dt \le x_0^T P_{t_2}(0) x_0\ (min.\ value\ of\ RHS)$$

And if u^* minimizes the LHS,

$$x_0^T P_{t_1}(0) x_0 = \int_0^{t_1} \big(\|u^*\|^2 + \|y^*\|^2 \big) dt \le x_0^T P_{t_2}(0) x_0$$

Noted earlier that, by using a u that drives the state to zero in finite time, can bound optimal cost above by some $M > 0$

Conclusion: For all $x_0 \in \mathbb{R}^n, t \mapsto x_0^T P_{t_1}(0) x_0$ is increasing in t_1 and bounded above by M, hence it has a limit!

Moreover, $P_{t_1}(0)$ symmetric $\Rightarrow$ every entry in $P_{t_1}(0)$ also approaches a limit as $t_1 \to \infty$ Let

$$P^* = \lim_{t_1 \to \infty} P_{t_1}(0)$$

- Note that P^* is positive definite because all $P_{t_1}(0)$'s are - in fact the matrix $P^* - P_{t_1}(0)$ is positive definite for all $t_1 > 0$.
- Furthermore, P^* is an <u>Equilibrium Point</u> of the Riccati Equation (follows from ODE theory

easily). Thus,

$$0_n = P^*A + A^TP^* - P^*BB^TP^* + C^TC \qquad (ARE)$$

Algebraic Riccati Equation

Claim: Solution to the infinite horizon problem is

$$u(t) = -B^TP^*x(t) \qquad all\ t \in [0,\infty)$$

Which is a constant gain state feedback control law, and the minimum cost is $x_0^TP^*x_0$. In open loop form

$$u(t) = B^TP^*e^{t(A-BB^TP^*)}x_0 \qquad t \in [0,\infty)$$

- Also, $A-BB^TP^*$ the *closed-loop A matrix* is Hurwitz. Show first that $A-BB^TP^*$is Hurwitz from (ARE). Rewrite it as

$$(A - BB^TP^*)^TP^*(A - BB^TP^*) = -P^*BB^TP^* - C^TC$$

Comment: Can almost use Lyapunov Lemma Part 2 to conclude that $A - BB^TP^*$ is Hurwitz, but not quite since RHS above is not in general negative definite.

- Suppose v is an eigenvector of $A - BB^TP^* \leftrightarrow$ eigenvalue λ_0
- Take $v^H(equation)v$, get

$$2Re\{\lambda_0\}v^HP^*v = -v^H(P^*BB^TP^* + C^TC)v \le 0$$

where $v^HP^*v > 0$ because P^* is positive definite.

Thus, $Re\{\lambda_0\} \le 0$. If zero, then

$$0 = -v^H(P^*BB^TP^* + C^TC)v$$

· In particular,

$$v^HC^TCv = 0 \Rightarrow Cv = 0$$

$$v^HP^*BB^TP^* = 0\ Rightarrow B^TP^*v = 0$$

· v is an eigenvector of $A - BB^TP^*$ in the nullspace of C

· v is an eigenvector of A in nullspace of A because $B^TP^*v = 0$

$\Rightarrow$ no such v can exist by (A,C) observable.

Bottom Line: $Re\{\lambda_0\} < 0 \rightarrow A - BB^TP^*$ Hurwitz

- Need to show that the optimal control is $u = -B^TP^*x(t)$ and optimal cost is $x_0^TP^*x_0$.

Suppose, without loss of generality that

$$u = -B^T P^* x + v$$

we'll see that optimal choice of v is $v \equiv 0$, cost will be $x_0^T P^* x_0$

$$\dot{x} = Ax - BB^T P^* x + Bv$$

$$Bv = \dot{x} - (A - BB^T P^*)x$$

And lets look at the cost integrand

$$\|u\|^2 + \|y\|^2 = \big\|+ - B^T P^* x + v\big\|^2 + \|Cx\|^2$$

First term:

$$x^T P^* BB^T P^* x - x^T P^* B^v - v^T B^T P^* x + \|v\|^2$$

$$= x^T P^* BB^T P^* x - x^t P^* (\dot{x} - (A - BB^T P^* x))^T - (\dot{x} - (A - BB^T P^* x))^T P^* x + \|v\|^2$$

$$= -x^T P^* x - \dot{x}^T P^* x + x^T (P^* A + A^T P^* - P^* BB^T P^*)x + \|v\|^2$$

Second term:

$$= x^T C^T C x$$

And together get:

$$-\frac{d}{dt}(x^T P^* x) + \|v\|^2 + x^T (P^* A + A^T P^* - P^* BB^T P^* + C^T C)x$$

$$P^* A + A^T P^* - P^* BB^T P^* + C^T C = 0_n \text{ by } (ARE)$$

Thus,

$$cost = \int_0^\infty \big(- \frac{d}{dt}(x^T P^* x) + \|v\|^2 \big) dt$$

$$cost = -x^T(t) P^* x(t)\Big|_{t=0}^{t=\infty} + \int_0^\infty \|v\|^2 \, dt$$

Making it clear that the best choice is $v \equiv 0$, making $\dot{x} = (A - BB^T P^*)x$ so $x(t) \to 0$ as $t \to \infty$ making $cost = x_0^T P^* x_0$!

A few Comments

1) P^* is the only nonnegative definite solution to (ARE). Not hard to show, but tedious
2) P^* is the only symmetric matrix $\hat{P}$ that solves (ARE) and makes $A - BB^T P^*$ Hurwitz. Thus, P^* is THE positive definite solution to (ARE) and THE stabilizing solution to (ARE)

3) The mapping

$$(A, B, C) \mapsto P^*(A, B, C) \qquad Delchamps's Lemma$$

on set of triples (A,B,C) where (A,B) reachable and (A,C) observable is super smooth / infinitely differentiable / real analytic.

5.2.2 Discrete Time Notes

Discrete Time Infinite Horizon Problem iSimilar:

- $P_{k_1}(0)$ is increasing in k_1 to P^* as $k_1 \to \infty$
- Optimal control is $u = -(I_n + B^T P^* B)^{-1} B^T P^* A x$
- Optimal cost is $x_0^T P^* x_0$
- P^* solves the discrete time (ARE)

$$0_n = F_*^T F_* + C^T C + (A - BF_*)^T P^* (A - BF_*)$$

$$F_* = (I_n + B^T P^* B)^{-1} B^T P^* A$$

Chapter 6

Appendix:

These sections are sourced from Claire Tomlin's EE221A at UCBerkeley

6.1 Linear Algebra Foundations

The main goals of this review is to cover:

- functions: injective, surjective, bijective, left inverse, right inverse
- field, ring
- vector space, subspace
- linear independence and dependance
- basic, coordinates
- Reference: *Callier & Desoer (C & D)*, Appendix A.1 - A.3

	Algebraic Concept		Ex from Linear Theory
1.	Linear Vector Space	$\leftrightarrow$	state space, input space, output space
2.	Linear Maps	$\leftrightarrow$	reachability map L_r, observability map L_0
3.	Normed Spaces	$\leftrightarrow$	stability analysis
4.	Inner product, adjoint	$\leftrightarrow$	controllability, observability grammians

6.1.1 Notation, Groundwork, and Abstract Math Concepts

Notation:

- $\mathbb{Z}$: Integers
- $\mathbb{N}$: Natural Numbers
- $\mathbb{R}$: Real Numbers
- $\mathbb{R}_+$: Positive Reals
- $\mathbb{C}$: Complex Numbers
- $\mathbb{C}_+$: Positive Complex

Quantifiers:

- $\in$: In, $\notin$: Not In
- $\forall$ For All
- $\exists$: Exists
- $\exists!$ Exists One and Only One
- $\exists?$: Does There Exists
- $\ni$: Owns

Implications:

- $\Rightarrow$
- $\Leftarrow$
- $\Leftrightarrow$

Please remember, order of operations is important!

$$\neg(\forall(P)\exists Q) \Leftrightarrow \exists(\neg P)\forall(\neg Q)$$

Cartesian Product: Given two sets X and Y, the cartesian product is the set of all ordered pairs (x, y) where $x \in X$ and $y \in Y$. The cartesian product of X and Y is denoted $X \times Y$. The set of all ordered n-tuples of real (complex) numbers is denoted by $\mathbb{R}^N$ or $\mathbb{C}^n$).

Functions: Given two sets X-domain and Y-codomain/range,by $f : X \mapsto Y$, we mean that $\forall x \in X,\ f$ assigns a unique $f(x) \in Y$. From here, we say f maps X *into* Y. Finally, $f(x) := \{f(x)|x \in X\}$ is the range of f. For these examples, consider a matrix function $A \in \mathbb{R}^{m\times n}$ (Being an m-row and n-column matrix) so that we have the mapping $f\ :\ f(x) = Ax$

- **Injectivity (one-to-one)**: iff A has rank n, or *full column rank*

$$\begin{aligned} f\ is\ injective &\iff f(x_1) = f(x_2) \Rightarrow x_1 = x_2 \\ &\iff x_1 \neq x_2 \Rightarrow f(x_1) \neq f(x_2) \end{aligned}$$

- **Surjectivity (onto)**: iff A has rank m, or *full row rank*

$$f\ is\ surjective \iff \forall y \in Y,\ \exists x \in X \ni y = f(x)$$

- **Bijectivity (Injective and Surjective)**

$$i.e.\ \forall y \in Y,\ \exists! x \in X,\ \ni y = f(x)$$

Example: Consider $f : X \mapsto Y$ and let $\mathbb{1}x$ be the identity map on X. $(\mathbb{1}_x : X \mapsto X,\ s.t.\ \mathbb{1}_x(x) = x)$. We define the left inverse of f as the map $g_L : Y \mapsto X$ such that $g_L \circ f = \mathbb{1}x$. In this equality, $\circ$ is the composition, $g_L \circ f : X \to X : x \to g_L(f(x))$.

Prove That: f has a left inverse $g_L \iff f$ is injective. **Proof**:

- Assume f is injective $\therefore f(x_1) = f(x_2) \Rightarrow x_1 = x_2$. Construct $g_L : Y \mapsto X$ s.t. on $f(x),\ g_L(f(x)) = x$. This is a well-defined function due to the injectivity of f.

$$\therefore g_l \circ f = \mathbb{1}x$$

- Assume f has a left inverse (proof by prooving that if if f has a left inverse g_L, and $\neg B$, gives us a contradiction) $g_L, \therefore g_L \circ f = \mathbb{1}x, \ \therefore \forall x \in X, \ g_L(f(x)) = x$
- Suppose f is not injective, then $\exists$ some $x_1 \neq x_2$ such that $f(x_1) = f(x_2)$
- If not injective, gives $g_L(f(x_1)) = x_1, \quad g_L(f(x_2)) = x_2$, but g_L is a function, thus $x_1 = x_2$, which contradicts our assumption

$\therefore f$ is injective

Similarly, we define the right inverse of f as the map $g_R : Y \mapsto X$ such that $f \circ g_R = \mathbb{1}y$.
$\star$ Exercise: prove that

$$f \text{ has a right inverse } g_R \iff f \text{ is surjective}$$

Finally, g is called a two-sided inverse, or simply an inverse of f, and is denoted f^{-1}

$$\begin{aligned} f^{-1} \textit{ is inverse} &\iff g \circ f = \mathbb{1}x \ \& \ f \circ g = \mathbb{1}y \\ &\iff f \textit{ invertible, } f^{-1} \textit{ exists} \end{aligned}$$

Field: A field $\mathbb{F}$ is an object consisting of a set of elements, and two binary operations: addition $(+)$ and multiplication $(\cdot)$, such that the following axioms are obeyed:

Addition is:

i. associative $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$
ii. commutative $\alpha + \beta = \beta + \alpha \qquad \forall \alpha, \beta, \gamma \in \mathbb{F}$
iii. $\exists$ *identity element* $0 \ : \ \alpha + 0 = \alpha$
iv. $\exists$ *inverse* $\ : \ \forall \alpha \ \ \exists (-\alpha) \ s.t. \ \alpha + (-\alpha) = 0$

Similarly, for Multiplications

v. associative
vi. commutative
vii. $\exists$ *identity* $1 \ : \ \alpha \cdot 1 = \alpha$
viii. $\exists$ *inverse* $\ : \forall \alpha \neq 0, \ \exists \ \alpha^{-1} \ s.t. \ \alpha \cdot \alpha^{-1} = 1$
ix. $(\cdot)$ distribution over $(+)$: $\alpha \cdot (\beta + \gamma) = \alpha \cdot \beta + \alpha \cdot \gamma$

Examples:

$$\mathbb{R} \qquad \mathbb{C} \qquad \mathbb{C}(s)$$

$\mathbb{R}(s)$: set of rational functions in s with coefficients in $\mathbb{R}$ ie $\dfrac{s^2 + 3s + 1}{s + 1}$

Similarly, $\mathbb{R}_{p,o}(s)$ (strictly proper rational functions), is not a field

$R[s]$ polynomials. Note a field

Ring: Same as a field, expect

$\cdot$ not necessarily commutative under $\cdot$

· no inverse for non-zero elements under ·

Commutative Rings:

Examples: $\mathbb{Z},\ \mathbb{R}[s],\ \mathbb{C}[s],\ \mathbb{R}_{p,o}(s),\ \mathbb{R}_p(s)$

Non-commutative Rings:

Examples: $\mathbb{R}^{n\times n},\ \mathbb{C}^{n\times n},\ \mathbb{C}^{n\times n}[s],\ \mathbb{C}^{n\times n}[s]$...

