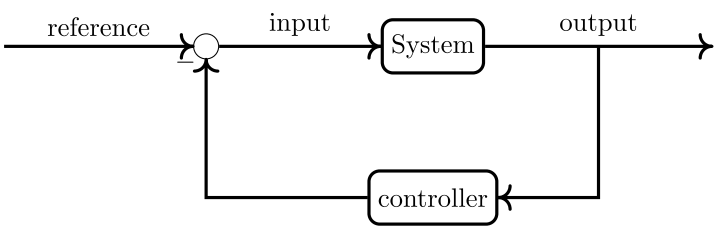
Control Theory: focuses on modeling systems and designing inputs to adjust behavior—e.g., stabilize, track a trajectory, etc.

Classical control (developed largely pre-1960s) largely adopts an input-output approach:



Key theoretical tool: Fourier/Laplace transform (i.e. Frequency domain analysis—root locus, frequency response)
Return to differential equations beginning in the '60s to address:

- numerical simulation
- many inputs/outputs
- ill-defined inputs/outputs
- non-linearities
- optimality

Modern control theory (∼1950's), i.e. state-space approach: Overcame some limitations of classical control enabling control of fighter jets, e.g. (related "state space" approach to ODE's is over 100 years old; control theorist just adopted it)

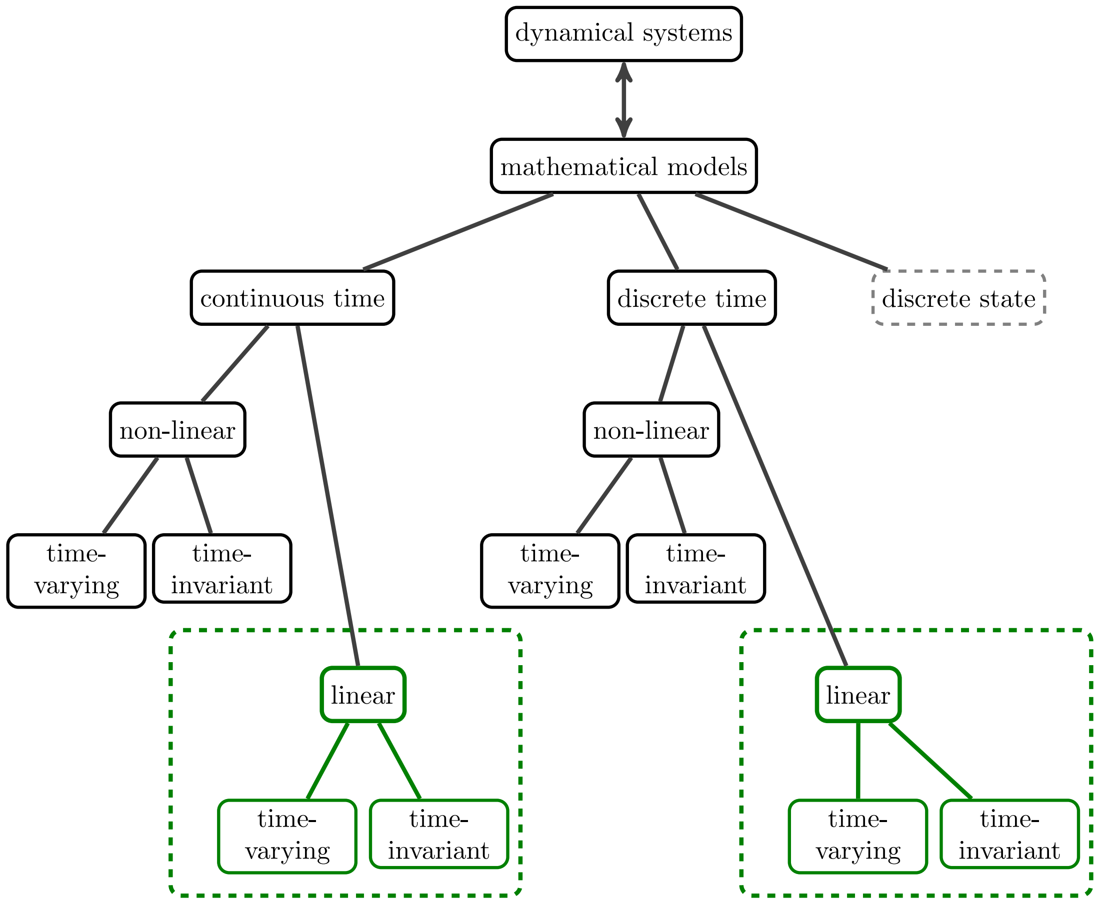
- System/model state is defined to capture all relevant info about past
- State often denoted by $x \in \mathbb{R}^n$, where n is state-space dimension.

What is the state for the above examples?

- engine speed/velocity
- position and yaw/pitch/roll and velocities (aircraft and quadrotor)
- disk speed/velocity
- species population, food supply, predator population, etc.
- temperature along cooling line

3 Dynamical Systems (non-comprehensive) Taxonomy

(This is not comprehensive as there are other types of systems combining various aspects in the diagram; however, this picture give a bit of a sense of broad categories of dynamical systems)



This course: given linear differential or difference equations, conduct rigorous analysis and design.

4 Finite Dimensional Systems as Models

How to describe a dynamical system? Some options:

- Database/look-up table containing all inputs and resulting outputs. *(What if output depends on input history? What if desired input is not in table?)*
- Function/routine in computer code
- Set of mathematical equations

As indicated in our diagram, we will study continuous-time and discrete-time finite-dimensional systems described by ordinary differential equations or difference equations.

So, in continuous time this might look like. . .

Continuous-Time: ($t \in \mathbb{R}_+ := [0, \infty)$)

$$\begin{aligned} \dot{x} &= f(t, x, u), \quad x \in \mathbb{R}^n, \ u \in \mathbb{R}^m \\ y &= g(t, x, u), \quad y \in \mathbb{R}^p \end{aligned}$$

We have the following nomenclature:

- x is the state,
- u is the (control) input,
- y is the output (observation)

Discrete-Time: ($k \in \mathbb{N} = \{0, 1, 2, \dots\}$)

$$\begin{aligned} x[k+1] &= f(k, x, u), \quad x \in \mathbb{R}^n, \ u \in \mathbb{R}^m \\ y[k] &= g(k, x, u), \quad y \in \mathbb{R}^p \end{aligned}$$

What about apparently more exotic systems with higher order derivatives?
For instance, consider

$$z^{(n)} = f(t, z, z^{(1)}, z^{(2)}, \dots, z^{(n-1)})$$

where $z^{(n)}$ indicates the n–th derivative of the function $z(t)$. For simplicity, assume $z \in \mathbb{R}$. Define new state variables

$$x_1 = z, \ x_2 = z^{(1)}, \ \dots, \ x_n = z^{(n-1)}$$

Then, we have

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= x_3 \\ &\vdots \\ \dot{x}_n &= f(t, x_1, x_2, \dots, x_n) \end{aligned}$$

Thus, without loss of generality (w.l.o.g), we study first-order differential equations.

Example Systems: Vibrating Springs

We consider the motion of an object with mass at the end of a spring that is either vertical (as in Figure 1) or horizontal on a level surface (as in Figure 2). Hooke's Law says that if the spring is stretched (or compressed) z units from its natural length, then it exerts a force that is proportional to z—that is,

$$\text{restoring force} = -kx$$

where k is a positive constant called the spring constant. If we ignore any external resisting forces (due to air resistance or friction) then, by Newton's Second Law (force equals mass times acceleration), we have

$$m\ddot{x} = -kx \quad \text{or} \quad m\ddot{x} + kx = 0$$

Let $x_1 = z$ and $x_2 = \dot{z}$. Then, $\dot{x}_1 = \dot{z} = x_2$ and $\dot{x}_2 = \ddot{z} = -\frac{k}{m}z = -\frac{k}{m}x_1$. Hence,

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -\frac{k}{m}x_1 \end{aligned}$$

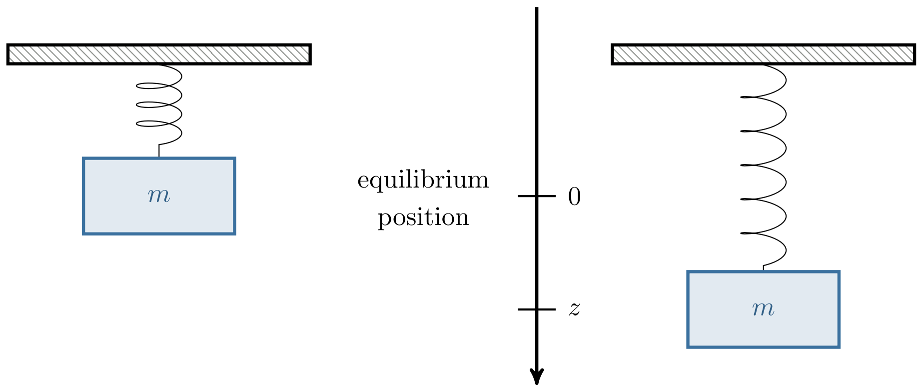


Figure 1: Vertical Pull

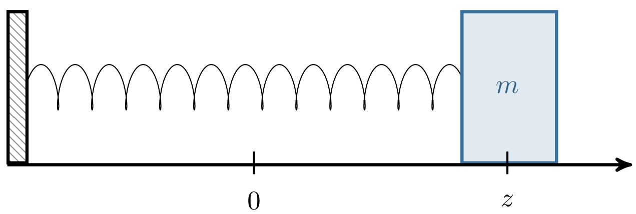


Figure 2: Horizontal Pull

In this course, we will focus on (finite-dimensional) **linear time-varying (LTV)** systems. So what do these look like notationally:

Continuous Time:

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

where

- $t \in \mathbb{R}$: time
- $x(t) \in \mathbb{R}^n$: state (vector)
- $u(t) \in \mathbb{R}^m$: input or control
- $y(t) \in \mathbb{R}^p$: output
- $A(t) \in \mathbb{R}^{n\times n}$: dynamics (matrix)
- $B(t) \in \mathbb{R}^{n\times m}$: input matrix
- $C(t) \in \mathbb{R}^{p\times n}$: output or sensor matrix
- $D(t) \in \mathbb{R}^{p\times m}$: Feedthrough matrix

Equations are often written as

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

- A CT LDS is a first order vector differential equation
- also called state equations or m-input, n-state, p-output LDS

Discrete Time:

$$\begin{aligned} x[k+1] &= A[k]x[k] + B[k]u[k], \quad x \in \mathbb{R}^n, \ u \in \mathbb{R}^m \\ y[k] &= C[k]x[k] + D[k]u[k], \quad y \in \mathbb{R}^p \end{aligned}$$

Finally, we will further specialize our results to **linear time-invariant (LTI)** systems.

Continuous Time LTI:

$$\begin{aligned} \dot{x} &= Ax + Bu, \quad x \in \mathbb{R}^n, \ u \in \mathbb{R}^m \\ y &= Cx + Du, \quad y \in \mathbb{R}^p \end{aligned}$$

where $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{p\times n}$, and $D \in \mathbb{R}^{p\times m}$ are static matrices.

Discrete Time LTI:

$$\begin{aligned} x[k+1] &= Ax[k] + Bu[k], \quad x \in \mathbb{R}^n, \ u \in \mathbb{R}^m \\ y[k] &= Cx[k] + Du[k], \quad y \in \mathbb{R}^p \end{aligned}$$

Some other points:

- most linear systems encountered are time-invariant: A, B, C, D are constant as above—i.e. don't depend on t
- when there is no input u (hence, no B or D) system is called autonomous
- very often there is no feedthrough—i.e. $D = 0$
- when $u(t)$ and $y(t)$ are scalar, the system is called single-input, single-output (SISO); when input & output signal dimensions are more than one, MIMO

History Lesson:

- parts of LDS theory can be traced to 19th century
- builds on classical circuits & systems (1920s on) (transfer functions ...) but with more emphasis on linear algebra
- first engineering application: aerospace, 1960s
- transitioned from specialized topic to ubiquitous in 1980s (just like digital signal processing, information theory, ...)

Many dynamical systems are nonlinear, yet

- most techniques for nonlinear systems are based on linear methods; e.g., linearization to determine stability or to construct extended Kalman Filters...
- methods for linear systems often work unreasonably well, in practice, for nonlinear systems
- if you don't understand linear dynamical systems you certainly can't understand nonlinear dynamical systems

AA/EE547: Wi 2020

Lecture 2: Introduction to Linear Dynamical Systems and ODE Review

Lecturer: L.J. Ratliff

1 General Systems Models

A continuous time finite dimensional dynamical system is specified by a systems of equations of the form

$$\begin{aligned}\dot{x}_i(t) &= f_i(t, x_1(t), \ldots, x_n(t), u_1(t), \ldots, u_m(t)),\ i = 1, \ldots, n\\ y_i(t) &= g_i(t, x_1(t), \ldots, x_n(t), u_1(t), \ldots, u_m(t)),\ i = 1, \ldots, p\end{aligned}$$

where u_i, $i = 1, \ldots, m$ denote the inputs and y_i, $i = 1, \ldots, p$ denotes the outputs (responses), x_i, $i = 1, \ldots, m$ denote the state variables, and t denotes time.

A complete description of such systems will usually also require[1] a set of initial conditions $x_i(t_0)$, $i = 1, \ldots, n$ where t_0 is the initial time.

In vector notation, we write the above as

$$\begin{aligned}\dot{x} &= f(t, x(t), u(t)),\ x \in \mathbb{R}^n,\ u \in \mathbb{R}^m\\ y &= g(t, x, u),\ y \in \mathbb{R}^p\end{aligned}$$

The discrete time counter part is given by

$$\begin{aligned}x(k+1) &= f(k, x(k), u(k)),\ x \in \mathbb{R}^n,\ u \in \mathbb{R}^m\\ y(k) &= g(t, x(k), u(k)),\ y \in \mathbb{R}^p\end{aligned}$$

Sometimes we write $x(k) = x_k$.

2 Linear Dynamical System Representation

We study dynamical systems of the form

$$\begin{aligned}\dot{x}(t) &= A(t)x(t) + B(t)u(t) \quad \text{(state DE)}\\ y(t) &= C(t)x(t) + D(t)u(t) \quad \text{(read-out eqn.)}\end{aligned}$$

where as usual $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^m$, and $y(t) \in \mathbb{R}^p$. $A(\cdot), B(\cdot), C(\cdot), D(\cdot)$ are matrix-valued functions on $\mathbb{R}_+$, assumed to be *piecewise continuous*. We are going to start with continuous time.

2.1 A brief review of analysis

[1]The exception being if other sufficient data are given such as boundary conditions.

Definition 2.1 (***Continuity***) Let $W \subset \mathbb{R}$ denote an open interval and consider a function $f : W \to \mathbb{R}$. The function f is said to be continuous at the point $t_0 \in U$ if

$$\lim_{t \to t_0} f(t) = f(t_0)$$

exists. That is, for every $\varepsilon > 0$, there exists $\delta(t_0, \varepsilon) > 0$ such that

$$|f(t) - f(t_0)| < \varepsilon$$

whenever $|t - t_0| < \delta(t_0, \varepsilon)$ and $t \in W$. If f is continuous at every point in W, then we just say f is a continuous function or we say f is continuous on W.

Note that δ depends on the choice of t_0 and ε in the above definition. If at each $t_0 \in W$, it is true that there is a $\delta > 0$ independent of t_0 (i.e., $\delta = \delta(\varepsilon)$), such that $|f(t) - f(t_0)| < \varepsilon$ whenever $|t - t_0| < \delta$ and $t \in W$, then we say f is *uniformly continuous* on W.

We denote by $C(W, \mathbb{R})$ the space of real-valued continuous functions on W.

Definition 2.2 (***Piecewise Continuity.***) A function or curve is piecewise continuous if it is continuous on all but a finite number of points in any compact (closed and bounded in $\mathbb{R}$) interval.

For a function $f : W \to \mathbb{R} : t \mapsto z$, let $f^{(0)} \equiv f$ and $f^{(k)} \equiv \partial^k f / \partial t^k$. For $W \subset \mathbb{R}$ an open set, given $r \in \mathbb{N}$ and W an open set we will use the notation

$$C^r(W, \mathbb{R}) = \{f : W \to \mathbb{R} |\ f^{(j)} \text{ exists on } W \text{ and } f^{(j)} \in C(W, \mathbb{R}),\ j = 0, 1, \ldots, r\}$$

for the C^r-functions. For $W \subset \mathbb{R}^n$ with non-empty interior, we can define $C(W, \mathbb{R})$ and $C^r(W, \mathbb{R})$ similarly. Indeed,

$$C^r(W, \mathbb{R}) = \left\{ f : W \to \mathbb{R} |\ f^{(j)} \text{ exists on } W \text{ and } f^{(j)} = \frac{\partial^j f}{(\partial x_1^{i_1} \cdots \partial x_n^{i_n}} \in C(W, \mathbb{R}), \sum_\ell i_\ell = j,\ j = 0, 1, \ldots, r \right\}$$

Let W be a subset of $\mathbb{R}^n$ with non-empty interior and let $f : W \to \mathbb{R}^m$. Then $f = (f_1, \ldots, f_m)^\top$ where $f_i : W \to \mathbb{R}$. We say that $f \in C(W, \mathbb{R}^m)$ if $f_i \in C(W, \mathbb{R})$ for each i and that for some r, $f \in C^r(W, \mathbb{R}^m)$ if $f_i \in C^r(W, \mathbb{R})$ for each i.

2.2 Back to systems theory

The linear time invariant counter part is given by

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) \quad \text{(state DE)} \\ y(t) &= Cx(t) + Du(t) \quad \text{(read-out eqn.)} \end{aligned}$$

Throughout the course I will present most things in continuous time and then either I will show you the analogs in discrete time, or you will read on your own about discrete time (and learn through homework examples). The notation we will use for discrete-time, finite dimensional dynamical systems is as follows:

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

This type of system characterization is called *state-space description* or internal description of finite-dimensional systems. Another way of describing continuous-time and discrete-time finite-dimensional dynamical systems involves operators that establish a relationship between the system inputs and outputs. Such characterization is called *input–output description* or external description of a system.

The input function $u(\cdot) \in U$ where

$$U = \{u(\cdot)|\ u : \mathbb{R}_+ \to \mathbb{R}^m,\ u(\cdot) \text{ is piecewise continuous}\} = \mathrm{PC}(\mathbb{R}_+, \mathbb{R}^m)$$

Let define the state transition map and the response map:

$$\begin{aligned} x(t) &= s(t, t_0, x_0, u) \\ y(t) &= \rho(t, t_0, x_0, u) \end{aligned}$$

We now study the state transition map $s(t, t_0, x_0, u)$ as the unique solution to the state DE given by

$$\dot{x}(t) = A(t)x(t) + B(t)u(t)$$

for some initial condition $(t_0, x_0) \in \mathbb{R}_+ \times \mathbb{R}^n$, $x(t_0) = x_0$ and $u(\cdot) \in U$. Under these conditions (u.t.c.), the above reduces to

$$\dot{x}(t) = f(x(t), t),\ t \in \mathbb{R},\ x(t_0) = x_0$$

where the right-hand side is a given function

$$f : \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R}^n : (x, t) \mapsto f(x, t) = A(t)x(t) + B(t)u(t).$$

3 Solutions to ODEs: The Practice

I will not cover this in class since it should be a review for you. Please read through on your own.

Reference: A good reference for basic ODEs is Kreyszig, 'Advanced Engineering Mathematics'. Chapter 1 is basic ODE stuff. It has loads of examples.

Let's just review how to solve a simple ordinary differential equation. The kinds of ODEs we will deal with in this class largely fall in the class of so-called *separable* equations of the form

$$dx = f(x)g(t)dt$$

which are solved by

$$\int \frac{1}{f(x)}\ dx = \int g(t)\ dt + C$$

Example 3.1 Consider the following ODE:

$$(t^2 - 1)\frac{dx}{dt} + 2x = 0$$

We can solve this using the above technique. Indeed,

$$\begin{aligned}
\frac{dx}{dt}(t^2-1) = -2x &\implies dx(t^2-1) = -2xdt \\
&\implies \int -\frac{1}{2x}\ dx = \int \frac{1}{t^2-1}\ dt \\
&\implies -\frac{1}{2}\ln|x| = \ln\left|\frac{t+1}{t-1}\right| + C \\
&\implies |x| = \exp\left(\ln\left|\frac{t+1}{t-1}\right|\right)\exp(C) \\
&\implies x = k\left(\frac{t+1}{t-1}\right)
\end{aligned}$$

Definition 3.1 (***Linear ODE***) A first order ODE is linear if it can be written as

$$x' + p(t)x = r(t) \tag{1}$$

and it is homogeneous if $r(t) = 0$.

Let us find a formula for the general solution of a linear ODE on some interval I and assuming p and r are continuous (we will get to the formal requirements for existence and uniqueness next). For the homogeneous equation

$$x' + p(t)x = 0$$

this is very simple. Indeed, separating variables we have

$$\frac{dx}{x} = -p(t)dt \implies \ln|x| = -\int p(t)\ dt + c^*$$

and by taking the exponential of both sides we get

$$x(t) = c\exp(-\int p(t)\ dt)$$

where $c = \pm\exp(c^*)$ when $x \gtrless 0$.

Now, we solve the non-homogeneous equation given this form of solution. It turns out that the general form ODE has an integrating factor depending only on x. First, let us recall integrating factors and the notion of exactness.

Definition 3.2 (***Exactness.***) A first-order differential equation of the form

$$P(t,x)dt + Q(t,x)dx = 0$$

is called an *exact differential equation* if the differential form

$$P(t,x)dt + Q(t,x)dx$$

is *exact*—that is, there exists some function h such that

$$dh = \frac{\partial h}{\partial t}dt + \frac{\partial h}{\partial x}dx$$

with $\frac{\partial h}{\partial t} = P(t,x)$ and $\frac{\partial h}{\partial x} = Q(t,x)$.