Example The set $\{0,1\}$ with $(\cdot)$ = binary AND, $(+)$ binary XOR, is a field. Show!

6.1.2 Vectors

A vector space (V,F) is a set of vectors v, and a field of scalars $\mathbb{F}$, and two binary operations: vector addition $+$ and scalar multiplications $\cdot$, such that:

Addition: $+\ : V\times X\mapsto X : (x,y)\mapsto x+y,\ \forall x,y\in V$

i. associative $(x+y)+z=x+(y+z)$

ii. commutative $x+y=y+x$

iii. $\exists$ *identity element* Θ ('Zero Vector') $\ni x+\Theta=\Theta+x=x$

iv. $\exists$ *inverse* : $\forall x\in V,\ \exists(-x)\ s.t.\ x+(-x)=0$

Scalar Multiplication: $\cdot\ :\mathbb{F}\times V\mapsto V\ :\ (\alpha,x)\mapsto\alpha x \qquad \forall x\in V,\ \alpha,\ \beta\in\mathbb{F}$

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

vi. $1\cdot x=x$: Multiplicative inverse of field

vii. $0\cdot x=\Theta$: Additive inverse of field

Distribution Laws:

$\forall\ x\in V,\ \forall\alpha,\beta\in\mathbb{F} \qquad (\alpha+\beta)x=\alpha x+\beta x$

$\forall\ x,y\in V,\ \forall\alpha\in\mathbb{F} \qquad \alpha(x+y)=\alpha x+\alpha y$

Examples:

1. $(\mathbb{F}^n,\mathbb{F})$ *i.e.* $(\mathbb{R}^n,\mathbb{R}),(\mathbb{C}^n,\mathbb{C})$, the space of n-tuples in F over the field F is a vector space. $\star$ show this! Also, $(\mathbb{R}^n,\mathbb{C})$ is not a vector space because it is not closed under scalar multiplication.
2. The function space $F(D,V)$ defined by
   (a) $(V,\mathbb{F})$ is a vector space, D is a set (ie $\mathbb{R},\mathbb{R}^n$)
   (b) $\therefore\ F(D,V)$ is the class of all functions which maps D to V
   (c) Then, $(F(D,V),\mathbb{F})$ is a vector space with the following
      i. addition: $(f+g)d=f(d)+g(d) \qquad f,g\in F(D,V),\ d\in D$
      ii. scalar mult : $(\alpha f)d=\alpha f(d), \qquad alpha\in F,...$
      iii. Note here, Both the $\mathbb{F}$ in the original definition and the $\mathbb{F}$ in $(F(D,B),\mathbb{F})$ must be the same. Same field for defining new function.
3. for example
   (a) continuous functions of $[t_0,t_1]\mapsto\mathbb{R}^n, \qquad \big(C([t_0,t_1],\mathbb{R}^n),\mathbb{R}\big)$
   (b) k times differentiable functions on $[t_0,t_1]\mapsto\mathbb{R}^n, \qquad \big(C^k([t_0,t_1],\mathbb{R}^n),\mathbb{R}\big)$

(c) Other kinds include Lebesgue functions,
$L_p[t_0,t_1] = \{f : [t_0,t_1] \mapsto \mathbb{R} \ : \ \int_{t_0}^{t_1} |f(t)|^p dt < \infty\}$ $\qquad ((L_p[t_0,t_1],\mathbb{R}),\mathbb{R})$

Vector Subspaces (or Linear Subspaces): Let $(V,\mathbb{F})$ be a linear space and W a subset of V. Then $(W,\mathbb{F})$ is called a subspace of $(V,\mathbb{F})$ if $(W,\mathbb{F})$ is itself a vector space.
How to check if something is a subspace?

i. verify that W is a subset of V (Thus W inherits the vector space axioms of V)
ii. verify closure under vector addition and scalar multiplication
ie $\forall w_1, w_2 \in W, \ \forall \alpha_1, \alpha_2 \in \mathbb{F}, \ \alpha_1 w_1 + \alpha_2 w_2 \in W$

An interesting example is that of the subspace of a plane in the space of $\mathbb{R}^3 \times \mathbb{R}$ or $(\mathbb{R}^3,\mathbb{R})$. What if this plane does not include the origin - zero vector issues?

Geometric Interpretation: A plane in $\mathbb{R}^3$ is a subspace!
Example: Prove that if W_1, W_2 are subspaces of V

i. $W_1 \cap W_2$ is a subspace
ii. $W_1 \cup W_2$ is not necessarily a subspace

Proof!:

i. Can decompose and vector as combination of vectors making up the two. Formally: $W_1 \subset V, \ W_2 \subset V \therefore W_1 \cap W_2 \subset V$. Then, can find $w_1, w_2 \in W_1 \cap W_2 \Rightarrow \alpha w_1 + \beta w_2 \in W_1 \cap W_2$. Both elements are clearly in the combo space as both originally were subspaces.
ii. A counterexample is sufficient Example: Prove that if W_1, W_2 are subspaces of V
- Let $W_1 = span\{\left[\begin{smallmatrix}1\\0\end{smallmatrix}\right]\}, \qquad W_2 = span\{\left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]\}, \qquad V = \mathbb{R}^2$
Consider, $w = \left[\begin{smallmatrix}1\\1\end{smallmatrix}\right] = \left[\begin{smallmatrix}1\\0\end{smallmatrix}\right] + \left[\begin{smallmatrix}0\\1\end{smallmatrix}\right]$, but $W \notin W_1 \cup W_2$
$\therefore W_1 \cup W_2$ is not closed under vector addition and thus is not a subspace

Linear Independence and Dependence: Suppose $(V,\mathbb{F})$ is a linear space. The set of vectors $\{v_1, v_2, \cdots v_p\}, v_i \in V$ is said to be linearly independent $iff \ \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_p v_p = \theta$ when $\alpha_i \in F \Rightarrow \alpha_1 = \alpha_2 = \cdots = \alpha_p = 0$. The set of vectors is said to be linearly dependent $iff \ \exists$ scalars $\alpha_1, \alpha_2 \cdots \alpha_p \in \mathbb{F}$ not all zero such that $\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_p v_p = \theta$

Example:

$$\mathbb{F} = \mathbb{R}, \ k \in \{0,1,2,3,\cdots\}, f_k : [-1,1] \mapsto \mathbb{R} \text{ such that } f_k(t) = t^k$$

Show that the set of vectors in $(F([-1,1],R),\mathbb{F}), \ \{(f_k)\} : \{1,t,t^2,t^3,\cdots,t^n\}, \ k = 0,\cdots n$ is linearly independent.

Solution:

$$(f_k)_0^n = \{f_0, f_1, \cdots f_n\} \therefore \ show \ (\alpha_0 + \alpha_1 t + \alpha_2 t^2 + \alpha_n t^n = \theta) \Rightarrow (\alpha_i = 0), \ \alpha_i \in \mathbb{R}, \forall t \in [-1,1]$$

Above, it is required that this is true for all t in the interval (that's what the dense notation says).

It is good to note that the zero vector θ is the *zero function*

Span: a set of vectors is said to span a space, if any vector in the space can be written as a linear combination of vectors in the set. Additionally, the span of a set of vectors is the set of all linear combinations of those vectors.

Basis: Suppose $(V, \mathbb{F})$ is a linear space. Then a set of vectors $B = \{b_1, b_2 \cdots b_n\}$ is called a basis of V if

i. $\{b_1, b_2 \cdots b_n\}$ spans V

ii. $\{b_1, b_2 \cdots b_n\}$ is a linearly independent set

Coordinate Representation: Any vector $x \in V$ may be written as a linear combination of the basis vectors:

$$x = \eta_1 b_1 + \eta_2 b_2 + \cdots + \eta_n b_n = \sum_{i=1}^{n} \eta_1 b_i$$

and hence, the vector η is called the Coordinate Vector of x with respect to the b_i

$$\eta = \begin{bmatrix} \eta_1 \\ \vdots \\ \eta_n \end{bmatrix} \in F^n$$

Where the η_i's are called the coordinates of x with respect to $\{b_i\}$.

Fact: The η_i's are uniquely defined in terms of x and the $\{b_i\}$.

Proof: Suppose not. Then $\exists \eta, \eta'$ such that

$$x = \eta_1 b_1 + \eta_2 b_2 + \cdots \eta_n b_n \qquad (1)$$

and

$$x = \eta_1' b_1 + \eta_2' b_2 + \cdots \eta_n' b_n \qquad (2)$$

subtracting (2) from (1) leads to

$$\theta = (\eta_1 - \eta_1')b_1 + (\eta_2 - \eta_2')b_2 + \cdots + (\eta_n - \eta_n')b_n$$

Which, for $\eta_i \neq \eta_i'$ implies that the $\{b_i\}$ are linearly dependent, contradicting the assumption that $\{b_i\}$ is a basis.

Notes:

i. a basis of a vector space is *not* unique, e.g. in $\mathbb{R}^3$

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

ii. if $\{b_1, b_2, \cdots b_n\}$ is a basis for $(V, \mathbb{F})$, then any other basis also has n elements. The number of elements in the basis is called the dimension of the vector space.

example: the linear space of polynomials with real coefficients defined over the field of reals, denoted $(\mathbb{R}[s], \mathbb{R})$ is an example of an infinite dimensional vector space. ie.

$$B = \{1, s, s^2, s^3, \cdots, s^k, \cdots\}$$

$\star$ exercise: show this infinite set of vectors is linearly independent over $\mathbb{R}$

example: A basis in $(\mathbb{R}^{2\times 2}, \mathbb{R})$

$$\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$$

6.1.3 Linear Maps

Linear Maps: Let $(V, \mathbb{F})$ and $(W, \mathbb{F})$ be linear spaces over the *same field* $\mathbb{F}$. Let $\mathcal{A}$ be a map from $V \mapsto W$:

$$\mathcal{A}: \ V \mapsto W \qquad s.t. \ A(v) = w \qquad v \in V, \ w \in W$$

Then, $\mathcal{A}$ is said to be a linear map (Equiv linear operation) iff it follows superposition, ie

$$\mathcal{A}(\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 \mathcal{A}(v_1) + \alpha_2 \mathcal{A}(v_2), \qquad \forall \alpha_1, \alpha_2 \in \mathbb{F}, \ \forall v_1, v_2 \in V$$

In text, *map* notation ... '$\mathcal{A}$ operates on an element $\alpha_1 v_1 + \alpha_2 v_2$ in V'. We will show that this operation is equivalent to pre-multiplication of $\alpha_1 v_1 + \alpha_2 v_2$ by a matrix, if $\mathcal{A}$ is linear.

Example: Consider the following mapping on the set of polynomials of degree 2:

$$\mathcal{A}: \ as^2 + bs + c \mapsto cs^2 + bs + a$$

Is this a linear map?