Fact 3.1 A necessary and sufficient condition for exactness is that

$$\frac{\partial P}{\partial x} = \frac{\partial^2 h}{\partial x \partial t}, \ \frac{\partial Q}{\partial t} = \frac{\partial^2 h}{\partial t \partial x}$$

since for *sufficiently smooth* functions—here we just need $h \in C^2(\mathbb{R}^n \times \mathbb{R}, \mathbb{R})$—the second-order mixed partial derivatives are equal. Equality of second derivatives is the Schwarz's theorem (or Clairaut's theorem on equality of mixed partials).

The niceness of exactness comes from the fact that if the ODE is exact then

$$dh = 0$$

can be integrated to directly get the general solution

$$h(t, x) = c$$

We can reduce an ODE to an exact form if there exists a so-called *integrating factor*.

Definition 3.3 (***Integrating Factor.***) If we have an ODE

$$P(t, x)dt + Q(t, x)dx = 0 \tag{2}$$

and there exists a function F—which in general will be a function of both x *and* t—such that

$$FPdt + FQdx = 0$$

is exact, then F is an *integrating factor*.

Then due to exactness, we have that

$$\frac{\partial}{\partial x}(FP) = \frac{\partial}{\partial t}FQ$$

and by the product rule

$$F_x P + F P_x = F_t Q + F Q_t$$

If we get lucky and the integrating factor only depends on one variable, say t, then $F_x \equiv 0$ and $F_t = F' = dF/dt$ so that

$$FP_x = F'Q + FQ_t$$

Dividing by FQ, we get that

$$\frac{1}{F}\frac{dF}{dt} = \frac{1}{Q}\left(\frac{\partial P}{\partial x} - \frac{\partial Q}{\partial t}\right) \tag{3}$$

It turns out that if (5) is such that the right-hand side of the above depends only on t, then (5) has an integrating factor $F = F(t)$ which is obtained by integrating the above and taking exponentials of both sides to get

$$F(t) = \exp\left(\int R(x)\ dx\right)$$

where

$$R(x) = \frac{1}{Q}\left(\frac{\partial P}{\partial x} - \frac{\partial Q}{\partial t}\right)$$

Now, coming back to our problem, we first write (1) as

$$(px - r)dt + dx = 0 \iff Pdt + Qdx = 0, \ P = px - r, \ Q = 1$$

Hence (3) becomes

$$\frac{1}{F}\frac{dF}{dt} = p(t)$$

Since this depends only on t, the ODE has integrating factor $F(t)$ which we obtain as

$$F(t) = \exp\left(\int p(t)\ dt\right)$$

Multiplying our ODE by this F, we get

$$\exp\left(\int p(t)\ dt\right)(x' + px) = \left(\exp\left(\int p(t)\ dt\right)x\right)' = \exp\left(\int p(t)\ dt\right)r$$

Integrating we get

$$\exp\left(\int p(t)\ dt\right)x = \int \exp\left(\int p(t)\ dt\right)rdt + c$$

Let $h = \int p(t)\ dt$ and divide on both sides by $\exp(h)$ to get

$$x(t) = \exp(-h)\left(\int \exp(h)r\ dt + c\right)$$

Example 3.2 Consider the following ODE:

$$\frac{dx}{dt} + 3\frac{x}{t} = \frac{e^t}{t^3}$$

It is of the form

$$\frac{dx}{dt} + Py = Q$$

so the integrating factor is

$$\exp\left(\int P\ dt\right) = \exp\left(\int \frac{3}{t}\ dt\right) = \exp(3\ln t) = \exp(\ln t^3) = t^3$$

Multiplying by the integrating factor we get

$$t^3\frac{dx}{dt} + 3t^2x = e^t$$

Now, we integrate to get

$$t^3x = e^t + c$$

where c is some constant (eventually we will see this is related to the initial condition). Thus,

$$x = \frac{e^t + c}{t^3}$$

4 Solutions to ODEs: The Theory

Before getting into the specifics of linear systems, we will review existence and uniqueness of solutions to general systems. Consider the general autonomous ordinary differential equation (ODE) given by

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

The function f must satisfy two assumptions.

(A1) Let $\mathcal{D}$ be a set in $\mathbb{R}_+$ which contains at most a finite number of points per unit interval. $\mathcal{D}$ is the set of possible discontinuity points; it may be empty. For each fixed, $x \in \mathbb{R}^n$, the function $t \in \mathbb{R}_+ \backslash \mathcal{D} \to f(x,t) \in \mathbb{R}^n$ is <u>continuous</u> and for any $\tau \in \mathcal{D}$, the left-hand and right-hand limits $f(x, \tau_-)$ and $f(x, \tau_+)$, respectively, are finite vectors in $\mathbb{R}^n$.

(A2) There is a piecewise continuous function $k(\cdot) : \mathbb{R}_+ \to \mathbb{R}_+$ such that

$$\|f(\xi, t) - f(\xi', t)\| \leq k(t)\|\xi - \xi'\| \quad \forall\, t \in \mathbb{R}_+,\ \forall \xi, \xi' \in \mathbb{R}^n$$

This is called a global Lipschitz condition because it must hold for all ξ and ξ'.

Often times it is easy to characterize whether or not a given function satisfies (A2) by checking an auxiliary condition.

Proposition 4.1 Let $f : \mathbb{R} \to \mathbb{R}$ be continuous and differentiable. Then

$$f \text{ Lipschitz} \iff \exists K,\ \text{s.t. } \forall x \in \mathbb{R},\ |f'(x)| \leq K$$

To prove this, we need the mean value theorem.

Theorem 4.1 (***Mean Value Theorem.***) If a function f is continuous on the closed interval $[a, b]$, and differentiable on the open interval (a, b), then there exists a point c in (a, b) such that

$$f'(c) = \frac{f(b) - f(a)}{b - a}$$

Proof. ($\Longleftarrow$). Suppose the derivative is bounded by some K. By the mean value theorem, we have that for $x, y \in \mathbb{R}$, there exists $c \in \mathbb{R}$ such that

$$f(x) - f(y) = (x - y)f'(c)$$

so that

$$|f(x) - f(y)| = |(x - y)f'(c)| \leq K|x - y|$$

Hence, f is Lipschtiz.

($\Longrightarrow$). Suppose f is K-Lipschitz so that $|f(x) - f(y)| \leq K|x - y|$ for all x, y and hence, in particular, $|f(x + h) - f(x)| \leq K|h|$ for all x and h. Then, taking the limit we have that

$$f'(x) = \lim_{h \to 0} \left| \frac{f(x + h) - f(x)}{h} \right| \leq K$$

Alternative argument:

$$f \text{ Lipschitz} \quad \Longrightarrow \quad |f(x) - f(y)| \leq K|x - y|, \forall x, y \quad \Longrightarrow \quad \lim_{x \to y} \left| \frac{f(x) - f(y)}{x - y} \right| \leq K$$

$\square$

Note: Lipschitz functions do not have to be differentiable. They have to be almost everywhere differentiable (except on a set of measure zero).

Proposition 4.2 Given $R > 0$, if there is a PC function $k(\cdot)$ s.t.

$$\|D_1 f(x, t)\| \leq k(t), \quad \forall x \in B_R(0),\ \forall t \in \mathbb{R}_+$$

then the Lipschitz condition in (A2) holds for all $\xi, \xi' \in B_R(0)$, $t \in \mathbb{R}_+$

Examples.

1. The function

$$f(x) = \sqrt{x^2 + 1}$$

defined for all real numbers is Lipschitz continuous with the Lipschitz constant $K = 1$, because it is everywhere differentiable and the absolute value of the derivative is bounded above by 1. Indeed,

$$f'(x) = x(x^2+1)^{-1/2}$$

so that

$$|f'(x)| = |x(x^2+1)^{-1/2}| \leq |x||(x^2+1)^{-1/2}|$$

Claim:

$$\frac{|x|}{|(x^2+1)^{1/2}|} \leq 1$$

This is true because

$$|x| = |(x^2)^{1/2}| \leq |(x^2+1)^{1/2}|$$

2. The functions $\sin(x)$ and $\cos(x)$ are Lipschitz with constant $K = 1$ since their derivatives are bounded by 1.
3. **Practice**[2]**.** is the function $\sin(x^2)$? what about $\sqrt{x}$? (hint: consider the derivatives)

5 Proof of FTO

Theorem 1 (Fundamental Theorem of Existence and Uniqueness.)**.** Consider

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

where initial condition (t_0, x_0) is such that $x(t_0) = x_0$. Suppose f satisfies (A1) and (A2). Then,

a. For each $(t_0, x_0) \in \mathbb{R}_+ \times \mathbb{R}^n$ there exists a continuous function $\phi : \mathbb{R}_+ \to \mathbb{R}^n$ such that

$$\phi(t_0) = x_0$$

and

$$\dot{\phi}(t) = f(\phi(t), t), \quad \forall t \in \mathbb{R}_+ \backslash \mathcal{D}$$

b. This function is unique. The function ϕ is called the solution through (t_0, x_0) of the differential equation.

Note that if the Lipschitz condition does not hold, it may be that the solution cannot be continued beyond a certain time. e.g., consider

$$\dot{\xi}(t) = \xi^2(t), \quad \xi(0) = \frac{1}{c}, \ c \neq 0$$

where $\xi : \mathbb{R}_+ \to \mathbb{R}$. This differential equation has the solution

$$\xi(t) = \frac{1}{c-t}$$

on $t \in (-\infty, c)$. As $t \to c$, $\|\xi(t)\| \to \infty$. This is called finite escape time at c.

We need the following notion of a Cauchy sequence—this is a stronger notion of convergence.

Definition 2 (Cauchy sequence.)**.** A sequence $(v_i)_{i=1}^{\infty}$ in $(V, F, \|\cdot\|)$ is said to be a Cauchy sequence in V if and only if for any $\varepsilon > 0$, $\exists\ N_\varepsilon \in \mathbb{N}$ such that for any pair $m, n > N_\varepsilon$,

$$\|v_m - v_n\| < \varepsilon \ \ \forall p \in \mathbb{N}$$

[2]Throughout the notes as with standard textbooks, I will place practice problems that you can consider solving on your own or that might show up in your homework. It worth mentioning that most of these problems have solutions which can be found in the references on the main page of canvas. Further, they are designed so that with a little critical thinking, you should be able to solve them on your own. If you cannot, then it is likely that you have not quite internalized the material that proceeds them. You are welcome to ask myself of Andrew (TA) if you need help, but please try on your own first.

And, we need the notion of a Banach space.

Definition 3. Banach Space. A Banach space X is a normed linear space that is complete with respect to that norm—that is, every Cauchy sequence $\{x_n\}$ in X converges in X.

Proof sketch for existence. Construct a sequence of continuous functions

$$x_{m+1}(t) = x_0 + \int_{t_0}^{t} f(x_m(\tau), \tau)\ d\tau$$

where $x_0(t_0) = x_0$ and $m = 0, 1, 2, \ldots$. The idea is to show that the sequence of continuous functions $\{x_m(\cdot)\}_0^\infty$ converges to (i) a continuous function $\phi : \mathbb{R}_+ \to \mathbb{R}^n$ which is (ii) a solution of $\dot{x} = f(x, t)$, $x(t_0) = x_0$.

To show (i), we show that $\{x_m(\cdot)\}_0^\infty$ is a Cauchy sequence in a Banach space $(C([t_1, t_2], \mathbb{R}^n), \mathbb{R}, \|\cdot\|_\infty)$, where $t_0 \in [t_1, t_2]$.

To show [(ii)—i.e. that $\phi(\cdot)$ is a solution of the differential equation—recall that

$$x_{m+1}(t) = x_0 + \int_{t_0}^{t} f(x_m(\tau), \tau)\ d\tau$$

By the above argument, $m \to \infty$, $x_m(\cdot) \to \phi(\cdot)$ on $[t_1, t_2]$. Hence, it suffices to show that

$$\int_{t_0}^{t} f(x_m(\tau), \tau)\ d\tau \to \int_{t_0}^{t} f(\phi(\tau), \tau)\ d\tau, \quad \text{as } m \to \infty$$

$\square$

To prove uniqueness, we need the so called Bellman-Gronwall Lemma.

Lemma 4 (Bellman-Gronwall)**.** Let $u(\cdot)$, $k(\cdot)$ be real-valued, piecewise continuous functions on $\mathbb{R}_+$ and assume $u(\cdot), k(\cdot) > 0$ on $\mathbb{R}_+$. Suppose $c_1 > 0$, $t_0 \in \mathbb{R}_+$. If

$$u(t) \le c_1 + \int_{t_0}^{t} k(\tau)u(\tau)\ d\tau$$

then

$$u(t) \le c_1 \exp\left(\int_{t_0}^{t} k(\tau)\ d\tau\right)$$

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,

$$u(t) \le U(t)$$

Multiply both sides of

$$u(t) \le c_1 + \int_{t_0}^{t} k(\tau)u(\tau)\ d\tau$$

by the non-negative function

$$k(t)\exp\left(-\int_{t_0}^{t} k(\tau)\ d\tau\right)$$

resulting in

$$\frac{d}{dt}\left(U(t)\exp\left(-\int_{t_0}^{t} k(\tau)\ d\tau\right)\right) \le 0$$

and thus

$$u(t) \leq U(t) \leq c_1 \exp\left(-\int_{t_0}^{t} k(\tau)\, d\tau\right)$$

□

Proof of uniqueness sketch: Invoke Bellman-Grownwall.

Let's consider an example.

Example. Consider

$$\begin{aligned} \dot{x}(t) &= A(t)x(t) + B(t)u(t) \\ x(t_0) &= x_0 \end{aligned}$$

Show the solution is unique.

Proof. Assume $\phi(t), \psi(t)$ are two solutions so that $\phi(t_0) = \psi(t_0) = x_0$ and

$$\begin{aligned} \dot{\phi}(t) &= A(t)\phi(t) + B(t)u(t) \\ \dot{\psi}(t) &= A(t)\psi(t) + B(t)u(t) \end{aligned}$$

Then

$$\phi(t) - \psi(t) = \int_{t_0}^{t} (A(\tau)\phi(\tau) - A(\tau)\psi(\tau))\, d\tau$$

so that

$$\|\phi(t) - \psi(t)\| \leq \|A(t)\|_{\infty,[t_0,t]} \int_{t_0}^{t} \|\phi(\tau) - \psi(\tau)\|\, d\tau$$

By Bellman-Gronwall,

$$\|\phi(t) - \psi(t)\| \leq c_1 + \|A(t)\|_{\infty,[t_0,t]} \int_{t_0}^{t} \|\phi(\tau) - \psi(\tau)\|\, d\tau$$

implies

$$\|\phi(t) - \psi(t)\| \leq c_1 \exp\left(\|A(t)\|_{\infty,[t_0,t]}(t - t_0)\right)$$

This is true for all $c_1 \geq 0$, so set $c_1 = 0 \ldots$ □

6 Solutions to Linear Systems

Recall

$$\begin{aligned} \dot{x}(t) &= A(t)x(t) + B(t)u(t) \quad \text{(state DE)} \\ y(t) &= C(t)x(t) + D(t)u(t) \quad \text{(read-out eqn.)} \end{aligned}$$

with initial data (t_0, x_0) and the assumptions on $A(\cdot), B(\cdot), C(\cdot), D(\cdot), u(\cdot)$ all being PC:

- $A(t) \in \mathbb{R}^{n \times n}$
- $B(t) \in \mathbb{R}^{n \times m}$
- $C(t) \in \mathbb{R}^{p \times n}$
- $D(t) \in \mathbb{R}^{p \times m}$

The input function $u(\cdot) \in \mathcal{U}$, where $\mathcal{U}$ is the set of piecewise continuous functions from $\mathbb{R}_+ \to \mathbb{R}^m$.

This system satisfies the assumptions of our existence and uniqueness theorem. Indeed,

(A1) For all fixed $x \in \mathbb{R}^n$, the function $t \in \mathbb{R}_+ \backslash \mathcal{D} \to f(x,t) \in \mathbb{R}^n$ is continuous where $\mathcal{D}$ contains all the points of discontinuity of $A(\cdot), B(\cdot), C(\cdot), D(\cdot), u(\cdot)$

(A2) There is a piecewise continuous function $k(\cdot) = \|A(\cdot)\|$ such that

$$\|f(\xi,t) - f(\xi',t)\| = \|A(t)(\xi - \xi')\| \leq k(t)\|\xi - \xi'\| \quad \forall t \in \mathbb{R}_+, \ \forall \xi, \xi' \in \mathbb{R}^n$$

Hence, by the above theorem, the differential equation has a unique continuous solution $x : \mathbb{R}_+ \to \mathbb{R}^n$ which is clearly defined by the parameters $(t_0, x_0, u) \in \mathbb{R}_+ \times \mathbb{R}^n \times U$. Therefore, recalling the state transition map s we have the following theorem.

Theorem 5. (Existence of the state transition map.) Under the assumptions and notation above, for every triple $(t_0, x_0, u) \in \mathbb{R}_+ \times \mathbb{R}^n \times U$, the state transition map

$$x(\cdot) = s(\cdot, t_0, x_0, u) : \mathbb{R}_+ \to \mathbb{R}^n$$

is a continuous map well-defined as the unique solution of the state differential equation

$$\dot{x}(t) = A(t)x(t) + B(t)u(t)$$

with (t_0, x_0) such that $x(t_0) = x_0$ and $u(\cdot) \in U$.

Remark: Since the state transition function being well-defined, so is the response map

$$y(t) = \rho(t, t_0, x_0, u)$$

It follows that the state transition function is <u>differentiable</u> at every $t \in \mathbb{R}_+ \backslash \mathcal{D}$. Moreover, with the state transition function being well-defined, so is the response map

$$y(t) = \rho(t, t_0, x_0, u)$$

as the composition of the state transition map and the output function g.

7 Zero-Input and Zero-State Response

The state transition function of a linear system is equal to its <u>zero-input</u> state transition function and its <u>zero-state</u> state transition map:

$$\begin{aligned} s(t, t_0, x_0, u) &= \underbrace{\xi(t, t_0, x_0, 0)}_{\text{zero-input state trans. func.}} + \underbrace{s(t, t_0, 0, u)}_{\text{zero-state state trans. func.}} \\ \rho(t, t_0, x_0, u) &= \underbrace{\rho(t, t_0, x_0, 0)}_{\text{zero-input respons}} + \underbrace{\rho(t, t_0, 0, u)}_{\text{zero-state response}} \end{aligned}$$

Due to the fact that the state transition map and the response map are <u>linear</u>, they have the property that for fixed $(t, t_0) \in \mathbb{R}_+ \times \mathbb{R}_+$ the maps

$$s(t, t_0, \cdot, 0) : \mathbb{R}^n \to \mathbb{R}^n : x_0 \mapsto \xi(t, t_0, x_0, 0)$$

and

$$\rho(t, t_0, \cdot, 0) : \mathbb{R}^n \to \mathbb{R}^p : x_0 \mapsto \rho(t, t_0, x_0, 0)$$

are linear.

Hence by the *Matrix Representation Theorem*, they are representable by matrices. Therefore there exists a matrix $\Phi(t, t_0) \in \mathbb{R}^{n\times n}$ such that

$$\xi(t, t_0, x_0, 0) = \Phi(t, t_0)x_0, \quad \forall x_0 \in \mathbb{R}^n$$

and

$$\rho(t, t_0, x_0, 0) = C(t)\Phi(t, t_0)x_0, \quad \forall x_0 \in \mathbb{R}^n$$

7.1 State Transition Matrix

(State transition matrix.) $\Phi(t, t_0)$ is called the state transition matrix.

Consider the matrix differential equation

$$\dot{X} = A(t)X, \quad X(\cdot) \in \mathbb{R}^{n\times n}$$

Let $X(t_0) = X_0$.

The state transition matrix $\Phi(t, t_0)$ is defined to be the solution of the above matrix differential equation starting from $\Phi(t_0, t_0) = I$. That is,

$$\frac{\partial}{\partial t}\Phi(t, t_0) = A(t)\Phi(t, t_0)$$

and $\Phi(t_0, t_0) = I$.

Definition 6. (LTI State transition matrix.) The state transition matrix for

$$\dot{x} = Ax, \ x \in \mathbb{R}^n, \ A \in \mathbb{R}^{n\times n}$$

is the *matrix exponential* e^{At} defined to be

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

where I is the $n \times n$ identity matrix.

Proof. It is easy to verify that

$$\Phi(t, 0) = e^{At} \quad \text{and} \quad \Phi(t, t_0) = e^{A(t-t_0)}$$

by checking that

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

satisfies the differential equation

$$\dot{x} = Ax, \quad x(t_0) = x_0$$

Indeed, by the fact that $\Phi(t, t_0)$ satisfies the ODE (by definition) we know that

$$\frac{\partial}{\partial t}\Phi(t, t_0) = A\Phi(t, t_0)$$

and

$$\frac{\partial}{\partial t}\exp(A(t - t_0))x_0 = A\exp(A(t - t_0))x_0$$

In addition, $\Phi(t_0, t_0) = I$ and $\exp(A(t_0 - t_0)) = I$. Hence, $\Phi(t, t_0)$ and $\exp(A(t - t_0))$ satisfy the same ODE so they are equal. □

Hence, the matrix exponential is our friend and we need computational approaches for expressing this little monster.

7.2 Properties of the Matrix Exponential

First, we note that the matrix exponential has several important properties.

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

Recall that
$$\exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!}$$
This is also true for the matrix exponential—i.e.
$$\exp(At) = \sum_{k=0}^{\infty} \frac{(At)^k}{k!}$$

Fact. Note also that Cayley-Hamilton implies that the matrix exponential is expressible as a polynomial of order $n-1$!

Using the series representation of e^{At} to compute e^{At} is difficult unless, e.g., the matrix A is nilpotent in which case the series yields a closed form solution—i.e., A nilpotent matrix is such that $A^k = 0$ for some k.

Example. Consider
$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$
Then
$$A^2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$
so that
$$e^{At} = I + At = \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}$$

Hence we need an alternative method to compute it. Methods of computation include (see 510 notes, lecture 10):

1. Summation expression for the function $\exp(\cdot)$.
2. Laplace transform
3. Diagonalization or Jordan form combined with functions of matrices.