Solution: let

$$\begin{cases} v_1 = a_1 s^2 + b_1 s + c_1 \\ v_2 = a_2 s^2 + b_2 s + c_2 \end{cases}$$

$$\begin{aligned} \mathcal{A}(\alpha_1 v_1 + \alpha_2 v_2) &= \mathcal{A}(\alpha_1 a_1 s^2 + \alpha_1 b_1 s + \alpha_1 c_1 + \alpha_2 a_2 s^2 + \alpha_2 b_2 s + \alpha_2 c_2) \\ &= \alpha_1 \mathcal{A}(v_1) + \alpha_2 \mathcal{A}(v_2) \\ \therefore \quad & \mathcal{A} \textit{ is linear map}! \end{aligned}$$

$\star$ Exercise: How about

$$\mathcal{A}: \ as^2 + bs + c \mapsto \int_0^s (bt + a)dt \ ?$$

$\star$ Exercise: How About (generally not linear unless $k = 0$

$$\mathcal{A}: \ v(t) \mapsto \int_0^1 v(t)dt + k \qquad for \ v(\cdot) \in C([0,1], \mathbb{R}), k \in \mathbb{R} \ ?$$

$\star$ Exercise: And,

$$\mathcal{A}: \mathbb{R}^3 \mapsto \mathbb{R}^3, \qquad \mathcal{A}(v) := Av \qquad where \ A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 5 \\ 7 & 0 & 16 \end{bmatrix} ?$$

$\star$ Exercise: Let $\mathcal{A}$ be a linear map from $(U, \mathbb{F})$ to $(V, \mathbb{F})$ with $dim \ U = n$ and $dim \ V = m$. Then show that

$$dim \ R(A) + dim \ N(A) = n$$

Note: What happens to the zero vector under this (or any) linear map?... It's mapped to 0!

Range Space or Image: of $\mathcal{A}$ is:

$$R(\mathcal{A}) = \{v \mid v = \mathcal{A}(u), u \in U\}$$

Nullspace or Kernel: of $\mathcal{A}$, (note θ_v is the zero vector of V)

$$N(\mathcal{A}) = \{u \in U \mid A(u) = \theta_v$$

From here, can **prove** $R(\mathcal{A}), N(\mathcal{A})$ are subspaces (exercise) $\star$.

Theorem: Range and Null Spaces of Linear Operators (likes to show up on Prelim). Consider $\mathcal{A} : U \mapsto V$ with $(U, \mathbb{F}), (V, \mathbb{F})$ linear spaces. Let $b \in V$ Then,

a) $\mathcal{A}(u) = b$ has at least one solution $\Leftrightarrow b \in R(\mathcal{A})$

b) If $b \in R(\mathcal{A})$, then

i $\mathcal{A}(u) = b$ has a unique solution $\Leftrightarrow N(\mathcal{A}) = \theta_u$

ii let x_0 be such that $\mathcal{A}(x_0) = b$. Then $\mathcal{A}(x) = b \Leftrightarrow x - x_0 \in N(\mathcal{A})$

Proof: Exercise.

Exercise: let $\mathcal{A}$ be a linear map from $(U, \mathbb{F})$ to $(V, \mathbb{F})$ with $dim\ U = n$ and $dim\ V = m$. Then show that

$$dim(R(\mathcal{A})) + dim(N(\mathcal{A})) = n$$

6.1.4 Matrix Representation

Any linear map between finite dimensional linear spaces can be represented as matrix multiplication.

Let $\mathcal{A} : u \mapsto v$ be a linear map from $(U, \mathbb{F})$ to $(V, \mathbb{F})$ where $dim\ U = n$ and $dim\ V = m$. Let $\{u_j\}_{j=1}^n$ be a basis for U and let $\{v_j\}_{j=1}^m$ be a basis for V. Thus, for any $x \in U, \exists!\ \eta = \{\eta_j\}_{j=1}^n \in \mathbb{F}^n$ such that $x = \sum_{j=1}^n \eta_j u_j$.

By linearity, $\mathcal{A}(x) = \mathcal{A}(\sum_{j=1}^n \eta_j u_j) = \sum_{j=1}^n \eta_j \cdot \mathcal{A}(u_j)$. Now, each $\mathcal{A}(u_j) \in V$, thus each $\mathcal{A}(u_j)$ has a unique representation in terms of the $\{v_j\}_{j=1}^m$:

$$\mathcal{A}(u_j) = \sum_{i=1}^m a_{ij} v_i \qquad \forall j \in 1 \cdots n, a_j : j^{th}\ column$$

$$ie \qquad \mathcal{A}(u_2) = \sum_{i=1}^m a_{i2} v_i,\ etc.$$

Thus, $\{a_{ij}\}_{i=1}^m$ is the representation of $\mathcal{A}(u_j)$ in terms of $\{v_1, v_2, \cdots, v_n\}$.

$$\begin{aligned} \therefore \mathcal{A}(x) &= \textstyle\sum_{j=1}^n \eta_j \cdot \sum_{i=1}^m a_{ij} v_i \\ &= \textstyle\sum_{i=1}^m \big(\sum_{j=1}^n a_{ij}\eta_j\big) v_i \\ &= \textstyle\sum_{i=1}^m \gamma_i v_i \end{aligned}$$

Thus, the representation of $\mathcal{A}(x)$ with respect to $\{v_1, v_2, \cdots, v_n\}$ is $[\gamma_1, \gamma_2, \cdots, \gamma_m]^T \in \mathbb{F}^m$ and the representation of x with respect to $\{u_1, u_2, \cdots, u_n\}$ is $[\eta_1, \eta_2, \cdots, \eta_n]^T \in \mathbb{F}^n$. By uniqueness of the representation:

$$\gamma_i = \sum_{j=1}^n a_{ij} \eta_j \qquad i \in \{1, \cdots, m\}$$

So, $\gamma = A\eta$, $A \in \mathbb{F}^{m \times n}$ where A is the matrix representation of the linear operator $\mathcal{A} :\ U \mapsto V$.

Good to Remember: the j^{th} column of the matrix A is $\mathcal{A}(u_j)$ expressed with respect to $\{v_j\}$.

Example: let $\mathcal{A} : (\mathbb{R}^n, \mathbb{R}) \mapsto (\mathbb{R}^n, \mathbb{R})$, $\mathcal{A}^n = -\alpha_1 \mathcal{A}^{n-1} - \alpha_2 \mathcal{A}^{n-2} \cdots - \alpha_{n-1}\mathcal{A} - \alpha_n I,\ \alpha_i \in \mathbb{R}, b \in \mathbb{R}^n$. Suppose $(b, \mathcal{A}(b), \cdots, \mathcal{A}^{n-1}(b))$ is a basis for $\mathbb{R}^n$. Show that, with respect to this basis, the

vector b and the linear map $\mathcal{A}$ (where A below is in companion form) are represented by:

$$\bar{b} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \qquad A = \begin{bmatrix} 0 & 0 & \cdots & 0 & -\alpha_n \\ 1 & 0 & \cdots & 0 & -\alpha_{n-1} \\ 0 & 1 & 0 & 0 & \vdots \\ \vdots & \vdots & \vdots & \ddots & -\alpha_2 \\ 0 & 0 & \cdots & 1 & -\alpha_1 \end{bmatrix}$$

Solution:

$$\begin{array}{rrcc} start & \bar{b} & = & 1 \cdot b + 0 \cdot \mathcal{A}(b) + \cdots + 0 \cdot \mathcal{A}^{n-1}(b) = [1\ 0\ \cdots\ 0]^T \\ then & \mathcal{A}(b) & = & 0 \cdot b + 1 \cdot \mathcal{A}(b) + \cdots + 0 \cdot \mathcal{A}^{n-1}(b) \\ \mathcal{A}(\mathcal{A}^{n-1}(b)) & \mathcal{A}^n(b) & = & -\alpha_n b - \alpha_{n-1}\mathcal{A}(b) + \cdots - \alpha_1 \mathcal{A}^{n-1}(b) \end{array}$$

Giving us the A Matrix as expected. Now, consider the maps,

$$\mathcal{A} : U \mapsto V \qquad \mathcal{B} : V \mapsto W$$

From here, let

$$\begin{array}{c} \{u_j\}_{j=1}^n \text{ be a basis for } (U, F) \\ \{v_j\}_{j=1}^m \text{ be a basis for } (V, F) \\ \{w_j\}_{j=1}^p \text{ be a basis for } (W, F) \end{array}$$

And let $\zeta,\ \eta,\ \xi$ be the corresponding component vectors:

$$\zeta \in F^n,\ \eta \in F^m,\ \xi \in F^p$$

Thus, with $A \in F^{m \times n}$ and $B \in f^{p \times m}$ the matrix representations of the linear maps $\mathcal{A}$ and $\mathcal{B}$ w.r.t the above bases. Which gives us,

$$\eta = A\zeta \qquad \xi = B\eta$$

and thus,

$$\xi = BA\zeta := C\zeta$$

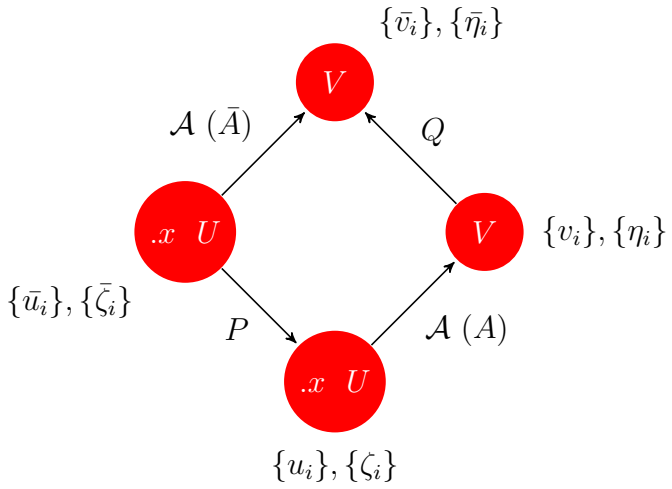
$\Rightarrow$ The composition of linear maps corresponds to matrix multiplication!

6.1.5 Matrix Operations

Change of Basis: Here, we will study the relationships between two matrix representations of the same linear map.

Let $(u_j)_1^n$ and $(\bar{u}_j)_1^n$ be two bases for (U, F), and $(v_i)_1^m$ and $(\bar{v}_i)_1^m$ be two bases for (V, F).

Let, A be the matrix representation of $\mathcal{A} : U \mapsto V$ w.r.t. the bases $(u_j)_1^n$ and $(v_i)_1^m$; and let, $\bar{A}$ be the matrix representation of $\mathcal{A} : U \mapsto V$ w.r.t. the bases $(\bar{u}_j)_1^n$ and $(\bar{v}_i)_1^m$:



Now, for $x \in U$,

$$x = \bar{\zeta}_1\bar{u}_1 + \bar{\zeta}_2\bar{u}_2 + \cdots + \bar{\zeta}_n\bar{u}_n = \begin{bmatrix}\bar{u}_1 & \bar{u}_2 & \cdots & \bar{u}_n\end{bmatrix}\bar{\zeta}$$

and,

$$x = \zeta_1 u_1 + \zeta_2 u_2 + \cdots + \zeta_n u_n = \begin{bmatrix} u_1 & u_2 & \cdots & u_n\end{bmatrix}\zeta$$

$$\therefore \begin{bmatrix}\bar{u}_1 & \bar{u}_2 & \cdots & \bar{u}_n\end{bmatrix}\bar{\zeta} = \begin{bmatrix} u_1 & u_2 & \cdots & u_n\end{bmatrix}\zeta$$

$$\Rightarrow \zeta = P\bar{\zeta} \quad where \quad P = \begin{bmatrix} u_1 & u_2 & \cdots & u_n\end{bmatrix}^{-1}\begin{bmatrix}\bar{u}_1 \cdots \bar{u}_n\end{bmatrix}$$
$$\bar{\eta} = Q\eta \quad where \quad Q = \begin{bmatrix}\bar{v}_1 & \bar{v}_2 & \cdots & \bar{v}_n\end{bmatrix}^{-1}\begin{bmatrix} v_1 \cdots v_m\end{bmatrix}$$

$$Giving,\ \eta = A\zeta \qquad \therefore \bar{\eta} = QA\zeta = QAP\bar{\zeta} := \bar{A}\bar{\zeta}, \qquad \bar{A} = QAP \Leftarrow$$

so, if A is the matrix representation of $\mathcal{A}$ w.r.t $\{u_i\}$, $\{v_j\}$, then $\bar{A} = QAP$ is the matrix representation of $\mathcal{A}$ w.r.t $\{\bar{u}_i\}$, $\{\bar{v}_j\}$.

Note: $\bar{A}$ and A are said to equivalent, and $\bar{A} = QAP$ is said to be a Similarity Transform.

Note: if $U = V$ and $(\bar{v}_J)_1^n = (\bar{u}_j)_1^n$, then $\bar{A} = P^{-1}AP \Leftarrow$

example: let $\mathcal{A} : \mathbb{R}^3 \mapsto \mathbb{R}^3$ be a linear map. Consider

$$B = \{b_1\ b_2\ b_3\} = \left\{ \begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\end{bmatrix} \right\}$$

$$C = \{c_1\ c_2\ c_3\} = \left\{ \begin{bmatrix}1\\1\\0\end{bmatrix}, \begin{bmatrix}0\\1\\1\end{bmatrix}, \begin{bmatrix}1\\0\\1\end{bmatrix} \right\}$$

Clearly, B, and C are bases for $\mathbb{R}^3$. Suppose $\mathcal{A}$ maps $\mathcal{A}(b_1) = [2\ -1\ 0]^T, \mathcal{A}(b_2) = [0\ 0\ 0]^T, \mathcal{A}(b_3) = [0\ 4\ 2]^T$. Write down the matrix representation of $\mathcal{A}$ w.r.t. B and then w.r.t. C.

We are back to more elementary matrix concepts. By rank of the matrix $A \in F^{m\times n}$ (denoted as $rank(A)$ or $rk(A)$) we mean $dim(R(A))$; and, by nullity of $A \in F^{m\times n}$ denoted as $nl(A)$ we mean $dim(N(A))$.

- row rank; full row rank
- column rank; full column rank

Sylvester's Inequality: let $A \in F^{m\times n}$ and $B \in F^{n\times p}$, then $AB \in F^{m\times p}$ and

$$rk(A) + rk(B) - n \leq rk(AB) \leq min(rk(A), rk(B))$$

Proof - exercise. Note that you can also formulate this as an A and B when B is a linear map applied after A, so the dimensions of the spaces change.

Finally, we will be interested in reducing matrices $A \in F^{m\times n}$ to row or column echelon forms; as these are well suited for discussing the construction of a basis for $R(A), N(A), ...$

Elementary Row Operations (ERO)

i) interchange two rows
ii) multiply row i by a nonzero $c \neq 0 \in \mathbb{F}$
iii) add to row i another row j

ERO's are equivalent to premultiplying A by left elementary matrix L, which is obtained from the identity matrix by performing the desired ERO upon it. ie $L \cdot A$

⋆ Show $N(l \cdot A) = N(A)$

Elementary column Operations (ECO)

i) interchange two columns
ii) multiply column j by a nonzero constant
iii) add a column to another

ECOs correspond to post multiplying A by right elementary matrix R, which is obtained from the identity matrix by performing desired ECO on it.

⋆ show $R(AR) = R(A)$

6.2 Linear Algebra Operations

6.2.1 Norms

Normed Linear Spaces: Let the field F be $\mathbb{R}$ or $\mathbb{C}$. $\mathcal{A}$ linear space (V, F) is said to be a normed linear space if $\exists$ mapping:$\|\cdot\| : V \mapsto \mathbb{R}_+$ satisfying the following axioms:

i) $\|v_1 + v_2\| \leq \|v_1\| + \|v_2\| \qquad \forall v_1, v_2 \in V$
ii) $\|\alpha v\| = |\alpha| \, \|v\| \qquad \forall \alpha \in F, v \in V$
iii) $\|v\| = 0 \Leftrightarrow v = \theta_V$

1. examples (F^n, F) with $x = [x_1 \; x_2 \; \cdots \; x_n]^T$
   (a) l_1 norm: $\|x\|_1 = \sum_{i=1}^n |x_i|$
   (b) l_2 norm: $\|x\|_2 = \left[\sum_{i=1}^n |x_i|^2\right]^{1/2}$
   (c) l_p norm: $\|x\|_p = \left[\sum_{i=1}^n |x_i|^p\right]^{1/p}$
   (d) l_∞ norm: $\|x\|_\infty = \max_i |x_i|$
2. examples matrices $A \in F^{m \times n}$
   (a) $\|A\|_a = \sum_{i=1}^m \sum_{j=1}^n |a_{ij}|$
   (b) $\|A\|_F = \left[\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2\right]^{1/2}$ *Frobenius norm*
   (c) $\|A\|_B = \max_{j \in n, i \in m}\{|a_{ij}|\}$
3. examples functions $C([t_0, t_1], F^n)$, C: continuous, $[t_0, t_1]$: domain, F^n: co-domain, F: field
   (a) l_1 norm: $\|f\|_1 = \int_{t_0}^{t_1} \|f(t)\| \, dt$ where $\|f(t)\|$ is any of the norms $a) - d)$
   (b) l_2 norm: $\|f\|_2 = \left[\int_{t_0}^{t_1} \|f(t)\|^2 \right]^{1/2} dt$
   (c) l_∞ norm: $\|f\|_\infty = max\left\{ \|f(t)\|, t \in [t_0, t_1]\right\}$

Equivalent Norms: Two norms, $\|\cdot\|_a$ and $\|\cdot\|_b$ on (V, F) are said to be equivalent if $\exists m_l, m_u \in \mathbb{R}_+$ such that $\forall v \in V$:

$$m_l \|v\|_a \leq \|v\|_b \leq m_u \|v\|_a$$

Example $\star$ For (F^n, F) verify that:

$$\begin{array}{ccccc} \|x\|_\infty & \leq & \|x\|_1 & \leq & n \|x\|_\infty \\ \|x\|_\infty & \leq & \|x\|_2 & \leq & \sqrt{n} \|x\|_\infty \\ \frac{1}{\sqrt{n}} \|x\|_1 & \leq & \|x\|_2 & \leq & \|x\|_1 \end{array}$$

Continuous Function: $f : U \mapsto V$ is continuous if $\forall \; \epsilon \;\; \exists \; \delta \; s.t. \; \|u_1 - u_2\|_u < \delta \Rightarrow \|f(u_1) - f(u_2)\|_v < \epsilon$. Note: all linear maps between finite dimensional vector spaces are continuous.

Induced Norms: Let $\mathcal{A} : (U, F) \mapsto (V, F)$ be a continuous linear operator. Let U and V be endowed with the norms $\|\cdot\|_u$ and $\|\cdot\|_v$ respectively. Then the induced norm of A is defined by:

$$\|A\|_i = \sup_{u \neq 0} \frac{\|Au\|_v}{\|u\|_u}$$

A brief aside, *sup* = supremum = *least upperbound*. The least element greater than or equal to all elements of a set. Consider and example where you divide the distance between the 0 and 1 in half, then the upper half again to infinity. It gets closer and closer to 1, but by definition will never reach it. Supremum is 1!

Theorem: Facts about induced norms.

Let $(U, \|\cdot\|_U),\ (V, \|\cdot\|_V),\ (W, \|\cdot\|_W)$ be normed linear spaces and let

$$\begin{aligned} A &: V \mapsto W \\ B &: U \mapsto W \\ \widetilde{A} &: V \mapsto W \end{aligned}$$

Then we can show that these induced norms comply with the definition of an induced norm $\star$:

$$\begin{array}{rcl} \|Av\|_W & \leq & \|A\|_i \|v\|_V \\ \|\alpha A\| & = & |\alpha| \|A\| \\ \left\|A + \widetilde{A}\right\|_i & \leq & \|A\|_i + \left\|\widetilde{A}\right\|_i \\ \|A\|_i = 0 & \Leftrightarrow & A = 0 \\ \|AB\|_i & \leq & \|A\|_i \|B\|_i \end{array}$$

All linear maps between finite dimensional vector spaces are continuous.

$$\|Au_1 - Au_2\| = \|A(u_1 - u_2)\| \leq \|A\|_i \|u_1 - u_2\|$$

want: $\|Au_1 - Au_2\| < \epsilon$, choose δ such that $\epsilon = \|A\|_i \delta$. Choose delta for accuracy that there is just a scaling factor on the difference of distance and bounds herein.

Examples: $\mathcal{A} : F^n \mapsto F^m,\ \|A\|_{p,i} := \sup_{x \neq 0} \frac{\|Ax\|_p}{\|x\|_p}$. Show that:

a) $\|A\|_{1,i} = \max_{j=1\cdots n}\{\sum_{i=1}^m |a_{ij}|\}$ <u>max column sum</u>

Proof!

$$\begin{array}{rcll} & \|A\|_{1,i} & = & \sup_{u \neq \theta} \frac{\|Ax\|_i}{\|x\|_1} \qquad x \in \mathbb{R}^n, Ax \in \mathbb{R}^m \\ Show: & \|A\|_{1,i} & = & \max_{j=1\cdots n} \left\{\{\sum_{i=1}^m |a_{ij}|\}\right. \\ & \|A\|_{1,i} & = & \sup_{x \neq \theta} \left(\frac{\sum_{i=1}^m |(Ax)_i|}{\|x\|_1}\right) \qquad (Ax)_i : i^{th}\ element\ of\ Ax \\ & & \leq & \sup_{x \neq \theta} \frac{\sum_{i=1}^m \sum_{j=1}^n |a_{ij} x_j|}{\sum_{j=1}^n |x_j|} \qquad by\ \underline{triangle}\ inequality \\ & & = & \sup_{x \neq \theta} \frac{\sum_{i=1}^m \sum_{i=1}^n |a_{ij}||x_j|}{\sum_{j=1}^n |x_j|} \\ & & = & \sup_{x \neq \theta} \frac{\sum_{i=1}^m |x_j| \sum_{i=1}^m |a_{ij}|}{\sum_{j=1}^n |x_j|} \qquad by\ factoring\ matrix\ mult. \\ & & \leq & \max_{j=1\cdots n} \sum_{i=1}^m |a_{ij}| \end{array}$$

Where equality holds because the value of x where the maximum occurs is guaranteed to exist by the vector $x = \begin{bmatrix} 0 & \cdots 0 & 1 & 0 \cdots 0 \end{bmatrix}^T$, where the 1 picks out the actual max of the

column sum!

b) $\|A\|_{2,i} = \max_{j=1\cdots n}\{\lambda_j(A^*A)\}^{1/2}$ Where A^* is the Hermitian transpose of A, it's the same as A^T if A is real. For complex, it is the transpose and complex conjugate.

c) $\|A\|_{\infty,i} = \max_{i=1\cdots m}\{\sum_{j=1}^{n}|a_{ij}|\}$ max row sum

Sensitivity Analysis: a measure of how the solution x to $Ax = b$ changes as $\mathcal{A}$ and b are perturbed.

Consider $Ax = b$; $\mathcal{A} : F^n \mapsto F^n,\ b \in F^n$. If A^{-1} exists, then $x = A^{-1}b$ is the unique solution. Let's denote this the nominal solution.

$$x_0 = A^{-1}b$$

Now suppose A is perturbed to $A + \delta A$ and b is perturbed to $b + \delta B$; call the new solution $x_0 + \delta x$. Get this from some algebra here:

$$\begin{aligned} (A+\delta A)(x_0+\delta x) &= (b+\delta b) \\ Ax_0 + A\delta x + \delta A x_0 + \delta A\delta x &= b+\delta b \\ A\delta x + \delta A x_0 &= \delta b \end{aligned}$$

Goal: relate the size of δx to that of $\delta A \& \delta b$:

$$\delta x = A^{-1}[-\delta A x_0 + \delta b]$$

$$\therefore \|\delta x\| \le \left\|A^{-1}\right\|_i \cdot \left[\, \|\delta A\|_i \, \|x_0\| + \|\delta b\| \,\right]$$

Where $\|\cdot\|$ is a norm in F^n, $\|\cdot\|_i$ is the corresponding induced norm.Thus,

$$\frac{\|\delta x\|}{\|x_0\|} \le \left\|A^{-1}\right\|_i \|\delta A\|_i + \frac{\|A^{-1}\|_i \, \|\delta b\|}{\|x_0\|}$$

And from $Ax_0 = b$, we have that $\|x_0\| \ge \frac{\|b\|}{\|A\|_i}$, and thus,

$$\frac{\|\delta x\|}{\|x_0\|} \le \left\|A^{-1}\right\|_i \|A\|_i \left[\frac{\|\delta A\|_i}{\|A\|_i} + \frac{\|\delta b\|}{\|b\|}\right]$$

The quantity $\|A^{-1}\|_i \, \|A\| =: K(A) \ge 1$ is called the condition number of A. If $K(A) >> 1$, small changes in $\delta b, \delta A$ can cause big changes in δx.