8 Properties of State Transition Function

Proposition 7. The following properties hold:

1. The solution of $\dot{x} = A(t)x$, $s(t, t_0, x_0, 0)$ is given by
$$s(t, t_0, x_0, 0) = \Phi(t, t_0)x_0 \tag{4}$$

2. For all $t, t_0, t_1 \in \mathbb{R}_+$,
$$\Phi(t, t_0) = \Phi(t, t_1)\Phi(t_1, t_0)$$
3. The inverse of the state transition matrix is
$$(\Phi(t, t_0))^{-1} = \Phi(t_0, t)$$
4. The determinant is give by
$$\det \Phi(t, t_0) = \exp\left(\int_{t_0}^{t} \mathrm{trace}(A(\tau))\, d\tau\right)$$

Proof. Call the left-hand side of (4) LHS, and the right-hand side RHS.

1. Check first that the LHS of (4) and the RHS are equal at t_0:
$$\mathrm{LHS}(t_0) = s(t_0, t_0, x_0) = x_0 \quad \text{and} \quad \mathrm{RHS}(t_0) = \Phi(t_0, t_0)x_0 = Ix_0 = x_0$$
Now, we check they satisfy the same differential equation:
$$\frac{d}{dt}\mathrm{LHS}(t) = A(t)\mathrm{LHS}(t) \quad \text{and} \quad \frac{d}{dt}\mathrm{RHS}(t) = A(t)\mathrm{RHS}(t)$$
so that $s(t, t_0, x_0) = \Phi(t, t_0)x_0$.
2. Again we use the same trick of checking the initial condition and the differential equation (and invoke the existence and uniqueness theorem).
$$\begin{aligned} \mathrm{RHS}(t_1) &= \mathrm{LHS}(t_1) \\ \frac{d}{dt}\mathrm{RHS}(t) &= A(t)\mathrm{RHS}(t) \\ \frac{d}{dt}\mathrm{LHS}(t) &= A(t)\mathrm{LHS}(t) \end{aligned}$$
Hence, LHS $\equiv$ RHS.
3. First, $\Phi(t, s) = \Phi(s, \tau)\Phi(\tau, t)$ for any t, s, τ since the following diagram commutes:
$$\begin{tikzcd} X \arrow[r, "\Phi(s,\tau)"] \arrow[dr, "\Phi(t,s)"'] & X \arrow[d, "\Phi(t,\tau)"] \\ & X \end{tikzcd}$$
Indeed, consider the unique solution to
$$\begin{cases} \dot{x}(\sigma) &= A(\sigma)x(\sigma) \\ x(s) &= a \end{cases}$$
Then, $x(t) = \Phi(t, s)a$, $x(t) = \Phi(t, \tau)x(\tau)$ and $x(\tau) = \Phi(\tau, s)a$, and hence
$$\Phi(t, \tau)\Phi(\tau, s)a = \Phi(t, s)a$$
that is
$$(\Phi(t, \tau)\Phi(\tau, s) - \Phi(t, s))a = 0$$
since this must hold for all $a \in \mathbb{R}^n$, the claim holds.

We claim that $\Phi(t, s)$ is invertible and that its inverse is given by $\Phi(s, t)$. Indeed, from $\Phi(t, s) = \Phi(s, \tau)\Phi(\tau, t)$ we have that
$$I = \Phi(t, s)\Phi(s, t)$$
Thus, $\Phi(t, t_0)$ is invertible for all t. Hence,
$$\Phi(t_0, t_0) = I = \Phi(t_0, t)\Phi(t, t_0) \implies \Phi(t, t_0)^{-1} = \Phi(t_0, t)$$
4. This is called the *Jacobi-Liouville* equation. We will take this one as given.

$\square$

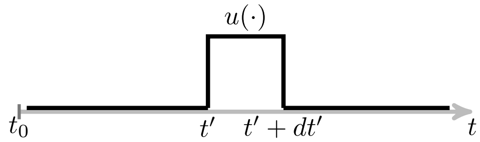


Figure 1: Input

9 Solving Linear ODE via $\Phi(t, t_0)$

First, let us consider a *heuristic* derivation of the zero-state transition. (page 35 of C& D) Consider the input in Fig. 1.

Then,

$$x(t') = \Phi(t', t_0)x_0$$

and

$$x(t' + dt') = x(t') + [A(t')x(t') + B(t')u(t')]dt'$$

$$\begin{aligned}
x(t) &= \Phi(t, t' + dt')x(t' + dt') \\
&= \Phi(t, t' + dt')[x(t') + A(t')x(t')dt' + B(t')u(t')dt'] \\
&= \Phi(t, t' + dt')[I + A(t')dt']x(t') + \Phi(t, t' + dt')B(t')u(t')dt' \\
&\simeq \Phi(t, t' + dt')\Phi(t' + dt', t')\Phi(t', t_0)x_0 + \Phi(t, t' + dt')B(t')u(t')dt'
\end{aligned}$$

Hence,

$$x(t) \simeq \Phi(t, t_0)x_0 + \Phi(t, t')B(t')u(t')dt'$$

Theorem 8 (Solution of Linear System)**.**

$$x(t) = \Phi(t, t_0)x_0 + \int_{t_0}^{t} \Phi(t, t')B(t')u(t')\ dt' \tag{5}$$

Proof idea: We will use the trick that checks the equality by showing the left and right hand sides of (5) satisfy the same ODE. That is, at t_0, they have the same value (initial condition) and the derivative of the left and right hand sides is the same. The key here is that since we have the existence and uniqueness theorem, we know then that the solution of the ODE is unique, so that means any two expressions that satisfy it have to be equal.

Proof. We will use the same trick as before where we check the initial condition and the differential equation and invoke the existence/uniqueness theorem for solutions to ODEs.

$$\frac{d}{dt}\text{LHS}(t) = A(t)\text{LHS}(t) + B(t)u(t)$$

$$\text{LHS}(t_0) = x_0$$

$$\text{RHS}(t_0) = x_0$$

$$
\begin{aligned}
\frac{d}{dt}\text{RHS}(t) &= \frac{d}{dt}(\Phi(t,t_0)x_0) \\
&\quad + \frac{d}{dt}\left(\int_{t_0}^{t}\Phi(t,t')B(t')u(t')\ dt'\right) \\
&= A(t)\Phi(t,t_0)x_0 + \frac{d}{dt}(t)\Phi(t,t)B(t)u(t) \\
&\quad - \cancelto{0}{\frac{d}{dt}(t_0)}\Phi(t,t_0)B(t_0)u(t_0) + \int_{t_0}^{t}\frac{d}{dt}\left(\Phi(t,t')B(t')u(t')\right)dt'
\end{aligned}
$$

so that

$$
\begin{aligned}
\frac{d}{dt}\text{RHS}(t) &= A(t)\Phi(t,t_0)x_0 + B(t)u(t) \\
&\quad + A(t)\int_{t_0}^{t}\Phi(t,t')B(t')u(t')\ dt' \\
&= A(t)\text{RHS}(t) + B(t)u(t)
\end{aligned}
$$

Thus LHS and RHS has the same initial condition and satisfy the same ODE $\square$

Thus, we have that the state transition function is given by

$$
\boxed{s(t,t_0,x_0,u_{[t_0,t]}) = \Phi(t,t_0)x_0 + \int_{t_0}^{t}\Phi(t,t')B(t')u(t')\ dt'}
$$

By definition, we have that it satisfies the <u>state transition axiom</u>. Check that it satisfies the semi-group property:

$$
\begin{aligned}
s(t,t_0,s(t_1,t_0,x_0,u_{[t_0,t_1]}),u_{[t_0,t]}) &= \Phi(t,t_1)\Big[\Phi(t_1,t_0)x_0 \\
&\quad + \int_{t_1}^{t}\Phi(t,t')B(t')u(t')\ dt' \\
&= \Phi(t,t_0)x_0 + \int_{t_0}^{t_1}\Phi(t,t')B(t')u(t')\ dt' \\
&\quad + \int_{t_1}^{t}\Phi(t,t')B(t')u(t')\ dt' \\
&= \Phi(t,t_0)x_0 + \int_{t_0}^{t}\Phi(t,t')B(t')u(t')\ dt' \\
&= s(t,t_0,x_0,u_{[t_0,t]})
\end{aligned}
$$

10 Solutions to LTI Systems with Inputs

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

where

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

Taking the Laplace transform, we have

$$
\begin{aligned}
(sI - A)\hat{x}(s) &= x_0 + B\hat{u}(s) \\
\hat{y}(s) &= C\hat{x}(s) + D\hat{u}(s)
\end{aligned}
$$

so that

$$\begin{aligned}
\hat{x}(s) &= (sI - A)^{-1}x_0 + (sI - A)^{-1}B\hat{u}(s) \\
\hat{y}(s) &= C(sI - A)^{-1}x_0 + \big(C(sI - A)^{-1}B + D\big)\,\hat{u}(s)
\end{aligned}$$

Note that

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

is what is normally referred to as the transfer function. It is a map from input to output.

Claim 1. The solution $x(t)$ to

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

with $x(t_0) = x_0$ and where $x \in \mathbb{R}^n$ and

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

is

$$x(t) = e^{A(t-t_0)}x_0 + \int_{t_0}^{t} e^{A(t-\tau)}Bu(\tau)\ d\tau \tag{6}$$

and the output is

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

Proof. All we need to do is the usual trick of checking that it satisfies the differential equation and initial condition (then invoke the existence/uniqueness theorem). Indeed,

$$\begin{aligned}
\dot{x} &= \frac{d}{dt}\left(e^{A(t-t_0)}x_0\right) + \frac{d}{dt}\left(\int_{t_0}^{t} e^{A(t-\tau)}Bu(\tau)\ d\tau\right) \\
&= Ae^{A(t-t_0)}x_0 + A(0)Bu(t) + \int_{t_0}^{t} Ae^{A(t-\tau)}Bu(\tau)\ d\tau \\
&= Ax(t) + Bu(t)
\end{aligned}$$

Now for the initial condition,

$$x(t_0) = e^{A(0)}x_0 + \cancelto{0}{\int_{t_0}^{t_0} e^{A(t_0-\tau)}Bu(\tau)\ d\tau} = x_0$$

$\square$

AA/EE547: Wi 2020

Lecture 3: Introduction to Stability

Lecturer: L.J. Ratliff

Disclaimer: *These notes have not been subjected to the usual scrutiny reserved for formal publications, meaning you should take your own notes in class and review the provided references as opposed to taking these notes as your sole resource. I provide the lecture notes to you as a courtesy; it is not required that I do this. They may be distributed outside this class only with the permission of the Instructor.*

References. Chapter 8 and 9, **[JH]**; Chapter 4 and Chapter 7, **[C&D]**. review your notes on norms including (induced) matrix norms.

1 Lyapunov Stability

You may recall from an undergrad controls or signals class input-output stability. In our study of stability, we will start with state related stability concepts (stability in the sense of Lyapunov).

Intuition. Let's consider the following examples.

a. Continuous Time. Recall that the solution to

$$\dot{x} = -\lambda x$$

is

$$x(t) = x_0 e^{-\lambda t},$$

and if $\lambda > 0$ solution decays to zero, otherwise it blows up. This λ is an 'eigenvalue' for this scalar system, and its sign can be used to characterize a notion of stability.

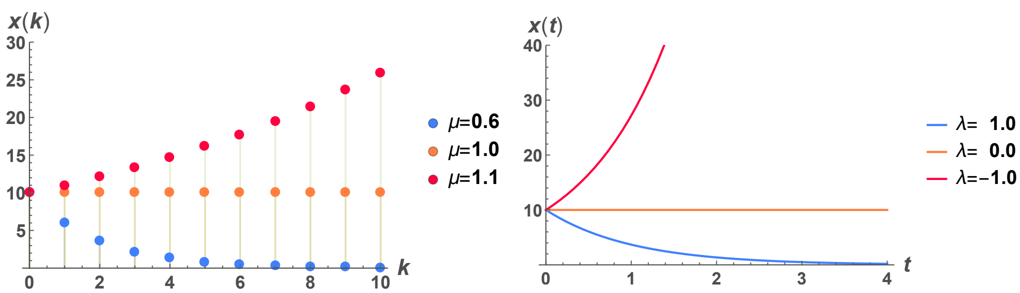
b. Discrete Time. Recall that the solution to

$$x_{k+1} = \mu x_k$$

is

$$x_k = \mu^k x_0$$

and if $|\mu| < 1$, then the solution decays to zero and otherwise, it blows up.



(a) discrete time (b) continuous time

1.1 Continuous Time

Recall that for a given linear system

$$\dot{x}(t) = A(t)x(t) + B(t)u(t),\ \ y(t) = C(t)x(t) + D(t)u(t)$$

the zero input response is given by

$$x(t) = \Phi(t, t_0)x_0$$

where $\Phi(t, t_0)$ is the state transition matrix and $x(t_0) = x_0$. Note that

$$x_0 = 0 \Longrightarrow x(t) = 0\ \forall t$$

The points $x_e = 0$ is called the *equilibrium point.*

The following are characterizations of stability à la Lyapunov.

Definition 1 ((Marginally) Stable Equilibrium)**.** Consider the equilibrium point $x_e = 0$.

$$x_e \text{ is stable } \iff \ \forall\, x_0 \in \mathbb{R}^n,\ \forall t_0 \in \mathbb{R}^n,\ t \mapsto x(t) = \Phi(t, t_0)x_0 \text{ is bounded } \forall\, t \geq t_0.$$

Note: the effect of initial conditions does not grow unbounded with time (but it may grow temporarily during a transient phase).

Definition 2 (Asymptotic Stability)**.** Consider the equilibrium point $x_e = 0$.

$$x_e = 0 \text{ is asymptotically stable } \iff\ x_0 = 0 \text{ is stable and } x(t) = \Phi(t, t_0)x_0 \longrightarrow 0 \text{ as } t \to \infty.$$

Note: the effect of initial conditions eventually disappears with time.

Definition 3 (Exponential Stability)**.** Consider the equilibrium point $x_e = 0$.

$$x_e = 0 \text{ is exponentially stable } \iff \exists M, \alpha > 0:\ \|x(t_0)\| \leq M\exp(-\alpha(t - t_0))\|x_0\|$$

We say an equilibrium point or the system is unstable if it is not marginally stable in the sense of Lyapunov. For such systems, the effect of initial conditions (may) grow over time (depending on the specific initial conditions and the value of $C(t)$).

Theorem 4 (Asymptotic Stability of Linear CT Systems)**.** $x_0 = 0$ is asymptotically stable $\iff \Phi(t, 0) \to 0$ as $t \to \infty$.

Proof. ($\Longleftarrow$)

$$x(t) = \Phi(t, t_0)x_0 = \Phi(t, 0)\Phi(0, t_0)x_0$$

since $\Phi(t, 0) \to 0$ as $t \to \infty$ then $\|\Phi(t, 0)\| \to 0$ as $t \to \infty$ and

$$\|x(t)\| \leq \|\Phi(t, 0)\|\|\Phi(0, t_0)\|\|x_0\|$$

thus $x(t) \to 0$ as $t \to \infty$.
($\Longrightarrow$) By contradiction; assume that $t \to \Phi(t, 0)$ does not tend to zero as $t \to \infty$, i.e. $\exists\, i, j$ such that

$$\Phi_{ij}(t, 0) \not\to 0 \ \text{ as } \ t \to \infty$$

choose

$$x_0 = \begin{bmatrix} 0 \\ \cdots \\ 0 \\ 1 \\ 0 \\ \cdots \\ 0 \end{bmatrix}$$

with 1 in the j–th spot. Thus,

$$x_i(t) = \Phi_{ij}(t,0) \not\to 0 \text{ as } t \to \infty$$

contradicting the asymptotic stability of 0. □

A stability concept that is equivalent to exponential stability for LTV systems is uniform asymptotic stability.

Definition 5. We say that the zero solution of $\dot{x} = A(t)x(t)$ on $t \geq 0$ is uniformly asymptotically stable if and only if

a. $t \mapsto \Phi(t, t_0)$ is bounded on $t \geq t_0$ uniformly in $t_0 \in \mathbb{R}_+$, i.e.,

$$\exists\, k < \infty : \ \forall t_0 \in \mathbb{R}_+, \ \|\Phi(t,t_0)\| \leq k, \ \forall t \geq t_0$$

b. $t \mapsto \Phi(t, t_0)$ tends to zero as $t \to \infty$ uniformly in $t_0 \in \mathbb{R}_+$, i.e.,

$$\forall \varepsilon > 0, \ \exists T(\varepsilon) > 0 : \ \forall t_0 \in \mathbb{R}_+, \|\Phi(t,t_0)\| \leq \varepsilon, \ \forall t \geq t_0 + T(\varepsilon) \tag{$*$}$$

In the LTI case, i.e. $\dot{x} = Ax$,

$$\text{asymptotic stability} \iff \text{exponential stability}$$

Indeed, $\Phi(t,t_0) = \exp(A(t-t_0))$ depends only on the elapsed time $t - t_0$, so that the zero solution of $\dot{x} = Ax$ is asymptotically stable iff the zero solution is uniformly asymptotically stable iff the zero solution is exponentially stable.

To see that in the LTV case that

$$A(\cdot) \text{ uniformly asymptotically stable} \iff A(\cdot) \text{ exponentially stable}$$

consider the following proof.

$\Leftarrow$ a. and b. hold with $k = m$ and $T(\varepsilon) > 0$ such that $\exp(-\alpha T(\varepsilon)) \leq \varepsilon m^{-1}$

$\Rightarrow$ Given any $T > 0$, we have that

$$\forall t \geq t_0, \ \exists!\, n \in \mathbb{N}, \ \exists!\, s \in [0, T) : \ t - t_0 = nT + s$$

Let $t_0 \in \mathbb{R}_+$ be arbitrary but fixed and for b., pick some $T(\varepsilon) > 0$ for $\varepsilon = 1/2$. Then, by ($*$), $\|\Phi(s + t_0 + T, t_0)\| \leq 1/2$ for all $s \geq 0$. Since

$$\begin{aligned} &\Phi(s + t_0 + 2T, t_0) = \Phi(s + t_0 + 2T, t_0 + T)\Phi(t_0 + T, t_0) \\ &\quad\implies \ \|\Phi(s + t_0 + 2T, t_0)\| \leq \|\Phi(s + t_0 + 2T, t_0 + T)\| \|\Phi(t_0 + T, t_0)\| \leq 2^{-2}, \end{aligned}$$

we have that (by induction),

$$\forall s \geq 0, \forall n = 1, 2, \ldots, \ \|\Phi(s + t_0 + nT, t_0)\| \leq 2^{-n} \tag{$**$}$$

Pick $\alpha > 0$ s.t. $\exp(\alpha T) = 2$. Then,

$$\forall s \in [0, T), 1 \leq 2\exp(-\alpha s) \implies \forall s \in [0, T), \ 2^{-n} \leq 2\exp(-\alpha(s + nT))$$

Combining this with the upper bound on the norm of the state transition matrix in ($**$), we have that

$$\forall s \in [0, T), \forall n = 1, 2, \ldots, \ \|\Phi(s + t_0 + nT, t_0)\| \leq 2\exp(-\alpha(s + nT))$$

Now, using a. and the fact that $1 \leq 2\exp(-\alpha s)$, we have that

$$\forall s \in [0, T), \ \|\Phi(s + t_0, t_0)\| \leq k \leq 2k\exp(-\alpha s)$$

so that

$$\forall t \geq t_0, \ \|\Phi(t, t_0)\| \leq 2k\exp(-\alpha(t - t_0))$$

1.2 Discrete Time

Consider
$$x_{k+1} = A_k x_k + B_k u_k,\ y_k = C_k x_k + D_k u_k$$
The zero state is again the equilibrium for this system—i.e., $x_e = 0$.

Discrete Time: Solutions. First, let us recall the solution form for the discrete time case. While I will not cover this in detail (please see Chapter 2d [C&D]), we note that there is an equivalent "fundamental theorem of odes" for linear difference equations and we can define completely analogous state transition functions and response maps which are linear and hence, have corresponding matrix representations. Indeed, the state transition matrix $\Phi(\cdot,\cdot) : \mathbb{N}_+ \times \mathbb{N}_+ \to \mathbb{R}^{n\times n}$ is defined as follows:
$$\forall\ k_0,\ \Phi(k+1, k_0) = A_k \Phi(k, k_0),\ k = k_0, k_0+1, \ldots,\ \Phi(k_0, k_0) = I$$
Note: It is only when A_k is non-singular for all $k \in \mathbb{N}$, that $\Phi(k+1, k_0) = A_k\Phi(k, k_0)$ can be solved for $\Phi(k, k_0)$ in terms of $\Phi(k+1, k_0)$.

It turns out that
$$\Phi(k, k_0) = A_{k-1}A_{k-2}\cdots A_{k_0}$$

Practice Problem. show this by induction.

The DT state transition matrix also has analogous properties as those for its continuous time counter part. Indeed,
$$\Phi(k, k_0) = \Phi(k, k_1)\Phi(k_1, k_0),\ \forall k_0 \le k_1 \le k$$

The solution and output can then be characterized by
$$\begin{aligned} x_k &= \Phi(k, k_0)x_0 + \sum_{\ell=k_0}^{k-1} \Phi(k, \ell+1)B_\ell u_\ell \\ y_k &= C_k\Phi(k, k_0)x_0 + C_k\left(\sum_{\ell=k_0}^{k-1} \Phi(k, \ell+1)B_\ell u_\ell\right) + D_k u_k \end{aligned}$$

Asymptotic stability can be characterized informally by the statement that every solution of $x_{k+1} = A_k x_k$ tends to zero as $k \to \infty$.

Definition 6 (Asymptotic Stability)**.** Consider $x_k \equiv 0$ (i.e., the zero solution of $x_{k+1} = A_k x_k$). The zero solution is asymptotically stable iff for all $x_0 \in \mathbb{R}^n$, for all $k_0 \in \mathbb{N}$,

a. $k \mapsto x_k = \Phi(k, k_0)x_0$ is bounded on $k \ge k_0$
b. $k \mapsto x_k = \Phi(k, k_0)x_0 \to 0$ as $k \to \infty$.

Note: any solution x_k on any $[k_0, k]$ is a finite set and hence, b. $\implies$ a. Due to linearity, we get the following theorem.