6.2.2 Inner Products and Orthogonality

let the field F be $\mathbb{R}$ or $\mathbb{C}$; and consider the linear space (H, F). The function

$$< \cdot, \cdot > : H \times H \to (x, y) \mapsto < x, y >$$

is called an inner product iff

1. $< x, y + z >=< x, y > + < x, z >$ $\quad \forall x, y, z \in H$
2. $< x, \alpha y >= \alpha < x, y >$ $\quad \alpha \in F$
3. $\|x\|^2 :=< x, x > > 0 \Leftrightarrow x \neq \theta_H$
4. $< x, y >=< \overline{y, x} >$

Where in 3. $\|\cdot\|$ is called the norm induced by the inner product; and in 4. the overbar denotes the complex conjugate of $< y, x >$. A space equipped with an inner product is called an Hilbert Space $(H, F, < \cdot, \cdot >)$. Having such a norm as an inner product space turns a Hilbert space into a complete metric space, which also involves the convergence of Cauchy Sequences.

Examples:

1. The easiest example is the space $(\mathbb{R}^n, \mathbb{R})$, where

$$< x, y > = \sum_{i=1}^{n} \bar{x_i} y_i = x \cdot y$$

2. Building from here, $(F^n, F, < \cdot, \cdot >)$ is a Hilbert space under the inner product:

$$< x, y > := \sum_{i=1}^{n} \bar{x_i} y_i : \ x^* y$$

   Where $x^* y$ is the Hermitian transpose

3. $L_2([t_0, t_1], F^n)$ (Space of square, integrable F^n values functions on $[t_0, t_1]$), define the inner product by:

$$< f, g > : \int_{t_0}^{t_1} f(t)^* g(t) dt$$

Orthogonality: on $(H, F, < \cdot, \cdot >)$

Definition: two vectors $x, y \in H$ are said to be orthogonal iff $< x, y >= 0$. This is often written as $x \perp y$.

Definition: if M is a subspace of a Hilbert Space, the subset $M^\perp = \{y \in H :< x, y >= 0 \ \forall \ x \in M\}$ is called the orthogonal completment of M.

Example: $a, b \in V$ $\quad < a, b >= 0$ $\quad \|a + b\|^2$ (norm induced by inner product).

$$< a + b, a + b >=< a, a > + < b, b > + \cancelto{0}{< a, b >} + \cancelto{0}{< b, a >} \Rightarrow \|a + b\|^2 = \|a\|^2 + \|b\|^2$$

Exercise: $\star$ prove that $M \cap M^\perp = \{\theta\}$

Solution: say $\exists x \neq \theta$ such that $x \in M \cap M^\perp \ \therefore < x, y >= 0 \ \forall \ y \in M$, but $x \in M$, so $< x, x >= 0$, but this ($\|x\| = 0$) implies that $x = \theta_M$.

Adjoints: Let F be $\mathbb{R}$ or $\mathbb{C}$ and let $(U, F, <\cdot,\cdot>_U)$ and $(V, F, <\cdot,\cdot>_V)$ be Hilbert spaces. Let $\mathcal{A} : U \mapsto V$ be a continuous and linear mapping. Then the adjoint of A, denoted A^* is the map

$$A^* : V \mapsto U \qquad s.t. \qquad < v, Au >_V = < A^*v, u >_U$$

Example (From old prelim): let $f(\cdot), g(\cdot) \in C([t_0, t_1], \mathbb{R}^n)$ and define $A : C([t_0, t_1], \mathbb{R}^n) \mapsto \mathbb{R}$, where $g(\cdot)$ is specified:

$$A : f(\cdot) \mapsto < g(\cdot), f(\cdot) >$$

Find the adjoint map of A.

Solution:

$$A^* : \mathbb{R} \mapsto (C[t_0, t_1], \mathbb{R}^n)$$

such that $< v, Af(\cdot) >_{\mathbb{R}} = < A^*v, f(\cdot) >_{C([t_0,t_1],\mathbb{R}^n)}$ where $v \in \mathbb{R}$ and $f(\cdot) \in C([t_0, t_1], \mathbb{R}^n)$.

Quick aside:

$$< \cdot,\cdot >_{(C[t_0,t_1],\mathbb{R}^n)} = \int_{t_0}^{t_1} \underbrace{g(t)^T f(t)}_{dot\ product\ in\ \mathbb{R}^n}\, dt \qquad < v_1, v_2 >_{\mathbb{R}} = v_1 \cdot v_2$$

$$\begin{array}{rcl} < v, Af(\cdot) >_{\mathbb{R}} & = & v^* < g(\cdot), f(\cdot) >_{C\cdots} \\ & = & < g(\cdot), v^* f(\cdot) >_{C\cdots} \\ & = & < v \cdot g(\cdot), f(\cdot) >_{C\cdots} \end{array} \quad \begin{array}{rcl} & = & v \cdot \int_{t_0}^{t_1} g(t)^T f(t) dt \\ & = & \int_{t_0}^{t_1} (vg(t))^T f(t) dt \\ & = & \int_{t_0}^{t_1} (\ \ )^T f(t) dt \end{array}$$

$$c.f.\ with\ < A^*v, f(\cdot) > \Rightarrow\ A^* : v \mapsto vg(\cdot)$$

This comes down to a multiplication by $g(\cdot)$. We will use the concept of adjoint with controllability and observability.

Back to orthogonality. We say x is orthogonal to the set of vectors S if $x \perp y\ \forall\ y \in S$. This is often written as $x \perp S$. Additionally, a set of vectors S is called an orthogonal set if $x \perp y\ \forall\ x \neq y,\ x, y \in S$, and this is called orthonormal if in addition $\|x\| = 1\ \forall\ x \in S$. Where, this norm $\|\cdot\|$ is the norm induced by the inner product, unless otherwise specified.

Geometric Interpretation of Inner Product: Let V be an inner product space and $v \in V$. Let $b \in V$ be a unit vector ($\|b\| = 1$) and consider the subspace $S = span\{b\}$. Define $\hat{s} = < v, b > b$. We wish to find an optimal approximation for v from the subspace S. More precisely, we wish to solve the following problem:

$$\min_{s \in S} \|v - s\|$$

Lemma: The vector $\hat{s} = < v, b > b$ is the optimal approximation of v in $S = span\{b\}$. Thus, $< v, b >$ may be regarded as the length of the projection of v on S.

Example: Gram-Schmidt Orthonormalization

Given a collection of vectors $B = \{b_1, b_2, \cdots b_n\}$ drawn from an inner product space V, we wish to construct a set of vectors $C = \{c_1, c_2, \cdots c_r\},\ r \leq n$ that forms an orthonormal basis for $span\{B\}$. This may be accomplished with the Gram-Schmidt orthonormalization procedure.

1. initialize $k = 2,\ v_1 = b_1$, and set $c_1 = \frac{y_1}{\|y_1\|}$
2. Define $y_k = b_k - \sum_{i=1}^{k-1} < c_i, b_k > c_i$
3. If $y_k \neq \theta$ set $c_k = \frac{y_k}{\|y_k\|}$ else increment k and repeat the second step until the set B is empty.

Numerical Example: Gram-Schmidt Orthonormalization: Consider the following set in $\mathbb{R}^3$:

$$B = \left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}, \begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix} \right\}$$

We use the Gram-Schmidt procedure to find an orthonormal basis for $span\{B\}$:

$$y_1 = b_1 = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}^T \text{ and } c_1 = \frac{1}{\sqrt{14}} \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}^T$$

$$y_2 = b_2 - < c_1, b_2 > c_1 = \frac{1}{7} \begin{bmatrix} 4 & 1 & -2 \end{bmatrix}^T \text{ and } c_2 = \frac{7}{\sqrt{21}} \begin{bmatrix} 4 & 1 & -2 \end{bmatrix}^T$$

$$y_3 = b_3 - < c_1, b_3 > c_1 - < c_2, b_3 > c_2 = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}^T$$

Thus, $C = \{c_1, c_2\}$ forms the desired orthonormal basis $\leftarrow$.

Self-Adjoint Maps: Given $(H, F, < \cdot, \cdot >_H)$, a Hilbert space. Let $\mathcal{A} : H \mapsto H$ be a continuous linear map with adjoint $\mathcal{A}^* : H \mapsto H$. We say that the map $\mathcal{A}$ is self-adjoint iff $A = A^*$, or equivalently

$$< x, Ay >_H = < Ax, y >_H \qquad \forall\ x, y \in H$$

Example: Hermitian Matrices. Let $H = F^n$ and let $\mathcal{A}$ be represented by a matrix $A = (a_{ij})\ i, j \in \{1 \cdots n\} \in F^{n \times n}$. Then $\mathcal{A}$ is self-adjoint iff the matrix A is Hermitian, or equivalently, $A = A^*$, meaning $a_{ij} = \overline{a_{ji}}\ \forall i, j \in \{1 \cdots n\}$, or that A is equal to its complex conjugate transpose.

Definiton: Unitary Matrix: A matrix $U \in F^{n \times n}$ is said to be unitary iff $U^*U = UU^* = I_n$ (equivalently, the n columns, or the n rows of U form orthonormal bases of F^n). If $\mathbb{F} = \mathbb{R}$, such a matrix is called orthogonal.

exercise: $\star$ Show why the above statements are equivalent

6.2.3 Singular Value Decomposition

Recall: eigenvalue of $A \in \mathbb{R}^{n\times n}$, is a complex number ($\lambda_i \in \mathbb{C}$) $\exists v_i \in \mathbb{C}^n$ such that

$$Av_i = \lambda_i v_i$$

where v_i is the eigenvector with it's corresponding eigenvalue λ_i. It is good to remember for $A \in \mathbb{R}^{m\times n}$, AA^* & A^*A for square matrices and have the same non-zero eigenvalues.

Singular Values:

$$A \in \mathbb{C}^{m\times n}, \quad AA^* \in \mathbb{C}^{m\times m}, \quad A^*A \in \mathbb{C}^{n\times n}$$

Let λ_i $i = 1 \cdots m$ be the eigenvalues of AA^* (note that the non-zero eigenvalues of AA^* are the same as those of AA^* - **exercise** $\star$). Also, the λ_i are all real and nonnegative. Why?

$$AA^*v_i = \lambda_i v_i \qquad \Rightarrow \qquad < v_i, AA^*v_i > = \lambda_i \|v_i\|^2$$
$$\Rightarrow v_i^* AA^* v_i = \lambda_i \|v_i\|^2$$
$$\Rightarrow \|A^*v_i\|^2 = \lambda_i \|v_i\|^2$$

AA^*, $\lambda_1, \lambda_2, \cdots, \lambda_m$: Assume here that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r \geq 0$ and $\lambda_{r+1}, \lambda_{r+2} \cdots \lambda_m = 0$, where $r = rank(AA^*)$. The non-zero singular values of A are $\sqrt{\lambda_1}, \sqrt{\lambda_2}, \cdots \sqrt{\lambda_r}$. The remaining singular values are zero.

Example

$$A = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad AA^* = \begin{bmatrix} 5 & 0 \\ 0 & 0 \end{bmatrix}, \qquad \Rightarrow \lambda_1 = 5, \lambda_2 = 0, \qquad \sigma_1 = \sqrt{5}, \sigma_2 = 0$$

Relate this back to the

$$\|A\|_{i,2} = \sup_{\|x\|_2 \neq 0} \frac{\|Ax\|_2}{\|x\|_2} = \max_j \left(\lambda_j (A^*A)^{1/2}\right) = \bar{\sigma}(A)$$

Theorem: Let $M \in \mathbb{C}^{m\times n}$ with $rank(M) = r \leq min(m, n)$. Then we can find unitary matrices $U \in \mathbb{C}^{m\times m}$ and $V \in \mathbb{C}^{n\times n}$ such that

$$M = U\Sigma V^* = U \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} V^*$$

where,

$$\Sigma_1 = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \cdots & \cdots \\ 0 & 0 & \cdots & \sigma_r \end{bmatrix}$$

the real numbers $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r \geq 0$ are called the singular values of M and the representations above is called the singular value decomposition of M.

Theorem: Let $M \in C^{m \times n}$ with $rank(M) = r$ and let $M = U\Sigma V^*$ be the singular value decomposition of M. Partition U and V as below, where U_1 is size $m \times r$ and V_1 is size $n \times r$, you can figure out U_2, V_2. This allows us to simplify the representation because of the design of Σ:

$$U = \begin{bmatrix} U_1 & U_2 \end{bmatrix}, \qquad V = \begin{bmatrix} V_1 & V_2 \end{bmatrix}$$

$$\therefore \quad A = U\Sigma V^* = U_1 \Sigma_r V_1^*$$

Also,

$$U^*U = I_m, U_1^*U_1 = I_r \qquad V^*V = I_n, V_1^*V_1 = I_R$$

$$\Rightarrow AA^* = U_1 \Sigma_r^2 U_1^*, \ where \ U_1 = \begin{bmatrix} u_1 & u_2 & \cdots & u_r \end{bmatrix}$$

$$AA^*U_1 = U_1\Sigma_r^2 \Rightarrow AA^*u_i = \sigma_i^2 u_i$$

Above, u_i are therefore called the left singular vectors of A, formed from the columns of U_1. A similar argument can be constructed from $V_1 = \begin{bmatrix} v_1 & v_2 & \cdots & v_r \end{bmatrix}$, giving us the right singular vectors of A, being that $A^*Av_i = \sigma_i^2 v_1$ (used a lot in optimization).

6.2.4 Least Squares Optimization

To introduce least squares optimization, we will build off of adjoints and build through the Projection Theorem. We will see a bit later in the course, as the theory is built up more completely.The domain and co-domain of $\mathcal{A}$ can be decomposed into two sets (each actually subspaces), which only overlap at the zero vector θ.

Working in a mapping from $\mathbb{R}^n$ to $\mathbb{R}^m$ with $\mathcal{A}$ and backwards with $\mathcal{A}^*$. Picture two sets in $\mathbb{R}^m$, as the $null(\mathcal{A})$ and $range(\mathcal{A}^*)$ that only meet at the tangentation of the subspaces, connected by the zero vector θ_n. These two sets map to similar sets in the space $\mathbb{R}^m$, but being the conjoint, zero vector θ_m joined sets, $null(\mathcal{A}^*)$ and $range(\mathcal{A})$.

Finite Rank Operator Lemma: (Partial and grouped in pairs that easily prove the other)

1. $R(A) \perp N(A^*)$
2. $R(A^*) \perp N(A)$
3. $R(AA^*) = R(A)$

4. $R(A^*A) = R(A^*)$
5. $N(AA^*) = N(A^*)$
6. $N(A^*A) = N(A)$

Lets *prove* 1. $N(A^*) = R(A)^\perp$, the general form of showing that two spaces are equal, we show that one vector is in the other for all vectors. Then we check the other direction, and both sets are a subset of the other, yielding the same set. In detail, consider $y \in N(A^*) \ \therefore A^*y = \theta_n$. Then take $x \in \mathbb{R}^n$.

$$< A^*y, x > = 0 = < y, Ax > \therefore \ y \in R(A)^\perp \ \therefore N(A^*) \subseteq R(A)^\perp$$

From here, consider $y \in R(A)^\perp, x \in \mathbb{R}^n$

$$< y, Ax > = 0 = < A^*y, x > \therefore \ y \in N(A^*) \ \therefore R(A)^\perp \subseteq N(A^*)$$

Combining these two statements, formed from properties of the adjoint

$$\Rightarrow N(A^*) = R(A)^\perp$$

Now, lets *prove* 3. which will prove 4. Choose $y \in R(A)$, ie $y = Ax \ \forall x \in \mathbb{R}^n$. Any $x \in \mathbb{R}^n$ can be written as:

$$x = \underbrace{A^*\bar{y}}_{\in R(A^*)} + \underbrace{z}_{\in N(A)}$$

$$\underbrace{Ax}_{y = AA^*\bar{y}} = AA^*\bar{y} \ \therefore \ R(A) \subseteq R(AA^*)$$

and the other direction, and for some $y' \in \mathbb{R}^m$:

$$\therefore \ y \in R(AA^*), \ y = \underbrace{AA^*y'}_{\therefore \ y \in R(A)}$$

$$\therefore \ R(AA^*) \subseteq R(A) \Rightarrow R(AA^*) = R(A)$$

After this theorem, actually onto the **Least Square Optimization**: ultimately, gets this name as it is broadly:

$$min \left\| y - Ax \right\|_2^2, \ y \notin R(A)$$

To paint the picture, we are mapping from a vector space x to y, and the range $R(A)$ for the mapping $\mathcal{A} : X \mapsto Y$, where the y we are looking at may not fall in $R(A)$. The picture in $\mathbb{R}^3$ comes as a mapping a vector onto a planar subspace that minimizes the norm, or a measure of distance, between the range / plane and the arbitrary vector $x \in X$.

More formaly, call it $\hat{y}$, to y in $R(A)$ Since we cannot solve $y = Ax$ we instead solve $\hat{y} = Ax$ where $y - \hat{y}$ is *orthogonal* to the $R(A)$.

Aside:

. . . . . .

$$\min_{\hat{y}} \ < \hat{y} - Ax, \hat{y} - Ax > = \min_{\hat{y}} \ < y - \hat{y}, y - \hat{y} > = \min_{\hat{y}} \ < y - \hat{y}, y > \ - \ < y - \hat{y}, \hat{y} >$$

. . . . . .

$$\therefore (y - \hat{y}) \in R(A)^{\perp}, \qquad y - \hat{y} \in N(A^*)$$

$$\therefore A^*(y - \hat{y}) = \theta, \qquad \underline{A^*y - A^*A\hat{x} = \theta}$$

$$\therefore A^*A\hat{x} = A^*y$$

Looking into weather the matrix A^*A is invertible for a closed form solution, $\hat{x}$ called the least squares solution minimizing the original equation:

$$\hat{x} = (A^*A)^{-1}A^*y \qquad if \ (A^*A)^{-1} \ exists!$$

If A (as the matrix representation of $\mathcal{A}$), has full column rank, then $null(A) = \theta$ (If not, then have $\tilde{x} \in N(A)$ s.t. $\tilde{x} \neq \theta$, and $A\tilde{x} = \theta$, which forms a contradiction as you try to find a linearly dependent solution to the column vectors of A as you expand it into coefficients and basis form). If $N(A) = \{\theta\}$, then $dim(R(A^*)) = n, \ \therefore dim(R(A^*A)) = dim(R(A^*)) = n$. Therefore, after much ado, matrix A^*A is invertible.

6.3 Differential Equations

6.3.1 Theorem & Definitions

$$\dot{x} = f(x,t); \qquad x(t) \in \mathbb{R}^n; \ x(t_0) = x_0; \ f(x,t) : \mathbb{R}^n \times \mathbb{R}_+ \mapsto \mathbb{R}^n$$

Under what conditions (a) does a solution exists, ie. meaning that $x(t_0) = x_0$ guarantees that $x(t)$ is defined for all $t \geq t_0$? and (b) is the solutions unique?

Definition: $f(x,t) : \mathbb{R}^n \times \mathbb{R}_+ \mapsto \mathbb{R}^n$ is piecewise continuous in t (PC) $\forall \ x$ if $f(x,\cdot) : \mathbb{R}_+ \mapsto \mathbb{R}^n$ is continuous except at points of discontinuity, and there can only be finitely many points of discontinuity in any compact interval.

Definition: $f(x,t) : \mathbb{R}^n \times \mathbb{R}_+ \mapsto \mathbb{R}^n$ is Lipshitz Continuous in x (LC) $\forall \ t$ if there exists a piecewise continuous function $k(\cdot) : \mathbb{R}_+ \mapsto \mathbb{R}_+$ such that

$$\|f(x,t) - f(y,t)\| \leq k(t) \|x - y\| \qquad \forall \ x, y \in \mathbb{R}^n, \forall \ t \in \mathbb{R}_+$$

This inequality is called the Lipschitz Condition.

aside

. . . . . .

Remark: This means that there exists one $k(t)$ which *works* for all $x, y \in \mathbb{R}^n,\ t \in \mathbb{R}_+$

Remark 2: Often, we are dealing with DE's that are not explicitly functions of time (ie. they are Time Invariant).

$$\dot{x} = f(x) = \begin{bmatrix} f_1(x) & f_2(x) & \cdots & f_n(x) \end{bmatrix}^T$$

$$ie. \dot{x} = \underbrace{Ax}_{\text{not an explicit function of time}} \qquad \underbrace{\dot{x} = 3x^2 + 2x + t^2}_{f(x,t) \text{ is an explicit function of time}}$$

in these cases, $f(x)$ is trivially a P.C. function of time. With this, one can come up with a nice method to estimate (a bound on) $k(t)$.

Definition: If $Df(x)$ exists, where $Df(x)$ is the Jacobian of f wrt x, then its norm provides a possible $k(t)$

$$Df(x) = \begin{bmatrix} \frac{df_1}{dx_1} & \frac{df_1}{dx_2} & \cdots & \frac{df_1}{dx_n} \\ \frac{df_2}{dx_1} & \frac{df_2}{dx_2} & \cdots & \frac{df_2}{dx_n} \\ \vdots & \vdots & & \vdots \\ \frac{df_n}{dx_1} & \frac{df_1}{dx_2} & \cdots & \frac{df_n}{dx_n} \end{bmatrix}$$

for $\dot{x} = f(x)$, if $Df(x)$ exists, then its norm provides a (local) Lipschitz function $k(t)$

Sketch of proof from Mean Value Theorem:

Consider an arbitrary function defined on interval $[a, b]$ that is continuous. $\exists$ at least one point, say $f'(c) = \frac{f(b)-f(a)}{b-a}$ on the interval $[a, b]$, where the tangent (instantaneous slope) of the curve is equal to the average slope in $[a, b]$. From here, we can generalize to $\mathbb{R}^n$, $\exists \lambda \in [0, 1]$ st

$$f(x) - f(y) = Df\big(\lambda x + (1-\lambda)y\big)\big(x - y\big)$$

and taking the norm of both sides

$$\therefore \|f(x) - f(y)\| \leq \|Df\|_i \|x - y\| \qquad \star \forall\ \lambda$$

This is often a good thing to try first, when showing that $f(x)$ is LC

i. given $\dot{x} = f(x)$
ii. take $Df(x)$
iii. take $\|Df(x)\|_i$

Good norms to try:

- $\|Df(x)\|_{i,1} \to k(x)$
- $\|Df(x)\|_{i,\infty} \to k_2$, which if constant shows its globally LC
- $\|Df(x)\|_{i,2} \to k_3(x)$In general, these will depends on x.

if $k_i(x) \leq M$ in a bounded region of $\mathbb{R}^n$, that is enough s.t. $f(x)$ is locally Lipschitz continuous

example: Here is a function that is locally, but not globally Lipschitz

$$x \in \mathbb{R}^1 \qquad \dot{x} = -x^2 \;\; (f(x) = -x^2), \;\; x(0) = -1$$

Basic calculus yields,

$$x(t) = \frac{1}{t-1} \leftarrow \text{satisfies DE, IC}$$

But, $x(t)$ is not a solution for all t since at $t = 1,\ x$ is not defined.

Let's look at the LC of $-x^2$:

$$\|f(x_1) - f(x_2)\| \leq \|-2x\| \, \|x_1 - x_2\|$$

Since I can say that $-2x \leq M$ for some bounded x, such as a ball around 0 of radius M in any dimension of reals. Thus, $-x^2$ is *locally Lipschitz continuous*. This will enable us to say that a solution to $\dot{x} = -x^2$ exists and is unique locally in t.

. . . . . .

Cauchy Sequences: Related to vector spaces eg V with a well defined inner product, or an inner product space - leading to Hilbert Space. Or a space V with a well defined norm, is a normed vector space - which leads to an Banach Space.

Definition: on a vector space V, the sequence $\{v_i\}_{i=0}^{\infty}$ is called a Cauchy Sequence if $\forall \epsilon > 0 \; \exists m \; st \; \forall i, j > m$

$$\|v_i - v_j\| < \epsilon$$

Definition: a vector space in which all Cauchy Sequences converge (each to a possibly different vector in the space) is called a complete vector space.