Theorem 7 (Asymptotic Stability of Linear DT Systems)**.** Let $\det(A_k) \ne 0$ for all $k \in \mathbb{N}$. The zero solution of $x_{k+1} = A_k x_k$ on $k \ge 0$ is asymptotically stable iff $\Phi(k, 0) \to 0$ as $k \to \infty$.

Practice Problem. Try to prove this theorem by following the same proof structure as Theorem 4. Hint: not that since $\det(A_k) \ne 0$, $\forall k \ge k_0 \ge 0$, $\Phi(k, k_0) = \Phi(k, 0)\Phi(0, k_0)$.

Analogous to the continuous time case, exponential stability is a property of the system $x_{k+1} = A_k x_k$ which if possessed, guarantees that every solution of the system is bounded by a decaying exponential depending on the elapsed time $k - k_0$. Indeed, we have the following formal definition.

Definition 8 (Exponential Stability (DT))**.** The zero solution of $x_{k+1} = A_k x_k$ on $k \geq 0$ is exponentially stable iff $\exists\ \rho \in [0, 1)$ and $m > 0$ such that for all $k_0 \in \mathbb{N}$,

$$\|\Phi(k, k_0)\| \leq m\rho^{k-k_0},\ \forall\ k \geq k_0$$

where the matrix norm is arbitrary (recall that finite dimensional norms are equivalent).

Notes:

- The constants $\rho \in [0, 1)$ and $m > 0$ are fixed, meaning that they are independent of k_0. Further, the constant $\alpha \geq 0$ such that $\rho = \exp(-\alpha)$ is the exponential decay rate.
- An equivalent statement: the zero solution is exponentially stable iff

$$\exists \rho \in [0, 1),\ m > 0 :\ \forall (x_0, k_0) \in \mathbb{R}^n \times \mathbb{N},\ \|x_k\| \leq m\|x_0\|\rho^{k-k_0},\ \forall k \geq k_0$$

We say that the zero solution is uniformly asymptotically stable iff

a. $k \mapsto x_k = \Phi(k, k_0)x_0$ on $k \geq k_0$ is uniformly bounded (bound is independent of k_0)
b. $k \mapsto x_k = \Phi(k, k_0)x_0 \to 0$ uniformly as $k \to \infty$.

Theorem 9 (Equivalence of Asymptotic and Exponenital Stability)**.**

$$A(\cdot) \text{ is uniformly asymptotically stable } \iff A(\cdot) \text{ is exponentially stable}$$

Practice Problem. Prove this theorem. Hint: use the same method as in the CT case. In the necessity direction use a. and b. with $2\varepsilon = 1$ and pick $\rho \in [0, 1)$ such that $2\rho^K > 1$ so that $\|\Phi(k, k_0)\| \leq 2\ell\rho^{k-k_0}$ for all $k \geq k_0$.

2 LTI: Spectral Conditions

In the case of LTI systems, we can reduce the stability characterization to spectral conditions.

Recall (Lecture 10/11, 510) from last quarter that we have the following result about representing any analytic function[1] as a sum of polynomials.

Let $A \in \mathbb{C}^{n\times n}$ and let $\sigma(A)$ denote the spectrum of A (containing distinct eigenvalues of A) with $p = |\sigma(A)|$. The minimal polynomial of A is given as above by

$$\psi_A(\lambda) = \prod_{i=1}^{p} (\lambda - \lambda_i)^{m_i}$$

Fact. (Functions of a matrix.) Let $f(s)$ be any function of s analytic on the spectrum of A and $q(s)$ be a polynomial such that

$$f^{(k)}(\lambda_\ell) = q^{(k)}(\lambda_\ell)$$

for $0 \leq k \leq m_\ell - 1$ and $1 \leq \ell \leq p$. Then

$$f(A) = q(A)$$

[1] An *analytic function* is a function that is locally given by a convergent power series—i.e., f is real analytic on an open set U in the real line if for any $x_0 \in U$ one can write

$$f(x) = \sum_{n=0}^{\infty} a_n(x - x_0)^n$$

where $a_n \in \mathbb{R}$ for each n and the series is convergent to $f(x)$ for $x$in a neighborhood of x_0.

where p is the number of distinct roots of the characteristic polynomial $\chi_A(s)$ and

$$m_k = \min\{\mu \geq 1 : \mathcal{N}((A - \lambda_k I)^{\mu}) = \mathcal{N}((A - \lambda_k I)^{\mu+1})\}$$

i.e., m_k is the ascent of $A - I\lambda_k$.

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

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

where $a_1, a_2, \ldots, a_n$ are functions of

$$(f(\lambda_1), f^{(1)}(\lambda_1), f^{(2)}(\lambda_1), \ldots, f^{(m_1)}(\lambda_1), f(\lambda_2), \ldots)$$

and hence

$$f(A) = a_1 A^{m-1} + \cdots + a_m A^0 = \sum_{\ell=1}^{p} \sum_{k=0}^{m_\ell - 1} q_{k,\ell}(A) f^{(k)}(\lambda_\ell)$$

where $q_{k,\ell}$'s are polynomials independent of f.

This leads to the following theorem.

Theorem. (General Form of $f(A)$.) Let $A \in \mathbb{C}^{n\times n}$ have a minimal polynomial ψ_A given by

$$\psi_A(s) = \prod_{k=1}^{p} (s - \lambda_k)^{m_k}$$

Let the domain Δ contain $\sigma(A)$, then for any analytic function $f : \Delta \to \mathbb{C}$. we have

$$f(A) = \sum_{k=1}^{p} \sum_{\ell=0}^{m_k - 1} f^{(\ell)}(\lambda_k) q_{k,\ell}(A)$$

where $q_{k,\ell}$'s are polynomials independent of f.

Recalling our derivation of functions of matrices from 510, we can show that

$$\exp(tA) = \sum_{k=1}^{p} \sum_{\ell=0}^{m_k - 1} t^{\ell} \exp(\lambda_k t) p_{k\ell}(A)$$

This gives rise to the following stability condition:

Proposition 10 (Continuous Time)**.** Consider the differential equation $\dot{x} = Ax$, $x(0) = x_0$. From the above expression:

$$\{\exp(At) \to 0 \ \text{ as } \ t \to \infty\} \iff \{\forall \lambda_k \in \sigma(A),\ \mathrm{Re}(\lambda_k) < 0\}$$

and

$$\{t \mapsto \exp(At) \text{ is bounded on } \mathbb{R}_+\} \iff \left\{ \begin{array}{ll} \forall \lambda_k \in \sigma(A), & \mathrm{Re}(\lambda_k) < 0 \ \& \\ m_k = 1 \text{ when} & \mathrm{Re}(\lambda_k) = 0 \end{array} \right\}$$

We have a similar situation for discrete time systems:

$$\forall\ \nu \in \mathbb{N},\ A^{\nu} = \sum_{k=1}^{p} \sum_{\ell=1}^{m_k - 1} \nu(\nu - 1) \cdots (\nu - \ell + 1) \lambda_k^{\nu - \ell} p_{k\ell}(A)$$

The above gives rise to the following stability condition:

Proposition 11 (Discrete Time)**.** Consider the discrete time system $x(k+1) = Ax(k),\ k \in \mathbb{N}$, with $x(0) = x_0$. Then for $k \in \mathbb{N}$, $x(k) = A^k x_0$. From the above equation, we have that

$$\left\{A^k \to 0 \text{ as } k \to \infty\right\} \Longleftrightarrow \{\forall\ \lambda_i \in \sigma(A),\ |\lambda_i| < 1\}$$

and

$$\left\{k \to A^k \text{ is bounded on } \ \mathbb{N}_+\right\} \Longleftrightarrow \left\{ \begin{array}{ll} \forall \lambda_i \in \sigma(A), & |\lambda_i| \leq 1 \ \& \\ m_i = 1 \ \text{ when } & |\lambda_i| = 1 \end{array} \right\}$$

LTI: Asymptotic Stability is Equivalent to Exponential Stability. Coming back to the CT case, we will show that asymptotic stability is equivalent to exponentially stable using the function of matrix expansion.

Let's prove the claim in Proposition 10:

$$\dot{x} = Ax \text{ is exponentially stable} \iff \sigma(A) \subset \mathbb{C}_-^\circ$$

Proof. The state transition matrix for an LTI system is

$$\Phi(t, t_0) = \exp(A(t - t_0))$$

And from above we know that

$$\exp(At) = \sum_{k=1}^{p} \pi_k(t) \exp(\lambda_k t),\ \{\lambda_k\}_1^p = \sigma(A)$$

where π_k are some matrix polynomials in t. Hence, by taking matrix norms, we have that

$$\|\exp(At)\| \leq \sum_{k=1}^{p} \|\pi_k(t)\| \exp(\mathrm{Re}(\lambda_k)t) \leq \sum_{k=1}^{p} p_k(t) \exp(\mathrm{Re}(\lambda_k)t) \leq p(t) \exp(-\mu t)$$

where $p_k(t)$ are polynomials such that $\|\pi_k(t)\| \leq p_k(t)$, $p(t) = \sum_{k=1}^{p} p_k(t) \geq 0$ and

$$\mu = -\max\{\mathrm{Re}(\lambda) :\ \lambda \in \sigma(A)\}$$

Since a polynomial is growing slower than any growing exponential we have

$$\forall \varepsilon > 0,\ \exists\ m(\varepsilon) > 0 :\ 0 \leq |p(t)| \leq m \exp(\varepsilon t),\ \forall t \geq 0$$

Hence combing this with the above bound on $\|\exp(At)\|$, we have that

$$\forall \varepsilon > 0\ \exists\ m(\varepsilon) > 0 :\ \|\exp(At)\| \leq m \exp(-(\mu - \varepsilon)t)\ \forall t \geq 0$$

Then, if $\sigma(A) \subset \mathbb{C}_-^\circ$, by the above $\mu > 0$. Hence picking $\varepsilon \in (0, \mu)$ we have that

$$\|\Phi(t, t_0)\| \leq m \exp(-\alpha(t - t_0))$$

with $\alpha = \mu - \varepsilon > 0$. On the other hand if $\sigma(A)$ is not included in $\mathbb{C}_-^\circ$, then by the polynomial expansion for $\exp(At)$, $\exp(At)$ does not tend to the zero matrix as $t \to \infty$ and the zero solution is not exponentially stable.

$\square$

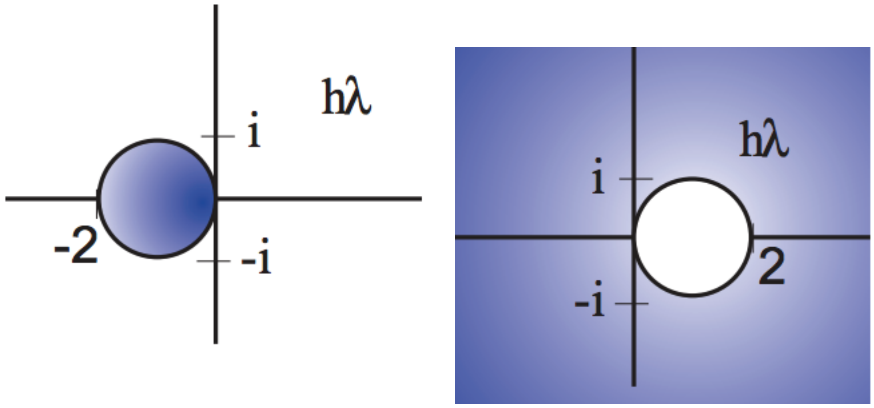


Figure 1: Left: Forward Euler; Right: Backward Euler.

2.1 Application to numerical integration of $\dot{x} = Ax$

Suppose we are given $\dot{x} = Ax$, $x(0) = x_0$, $A \in \mathbb{C}^{n\times n}$, $x \in \mathbb{C}^n$. Call $t \mapsto x(t)$ the exact solution $x(t) = \exp(At)x_0$. Note that $t \mapsto x(t)$ is analytic in t. Call $(\xi_0, \xi_1, \ldots)$ the sequence of computed values.

Forward Euler Method.
For small $h > 0$, we have for any $t_k \in \mathbb{R}_+$,

$$\begin{aligned} x(t_k + h) &= x(t_k) + h\dot{x}(t_k) + \mathrm{O}(h^2) \\ &= x(t_k) + hAx(t_k) + \mathrm{O}(h^2) \end{aligned}$$

In other words, we have approximately

$$x(t_k + h) \simeq (I + hA)x(t_k)$$

So if we perform repeatedly this step starting at $t_0 = 0$, we have the computed sequence $\{\xi_i\}_0^\infty$ by

$$\xi_m = (I + hA)^m x_0, \; m = 0, 1, 2, \ldots$$

From the spectral mapping theorem and the above equation for ξ_m we have the following.

Example. Consider

$$\dot{x}(t) = \lambda x(t)$$

with $\lambda \in \mathbb{C}$. Then the equation is stable if $\mathrm{Re}(\lambda) \leq 0$. In this case the system is exponentially decaying

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

When is the numerically solution x_i also decaying, $\lim_{i\to\infty} x_i = 0$?

$$x_{i+1} = x_i + h\lambda x_i \iff x_{i+1} = (1 + h\lambda)^{i+1} x_0$$

The solution is decaying (stable) if $|1 + h\lambda| \leq 1$

Fact. Suppose $\sigma(A) \subset \mathbb{C}_-^\circ$ (equiv., the origin is exponentially stable). Let h_0 be the largest positive h such that

$$\max_i |1 + h\lambda_i| = 1$$

Under these conditions,

1. $\{\xi_m\}_0^\infty \to 0$ exponentially for all ξ_0 iff $h \in (0, h_0)$.
2. if $h > h_0$, then for almost all x_0, the sequence of computed values $\{\xi_k\}_0^\infty$ is such that $\{\|\xi_m\}_0^\infty$ grows exponentially.

Interpretation. Thus even if $\sigma(A) \subset \mathbb{C}_-^\circ$ (and hence the exact solution $x(t) \to 0$ exponentially), for $h > h_0$, for almost all x_0, the sequence of computed vectors $\{\|\xi_m\|\}_0^\infty$ blows up. It is for this reason that in practice we often prefer the backward Euler method.

Backward Euler:
For small $h > 0$, we have that for any $t_k \in \mathbb{R}$,

$$\begin{aligned} x(t_k) &= x(t_k + h) - h\dot{x}(t_k + h) + \mathrm{O}(h^2) \\ &= x(t_k + h) - hAx(t_k + h) + \mathrm{O}(h^2) \end{aligned}$$

Thus we have approximately

$$x(t_k + h) \simeq (I - hA)^{-1} x(t_k)$$

So if we perform repeatedly this step, starting from $t_0 = 0$, we get the computed sequence $\{\xi_i\}_0^\infty$ given by

$$\xi_m = (I - hA)^{-m} x_0, \ m = 0, 1, 2, \dots$$

Now, the spectrum of $(I - hA)^{-1}$ is $\{(1 - h\lambda_i)^{-1}\}_{i=1}^\sigma$. Hence by the above expression for ξ_m, we have

$$\begin{aligned} & \xi_m \to 0 \text{ as } m \to \infty \\ \iff & \forall \lambda_i \in \sigma(A), \ |(1 - h\lambda_i)^{-1}| < 1 \\ \iff & \forall \lambda_i \in \sigma(A), \ |1 - h\lambda_i| > 1 \end{aligned}$$

Note that if $\mathrm{Re}(\lambda_i) < 0$, then $|1 - h\lambda_i| > 1$, since $h > 0$. Thus we have shown the following result.

Example. Consider

$$\dot{x}(t) = \lambda x(t)$$

with $\lambda \in \mathbb{C}$. Then the equation is stable if $\mathrm{Re}(\lambda) \leq 0$. In this case the system is exponentially decaying

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

When is the numerically solution x_i also decaying, $\lim_{i\to\infty} x_i = 0$?

$$x_{i+1} = x_i + h\lambda x_{i+1} \iff x_{i+1} = \left(\frac{1}{1 - h\lambda}\right)^{i+1} x_0$$

The solution is decaying (stable) if $|1 + h\lambda| \geq 1$.

Fact. If $\sigma(A) \subset \mathbb{C}_-^\circ$, then for all $h > 0$, for all $x_0 \in \mathbb{C}^n$, the computed sequence $\{\xi_m\}_0^\infty$ obtained via backward Euler goes to zero exponentially.

This is very important in practice, because if h is unfortunately chosen too large the computed sequence may lose accuracy but at least it will never blow up!

AA/EE547: Wi 2020

Lecture 4: Lyapunov Stability and Lyapunov Equation

Lecturer: L.J. Ratliff

Disclaimer: *These notes have not been subjected to the usual scrutiny reserved for formal publications, meaning you should take your own notes in class and review the provided references as opposed to taking these notes as your sole resource. I provide the lecture notes to you as a courtesy; it is not required that I do this. They may be distributed outside this class only with the permission of the Instructor.*

References. Chapter 8 **[JH]**; Chapter 7/7d, **[C&D]**. review your notes on norms including (induced) matrix norms.

1 Lyapunov Theory

Last time we talked about internal or Lyapunov stability. Our focus was on linear dynamical systems but this is an important concept for nonlinear systems as well.

Consider the nonlinear system

$$\dot{x} = f(x) \tag{1}$$

and its linearization around x_e

$$\dot{x} = Df(x_e)x \tag{2}$$

Reminder:

Definition 1. An equilibrium point x_e of (1) is stable if for all $\varepsilon > 0$,

$$\exists\ \delta > 0 : \quad \forall\ x \in B_\delta(x_e),\ t \geq 0,\ \phi_t(x) \in B_\varepsilon(x_0)$$

An equilibrium point x_e is asymptotically stable if it is stable and if

$$\exists\ \delta > 0 : \quad \forall x \in B_\delta(x_e), \lim_{t\to\infty} \phi_t(x) = x_e$$

Definition 2. Let $f \in C^1(\mathbb{R}^n, \mathbb{R})$, $V \in C^1(\mathbb{R}^n, \mathbb{R})$ and ϕ_t the flow of (1). Then for $x \in \mathbb{R}^n$, the derivative of the function $V(x)$ along trajectories (the solution) $\phi_t(x)$ is

$$\dot{V}(x) = \frac{d}{dt}V(\phi_t(x)) = \frac{\partial V(\phi_t)}{\partial \phi_t}\frac{d}{dt}\phi_t(x) = DV(x)f(x)$$

The existence of a Lyapunov function is an alternative way to prove stability. For a dynamical system, $\dot{x} = f(x)$ with $x = 0$ as an equilibrium point, a scalar function $V : \mathbb{R}^n \to \mathbb{R}$ is a Lyapunov function if it is C^1, locally positive definite, and $\dot{V} \leq 0$. Positive definite means

$$V(z) \geq 0,$$

all sublevel sets are bounded (i.e. $V(z) \to \infty$ as $z \to \infty$), and

$$V(z) = 0 \iff z = 0$$

Theorem 3. Let W be an open subset of $\mathbb{R}^n$ containing x_e. Suppose that $f \in C^1(\mathbb{R}^n, \mathbb{R})$ and that $f(x_e) = 0$. Suppose further that there exists a real valued function $V \in C^1$ such that $V(x_e) = 0$ and $V(x) > 0$ when $x \neq x_e$. Then

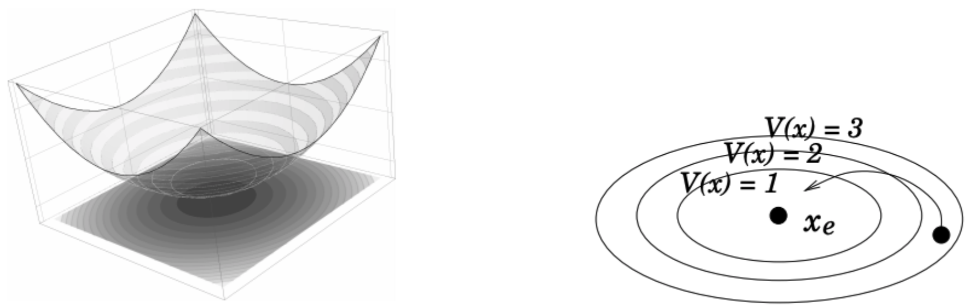


Figure 1: Example: Lyapunov Function; Level sets $V(x) = 1$, $V(x) = 2$, and $V(x) = 3$ for a Lyapunov function V; thus if a state trajectory enters one of these sets, it has to stay inside it since $\dot{V}(x) \leq 0$.

a.
$$\dot{V}(x) \leq 0 \ \forall x \in W \implies x_e \text{ is stable}$$

b.
$$\dot{V}(x) < 0 \ \forall x \in W \backslash \{x_e\} \implies x_e \text{ is asy. stable}$$

c.
$$\dot{V}(x) > 0 \ \forall x \in W \backslash \{x_e\} \implies x_e \text{ is unstable}$$

2 Stability of Linearized Systems

Consider a general non-linear system

$$\dot{x} = f(x), \ x \in \mathbb{R}^n$$

with an equilibrium point x_e such that $f(x_e) = 0$. Recall that the local linearization around x_e is given by

$$\delta\dot{x} = A\delta x$$

with $\delta x = x - x_e$ and

$$A = \frac{\partial f(x_e)}{\partial x}$$

Theorem 4. Suppose that $f \in C^2(\mathbb{R}^n, \mathbb{R})$. If the linearized system is exponentially stable, then there exists a ball $B \subset \mathbb{R}^n$ around x_e and constants $c, \lambda > 0$ such that for every solution $x(t)$ to the nonlinear system that starts at $x(t_0) \in B$, we have

$$\|x(t) - x_e\| \leq ce^{\lambda(t-t_0)}\|x(t_0) - x_e\|$$

This means that the properties of the linearized system are preserved in the nonlinear system.

Conversely, instability is also a property that is preserved.

Theorem 5. Suppose that $f \in C^2(\mathbb{R}^n, \mathbb{R})$. If the linearized system is unstable, then there are solutions that start arbitrarily close to x_e but do not converge to this point as $t \to \infty$.

Note: marginal stability is not a transferable property.

Example 6. The two systems

$$\dot{x} = \pm x^3$$

(one for "+" one for "−") have the same local linearization $\delta\dot{x} = 0$ around $x_e = 0$ which is only marginally stable. However $\dot{x} = -x^3$ is such that x always converges to zero, while for $\dot{x} = x^3$, x always diverges away from the equilibrium point.

Let's consider the inverted pendulum. I encourage you to play around with this in Python as well.

The inverted pendulum has dynamics

$$m\ell^2\ddot{\theta} = mg\ell\sin\theta - b\dot{\theta} + \tau$$

Let $u = \tau$ be the input and let $y = \theta$ the observed quantity. The local linearization of this system around the equilibrium point for which $\theta = \pi$ is given by

$$\delta\dot{x} = A\delta x + Bu, \;\; \delta y = C\delta x$$

with

$$A = \begin{bmatrix} 0 & 1 \\ -\frac{g}{\ell} & -\frac{b}{m\ell^2} \end{bmatrix}, \; B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \; C = \begin{bmatrix} 1 & 0 \end{bmatrix}$$

The eigenvalues of A are given by

$$\det(\lambda I - A) = \lambda\left(\lambda + \frac{b}{m\ell^2}\right) + \frac{g}{\ell} \iff \lambda = -\frac{b}{2m\ell^2} \pm \sqrt{\frac{b}{2m\ell^2} - \frac{g}{\ell}}$$

So, the linearized system is exponentially stable since the λ's have negative real part.

What about for the equilibrium $x_e = 0$? The eigenvalues in this case are

$$\lambda = -\frac{b}{2m\ell^2} \pm \sqrt{\frac{b}{2m\ell^2} + \frac{g}{\ell}}$$

so that

$$-\frac{b}{2m\ell^2} + \sqrt{\frac{b}{2m\ell^2} + \frac{g}{\ell}} > 0 \implies \text{unstable}$$

3 Lyapunov Stability

For linear systems, it turns out that Lyapunov functions take the form $z^\top P z$ for some positive definite matrix.

Recall the definition of a positive definite matrix.

Definition 7. Positive Definite Matrix A symmetric matrix $Q \in \mathbb{R}^{n\times n}$ is *positive definite* if

$$x^\top Q x \geq 0, \;\; \forall\, x \in \mathbb{R}^n/\{0\}$$

Proposition 8. Consider a symmetric matrix Q. The following are equivalent:

a. Q is positive definite
b. $\sigma(Q) \subset \mathbb{R}_+$ (i.e., the eigenvalues of Q are all strictly positive)
c. The determinants of all upper left submatrices of Q are positive
d. There exists an $n \times n$ non-singular real matrix H such that $Q = H^\top H$.

For a positive definite matrix Q, we also have that

$$0 < \lambda_{\min}(Q)\|x\|^2 \leq x^\top Qx \leq \lambda_{\max}(Q)\|x\|^2, \quad \forall x \neq 0$$

The Lyapunov stability theorem provides an alternative condition to check whether or not the continuous-time homogeneous LTI system

$$\dot{x} = Ax \tag{3}$$

is asymptotically stable.

The equation

$$A^\top P + PA + Q = 0$$

is called the **Lyapunov equation**. Here, $P, Q \in \mathbb{R}^{n\times n}$ are symmetric matrices. For a linear system $\dot{x} = Ax$, if

$$V(z) = x^\top Pz$$

then

$$\dot{V}(z) = (Az)^\top Pz + z^\top P(Az) = -z^\top Qz$$

That is, if $z^\top Pz$ is the (generalized) energy, then $z^\top Qz$ is the associated (generalized) dissipation.

If $P > 0$, then the sublevel sets[1] of this function are ellipsoids and bounded. Further, $V(z) = z^\top Pz = 0 \iff z = 0$.

If $P > 0, Q \geq 0$, then all the trajectories of $\dot{x} = Ax$ are bounded (i.e., $\mathrm{Re}(\lambda_i) \leq 0$ and if $\mathrm{Re}(\lambda_i) = 0$, then λ_i corresponds to a Jordan block of size one). Further, the ellipsoids $\{z | \; z^\top Pz \leq a\}$ are invariant sets.

Theorem 9. The following conditions are equivalent:

a. The system (3) is asymptotically (equiv. exponentially) stable
b. All the eigenvalues of A have strictly negative real parts
c. For every symmetric positive definite matrix Q, there exists a unique solution P to the following *Lyapunov equation*

$$A^\top P + PA = -Q$$

   Moreover, P is symmetric and positive-definite.
d. There exists a symmetric positive-definite matrix P for which the following Lyapunov matrix inequality holds:

$$A^\top P + PA < 0$$

Proof. Proof Sketch. We have already seen that 1 and 2 are equivalent. To show that 2 implies 3, we need to show that

$$P = \int_0^\infty e^{A^\top t} Q e^{At} \, dt$$

is the unique solution to $A^\top P + PA = -Q$. This can be done by showing that 1. the integral is well-defined (i.e. finite), 2. P as defined solves $A^\top P + PA = -Q$, 3. P as defined is symmetric and positive definite, and lastly, 4. no other matrix solves the equation. Indeed,

1. This follows from exponential stability—i.e.

$$\|e^{A^\top t} Q e^{At}\| \to 0$$

   exponentially fast as $t \to \infty$. Hence, the integral is absolutely convergent.

[1] i.e., $\{x | V(x) < a\}$

2. We simply need to compute by direct verification

$$A^\top P + PA = \int_0^\infty A^\top e^{A^\top t} Q e^{At} + e^{A^\top t} Q e^{At} A \; dt$$

But,

$$\frac{d}{dt}\left(e^{A^\top t} Q e^{At}\right) = A^\top e^{A^\top t} Q e^{At} + e^{A^\top t} Q e^{At} A$$

so that

$$\begin{aligned} A^\top P + PA &= \int_0^\infty \frac{d}{dt}\left(e^{A^\top t} Q e^{At}\right) \; dt \\ &= \left(e^{A^\top t} Q e^{At}\right)\Big|_{t=0}^{\infty} \\ &= \left(\lim_{t\to\infty} e^{A^\top t} Q e^{At}\right) - e^{A^\top 0} Q e^{A0} \end{aligned}$$

And, the right-hand side is equal to $-Q$ since $\lim_{t\to\infty} e^{At} = 0$ (by asymptotic stability) and $e^{A0} = I$.

3. this follows by direct computation. Symmetry easily follows:

$$P^T = \int_0^\infty (e^{A^\top t} Q e^{At})^\top \; dt = \int_0^\infty \left(e^{At}\right)^\top Q^\top \left(e^{A^\top t}\right)^\top \; dt = \int_0^\infty e^{A^\top t} Q e^{At} \; dt = P$$

To check positive definiteness, pick an arbitrary vector z and compute:

$$z^\top P z = \int_0^\infty z^\top e^{A^\top t} Q e^{At} z \; dt = \int_0^\infty w(t)^\top Q w(t) \; dt$$

where $w(t) = e^{At} z$. Since Q is positive definite, we get that $z^\top P z \geq 0$. Moreover

$$z^\top P z = 0 \implies \int_0^\infty w(t)^\top Q w(t) \; dt = 0$$

which only happens if $w(t) = e^{At} z = 0$ for all $t \geq 0$, from which one concludes that $z = 0$, because e^{At} is non-singular (recall all state transition matrices are!). Thus P is positive definite.

4. *prove by contradiction.* assume there is some $\bar{P}$ that solves it:

$$A^\top P + PA = -Q, \quad \text{and} \quad A^\top \bar{P} + \bar{P} A = -Q$$

Then

$$A^\top (P - \bar{P}) + (P - \bar{P}) A = 0$$

Multiplying by $e^{A^\top t}$ and e^{At} on the left and right, respectively, we conclude that

$$e^{A^\top t} A^\top (P - \bar{P}) e^{At} + e^{A^\top t} (P - \bar{P}) A e^{At} = 0, \quad \forall \; t \geq 0$$

Yet,

$$\frac{d}{dt}\left(e^{A^\top t} (P - \bar{P}) e^{At}\right) = e^{A^\top t} A^\top (P - \bar{P}) e^{At} + e^{A^\top t} (P - \bar{P}) A e^{At} = 0$$

implying that $e^{A^\top t}(P - \bar{P}) e^{At}$ must be constant. But,because of stability, this quantity must converge to zero as $t \to \infty$, so it must be always zero. Since e^{At} is nonsingular, this is possible only if $P = \bar{P}$.

The implication that condition 3. $\implies$ condition 4 follows immediately, because if we select $Q = -I$ in condition 3., then the matrix P that solves the Lyapunov equation also satisfies $A^\top P + PA < 0$.

Now to complete the proof we need to show that condition 4. implies condition 2. Let P be a symmetric positive-definite matrix for which

$$A^\top P + PA < 0 \tag{4}$$

holds and define

$$Q = -(A^\top P + PA)$$

Consider an arbitrary solution to the LTI system and define the scalar time-dependent map

$$v(t) = x^\top(t)Px(t) \geq 0$$

Taking derivatives, we have

$$\dot{v} = \dot{x}^\top Px + x^\top P\dot{x} = x^\top(A^\top P + PA)x = -x^\top Qx \leq 0$$

Thus, $v(t)$ is nonincreasing and we conclude that

$$v(t) = x^\top(t)Px(t) \leq v(0) = x^\top(0)Px(0)$$

But since $v = x^\top Px \geq \lambda_{\min}(P)\|x\|^2$ we have that

$$\|x\|^2 \leq \frac{x^\top(t)Px(t)}{\lambda_{\min}(P)} = \frac{v(t)}{\lambda_{\min}(P)} \leq \frac{v(0)}{\lambda_{\min}(P)}$$

which means that the system is stable. To verify that it is actually exponentially stable, we go back to the derivative of v, and using the facts that $x^\top Qx \geq \lambda_{\min}(Q)\|x\|^2$ and $v = x^\top Px \leq \lambda_{\max}(P)\|x\|^2$, we get that

$$\dot{v} = -x^\top Qx \leq -\lambda_{\min}(Q)\|x\|^2 \leq -\frac{\lambda_{\min}(Q)}{\lambda_{\max}(P)}v, \quad \forall\ t \geq 0$$

To finish we need a comparison lemma. Comparison results are commonly used analysis tools in control theory and hence, it is worth think a bit about them to better understand their utility.

Lemma 10 (Comparison Lemma)**.** Let $v(t)$ be a differentiable scalar signal. For some constant $\mu \in \mathbb{R}$,

$$\dot{v} \leq \mu v(t),\ \forall t \geq t_0 \implies v(t) \leq e^{\mu(t-t_0)}v(t_0),\ \forall t \geq t_0$$

Applying this lemma, we get that

$$v(t) \leq e^{-\lambda(t-t_0)}v(t_0),\ \forall t \geq 0, \quad \lambda = -\frac{\lambda_{\min}(Q)}{\lambda_{\max}(P)}$$

which shows that $v(t)$ converges to zero exponentially fast and so does $\|x(t)\|$. □

There are discrete time versions of the above results. Indeed, consider

$$x_{k+1} = Ax_k$$

then we have a similar theorem.

Theorem 11. The following four conditions are equivalent:

1. The DT LTI system is asymptotically (exponentially) stable.
2. All the eigenvalues of A have magnitude strictly smaller than 1.
3. For every symmetric positive definite Q, there exists a unique solution P to the following Stein equation (aka the discrete-time Lyapunov equation):

   $$A^\top PA - P = -Q$$

   Moreover, P is symmetric and positive-definite.
4. There exists a symmetric positive-definite matrix P for which the following Lyapunov matrix inequality holds:

   $$A^\top PA - P < 0$$

4 The Lyapunov Operator

The Lypunov operator is defined by

$$\mathcal{L}(P) = A^\top P + PA$$

Recall that we showed last quarter that the eigenvalues of this operator are $(\lambda_i + \lambda_j)$ where λ_i, λ_j are eigenvalues of A. Hence, we have the following result.

Proposition 12. $\mathcal{L}$ is non-singular if and only if A and $-A$ share no common eigenvalues.

As a consequence of this proposition, we have the following facts.

Fact. If A is stable, then the Lyapunov operator is non-singular. If A has imaginary eigenvalues, then the Lyapunov operator is singular (since they have to come in complex conjugate pairs). Thus if A is stable, for any Q there is exactly one solution P of Lyapunov equation $A^\top P + PA + Q = 0$.

If A is stable, and P is a unique solution of $A^\top P + PA + Q = 0$, then

$$V(z) = z^\top P z = z^\top \left(\int_0^\infty e^{tA^\top} Q e^{tA} \, dt \right) z = \int_0^\infty x(t)^\top Q x(t) \, dt$$

The interpretation of this is that $V(z)$ is a cost-to-go from point z given no input.

As in the above proof, if A is stable and $Q > 0$, then for each t

$$e^{tA^\top} Q e^{tA} > 0$$

so that

$$P = \int_0^\infty e^{tA^\top} Q e^{tA} \, dt > 0$$

This means that if A is stable, we can choose any positive definite quadratic form $z^\top Q z$ as the dissipation, i.e. $-\dot{V} = z^\top Q z$. Then, we can solve a set of linear equations to find the unique quadratic function form $V(z) = z^\top P z$. And, V will be positive definite so it is a Lyapunov function that proves A is stable.

That is, a linear system is stable iff there is a quadratic Lyapunov function that proves it.

5 Other applications of Lyapunov functions

We can use the Lyapunov equation to assess state feedback. Consider

$$\dot{x} = Ax + Bu, \; y = Cx, \; u = Kx, \; x(0) = x_0$$

and suppose the closed loop dynamics $\dot{x} = (A + BK)x$ are stable. Then to evaluate quadratic integral performance measures

$$J_u = \int_0^\infty u(t)^\top u(t) \, dt, \; J_y = \int_0^\infty y(t)^\top y(t) \, dt$$

we can solve Lyapunov equations

$$(A + BK)^\top P_u + P_u(A + BK) + K^\top K = 0$$

and

$$(A + BK)^\top P_y + P_y(A + BK) + C^\top C = 0$$

so that we have

$$J_u = x_0^\top P_u x_0, \; J_y = x_0^\top P_y x_0$$

AA/EE547: Wi 2020

Lecture 5: Introduction to Controllability and Observability

Lecturer: L.J. Ratliff

References. Chapter 11 and 15 **[JH]**; Chapter 8/8d, **[C&D]**.

1 Controllability

Let $\mathcal{D} = (\mathcal{U}, \mathcal{X}, \mathcal{Y}, s, \rho)$ be a dynamical system. Let $t_0 < t_1$. The input $u_{[t_0,t_1]}(\cdot)$ steers x_0 at t_0 to x_1 at t_1 if

$$x_1 = s(t_1, t_0, x_0, u_{[t_0,t_1]})$$

Definition 1 (Controllable)**.** The system representation $\mathcal{D}$ is controllable on $[t_0, t_1]$ if for all $(x_0, x_1) \in \mathcal{X}$, there exists $u_{[t_0,t_1]} \in \mathcal{U}$ which steers x_0 at t_0 to x_1 at t_1.

Proposition 2. $\mathcal{D}$ is controllable on $[t_0, t_1] \iff$ for all $x_0 \in \mathcal{X}$, the map

$$s(t_1, t_0, x_0, u_{[t_0,t_1]}(\cdot)) : \mathcal{U} \to \mathcal{X}$$

is **surjective**, that is it maps $\mathcal{U}$ onto $\mathcal{X}$.

1.1 Memoryless Feedback and Controllability

Consider two memoryless maps

$$\begin{aligned} F_s &: \mathcal{X} \to \mathcal{U} && \text{(memoryless state feedback)} \\ F_o &: \mathcal{Y} \to \mathcal{U} && \text{(memoryless output feedback)} \end{aligned}$$

Applying F_s and F_o to our dynamical system $\mathcal{D} = (\mathcal{U}, \mathcal{X}, \mathcal{Y}, s, r)$ (where s: state transition map, r: read out map) we get resulting systems $\mathcal{D}_s$ and $\mathcal{D}_o$, respectively, which are depicted in Figures 1 and 2, respectively.

Thus, for all t, we have

$$u(t) = v(t) - F_s(x(t))$$

and

$$u(t) = v(t) - F_o(y(t))$$

Assumption. (well-posedness) For $\mathcal{D}_s$ and $\mathcal{D}_o$, assume that for all (x_0, t_0), for all exogenous inputs $v(\cdot)$ there is one and only one state response $x(\cdot)$ and one output response $y(\cdot)$.

Example 3. (of a system violating well-posedness). Consider $A = 0, B = 0, C = 0$, and $D = 1$ (SISO linear system). Consider $u = v + y$. The closed loop system is ill-posed. Roughly speaking, well-posedness calls for some delay around the feedback loop.

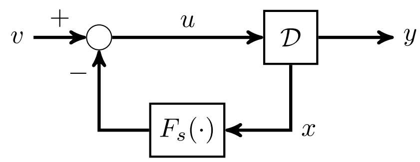


Figure 1: System $\mathcal{D}_s$ with memoryless state feedback

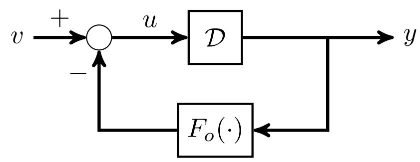


Figure 2: System $\mathcal{D}_o$ with memoryless state feedback

Theorem 4. Let $\mathcal{D}_s, \mathcal{D}_o$ be well-posed. Then

$$\begin{aligned} &\mathcal{D} \text{ is controllable on } [t_0, t_1] \\ &\quad \iff \mathcal{D}_s \text{ is controllable on } [t_0, t_1] && (1) \\ &\quad \iff \mathcal{D}_o \text{ is controllable on } [t_0, t_1] && (2) \end{aligned}$$

Remark. Roughly speaking: nonlinear memoryless state-feedback and output-feedback do not affect controllability. Moreover, the memoryless assumption is crucial: it allows us to use the same state space for $\mathcal{D}$, $\mathcal{D}_s$ and $\mathcal{D}_o$.

Proof. (1) ($\Longrightarrow$) By assumption $\mathcal{D}$ is controllable on $[t_0, t_1]$. Consider arbitrary (x_0, t_0) and (x_1, t_1). Since $\mathcal{D}$ is controllable on $[t_0, t_1]$, $\exists \tilde{u}_{[t_0,t_1]}(\cdot)$ steering x_0 at t_0 to x_1 at t_1. The state space of $\mathcal{D}$ is identical to $\mathcal{D}_s$ since F_s is memoryless. Apply to $\mathcal{D}_s$ the exogenous input defined

$$\begin{aligned} \tilde{v}(t) &= \tilde{u}(t) + F_s(x(t)) \\ &= \tilde{u}(t) + F_s(s(t, t_0, x_0, \tilde{u}_{[t_0,t]})) \end{aligned}$$

Then $\tilde{v}(t)$ steers x_0 at t_0 to x_1 at t_1 by the well-posedness assumption.

($\Longleftarrow$) By controllability of $\mathcal{D}_s$, for all $x_0, x_1 \in \mathcal{X}$, $\exists\ \tilde{v}_{[t_0,t_1]}$ that steers (x_0, t_0) to (x_1, t_1) on $\mathcal{D}_s$. Since $\mathcal{D}$ and $\mathcal{D}_s$ have the same state space, by the well-posedness assumption, $\tilde{v}$ will produce a unique input $\tilde{u}$ of $\mathcal{D}$ which steers x_0 at t_0 to x_1 at t_1.

(2) follows similarly. $\square$

Key Take-Away. Roughly speaking, *nonlinear memoryless* state-feedback and output-feedback do not affect controllability. The memeoryless assumption is *crucial* since it allows for us to use the *same state space* for $\mathcal{D}, \mathcal{D}_s, \mathcal{D}_o$.

2 Observability

Definition 5 (Observable)**.** The dynamical system $\mathcal{D}$ is called *observable* on $[t_0, t_1]$ if and only if, given, $\mathcal{D}$, for all inputs $u_{[t_0,t_1]}$ and for all corresponding outputs $y_{[t_0,t_1]} \in \mathcal{Y}$ the state x_0 at time t_0 is *uniquely* determined.

Of course, once x_0 is calculated, from $u_{[t_0,t_1]}$ and the state transition map s we can calculate the state trajectory $x(\cdot)$; indeed,

$$x(t) = s(t, t_0, x_0, u_{[t_0,t_1]}),\ \forall\, t \in [t_0, t_1]$$

Fact. $\mathcal{D}$ is observable on $[t_0, t_1] \iff$ for each fixed $u_{[t_0,t_1]}$ the partial response map

$$x_0 \mapsto y_{[t_0,t_1]} = \rho(\cdot, t_0, x_0, u_{[t_0,t_1]})$$

is ***injective***, that is the partial response map is a one to one map from $\mathcal{X}$ to $\mathcal{Y}$.

2.1 Memoryless Feedback and Feedforward, and observability

We consider a given dynamical system $\mathcal{D}$, a map $F_o : \mathcal{Y} \to \mathcal{U}$ and a map $F_f : \mathcal{U} \to \mathcal{Y}$. We use F_o to apply to $\mathcal{D}$ a memoryless output feedback and F_f to apply to $\mathcal{D}$ a memoryless feedforward. Call the resulting system $\mathcal{D}_o$ and $\mathcal{D}_f$, resp. (see Figs. 3 and 4).

Then for $\mathcal{D}_f$,

$$\eta(t) = y(t) + F_f(u(t))$$

For $\mathcal{D}_o$, we make the uniqueness (well-posedness) Assumption 1.1.

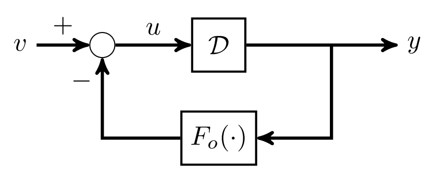


Figure 3: System $\mathcal{D}_o$ with memoryless output feedback

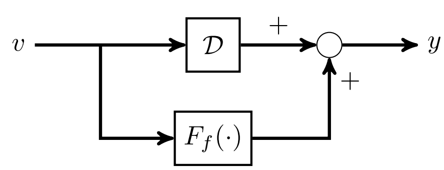


Figure 4: System $\mathcal{D}_f$ with memoryless feedforward

Theorem 6. For the system $\mathcal{D}_f$ and the system $\mathcal{D}_o$ satisfying Assumption 1.1, we have

$$\begin{aligned} &\mathcal{D} \text{ is observable on } [t_0, t_1] \\ &\quad \iff \mathcal{D}_o \text{ is observable on } [t_0, t_1] \qquad (1) \\ &\quad \iff \mathcal{D}_f \text{ is observable on } [t_0, t_1] \qquad (2) \end{aligned}$$

Remark. Memoryless state feedback may affect observability. For example, for a linear time-invariant system representation $R = [A, B, C, D]$, there may exist a linear state feedback F_s such that for some states x_0 and for some inputs $u(\cdot)$, the state trajectory remains, for all t, in the nullspace of C.