- A complete inner product space is called a Hilbert Space
- A complete normed vector space is called a Banach Space

example: consider V to be the space of rational numbers. Start with the $\sqrt{2} = 1.4142$, and the sequence

$$1, \; 1.4, \; 1.41, \; 1.414, \; \cdots$$

which is a Cauchy sequence, yet it is not converging to the space of rational numbers.

example: $f(x) = \frac{1}{x}$ is continuous, but not Lipschitz continuous.

$$\|f(x_1) - f(x_2)\| \stackrel{?}{\leq} K \, \|x_1 - x_2\|$$

6.3.2 Fundamental Theorem

Fundamental Theorem of Differential Equations (often called the Existence and Uniqueness

Theorem): Consider $\dot{x} = f(x,t),\ x(t_0) = x_0$ with $f(x,t)$ piecewise continuous in t and Lipschitz continuous in x. Then there exists a unique function of time $\phi(\cdot) : \mathbb{R}_+ \mapsto \mathbb{R}^n$ which is C^1 (continuous almost everywhere, and its derivatives are continuous almost everywhere) satisfying:

$$\phi(t_0) = x_0 \qquad \dot{\phi}(t) = f\big(\phi(t), t\big) \qquad \forall\ t \in [t_1, t_2] \setminus D$$

Where D is the set of discontinuous points for f as a function of t. $\setminus$ is the given set minus the following subset. We are working in arbitrarily large but finite interval of $\mathbb{R}$.

Using this, we will be able to generate mathematical models of input-output systems. Consider, for example:

$$\begin{aligned} \dot{x} &= f(x,u,t), & f &: \mathbb{R}^n \times \mathbb{R}^{n_i} \times \mathbb{R}_+ \mapsto \mathbb{R}^n \\ y &= h(x,u,t), & f &: \mathbb{R}^n \times \mathbb{R}^{n_i} \times \mathbb{R}_+ \mapsto \mathbb{R}^{n_0} \end{aligned}$$

If f is Lipschitz continuous in x, continuous in u, and piecewise continuous in t, we are guaranteed that given $x(t_0) = x_0\ \exists!\ x(t) \in \mathbb{R}^n$ satisfying the differential equation. With this, $\exists!\ y(t) \in \mathbb{R}^{n_0}$, called the output of the system.

Note: if the Lipschitz condition does not hold, it may be that the solution cannot be continuous beyond a certain time - **example**, consider:

$$\dot{\zeta}(t) = \zeta(t)^2, \qquad \zeta(0) = \frac{1}{c},\ x \neq 0,\ \zeta(t) : \mathbb{R}_+ \mapsto \mathbb{R}$$

the differential equation above has the following solution on $t \in (-\infty, c)$

$$\zeta(t) = \frac{1}{c-t}, \qquad as\ t \to c,\ \|\zeta(t)\| \to \infty$$

Above is an example of *finite escape time at c.*

Proof: of the Fundamental Theorem.

A. Existence: Construct a sequence of continuous functions

$$x_{m+1}(t) := x_0 + \int_{t_0}^{t} f(x_m(\tau), \tau) d\tau \qquad x_0(t_0) := x_0 \qquad m = 0, 1, 2, \cdots$$

The idea is to show that the sequence of continuous functions

$$\big\{x_m(\cdot)\big\}_0^\infty$$

Converges to:

1. a continuous function $\phi(t) : \mathbb{R}_+ \mapsto \mathbb{R}^n$.
2. which is a solution of $\dot{x} = f(x,t),\ \ x(t_0) = x_0$.

Construction of a solution by iteration

To show 1. we show that $\{x_m(\cdot)_0^\infty$ is a Cauchy sequence in a Banach Space $\Big(C\big([t_1,t_2],\mathbb{R}^n\big),\mathbb{R},\|\cdot\|_\infty\Big)$; where $t_0 \in [t_1,t_2]$ (be careful because the space of continuous functions given above C... but with the $\|\cdot\|_1$ is not Banach - Math 104) :

$$\begin{aligned}\|x_{m+1}(t)+x_m(t)\| &= \Big\|\int_{t_0}^t \Big(f\big(x_m(\tau),\tau\big)-f\big(x_{m-1}(\tau),\tau\big)\Big)d\tau\Big\| \\ &\le \int_{t_0}^t \Big\|\Big(f\big(x_m(\tau),\tau\big)-f\big(x_{m-1}(\tau),\tau\big)\Big)\Big\|\,d\tau \\ &\le \int_{t_0}^t k(\tau)\,\big\|\big(x_m(\tau),\tau-x_{m-1}(\tau),\tau\big)\big\|\,d\tau \qquad \textit{by L.C. of } f\end{aligned}$$

let $\bar{k}$ =$\sup_{[t_1,t_2]} k(t)$, then you get

$$\|x_{m+1}(t)+x_m(t)\| \le \bar{k}\int_{t_0}^t \|x_m(\tau)-x_{m-1}(\tau)\|\,d\tau$$

Now, we know by the definition of $\{x_m(\cdot)\}_0^\infty$ that

$$x_1(t) := x_0 + \int_{t_0}^t f(x_0,\tau)d\tau, \qquad t\in[t_1,t_2]$$

$$\therefore \|x_1(t)-x_0\| \le \int_{t_0}^t \|f(x_0,\tau)\|\,d\tau \le \int_{t_1}^{t_2} \|f(x_0,\tau)\|\,d\tau =: M$$

since we know $x_0, f, t_1, t_2, \cdots\ M$ is known.

$$\therefore \|x_2(t)-x_1(t)\| \le M\bar{k}|t-t_0|$$

and continuous recursively

$$\|x_3(t)-x_2(t)\| \le \frac{M\bar{k}^2|t-t_0|^2}{2!} \qquad \star\, check$$

$$\vdots$$

$$\|x_{m+1}(t)-x_m(t)\| \le \frac{M(\bar{k}|t-t_0|)^m}{m!}$$

here, recall that

$$\|f(\cdot)\|_\infty = max\big\{\,\|f(t)\|\,, t\in[t_1,t_2]\big\}$$

Continuing from above, and defining $T = t_2 - t_1$. Replaced point-wise vector norm with function norm.

$$\therefore \|x_{m+1}(\cdot)-x_m(\cdot)\|_\infty \le \frac{M(\bar{k}T)^m}{m!} \qquad m=0,1,2,\cdots$$

Next, to see that $\{x_m(\cdot)\}_0^\infty$ *is a Cauchy Sequence in* $\Big(C\big([t_1,t_2],\mathbb{R}^n\big),\mathbb{R},\|\cdot\|_\infty\Big)$:

$$\begin{array}{rcc}
\|x_{m+p}(\cdot)-x_m(\cdot)\|_\infty & = & \|x_{m+p}(\cdot)-x_{m+p-1}(\cdot)+x_{m+p-1}(\cdot)-x_{m+p-2}(\cdot)\cdots x_m(\cdot)\|_\infty \\
& = & \big\|\sum_{k=0}^{p-1}\big(x_{m+k+1}(\cdot)-x_{m+k}(\cdot)\big\|_\infty \\
& \le & \sum_{k=0}^{p-1}\|x_{m+k+1}(\cdot)-x_{m+k}(\cdot)\|_\infty \\
& \le & M\sum_{k=0}^{p-1}\frac{(\bar{k}T)^{m+k}}{(m+k)!} \\
& \le & M\frac{(\bar{k}T)^m}{m!}\sum_{k=0}^{p-1}\frac{(\bar{k}T)^k}{k!}
\end{array}$$

And getting to the end, since $(m+k)! \ge m!k!$ via Stirling's Formula

$$\le M\frac{(\bar{k}T)^m}{m!}e^{\bar{k}T} \qquad \star$$

And, because factorial in m grows faster than exponential in m, and $e^{\bar{k}T}=\sum_{k=0}^\infty \frac{(\bar{k}T)^k}{k!}$

$$\therefore \{x_m(\cdot)\}_0^\infty \text{ is Cauchy}$$

Finally, because a Cauchy Sequence in Banach Space, converges to a continuous function $\phi(\cdot)$ in the space $\Big(C\big([t_1,t_2],\mathbb{R}^n\big),\mathbb{R},\|\cdot\|_\infty\Big)$.

To show ($\phi(\cdot)$ is a solution of the differential equation:

$$x_{m+1}(t) := x_0 + \int_{t_0}^t f\big(x_m(\tau),\tau\big)d\tau$$

as $m\to\infty,\ x_m(\cdot)\to\phi(\cdot)$ (on $[t_1,t_2]$) we've just proved this

$$\therefore \textit{ need to show } \int_{t_0}^t f\big(x_m(\tau),\tau\big)d\tau \to \int_{t_0}^t f\big(\phi(\tau),\tau\big)d\tau \textit{ as } m\to\infty$$

Show: $\phi(\cdot)$ solves the D.E., so that $\dot{\phi}=f(\phi,t)$ Indeed,

$$\left\|\int_{t_0}^t\Big(f\big(x_m(\tau),\tau\big)d\tau - f\big(\phi(\tau),\tau\big)\Big)d\tau\right\|$$

By bringing norm under integral. Then by Lipschitz condition.

$$\le \int_{t_0}^t k(\tau)\,\|x_m(\tau)-\phi(\tau)\|\,d\tau$$

Where $\bar{k}=\max_{t\in[t_1,t_2]}k(t)$ and $\|x_m(\cdot)-\phi(\cdot)\|_\infty=\max_{t\in[t_1,t_2]}\|x_m(\tau)-\phi(\tau)\|$

$$\le \bar{k}\,\|x_m(\cdot)-\phi(\cdot)\|_\infty\cdot T$$

And by letting $m \to \infty$ in $\star$, which gives us an arbitrarily small upper bound.

$$\leq \bar{k} M e^{\bar{k}T} \frac{(\bar{k}T)^m}{m!} \cdot T$$

$$\therefore \phi(t) = x_0 + \in_{t_0}^{t} f\big(\big(\phi(\tau), \tau\big) d\tau \qquad \forall t \in [t_1, t_2]$$

$$\therefore \dot{\phi}(t) = f\big(\phi(t), t\big), \phi(t_0) = x_0 \leftarrow \qquad \forall t \in [t_1, t_2], t \notin D$$

Since $[t_1, t_2]$ is arbitrarily (containing t_0), then we can conclude that the proposed iterative scheme converges to a solution ϕ on $\mathbb{R}_+$.

We have constructed a solution on $\mathbb{R}_+$. Conceivably, a different construction might lead to another solution. Thus, we have to verify that ϕ is the unique solution:

B. Uniqueness:

To prove uniqueness, we will need the Bellman-Gronwall Lemma: let $u(\cdot), k(\cdot)$ be real-valued, piecewise continuous functions on $\mathbb{R}_+$; and assume $u(\cdot), k(\cdot) > 0$ on $\mathbb{R}_+$. Assume $c_1 > 0, t_0 \in \mathbb{R}_+$. Then, if

$$u(t) \leq c_1 + \int_{t_0}^{t} k(\tau) u(\tau) d\tau \qquad (**)$$

Then,

$$u(t) \leq c_1 e^{\int_{t_0}^{t} k(\tau) d\tau}$$

Proof:

Without loss of generality, assume $t > t_0$. Let $U(t) = c_1 + \int_{t_0}^{t} k(\tau) u(\tau) d\tau$.

Thus, $u(t) \leq U(t) (***)$.

Begin multiplying both sides of $(***)$ by the non-negative function $k(t) e^{-\int_{t_0}^{t} d\tau}$. Resulting in

$$\frac{d}{dt} \Big\{ U(t) e^{-\int_{t_0}^{t} d\tau} \Big\} \leq 0$$

And then integrate between t_0 and t, which is the integral of a derivative at specific points

$$u(t) \leq U(t) \leq c_1 e^{-\int_{t_0}^{t} d\tau}$$

■

Finally, using Bellman-Gromwall to show uniqueness:

$$\dot{x} = f\big(x(t), t\big), \qquad x(t_0) = x_0$$

Where f is piecewise continuous in t and Lipschitz continuous in x. We have shown there exists a

solution $\phi(t)$ to the above; suppose there are two solutions $\phi\&\zeta$ satisfying the above:

$$\dot{\phi}(t) = f\big(\phi(t), t\big),\ \phi(t_0) = x_0 \qquad \dot{\xi}(t) = f\big(\xi(t), t\big),\ \xi(t_0) = x_0$$

Start by looking at the difference of the solutions:

$$\therefore \phi(t) - \xi(t) = \int_{t_0}^{t} \Big(f\big(\phi(t), t\big) - f\big(\xi(t), t\big) \Big) d\tau \qquad \forall t \in \mathbb{R}_+$$

Bringing norm in integral, using Lipschitz Continuity ...

$$\underbrace{\|\phi(t) - \xi(t)\|}_{u(t)} \leq \bar{K} \int_{t_0}^{t} \underbrace{\|\phi(\tau) - \xi(\tau)\|}_{U(T)} d\tau \qquad \forall t \in [t_1, t_2]$$

From Bellman-Gromwall with $c_1 = 0, k(t) = \bar{K}$:

$$if\ \ \|\phi(t) - \xi(t)\| \leq c_1 + \bar{K} \int_{t_0}^{t} \|\phi(\tau) - \xi(\tau)\|\, d\tau$$

$$then\ \ \|\phi(t) - \xi(t)\| \leq c_1 e^{\bar{K}(t - t_0)}$$

but here, $c_1 = 0$, thus,

$$\|\phi(t) - \xi(t)\| = 0 \qquad \Rightarrow \qquad \phi(t) = \xi(t)$$

Above is a contradiction, therefore we have proved uniqueness! Finishing this proof.