3 Controllability & Observability of LTV Systems

Reference: **[C&D]** 8.2,8.3,8.4; **[JH]** Chapter 11, 15

Consider a LTV system

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

Then for an input $u(t)$ defined on the interval $[t_0, t]$, the solution is given by

$$x(t) = s(t, t_0, x_0, u_{[t_0,t]}) = \Phi(t, t_0)x_0 + \int_{t_0}^{t} \Phi(t, \tau)B(\tau)u(\tau)\ d\tau$$

3.1 Controllability and Reachability of LTV

Definition 7 (Controllable)**.** The dynamical system is controllable on $[t_0, t_1]$ iff $\forall(x_0, t_0)$ and $\forall(x_1, t_1)$, there exists $u(\cdot)$ that steers (x_0, t_0) to (x_1, t_1)—i.e., $x(t_1) = s(t_1, t_0, x_0, u) = x_1$.

Writing out $x(t_1)$ using the above expression for the solution, we can define the so-called reachbility map:

$$\begin{aligned}x_1 = x(t_1) = s(t_1, t_0, x_0, u_{[t_0,t_1]}) &= \Phi(t_1, t_0)x_0 + \int_{t_0}^{t_1} \Phi(t_1, \tau)B(\tau)u(\tau)\ d\tau\\ &= \Phi(t_1, t_0)x_0 + \mathcal{L}_r u\end{aligned}$$

where the reachability map $\mathcal{L}_{r,[t_0,t_1]} : PC([t_0, t_1]) \to \mathbb{C}^n$ is defined by

$$\mathcal{L}_{r,[t_0,t_1]}(u(\cdot)) = \int_{t_0}^{t_1} \Phi(t_1, \tau)B(\tau)u(\tau)\ d\tau$$

Note we will drop the dependence on $[t_0, t_1]$ when clear from context.

Then the expression for x_1 shows that there will be an input $u_{[t_0,t_1]}$ that transfers an arbitrary (x_0, t_0) to an arbitrary (x_1, t_1) if and only if the map $\mathcal{L}_r : PC([t_0, t_1]) \to \mathbb{C}^n$ is surjective.

Fact. Indeed,

$$(A(\cdot), B(\cdot)) \text{ is controllable on } [t_0, t_1] \iff \mathcal{L}_{r,[t_0,t_1]}(u(\cdot)) \text{ is surjective.}$$

The map $\mathcal{L}_{r,[t_0,t_1]}$ determines the set of states that can be reached from the origin at some time $t = t_1$. In short, the study of the range of $\mathcal{L}_r$ is central to the study of controllability.

Sometimes we do not want to specify t_1 in which case we say that the pair $(A(\cdot), B(\cdot))$ is controllable at to iff for some $t_1 > t_0$, the pair is controllable on $[t_0, t_1]$.

Given that the state transition matrix $\Phi(t, t_0)$ is non-singular for all (t, t_0), it is easy to prove the following result.

Theorem 8. The pair $(A(\cdot), B(\cdot))$ is completely controllable (CC) on $[t_0, t_1]$

$$\begin{aligned}&\iff \forall\ x_0 \in \mathbb{R}^n,\ \exists u_{[t_0,t_1]} \text{ that steers } (x_0, t_0) \text{ to } (0, t_1) \text{ (steering to origin)}\\ &\iff \forall x_1 \in \mathbb{R}^n,\ \exists u_{[t_0,t_1]} \text{ that steers } (0, t_0) \text{ to } (x_1, t_1) \text{ (reaching from origin)}\end{aligned}$$

Proof.

$$x_1 = \Phi(t_1, t_0)x_0 + \mathcal{L}_r u$$

since x_0 can be zero and x_1 arbitrary, a necessary condition for CC is that $\mathbb{R}(\mathcal{L}_r) = \mathbb{C}^n$. But this is sufficient too because if $\mathbb{R}(\mathcal{L}_r) = \mathbb{C}^n$, given x_1 at t_1 and x_0 at t_0, $\exists\ u_{[t_0,t_1]}$ such that

$$x_1 - \Phi(t_1, t_0)x_0 = \mathcal{L}_r u$$

Both implications follow. $\square$

Given $(A(\cdot), B(\cdot))$, we say that the state x_0 is *controllable to zero* on $[t_0, t_1]$ iff $\exists u_{[t_0,t_1]}$ that steers (x_0, t_0) to $(0, t_1)$. Analogously, we say that the state x_1 is *reachable from the origin* on $[t_0, t_1]$ iff $\exists u_{[t_0,t_1]}$ that steers $(0, t_0)$ to (x_1, t_1).

Remark. The essence of the reduction theorem (5) is that, for linear system representations, controllability on $[t_0, t_1]$, controllability to zero on $[t_0, t_1]$ of all states, and reachability on $[t_0, t_1]$ of all states are equivalent. The reader should construct a one-dimensional nonlinear example to show that this is not so for nonlinear systems.

Definition 9 (Reachable Subspace)**.** Given two times $t_1 > t_0 \geq 0$, the reachable or controllable from the origin on $[t_0, t_1]$ subspace $\mathcal{R}(\mathcal{L}_r)$ consists of all states x_1 for which there exists an input $u : [t_0, t_1] \to \mathbb{C}^m$ that transfers the state from $x(t_0) = 0$ to $x(t_1) = x_1$—i.e.,

$$\mathcal{R}(\mathcal{L}_r) = \left\{ x_1 \in \mathbb{C}^n :\ \exists u(\cdot),\ x_1 = \int_{t_0}^{t_1} \Phi(t_1, \tau)B(\tau)u(\tau)\ d\tau \right\}$$

Theorem 10 (Controllability in terms of Reachability)**.** Let $(A(\cdot), B(\cdot))$ be given and be piecewise continuous. Then,

$$\begin{aligned} (A(\cdot), B(\cdot)) \text{ controllable on } [t_0, t_1] \quad &\iff \quad \mathcal{R}(\mathcal{L}_r) = \mathbb{C}^n && (1) \\ &\iff \quad \mathcal{R}(\mathcal{L}_r\mathcal{L}_r^*) = \mathbb{C}^n && (2) \\ &\iff \quad \det(W_r(t_0, t_1)) \neq 0 && (3) \end{aligned}$$

where W_r is the reachability grammian

$$W_r(t_0, t_1) = \int_{t_0}^{t_1} \Phi(t_1, \tau)B(\tau)B(\tau)^*\Phi(t_1, \tau)^*\ d\tau$$

Further, the set of reachable states on $[t_0, t_1]$ is the subspace $\mathcal{R}(\mathcal{L}_r)$ which is equal to $\mathcal{R}(W_r(t_0, t_1))$.

Proof. To show the equivalence in (1), note that the left-hand side is equivalent to $\mathcal{L}_r$ being surjective which is by definition equivalent to $\mathcal{R}(\mathcal{L}_r) = \mathbb{C}^n$. To show the equivalence between (1) and (2), note that this is simply Finite Rank Operator Lemma (FROL) applied to $A = \mathcal{L}_r$ viewed as a map from the Hilbert space $L_2^m([t_0, t_1])$ into $\mathbb{C}^n$ (Note that I expect you to be able to argue this without invoking FROL). To show that (2) is equivalent to (3), note that $\mathcal{L}_r\mathcal{L}_r^* : \mathbb{C}^n \to \mathbb{C}^n$, $\mathcal{L}_r\mathcal{L}_r^*$ is surjective iff it is a bijection so using its matrix representation,

$$W_r(t_0, t_1) = \int_{t_0}^{t_1} \Phi(t_1, \tau)B(\tau)B(\tau)^*\Phi(t_1, \tau)^*\ d\tau,$$

we get that (2) is equivalent to (3).

The fact that the set of reachable states on $[t_0, t_1]$ is the subspace $\mathcal{R}(\mathcal{L}_r)$ follows directly from the expression for $x_1 = s(t_1, t_0, x_0, u)$ with $x_0 = 0$. $\square$

Note that by definition, $W_r(t_0, t_1)$ is the integral of a semi-definite Hermitian matrix so that $\forall z \in \mathbb{C}^n$, $z^*W_t(t_0, t_1)z \geq 0$.

Practice Problem. Show that $t_1 \mapsto W_t(t_0, t_1)$ solves the linear matrix differential equation

$$\dot{X}(t) = A(t)X(t) + X(t)A^*(t) + B(t)B(t)^*$$

with $X(t_0) = 0$.

The equivalence between controllability and reachability as described in Theorem 10 let's us define an ostensibly equivalent *controllability map*:

$$\mathcal{L}_c : u_{[t_0,t_1]} \mapsto \int_{t_0}^{t_1} \Phi(t_0, \tau)B(\tau)u(\tau)\ d\tau$$

Note: $\mathcal{L}_c$ surjective $\iff \exists\ u_{[t_0,t_1]}$ that steers arbitrary (x_0, t_0) to arbitrary (x_1, t_1).

Definition 11 (Controllable Subspace)**.** Given two times $t_1 > t_0 \geq 0$, the reachable or controllable to the origin on $[t_0, t_1]$ subspace $\mathcal{R}(\mathcal{L}_c)$ consists of all states x_0 for which there exists an input $u : [t_0, t_1] \to \mathbb{C}^m$ that transfers the state from $x(t_0) = x_0$ to $x(t_1) = 0$—i.e.,

$$\mathcal{R}(\mathcal{L}_c) = \{x_0 \in \mathbb{C}^n :\ \exists u(\cdot),\ 0 = \Phi(t_1, t_0)x_0 + \int_{t_0}^{t_1} \Phi(t_1, \tau)B(\tau)u(\tau)\ d\tau\}$$

Note that

$$0 = \Phi(t_1, t_0)x_0 + \int_{t_0}^{t_1} \Phi(t_1, \tau)B(\tau)u(\tau)\ d\tau \iff x_0 = \int_{t_0}^{t_1} \Phi(t_0, \tau)B(\tau)v(\tau)\ d\tau,\ u(\cdot) = -v(\cdot)$$

This in turn gives us an analogous theorem to the reachability theorem above.

Theorem 12. Let $(A(\cdot), B(\cdot))$ be given and be piecewise continuous. Then,

$$\begin{aligned} (A(\cdot), B(\cdot)) \text{ controllable on } [t_0, t_1] &\iff \mathcal{R}(\mathcal{L}_c) = \mathbb{C}^n && (4)\\ &\iff \mathcal{R}(\mathcal{L}_c\mathcal{L}_c^*) = \mathbb{C}^n && (5)\\ &\iff \det(W_c(t_0, t_1)) \neq 0 && (6) \end{aligned}$$

where W_c is the reachability grammian

$$W_c(t_0, t_1) = \int_{t_0}^{t_1} \Phi(t_0, \tau)B(\tau)B(\tau)^*\Phi(t_0, \tau)^*\ d\tau$$

Further, the set of reachable states on $[t_0, t_1]$ is the subspace $\mathcal{R}(\mathcal{L}_c)$ which is equal to $\mathcal{R}(W_c(t_0, t_1))$.

Practice Problem. Show that given $t_1 > t_0$,

$$\mathcal{R}(\mathcal{L}_r) = \Phi(t_1, t_0)\mathcal{R}(\mathcal{L}_c)$$

and derive a matrix differential equation that $t \mapsto W_c(t, t_1)$ solves.

One interesting application is to the problem of finding the minimum cost control. Consider the cost of control to be the given by the L_2-norm of $u(\cdot)$:

$$\langle u, u\rangle = \int_{t_0}^{t_1} u(t)^*u(t)\ dt = \|u\|_2^2$$

Then if $(A(\cdot), B(\cdot))$ is controllable on $[t_0, t_1]$, then for all $x_0, x_1 \in \mathbb{C}^n$, the input $\tilde{u} : [t_0, t_1] \to \mathbb{C}^m$ defined by

$$\tilde{u}(t) = B(t)^*\Phi(t_1, t)^*W_r(t_0, t_1)^{-1}(x_1 - \Phi(t_1, t_0)x_0)$$

steers (x_0, t_0) to (x_1, t_1). Note this is *one* such control that gets the job done. There are potentially infinitely many others, since controllability is about 'surjectivity' of a particular linear map. To better understand this, the condition

$$x_1 = s(t_1, t_0, x_0, u) = \Phi(t_1, t_0)x_0 + \int_{t_0}^{t_1} \Phi(t_1, \tau)B(\tau)u(\tau)\ d\tau$$

imposes n independent constraints on an infinitely dimensional space (i.e., the space where $u(\cdot)$ lies is an infinite dimensional space). What this implies is that the set of controls that satisfy the above "constraint" forms a linear variety (or affine subspace) of codimension n. Indeed, any $u = \tilde{u} + v$ with $v \in \mathcal{N}(\mathcal{L}_r)$ also gets the job done.

Further, geometrically, $\tilde{u}$ is the least-cost L_2-solution iff $\|\tilde{u}\|_2$ is the minimum distance between the origin and the linear variety $\tilde{u} + \mathcal{N}(\mathcal{L}_r)$. Recall from last quarter that this means that $\tilde{u}$ is the least cost L_2 solution iff $\tilde{u}$ is orthogonal to the variety $\tilde{u} + \mathcal{N}(\mathcal{L}_r)$ which is, in turn, equivalent to $\tilde{u} \perp \mathcal{N}(\mathcal{L}_r)$.

Now, we get to use our friend FROL once again! By FROL, this means that

$$\tilde{u} \text{ is the least cost solution } \iff \tilde{u} \in \mathcal{R}(\mathcal{L}_r^*) \iff \tilde{u} = \mathcal{L}_r^*\xi,\ \xi \in \mathbb{C}^n$$

And, the minimal cost for going for reaching (x_1, t_1) from $(0, t_0)$ is given by

$$\|u\|_2^2 = x_1^* W_r(t_0, t_1)^{-1} x_1$$

This can easily be checked by the following reasoning:

$$u_{[t_0,t_1]} \text{ transfers } (x_0, t_0) \text{ to } (x_1, t_1) \iff x_1 - \Phi(t_1, t_0)x_0 = \mathcal{L}_r u$$

Since $\tilde{u} = \mathcal{L}_r^*\xi$, we have that

$$\xi = (\mathcal{L}_r\mathcal{L}_r^*)^{-1}(x_1 - \Phi(t_1, t_0)x_0) \implies \tilde{u} = \mathcal{L}_r^*\xi = \mathcal{L}_r^*(\mathcal{L}_r\mathcal{L}_r^*)^{-1}(x_1 - \Phi(t_1, t_0)x_0)$$

Practice Problem. How is this connected to the minimum energy J_u from Lecture 4?

Recall again from 510 (in particular, exam 1), that for diagonalizable PSD Hermitian matrices, one can select an orthonormal eigenbasis. Hence, since $W_r(t_0, t_1)$ is a PSD Hermitian matrix, we can expand it as

$$W_r(t_0, t_1) = \sum_{i=1}^{n} \lambda_i v_i v_i^*$$

where (λ_i, v_i) are eigenpairs for W_r and the v_i are orthonormal. Recall that

$$\langle u, u \rangle = \int_{t_0}^{t_1} u(t)^* u(t)\ dt = \|u\|_2^2$$

(L_2 norm), so that for a unit cost $\|u\|_2 = 1$, we can reach any of the points $v_1/\sqrt{\lambda_i}$, $i = 1, \ldots, n$. Moreover, from $(0, t_0)$ we can reach any point on the ellipsoid whose semixes are $v_i/\sqrt{\lambda_i}$. Hence, if we order the eigevalues

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n > 0$$

then the direction v_n is the most expensive to reach and the direction v_1 is teh cheapest, so that the eigenvalues of W_r measure the effectiveness of the actuators in the task of reaching states. Furthermore, thinking along these lines was the origin for reachability computation in linear systems theory.

3.2 Observability of LTV

Consider $A(\cdot), C(\cdot)$. Recall

$$y(t) = \rho(t, t_0, x_0, u_{[t_0,t_1]}) = C(t_1)\Phi(t_1, t_0)x_0 + \int_{t_0}^{t_1} C(t_1)\Phi(t_1, \tau)B(\tau)u(\tau)\ d\tau$$

Let $\mathcal{L}_0 : \mathbb{R}^n \to \mathcal{Y}_{[t_0,t_1]}$ be defined by

$$\mathcal{L}_o x_0 = C(\cdot)\Phi(\cdot, t_0)x_0$$

(that is, $\mathcal{L}_o x_0$ is an operator in $PC([t_0, t_1])$) such that

$$(\mathcal{L}_o x_0)(t) = y(t) - \int_{t_0}^{t} C(t)\Phi(t, \tau)B(\tau)u(\tau)\ d\tau$$

Definition 13. The state x_0 is unobservable on $[t_0, t_1]$ iff its zero input response is zero on $[t_0, t_1]$.

Hence,

$$x_0 \text{ is unobservable on } [t_0, t_1] \iff x_0 \in \mathcal{N}(\mathcal{L}_o)$$

Theorem 14. Given $(A(\cdot), C(\cdot))$ (piecewise continuous on $\mathbb{R}_+$), the following are equivalent:

$$\begin{aligned} (A(\cdot), C(\cdot)) \text{ is completely observable (CO) on } [t_0, t_1] &\iff \mathcal{N}(\mathcal{L}_o) = \{0\} \\ &\iff \mathcal{N}(\mathcal{L}_o^*\mathcal{L}_o) = \{0\} \\ &\iff \det(W_o(t_0, t_1)) \neq 0 \end{aligned}$$

where

$$W_o(t_0, t_1) = \int_{t_0}^{t_1} \Phi(\tau, t_0)^* C(\tau)^* C(\tau)\Phi(\tau, t_0)\ d\tau$$

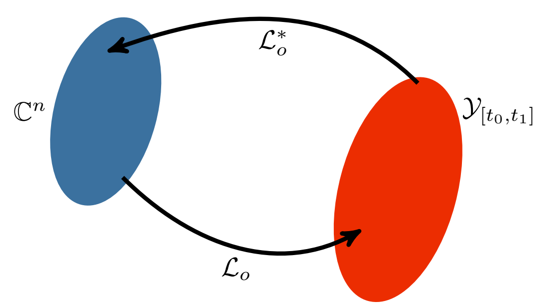


Figure 5: Graphic of Observability Map Operation

The proof of this result follows directly from the definitions of the observability map and FROL. Note: you should know how to argue the results of FROL in the context of the observability map.

Analogous to the results we had for controllability, a consequence of the above theorem is the following.

Corollary 15. Suppose that $(A(\cdot), C(\cdot))$ is observable on $[t_0, t_1]$. Then we have the following results:

a. Let y be the zero-input response due to x_0 so that

$$\langle y, y\rangle = x_0^* W_o(t_0, t_1)x_0$$

b. Given $y_{[t_0,t_1]}$, x_0 is restricted by

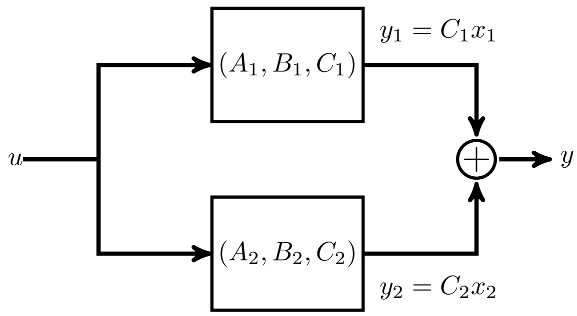
$$x_0 = (\mathcal{L}_o^*\mathcal{L}_o)^{-1}\mathcal{L}_o^* y = W_o(t_0, t_1)^{-1} \int_{t_0}^{t_1} \Phi(\tau, t_0)^* C(\tau)^* y(\tau)\ d\tau$$

And as in the case of the controllability map, we can characterize observability in terms of the eigenstructure of W_o. Indeed, let $\lambda_n > 0$ be the smallest eigenvalue of the positive definite Hermitian matrix $W_o(t_0, t_1)$ and en its corresponding normalized eigenvector. Then for $x_0 = e_n$, $\|x_0\|_2 = 1$ and its zero-input response is such that $\langle y, y\rangle = \lambda_n$. So, if $\lambda_n << 1$, some states are barely observable in case of noisy observations.

Practice Problem. Show that $t_0 \mapsto W_o(t_0, t_1)$ is the solution to the linear matrix differential equation

$$\dot{X}(t) = -A(t)^* X(t) - X(t)A(t) - C(t)^* C(t), \; X(t_1) = 0$$

Example 16. Consider two systems connected in parallel.



The overall system has state-space model

$$\dot{x} = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix} x + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} u, \;\; y = \begin{bmatrix} C_1 & C_2 \end{bmatrix} x$$

The output is

$$y(t) = C_1 e^{A_1 t} x_1(0) + C_2 e^{A_2 t} x_2(t) + \int_0^t (C_1 e^{A_1(t-\tau)} B_1 + C_2 e^{A_2(t-\tau)} B_2) u(\tau)\; d\tau$$

When $A = A_1 = A_2$ and $C_1 = C_2 = C$, this reduces to

$$y(t) = Ce^{At}(x_1(0) + x_2(0)) + \int_0^t Ce^{A(t-\tau)}(B_1 + B_2)u(\tau)\; d\tau$$

This example demonstrates that simply knowing the input and output of the system, we cannot necessarily distinguish between initial states for which $x_1(0) + x_2(0)$ is the same value.

Theorem 17. Given t_1, t_0 with $t_1 > t_0 \geq 0$, the unobservable subspace is such that

$$\mathcal{UO}(t_0, t_1) = \ker W_o(t_0, t_1)$$

Proof. For every $x_0 \in \mathbb{R}^n$, we have that

$$x_0^\top W_o(t_0, t_1) x_0 = \int_{t_0}^{t_1} x_0^\top \Phi(\tau, t_0)^\top C(\tau)^\top C(\tau)\Phi(\tau, t_0) x_0 \; d\tau = \int_{t_0}^{t_1} \|C(\tau)\Phi(\tau, t_0)x_0\|^2 \; d\tau$$

so that

$$x_0 \in \ker W_o(t_0, t_1) \implies C(\tau)\Phi(\tau, t)x_0 = 0, \; \forall \tau \in [t_0, t_1] \implies x_0 \in \mathcal{UO}(t_0, t_1)$$

On the other hand,

$$x_0 \in \mathcal{UO}(t_0, t_1) \implies C(\tau)\Phi(\tau, t_0)x_0 = 0, \; \forall \tau \in [t_0, t_1] \implies x_0 \in \ker(W_o(t_0, t_1))$$

where we have used the fact that

$$x^\top W x = 0 \implies Wx = 0$$

for W PSD. $\square$

Remark. The key to giving constructive tests for controllability and observability is to give conditions under which $\mathcal{R}(\mathcal{L}_c) = \mathbb{R}^n$ and $\mathcal{N}(\mathcal{L}_o) = \{0\}$.

4 Duality

The controllability and observability theorems we have stated so far are intimately related. Consider the system

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

Recalling our knowledge about computing adjoints, we can write the dual representation as

$$\begin{aligned} -\dot{\tilde{x}}(t) &= A^*(t)\tilde{x}(t) + C^*(t)\tilde{u}(t) \\ \tilde{y}(t) &= B^*(t)\tilde{x}(t) + D^*(t)\tilde{u}(t) \end{aligned}$$

where here $\tilde{x}(t) \in \mathbb{C}^n, \tilde{u}(t) \in \mathbb{C}^p, \tilde{y}(t) \in \mathbb{C}^m$. The state transition matrix is

$$\Psi(t,\tau) = \Phi(\tau,t)^*$$

The minus sign on the dynamics essentially captures that the dual runs in reverse time. If we take the dual of the dual, we get

$$(A(\cdot), -B(\cdot), -C(\cdot), D(\cdot))$$

so that the original system is equal to the dual of the dual modulo a sign change for the state. And thus

$$(\mathcal{L}_r^*)^* = -\mathcal{L}_r,\ (\mathcal{L}_c^*)^* = -\mathcal{L}_c,\ (\mathcal{L}_o^*)^* = -\mathcal{L}_o$$

It turns out that controllability to zero on $[t_0, t_1]$ is the dual of unobservability on $[t_0, t_1]$ and vice versa.

Theorem 18. The subspace of all states of $(A(\cdot), B(\cdot), C(\cdot), D(\cdot))$ that are controllable to zero (unobservable) on $[t_0, t_1]$ is the orthogonal complement of the subspace of all states of its dual that are unobservable (controllable to zero, resp.) on $[t_0, t_1]$. That is,

$$\mathcal{R}(\mathcal{L}_c) = \mathcal{N}(\mathcal{L}_o^*)^{\perp},\ \mathcal{N}(\mathcal{L}_o) = \mathcal{R}(\mathcal{L}_c^*)^{\perp}$$

Corollary 19. The system $(A(\cdot), B(\cdot), C(\cdot), D(\cdot))$ is controllable (observable) on $[t_0, t_1]$ iff its dual is observable (controllable, resp.) on $[t_0, t_1]$.

5 Discrete Time Controllability and Reachability

Remark. The fundamental difference between CT and DT is that DT controllability to zero does not necessarily imply reachability from zero.

Definition 20. We say that the pair $(A(\cdot), B(\cdot))$ is controllable on $[k_0, k_1]$ iff for all (x_0, k_0) and for all (x_1, k_1) there exists a control sequence $u_{[k_0,k_1-1]} = (u(k_0), \ldots, u(k_1 - 1))$ that transfers the (x_0, t_0) to the (x_1, t_1).

14 Proof of Theorem (10). a) Calculations show that the state x_0 of $R(\cdot)$ is *controllable to zero* on $[t_0,t_1]$ iff $\exists\ u_{[t_0,t_1]}$ s.t.

15 $$x_0=-\int_{t_0}^{t_1}\Phi(t_0,\tau)B(\tau)u(\tau)d\tau\ ,$$

equivalently, iff $x_0\in R(L_c)$ where (see (8.2.25))

16 $$L_c: u_{[t_0,t_1]}\ \rightarrow\ \int_{t_0}^{t_1}\Phi(t_0,\tau)B(\tau)u(\tau)d\tau\ .$$

Now by (8.3.6), the state $\tilde{x}_0$ of $\tilde{R}(\cdot)$ is *unobservable* on $[t_0,t_1]$ iff $\tilde{L}_o\tilde{x}_0=\theta$, more precisely iff

17 $$B(t)^*\Psi(t,t_0)\tilde{x}_0=B(t)^*\Phi(t_0,t)^*\tilde{x}_0=\theta_{n_i}\qquad\forall\, t\in[t_0,t_1]\ .$$

But by (8.2.26), (17) is equivalent to: $\tilde{x}_0\in N(W_c(t_0,t_1))=N(L_c^*)$. Now by (A.7.57), $R(L_c)=N(L_c^*)^\perp$. Thus we have established (since $\tilde{L}_o=L_c^*$), $R(L_c)=N(\tilde{L}_o)^\perp$. Hence the first equality in (11) and in (12) hold.

b) To obtain the others repeat the reasoning of a) to the dual $\tilde{R}(\cdot)$, (4), and the dual of the dual $\tilde{\tilde{R}}(\cdot)$, (8), for which (9) holds. Hence by the first equality (11)

$$[R(\tilde{L}_c)]^\perp=N(\tilde{\tilde{L}}_o)=N(L_o),$$

i.e. the second equality (11) holds. The second equality (12) follows similarly. ∎

Given the system $(A(\cdot),B(\cdot))$, we know that $u_{[k_0,k_1-1]}$ transfers x_0 to x_1 iff

$$x_1=s(k_1,k_0,x_0,u_0)=\Phi(k_1,k_0)x_0+\sum_{\ell=k_0}^{k_1-1}\Phi(k_1,\ell+1)B(\ell)u(\ell)$$

This expression, in turn, shows that there will be such an input taking arbitrary x_0 to arbitrary x_1 if and only if the linear map

$$\mathcal{L}_r(k_0,k_1):\mathcal{U}_d(k_0,k_1-1)\to\mathbb{C}^n:u_{[k_0,k_1-1]}\to\sum_{\ell=k_0}^{k_1-1}\Phi(k_1,\ell+1)B(\ell)u(\ell)$$

is surjective. This map is the reachability map and since it is linear, we can invoke the matrix representation theorem to note that it has a matrix representation L_r given by

$$L_r(k_0,k_1)=\begin{bmatrix}B(k_1-1) & \Phi(k_1,k_1-1)B(k_1-2) & \cdots & \Phi(k_1,k_0+1)B(k_0)\end{bmatrix}$$

Unlike the CT case,

$$\Phi(k_1,k_0)=A(k_1-1)A(k_1-2)\cdots A(k_0)$$

has an inverse $\Phi(k_0,k_1)=(\Phi(k_1,k_0))^{-1}$ iff $\det(A(k))\neq 0$ for all $k\in[k_0,k_1-1]$. Because of this fact, we have the following result.

Theorem 21. The following implications hold:

$$\begin{aligned} &(A(\cdot), B(\cdot)) \text{ is controllable on } [k_0, k_1] \\ \iff &\forall x_1 \in \mathbb{C}^n,\ \exists u_{[k_0,k_1-1]} \text{ that steers } (0, k_0) \text{ to } (x_1, k_1) \text{ (reachable)} \\ \implies &\forall x_0 \in \mathbb{C}^n,\ \exists u_{[k_0,k_1-1]} \text{ that steers } (x_0, k_0) \text{ to } (0, k_1) \text{ (controllable to zero)} \end{aligned}$$

where the last statement is actually an equivalence if $\det(A(k)) \neq 0$ for all $k \in [k_0, k_1 - 1]$.

Example 22. Show that the constant pair (A, b) with

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix},\ b = e_2$$

is not controllable on $[0, 3]$ yet every state x_0 is driven to zero at $k = 3$. Indeed, this matrix A is nilpotent with $k = 3$ and hence with the zero control input every state goes to zero. On the other hand there are clearly states x_1 which are not reachable by any control.

AA/EE547: Wi 2020

Lecture 7: LTI Controllability and Observability

Lecturer: L.J. Ratliff

Disclaimer: *These notes have not been subjected to the usual scrutiny reserved for formal publications, meaning you should take your own notes in class and review the provided references as opposed to taking these notes as your sole resource. I provide the lecture notes to you as a courtesy; it is not required that I do this. They may be distributed outside this class only with the permission of the Instructor.*

References. Chapter 11 and 15 **[JH]**; Chapter 8/8d, **[C&D]**.

The goal today is to reduce the more complicated and abstract conditions for controllability/observability of LTV systems to the LTI case, and generate several tests for controllability and observability.

1 The Basics for Controllability and Observability of LTI Systems

Consider an LTI system defined by

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

with $x \in \mathbb{R}^n$.

Recall from Lecture 5/6 that

$$\begin{aligned} W_r(t_0, t_1) &= \int_{t_0}^{t_1} \Phi(t_1, \tau) B(\tau) B^*(\tau) \Phi^*(t_1, \tau)\, d\tau \\ &= \int_{t_0}^{t_1} e^{A(t_1-\tau)} B B^* e^{A^*(t_1-\tau)}\, d\tau \\ &= \int_0^{t_1-t_0} e^{At} B B^* e^{A^* t}\, dt \end{aligned}$$

Similarly,

$$W_c(t_0, t_1) = \int_0^{t_1-t_0} e^{-At} B B^* e^{-A^* t}\, dt$$

The so-called controllability matrix is given by

$$\mathcal{C} = \begin{bmatrix} B & AB & A^2B & \cdots & A^{n-1}B \end{bmatrix} \in \mathbb{C}^{n \times nm}$$

Analogously, the observability matrix is given by

$$\mathcal{O} = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} \in \mathbb{C}^{np \times n}$$

We can derive “tests” based on these two matrices in order to check observability and controllability properties of the above LTI system.

1.1 Observability and Controllability Tests for LTI

Theorem 1. The following are equivalent:

$$\begin{aligned} & \text{The LTI system is CC on some } [0, \Delta] && (1) \\ \iff & \operatorname{rank}\left(\begin{bmatrix} B & AB & \cdots & A^{n-1}B \end{bmatrix}\right) = n && (2) \\ \iff & \operatorname{rank}\left(\begin{bmatrix} sI - A & B \end{bmatrix}\right) = n, \quad \forall\, s \in \mathbb{C} && (3) \end{aligned}$$

Proof. [(1)$\Longrightarrow$ (2)]. We know that if a system is CC then the Gramian W_r is positive definite—indeed, by its construction its positive semi-definite and if it were to actually be zero at for some vector x (i.e. $x^T W_r x = 0$) then this means it drops rank which can be true if $\operatorname{rank}(W_r) = n$:

$$\begin{aligned} W_r[t_0, t_1] &= \int_{t_0}^{t_1} e^{A(t_1 - \tau)} B B^* e^{A^*(t_1 - \tau)}\, d\tau \\ &= \int_0^{t_1 - t_0} e^{A\tau} B B^* e^{A^*\tau}\, d\tau \\ &= \int_0^{\Delta} e^{A\tau} B B^* e^{A^*\tau}\, d\tau > 0 \end{aligned}$$

Now, suppose (2) is false. That is, $\exists\, v \in \mathbb{R}^n$ such that

$$\begin{aligned} & v^\top \begin{bmatrix} B & AB & \cdots & A^{n-1}B \end{bmatrix} = 0_{nm}^\top \\ \Longrightarrow\ & v^\top B = 0_m^\top,\ v^T AB = 0_m^\top, \ldots, v^\top A^{n-1} B = 0_m^\top \\ \Longrightarrow\ & v^\top f(A) B = 0_m^\top \text{ (by Cayley Hamilton)} \\ \Longrightarrow\ & v^\top e^{At} B = 0_m^\top \end{aligned}$$

Hence

$$v^\top W_r v = 0$$

which contradicts the positive definiteness of W_r.
Can you think of an alternative proof—e.g., by contrapositive?

[(2) $\Longrightarrow$ (1)]. Assume (2) and suppose that (1) is false. Then $\exists\, v \neq 0$ such that

$$\begin{aligned} & v^T \left(\int_0^{\Delta} e^{A\tau} B B^* e^{A^*\tau}\, d\tau \right) v = 0 \\ \Longrightarrow\ & \int_0^{\Delta} \| B^* e^{A^*\tau} v \|^2\, d\tau = 0 \\ \Longrightarrow\ & B^* e^{A^*\tau} v \equiv 0, \quad \forall\, \tau \in (0, \Delta) \end{aligned}$$

That is,

$$\begin{aligned} B^* v &= 0, && \text{at } t = 0 \\ B^* A^* v &= 0, && \text{derivative at } t = 0 \\ &\ \vdots && \vdots \\ B^* (A^{n-1})^* v &= 0, && n-1\text{–th derivative at } t = 0 \end{aligned}$$

Thus,

$$v^\top \begin{bmatrix} B & AB & \cdots & A^{n-1}B \end{bmatrix} = 0_{nm}^\top$$

which contradicts (2).

$[(2) \implies (3)]$. Suppose (2) holds and that (3) is false. Then, $\exists\ \lambda \in \sigma(A)$ such that

$$v^\top \begin{bmatrix} \lambda I - A & B \end{bmatrix} = 0_{n+m}^\top$$

That is,

$$\lambda v^\top = v^\top A \ \text{ and } \ v^\top B = 0_m^\top$$

Hence,

$$\begin{aligned} v^\top AB &= \lambda v^\top B = 0_m^\top \\ v^\top A^2 B &= \lambda v^\top AB = 0_m^\top \\ &\vdots \\ v^\top A^{n-1} B &= \lambda v^\top A^{n-1} B = 0_m^\top \end{aligned}$$

so that

$$v^\top \begin{bmatrix} B & AB & \cdots & A^{n-1}B \end{bmatrix} = 0_{nm}^\top$$

contradicting (2).

$[(3) \implies (2)]$. Suppose (3) holds and (2) does not. Consider $\mathcal{R}(\mathcal{C})$. Since (2) does not hold, $\mathcal{R}(\mathcal{C}) \subsetneq \mathbb{R}^n$. Note that $\mathcal{R}(\mathcal{C})$ is an A–invariant subspace containing $\mathcal{R}(B)$. Let V_1 be any subspace of $\mathbb{R}^n$ such that

$$\mathcal{R}(\mathcal{C}) \oplus V_1 = \mathbb{R}^n$$

Then, by the second representation theorem, there exists a representation of A, B w.r.t. $\mathcal{R}(\mathcal{C}), V_1$ given by

$$\tilde{A} = \begin{bmatrix} \tilde{A}_{11} & \tilde{A}_{12} \\ 0 & \tilde{A}_{22} \end{bmatrix}, \quad \tilde{B} = \begin{bmatrix} \tilde{B}_1 \\ 0 \end{bmatrix} \tag{1}$$

Thus, $\exists\ T \in \mathbb{R}^{n\times n}$ such that

$$T^{-1}AT = \tilde{A} \ \text{ and } \ T^{-1}B = \tilde{B}$$

Now

$$\operatorname{rank} \begin{bmatrix} sI - A & B \end{bmatrix} = \operatorname{rank} \begin{bmatrix} sI - \tilde{A} & \tilde{B} \end{bmatrix}$$

since

$$\begin{bmatrix} sI - \tilde{A} & \tilde{B} \end{bmatrix} = T^{-1} \begin{bmatrix} sI - A & B \end{bmatrix} \begin{bmatrix} T & 0 \\ 0 & I \end{bmatrix}$$

Further,

$$\operatorname{rank} \begin{bmatrix} sI - \tilde{A} & \tilde{B} \end{bmatrix} = \operatorname{rank} \begin{bmatrix} sI - \tilde{A}_{11} & -\tilde{A}_{12} & \tilde{B}_1 \\ 0 & sI - \tilde{A}_{22} & 0 \end{bmatrix}$$

But this has rank less than n for all $s \in \sigma(\tilde{A}_{22})$, contradicting (3). $\square$

Example 2. The equations of motion of a satellite, linearized around a steady-state solution, are given by $\dot{x} = Ax + Bu$, where x_1 and x_2 denote the perturbations in the radius and the radial velocity, respectively, x_3 and x_4 denote the perturbations in the angle and the angular velocity, and

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 3\omega^2 & 0 & 0 & 2\omega \\ 0 & 0 & 0 & 1 \\ 0 & -2\omega & 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}$$

The input is a radial thruster u_1 combined with a tangential thruster u_2.

a. *is the system controllable?* Yes, this is easy to check by just computing

$$\mathrm{rank}(\mathcal{C}) = \mathrm{rank}\left(\begin{bmatrix} B & AB \end{bmatrix}\right) = \mathrm{rank}\left(\begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 2 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & -2 & 1 \end{bmatrix}\right) = 4$$

That is we did not need to consider higher powers of A.

b. *What is the thrusters individually fail*?

If u_2 fails we have

$$\mathrm{rank}(\mathcal{C}) = \mathrm{rank}\left(\begin{bmatrix} B_2 & AB_2 & A^2B_2 \end{bmatrix}\right) = \mathrm{rank}\left(\begin{bmatrix} 0 & 1 & 0 & -1 \\ 1 & 0 & -1 & -4 \\ 0 & 0 & -2 & -2 \\ 0 & -2 & -2 & 0 \end{bmatrix}\right) = 3$$

so its not controllable.

If u_1 fails we have

$$\mathrm{rank}(\mathcal{C}) = \mathrm{rank}\left(\begin{bmatrix} B_2 & AB_2 & A^2B_2 \end{bmatrix}\right) = \mathrm{rank}\left(\begin{bmatrix} 0 & 0 & 2 & 2 \\ 0 & 2 & 2 & 0 \\ 0 & 1 & 1 & -3 \\ 1 & 1 & -3 & -7 \end{bmatrix}\right) = 4$$

Fact. Some facts:

1. Since the controllability matrix does not depend on time, if the LTI system is CC for *some* $\Delta > 0$ then it is CC for *all* $\Delta > 0$. Because of this fact, we often say that the pair (A, B) is controllable.
2. Controllability test can be done by just examining A and B without computing the grammian. The matrix-rank test is attractive in that it enumerates the vectors in the controllability subspace. However, numerically, since it involves powers of A, numerical stability needs to be considered.
3. The PBH test involves simply checking the condition at the eigenvalues. It is because for $(sI - A, B)$ to have rank less than n, s must be an eigenvalue.
4. The range space of the controllability matrix is of special interests. It is called the controllable subspace and is the set of all states that can be reached from zero-initial condition. This is A–invariant.
5. Using the basis for the controllable subspace as part of the basis for $\mathbb{R}_n$, the controllability property can be easily seen in the transformed representation in (1).

1.2 LTI Observability

The dual of the controllability theorem for LTI gives a similar theorem for observability.

Theorem 1.1 The following are equivalent:

$$\text{The LTI system is CO on some } [0, \Delta] \tag{1}$$

$$\iff \mathrm{rank}\left(\begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix}\right) = n \tag{2}$$

$$\iff \mathrm{rank}\left(\begin{bmatrix} sI - A \\ C \end{bmatrix}\right) = n, \quad \forall\, s \in \mathbb{C} \tag{3}$$

The proof is very similar to the one for controllability. (see C& D for details and alternate proof for controllability using observability proof). This being said, it is perhaps instructive to see the proof sketch for the equivalence of (3) with (2).

Proof. Instead of considering the range space of the controllability matrix, we consider the NULL space of the observability matrix. Its null space is also A-invariant. Hence if the observability matrix is not full rank, then using basis for its null space as the last k basis vectors of $\mathbb{R}^n$, the system can be represented as

$$\begin{aligned} \dot{z} &= \begin{bmatrix} \tilde{A}_{11} & 0 \\ \tilde{A}_{21} & \tilde{A}_{22} \end{bmatrix} z + \begin{bmatrix} \tilde{B}_1 \\ \tilde{B}_2 \end{bmatrix} u \\ y &= \begin{bmatrix} \tilde{C} & 0 \end{bmatrix} z \end{aligned}$$

where

$$C = \begin{bmatrix} \tilde{C} & 0 \end{bmatrix} T^{-1}, \quad \text{and} \quad A = T \begin{bmatrix} \tilde{A}_{11} & 0 \\ \tilde{A}_{21} & \tilde{A}_{22} \end{bmatrix} T$$

and the dim of $\tilde{A}_{22}$ is non-zero (and is the dim of the null space of the observability matrix). □

Fact. Some facts:

1. Observability of a LTI system does not depend on the time interval. So, theoretically speaking, if observing the output and input for an arbitrary amount of time will be sufficient to figure out x_0. In reality, when more data is available, one can do more averaging to eliminate effects of noise (e.g. using the Least squares Kalman Filter approach).
2. The subspace of particular interest is the null space of the controllability matrix. An initial state lying in this set will generate identically 0 zero-input response. This subspace is called the unobservable subspace.
3. Using the basis of the unobservable subspace as part of the basis of $\mathbb{R}_n$, the observability property can be easily seen.

Remark. An easy extension to the proofs of the above theorems is that

$$\begin{aligned} \mathcal{R}(W_c[0,\Delta]) &= \mathcal{R}(\mathcal{L}_c) &= \mathcal{R}(\mathcal{C}) \subset \mathbb{R}^n \\ \mathcal{N}(W_o[0,\Delta]) &= \mathcal{N}(\mathcal{L}_o) &= \mathcal{N}(\mathcal{O}) \subset \mathbb{R}^n \end{aligned}$$

2 Lyapunov Tests for Controllability/Observability

Recall for LTV systems the following facts.

Fact. The function $t_1 \mapsto W_r(t_0, t_1)$ is teh solution of the linear matrix equation

$$\dot{X}(t) = A(t)X(t) + X(t)A^*(t) + B(t)B^*(t)$$

with $X(t_0) = 0$.