6.4 Alternate State Space Definitions:

6.4.1 Dynamical Systems:

We abstract into a formal definitions of dynamical systems:

- *time* T: $(-\infty, \infty)$ on $[0, \infty)$ in the continuous case; on $\{nT, n \in \mathbb{Z}\}$ or $\{nT, n \in N\}$ in the discrete-time case.
- The inputs, outputs, and state variables are functions of time defined on T. A dynamical system is represented by:

$$(u, \Sigma, Y, s, r)\ where$$

1. Inputs: U is a set of input functions from $\tau \mapsto U$; typically U is $\mathbb{R}^{n_i}$.
2. Outputs: Y is a set of output functions from $\tau \mapsto Y$; typically Y is $\mathbb{R}^{n_o}$.

3. States: Σ is a set called the state space. Typically $\Sigma = \mathbb{R}^n$. The map $x(\cdot) : \tau \mapsto \Sigma$ is called a state trajectory.

4. State Transition Function:

$$s : \tau \times \tau \times \Sigma \times U \mapsto \Sigma$$

for x_0 at t_0 and input $u \in U$, and $t_1 \in \tau$ given (usually $t_1 \geq t_0$) the state x at time t_1 is given by $x(t_1) = s(t_1, t_0, x_0, u)$.

5. Output Read-out Map:

$$r : \tau \times \Sigma \times U \mapsto Y$$

for $t \in \tau$, $x(t) \in \Sigma$, $u(t) \in U$, $\exists$ an output $y(t)$ given by

$$y(t) = r\big(t, x(t), u(t)\big)$$

Note that the read-out function r is "memoryless" - ie all arguments are evaluated at the same time t.

The state transition map is required to satisfy two axioms:

1. State Transition Axiom: for all $t_1 \geq t_0 \in \tau$, if $u, \bar{u} \in U$ with

$$u(t) \equiv \bar{u}(t) \qquad \forall t \in [t_0, t_1] \cap \tau$$

then $s(t_1, t_0, x_0, u) = s(t_1, t_0, x_0, \bar{u})$. This property is suggested by writing

$$x(t_1) = s(t_1, t_0, x_0, u_{[t_0, t_1]})$$

where the notation $u_{[t,0,t_1]}$ denotes the restriction of u to the interval $[t_0, t_1]$

2. Semi-Group Axiom

$$\forall t_0 \leq t_1 \leq t_2 \in \tau, \qquad \forall x_0 \in \Sigma, \qquad \forall u \in U$$

$$s(t_2, t_1, s(t_1, t_0, x_0, u), u) = s(t_2, t_0, x_0, u)$$

Some remarks on these definitions:

1. Given x_0 and t_0, $x(t_1)$ does not depend on the input u prior to t_0; the state x_0 summarizes all of the effects of the input prior to t_0.

2. Given x_0 and t_0, $x(t_1)$ does not depend on the values of the input after t_1; ie the system is not anticipative.

The composition of the state transition function and the output read-out map is called the response function:

$$y(t) = r\big(t, s(t, t_0, x_0, u), u(t)\big)$$

$$=: \rho(t, t_0, x_0, u_{[t_0,t]})$$

6.4.2 Time-Invariant Dynamical Systems

Define the shift operator: $T_\tau : U \mapsto U$ as

$$(T_\tau u)(t) = u(t - \tau)$$

Definition: A dynamical system is said to be time-invariant if

1. U is closed under $T_\tau \ \ \forall \tau$
2. $\forall t_0, t_1 \geq t_0, \tau \in T, \forall x_0 \in \Sigma, \forall u \in U$

$$\rho(t_1, t_0, x_0, u) = \rho(t_1 + \tau, t_0 + \tau, x_0, T_\tau u)$$

Linear Dynamical Systems: A dynamical system is said to be *linear* if:

(a) U, Σ, Y are all linear spaces over the same field $\mathbb{F}$

(b) $\forall t \geq t_0, t_0 \in \tau$, the response map ρ is a linear map of $\Sigma \times U$ into Y (ie closed under linear combinations of states and inputs):

$$\rho(t, t_0, \alpha_1 x_1 + \alpha_2 x_2, \alpha_1 u_1 + \alpha_2 u_2 = \alpha_1 \rho(t, t_0, x_1, u_1) + \alpha_2 \rho(t, t_0, x_2, u_2)$$

Remarks

1. if θ_Σ is the zero-element of Σ and θ_u is the zero element of U, then

$$\rho(t, t_0, x_0, u) = \underbrace{\rho(t, t_0, \theta_\Sigma, u)}_{zero\ state\ response} + \underbrace{\rho(t, t_0, x_0, \theta_u)}_{zero\ input\ response}$$

2. Linearity of the zero state response (superposition):

$$\rho(t, t_0, \theta_\Sigma, \alpha_1 u_1 + \alpha_2 u_2) = \alpha_1 \rho(t, t_0, \theta_\Sigma, u_1) + \alpha_2 \rho(t, t_0, x_1, \theta_\Sigma, u_2)$$

3. Linearity of the zero-input response:

$$\rho(t, t_0, \alpha_1 x_1 + \alpha_2 x_2, \theta_u) = \alpha_1 \rho(t, t_0, x_1, \theta_u) + \alpha_2 \rho(t, t_0, x_1, x_2, \theta_u)$$

6.5 Other System Representation Examples

6.5.1 Jacobian Linearization

Consider the following differential equation:

$$\dot{x} = f(x, u, t), \qquad x(t_0) = x_0$$

Let the input $u^o(\cdot)$ results in the state $x^o(\cdot)$. Now let $u^o(\cdot)$ be perturbed to

$$u^o(\cdot) + \delta u(\cdot)$$

With resultant state perturbation

$$x^o(\cdot) + \delta x(\cdot)$$

Also, let the initial condition be perturbed to $x_0 + \delta x_0$.

$$\therefore \dot{x}^o = f(x^o, u^o, t); \qquad x^o(t_0) = x_0$$

$$\dot{x}^o + \delta\dot{x} = f(x^o + \delta x, u^o + \delta u, t); \qquad x^o + \delta x(t_0) = x_0 + \delta x_0$$

$$f(x^o + \delta x, u^o + \delta u, t) = f(x^o, u^o, t) + \frac{d}{dx} f(x, u, t)\Big|_{(x^o, u^o)^T} \delta x + \frac{d}{du} f(x, u, t)\Big|_{(x^o, u^o)^T} \delta u + h.o.t.$$

Putting the above in matrix form yields a lienarization:

$$\delta\dot{x} = \underbrace{D_1 f(x^o, u^o, t)}_{A(t) \in \mathbb{R}^{n \times n}} \delta x + \underbrace{D_2 f(x^o, u^o, t)}_{B(t) \in \mathbb{R}^{n \times n_i}} \delta u$$

$$\therefore \delta x(t) = \Phi(t_1, t_0)\delta x_0 + \int_{t_0}^{t_1} \Phi(t, t')B(t')\delta u(t')dt'$$

6.5.2 Affine Systems

Suppose the system dynamics are now given by:

$$\dot{x}(t) = Ax(t) + Bu(t) + c$$
$$x(0) = x_0$$

Define a new state variable $z = [x(t), 1]^T$, yielding

$$\dot{(z)} = \begin{bmatrix} \dot{(x(t))} \\ 1 \end{bmatrix} = \begin{bmatrix} A & c \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x(t) \\ 1 \end{bmatrix} + \begin{bmatrix} B \\ 0 \end{bmatrix} u(t) = A'z(t) + B'u(t)$$

6.6 Eigenvalue Placement by State Feedback

6.6.1 SISO Systems, alternate derivation

$$\dot{x} = Ax + bu, \qquad b \in \mathbb{R}^n,\ u \in \mathbb{R}$$

For the single input, single output system, propose that for the following $\bar{A}$ and $\bar{B}$ in canonical controllable fo

$$\exists T \in \mathbb{R}^{n\times n}\ s.t.\ \ \bar{A} = T^{-1}AT = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & 0 & 1 & & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & & & \ddots & 1 \\ -\alpha_n & \cdots & \cdots & \cdots & -\alpha_1 \end{bmatrix} \text{ and } \bar{b} = T^{-1}b = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

$$\Leftrightarrow (A, B) \text{ is completely controllable}$$

Proof ($\Leftarrow$):

$$I = \begin{bmatrix} | & | & & & | & | \\ e_1 & e_2 & \cdots & \cdots & e_{n-1} & e_n \\ | & | & & & | & | \end{bmatrix}$$

Then, $T^{-1}b = e_n$ from above.

$$\bar{A}e_n = T^{-1}ATe_n = e_{n-1} - \alpha_1 e_n \qquad \Rightarrow e_{n-1} = T^{-1}(Ab + \alpha_1 b)$$

$$\bar{A}e_{n-1} = T^{-1}ATe_{n-1} = e_{n-2} - \alpha_2 e_n \qquad \Rightarrow e_{n-2} = T^{-1}(Ab + \alpha_1 Ab + \alpha_2 b)$$

$$\therefore \begin{bmatrix} | & | & & & | & | \\ e_1 & e_2 & \cdots & \cdots & e_{n-1} & e_n \\ | & | & & & | & | \end{bmatrix} = \begin{bmatrix} | & | & & & | & | \\ b & Ab & \cdots & \cdots & A^{n-2}b & A^{n-1}b \\ | & | & & & | & | \end{bmatrix} \begin{bmatrix} \alpha_{n-1} & \cdots & \alpha_1 & 1 \\ \vdots & \iddots & \iddots & 0 \\ \alpha_1 & \iddots & \iddots & 0 \\ 1 & 0 & \cdots & 0 \end{bmatrix}$$

Thus, $b, Ab, A^2b, \cdots, A^{n-1}b$ are linearly independent $\therefore T^{-1}$ exists.

Proof ($\Rightarrow$):Note that if, as above

$$\bar{A} = T^{-1}AT \qquad \bar{b} = T^{-1}B$$

$$\bar{A}\bar{b} = \begin{bmatrix} 0 \\ \vdots \\ \vdots \\ 1 \\ -\alpha_1 \end{bmatrix}; \qquad \bar{A}^2\bar{b} = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ -\alpha_1 \\ \alpha_2 + \alpha_1^2 \end{bmatrix}$$

Thus, $\bar{b}, \bar{A}\bar{b}, \cdots, \bar{A}^{n-1}\bar{b}$are linearly independent.

Now, given a set of eigenvalues $\lambda_1, \lambda_2 \cdots, \lambda_n \in \mathbb{C}$ with complex conjugate symmetry, define:

$$\hat{\pi} = \prod_{i=1}^{n}(s - \lambda_i) = s^n + \pi_1 s^{n-1} + \cdots + \pi_n \qquad \pi_i \in \mathbb{R}$$

Theorem: let (A, b) be completely controllable. Let $\pi\hat{(}s)$ be any monic polynomial of degree n with real coefficients. Then $\exists! f^T \in \mathbb{R}^n$ such that

$$\hat{X}_{A+bf^T}(s) = \hat{\pi}(s)$$

where $f^T = -\begin{bmatrix}0 & 0 & \cdots & 0 & 1\end{bmatrix}^T \begin{bmatrix}b & Ab & \cdots & A^{n-1}b\end{bmatrix}^{-1}\hat{\pi}(A)$. Or, in another form,

$$(A, B)\ c.c. \Leftrightarrow \exists T \in \mathbb{R}^{n\times n}\ s.t.\ \bar{A} = T^{-1}AT = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & 0 & 1 & & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & & & \ddots & 1 \\ -\alpha_n & \cdots & \cdots & \cdots & -\alpha_1 \end{bmatrix} and\ \bar{b} = T^{-1}b = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

Which gives us (after omitted proof),

$$\bar{f}^T = f^T T = \begin{bmatrix}(\alpha_n - \pi_n) & \cdots & \cdots & (\alpha_1 - \pi_1)\end{bmatrix}$$