Fact. The map $t_0 \mapsto W_o(t_0, t_1)$ is the solution to the linear matrix equation

$$\dot{X}(t) = -A^*(t)X(t) - X(t)A(t) - C^*(t)C(t)$$

with $X(t_1) = 0$.

For LTI, this gives rise to the Lyapunov tests for controllability (resp. observability). The utility of this is in the synthesis of feedback controllers that stablilize the system. You have a homework on this.

Consider

$$\dot{x} = Ax + Bu, \quad (x^+ = Ax + Bu, \text{ DT})$$

Proposition 2.1 Assume that A is a stability matrix (i.e. $\sigma(A) \subset \mathbb{C}^\circ_-$). The LTI system is controllable if and only if there is a unique positive-definite solution W to the following Lyapunov equation

$$AW + WA^\top = -BB^\top, \quad (AWA^\top - W = -BB^\top, \text{ DT})$$

Moreover, the unique solution is

$$W = \int_0^\infty e^{A\tau} BB^\top e^{A^\top \tau}\, d\tau$$

resp. in DT case

$$W = \sum_{t=0}^{\infty} A^\tau BB^\top (A^\top)^\tau$$

3 Stabilizing via Feedback

Proposition 3.1 For real A, B, for any monoic real polynomial π of degree n, there exists $F \in \mathbb{R}^{m\times n}$ such that

$$\chi_{A+BF} = \pi$$

if and only if the pair (A, B) is controllable.

The interpretation of the above proposition is through the idea of *constant state feedback*. Suppose that the state variables are available, (from, say, measurements), that we calculate Fx for a given $F \in \mathbb{R}^{m\times n}$ and that we feedback Fx to the input: the resulting feedback system is, with $u = 0$,

$$x = (A + BF)x$$

Thus the proposition asserts that the pair (A, B) is controllable $\iff$ we can always choose F so that the closed-loop characteristic polynomial χ_{A+BF} has as roots a list of n preassigned points in $\mathcal{C}$; of course, these n points must be located symmetrically with respect to the real axis because the polynomial χ_{A+BF} has real coefficients.

Key Take Away: This is to say that given any unstable A with (A, B) controllable we can always stabilize it by constant state feedback.

3.1 Controllable Canonical Form

Consider a SISO LTI system that is CC. We claim that there exists a similarity transformation that converts the system to the form

$$\dot{x} = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & 0 & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & -a_{n-3} & \cdots & -a_1 \end{bmatrix} x + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} u$$

Indeed, since the system is CC, $\mathcal{C}$ has rank n and is invertible so that

$$\mathcal{C}^{-1} = \begin{bmatrix} \tilde{\mathcal{C}} \\ q \end{bmatrix}$$

where q is the last row of the matrix inverse. That is,

$$\mathcal{C}^{-1} = \begin{bmatrix} \tilde{\mathcal{C}} \\ q \end{bmatrix} \begin{bmatrix} b & Ab & \cdots & A^{n-1}b \end{bmatrix} = I$$

so that

$$qA^{i-1}b = \begin{cases} 0, & i = 1, \ldots, n-1 \\ 1, & i = n \end{cases}$$

Then by Cayley-Hamilton we have that

$$\begin{bmatrix} qb \\ qAb \\ \vdots \\ qA^{n-1}b \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} = \bar{B}$$

and

$$\begin{aligned}
\begin{bmatrix} qA \\ qA^2 \\ \vdots \\ qA^{n-1} \\ qA^n \end{bmatrix} &= \begin{bmatrix} qA \\ \vdots \\ qA^{n-1} \\ -q\sum_{i=1}^{n} a_{n-i+1}A^{i-1} \end{bmatrix} \\
&= \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & 0 & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & -a_{n-3} & \cdots & -a_1 \end{bmatrix} \begin{bmatrix} q \\ qA \\ \vdots \\ qA^{n-2} \\ qA^{n-1} \end{bmatrix} \\
&= \bar{A}T
\end{aligned}$$

where

$$T = \begin{bmatrix} q \\ qA \\ \vdots \\ qA^{n-1} \end{bmatrix}$$

Hence, $A = T^{-1}AT$ and $b = T^{-1}\bar{B}$.

Note. Any system that can be placed in controllable canonical form can be stabilized by state feedback.

Example. Consider

$$\dot{x} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -\alpha_3 & -\alpha_2 & -\alpha_1 \end{bmatrix} x + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} u$$

With state feedback the input is $u = -Kx$ where K is a constant row vector. Consider the polynomial $p(a) = \sum_{k=0}^{3} a_k s^{3-k} = a_0 s^3 + a_1 s^2 + a_2 s + a_3$ with $a_0 = 1$. For the closed loop feedback system with $K = \begin{bmatrix} k_1 & k_2 & k_3 \end{bmatrix}$, we have

$$A - BK = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -\alpha_3 - k_1 & -\alpha_2 - k_2 & -\alpha_1 - k_3 \end{bmatrix}$$

Then, $\chi(s) = s^3+(\alpha_1+k_3)s^2+(\alpha_2+k_2)s+\alpha_3+k_1$. Equating the coefficients with $p(s)$ we get $a_1 = \alpha_1+k_3$, $a_2 = \alpha_2 + k_2$, and $a_3 = \alpha_3 + k_1$. So, $K = [a_3 - \alpha_3 \;\; a_2 - \alpha_2 \;\; a_1 - \alpha_1]^T$.

Example. Consider

$$\dot{x} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} x + \begin{bmatrix} 1 \\ 1 \end{bmatrix} u$$

and the desired characteristic polynomial $p(s) = (s+1)(s+3)$. First,

$$\operatorname{rank} \mathcal{C} = \operatorname{rank}\left(\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}\right) = 2$$

Then for $u = -kx$,

$$\begin{aligned}
\det(sI - A + bk) &= \det\left(\begin{bmatrix} s & 0 \\ 0 & s \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix}\begin{bmatrix} k_1 & k_2 \end{bmatrix}\right) \\
&= \det\left(\begin{bmatrix} s-1+k_1 & k_2 \\ k_1 & s-2+k_2 \end{bmatrix}\right) \\
&= (s-1+k_1)(s-2+k_2) - k_2k_1 \\
&= (s-1)(s-2) + k_1(s-2) + k_2(s-1) \\
&= s^2 - 3s + 2 + k_1 s - 2k_1 + k_2 s - k_2 \\
&= s^2 + (k_1+k_2-3)s + 2 - 2k_1 - k_2
\end{aligned}$$

So then by equating coefficients of the above and

$$p(s) = s^2 + 4s + 3$$

we get

$$\begin{aligned}
4 &= k_1 + k_2 - 3 \implies 7 - k_2 = k_1 \\
3 &= 2 - 2k_1 - k_2 \implies 1 = -2k_1 - k_2 \implies 1 = -2(7-k_2) - k_2 = -14 + k_2
\end{aligned}$$

so that

$$k_1 = -8 \text{ and } k_2 = 15$$

and the closed loop system is thus

$$\dot{x} = (A - BK)x = \begin{bmatrix} 9 & 15 \\ 8 & 17 \end{bmatrix} x$$

AA/EE547:	Wi 2020

Lecture 6: LTV Controllability/Observability: Gramians and DT vs. CT

Lecturer: L.J. Ratliff

Disclaimer: *These notes have not been subjected to the usual scrutiny reserved for formal publications, meaning you should take your own notes in class and review the provided references as opposed to taking these notes as your sole resource. I provide the lecture notes to you as a courtesy; it is not required that I do this. They may be distributed outside this class only with the permission of the Instructor.*

References. Chapter 11 and 15 **[JH]**; Chapter 8/8d, **[C&D]**.

1 Gramians

We can characterize the controllability and observability in terms of operators.

Remark. This is where the Finite Rank Operator Lemma (FROL) and adjoint map definition will play a role.

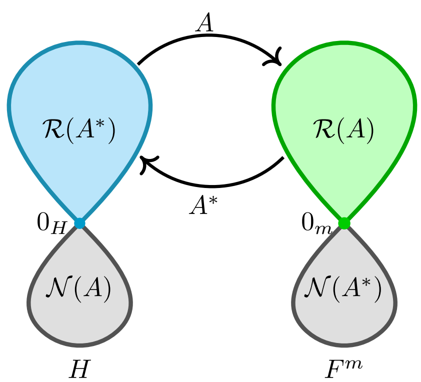


Figure 1: The orthogonal decomposition of the domain and the co-domain of a finite rank operator $A : H \to F^m$ and its associated bijections.

Recall FROL: It tells us that for a map linear map $A : \mathcal{U} \to \mathcal{V}$ that

$$\mathcal{V} = \mathcal{R}(A) \overset{\perp}{\oplus} \mathcal{N}(A^*)$$

and

$$\mathcal{U} = \mathcal{R}(A^*) \overset{\perp}{\oplus} \mathcal{N}(A)$$

Moreover,

$$\begin{aligned} \mathcal{N}(AA^*) &= \mathcal{N}(A^*), \quad \mathcal{R}(AA^*) = \mathcal{R}(A) \\ \mathcal{N}(A^*A) &= \mathcal{N}(A), \quad \mathcal{R}(A^*A) = \mathcal{R}(A^*) \end{aligned}$$

and

$$AA^*|_{\mathcal{R}(A)} \to \mathcal{R}(A) \text{ and } A^*A|_{\mathcal{R}(A^*)} \to \mathcal{R}(A^*) \text{ are one-to-one and onto}$$

Claim: the adjoint of $\mathcal{L}_r : \mathcal{U}_{[t_0,t_1]} \to \mathbb{R}^n$, which is defined by

$$\mathcal{L}_r(u(\cdot)) = \int_{t_0}^{t_1} \Phi(t_1,\tau)B(\tau)u(\tau)\ d\tau,$$

is

$$\mathcal{L}_r^* x = B^*(\cdot)\Phi^*(t_1,\cdot)x$$

Proof. $\mathcal{L}_r : u_{[t_0,t_1]} \mapsto \int_{t_0}^{t_1} \Phi(t_1,\tau)B(\tau)u(\tau)d\tau$ so that

$$\langle \mathcal{L}_r u, z\rangle_{\mathbb{R}^n} = \langle u, \mathcal{L}_r^* z\rangle_{\mathcal{U}_{[t_0,t_1]}}$$

where

$$\begin{aligned}
\langle \mathcal{L}_r u, z\rangle_{\mathbb{R}^n} &= z^* \int_{t_0}^{t_1} \Phi(t_1,\tau)B(\tau)u(\tau)d\tau \\
&= \int_{t_0}^{t_1} (B^*(\tau)\Phi^*(t_1,\tau)z)^* u(\tau)d\tau \\
&= \langle u, B^*(\tau)\Phi^*(t_1,\tau)z\rangle_{\mathcal{U}_{[t_0,t_1]}}
\end{aligned}$$

so that

$$\mathcal{L}_r^* z = B^*(\cdot)\Phi^*(t_1,\cdot)z$$

□

The adjoint $\mathcal{L}_c^*$ is derived similarly:

$$\mathcal{L}_c^* z = B^*(\cdot)\Phi^*(t_0,\cdot)z$$

Note that

$$\mathcal{R}(\mathcal{L}_r) = \Phi(t_1,t_0)\mathcal{R}(\mathcal{L}_c)$$

Indeed, let $x \in \mathcal{L}_r$ then

$$x = \mathcal{L}_r u = \int_{t_0}^{t_1} \Phi(t_1,\tau)B(\tau)u(\tau)\ d\tau$$

But, $\Phi(t_1,t_0)\Phi(t_0,\tau) = \Phi(t_1,\tau)$ so that

$$x = \mathcal{L}_r u = \Phi(t_1,t_0)\int_{t_0}^{t_1} \Phi(t_0,\tau)B(\tau)u(\tau)\ d\tau$$

Since $\mathcal{L}_r\mathcal{L}_r^* : \mathbb{R}^n \to \mathbb{R}^n$ is represented by

$$\int_{t_0}^{t_1} \Phi(t_1,\tau)B(\tau)B^*(\tau)\Phi^*(t_1,\tau)\ d\tau$$

we have our definitions of the reachability and controllability Gramians.

Definition 1 (Reachability and Controllability Gramian.)**.** The reachability Gramian is given by the symmetric positive semi-definite matrix

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

The controllability Gramian is given by the symmetric positive semi-definite matrix

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

Fact. The set of reachable states on $[t_0, t_1]$ is the subspace $\mathcal{R}(L_r)$, which is equal to $\mathcal{R}(W_r[t_0, t_1])$.

Definition 2. Controllability test

$$(A(\cdot), B(\cdot)) \text{ CC } \iff \text{rank}(W_c) = n$$

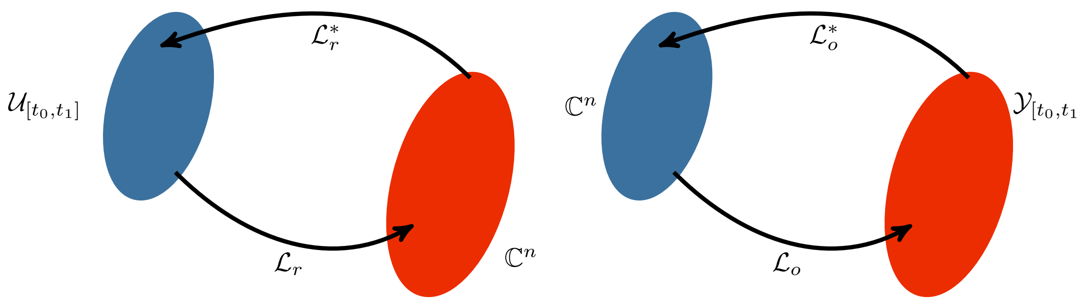


Figure 2: Graphic of Reachability and Observability Map Operation

Analogous construction for observability. The observability operator is denoted by $\mathcal{L}_o : \mathbb{R}^n \to \mathcal{Y}_{[t_0,t_1]}$ where

$$\mathcal{L}_o x_0 = C(\cdot)\Phi(\cdot, t_0)x_0$$

and

$$\mathcal{L}_o^* y(\cdot) = \int_{t_0}^{t_1} \Phi^*(\tau, t_0)C^*(\tau)y(\tau)\ d\tau$$

and

$$\mathcal{L}_o^*\mathcal{L}_o = \int_{t_0}^{t_1} \Phi(\tau, t_0)C^*(\tau)C(\tau)\Phi(\tau, t_0)\ d\tau$$

Definition 3 (Observability Gramian.)**.** The observability Gramian is the matrix

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

Definition 4 (Observability test.)**.**

$$(A(\cdot), C(\cdot)) \text{ CO } \iff \text{rank}(W_0) = n$$

1.1 Duality

Fact. Observability and controllability are dual to each other.

The dual of $R(\cdot) = [A(\cdot), B(\cdot), C(\cdot), D(\cdot)]$ is $\tilde{R}(\cdot) = [A(\cdot), -B(\cdot), -C(\cdot), D(\cdot)]$ and

$$\tilde{\mathcal{L}}_c = -\mathcal{L}_c \text{ and } \tilde{\mathcal{L}}_o = -\mathcal{L}_o$$

Theorem 1.1 (Duality: controllability to zero versus unobservability.) The subspace of all states of $R(\cdot)$ that are controllable to zero (unobservable) on $[t_0, t_1]$ is the *orthogonal complement* of the subspace of all states of its dual $\tilde{R}(\cdot)$ that are unobservable (controllable to zero, resp.) on $[t_0, t_1]$. That is,

$$\mathcal{R}(\mathcal{L}_c) = \mathcal{N}(\tilde{\mathcal{L}}_o)^\perp \text{ and } \mathcal{N}(\mathcal{L}_o) = \mathcal{R}(\tilde{\mathcal{L}}_c)^\perp$$

or equivalently

$$\mathcal{R}(W_c[t_0,t_1]) = \mathcal{N}(\tilde{W}_o(t_0,t_1))^{\perp} \;\text{ and }\; \mathcal{N}(W_o(t_0,t_1)) = \mathcal{R}(\tilde{W}_c[t_0,t_1])^{\perp}$$

2 Duality

The controllability and observability theorems we have stated so far are intimately related. Consider the system

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

Recalling our knowledge about computing adjoints, we can write the dual representation as

$$\begin{aligned} -\dot{\tilde{x}}(t) &= A^*(t)\tilde{x}(t) + C^*(t)\tilde{u}(t) \\ \tilde{y}(t) &= B^*(t)\tilde{x}(t) + D^*(t)\tilde{u}(t) \end{aligned}$$

where here $\tilde{x}(t) \in \mathbb{C}^n, \tilde{u}(t) \in \mathbb{C}^p, \tilde{y}(t) \in \mathbb{C}^m$. The state transition matrix is

$$\Psi(t,\tau) = \Phi(\tau,t)^*$$

The minus sign on the dynamics essentially captures that the dual runs in reverse time. If we take the dual of the dual, we get

$$(A(\cdot), -B(\cdot), -C(\cdot), D(\cdot))$$

so that the original system is equal to the dual of the dual modulo a sign change for the state. And thus

$$(\mathcal{L}_r^*)^* = -\mathcal{L}_r,\; (\mathcal{L}_c^*)^* = -\mathcal{L}_c,\; (\mathcal{L}_o^*)^* = -\mathcal{L}_o$$

It turns out that controllability to zero on $[t_0,t_1]$ is the dual of unobservability on $[t_0,t_1]$ and vice versa.

Theorem 5. The subspace of all states of $(A(\cdot), B(\cdot), C(\cdot), D(\cdot))$ that are controllable to zero (unobservable) on $[t_0,t_1]$ is the orthogonal complement of the subspace of all states of its dual that are unobservable (controllable to zero, resp.) on $[t_0,t_1]$. That is,

$$\mathcal{R}(\mathcal{L}_c) = \mathcal{N}(\mathcal{L}_o^*)^{\perp},\; \mathcal{N}(\mathcal{L}_o) = \mathcal{R}(\mathcal{L}_c^*)^{\perp}$$

Corollary 6. The system $(A(\cdot), B(\cdot), C(\cdot), D(\cdot))$ is controllable (observable) on $[t_0,t_1]$ iff its dual is observable (controllable, resp.) on $[t_0,t_1]$.

3 Discrete Time Controllability and Reachability

Remark. The fundamental difference between CT and DT is that DT controllability to zero does not necessarily imply reachability from zero.

Definition 7. We say that the pair $(A(\cdot), B(\cdot))$ is controllable on $[k_0,k_1]$ iff for all (x_0,k_0) and for all (x_1,k_1) there exists a control sequence $u_{[k_0,k_1-1]} = (u(k_0),\dots,u(k_1-1))$ that transfers the (x_0,t_0) to the (x_1,t_1).

Given the system $(A(\cdot), B(\cdot))$, we know that $u_{[k_0,k_1-1]}$ transfers x_0 to x_1 iff

$$x_1 = s(k_1,k_0,x_0,u_0) = \Phi(k_1,k_0)x_0 + \sum_{\ell=k_0}^{k_1-1} \Phi(k_1,\ell+1)B(\ell)u(\ell)$$

This expression, in turn, shows that there will be such an input taking arbitrary x_0 to arbitrary x_1 if and only if the linear map

$$\mathcal{L}_r(k_0, k_1) : \mathcal{U}_d(k_0, k_1 - 1) \to \mathbb{C}^n : u_{[k_0,k_1-1]} \to \sum_{\ell=k_0}^{k_1-1} \Phi(k_1, \ell + 1)B(\ell)u(\ell)$$

is surjective. This map is the reachability map and since it is linear, we can invoke the matrix representation theorem to note that it has a matrix representation L_r given by

$$L_r(k_0, k_1) = \begin{bmatrix} B(k_1 - 1) & \Phi(k_1, k_1 - 1)B(k_1 - 2) & \cdots & \Phi(k_1, k_0 + 1)B(k_0) \end{bmatrix}$$

Unlike the CT case,

$$\Phi(k_1, k_0) = A(k_1 - 1)A(k_1 - 2) \cdots A(k_0)$$

has an inverse $\Phi(k_0, k_1) = (\Phi(k_1, k_0))^{-1}$ iff $\det(A(k)) \neq 0$ for all $k \in [k_0, k_1 - 1]$. Because of this fact, we have the following result.

Theorem 8. The following implications hold:

$$\begin{aligned}
&(A(\cdot), B(\cdot)) \text{ is controllable on } [k_0, k_1] \\
&\iff \forall x_1 \in \mathbb{C}^n,\ \exists u_{[k_0,k_1-1]} \text{ that steers } (0, k_0) \text{ to } (x_1, k_1) \text{ (reachable)} \\
&\implies \forall x_0 \in \mathbb{C}^n,\ \exists u_{[k_0,k_1-1]} \text{ that steers } (x_0, k_0) \text{ to } (0, k_1) \text{ (controllable to zero)}
\end{aligned}$$

where the last statement is actually an equivalence if $\det(A(k)) \neq 0$ for all $k \in [k_0, k_1 - 1]$.

Example 9. Show that the constant pair (A, b) with

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix},\ b = e_2$$

is not controllable on $[0, 3]$ yet every state x_0 is driven to zero at $k = 3$. Indeed, this matrix A is nilpotent with $k = 3$ and hence with the zero control input every state goes to zero. On the other hand there are clearly states x_1 which are not reachable by any control.

EE548: **Winter 2018**

Lecture 6: Dynamic Programming and Discrete Time LQR

Lecturer: L.J. Ratliff

Disclaimer: *These notes have not been subjected to the usual scrutiny reserved for formal publications. They may be distributed outside this class only with the permission of the Instructor.*

Goal: Learn the basics of dynamic programming via discrete time linear quadratic regulator (LQR)

References: Stengel §3.4; Lewis §6

Remark. If you are somehow at the point not aware, there are tons of Python tutorials and references linked on canvas. Moreover, there is a virtual machine you can download with everything pre-installed. Further yet, one of the links provided maps Matlab commands to Python commands!!

6.1 Bellman's Principle

In previous lectures, we have discussed the Lyapunov equation and NLP. We will see how these two things combine to enable us to solve optimal control problems.

We will start with discrete time in order to give some intuition for dynamic programming and to better connect with the NLP framework we recently discussed.

Suppose we have a DT difference equation

$$x^+ = f(x,u),\ x \in \mathbb{R}^n,\ u \in \mathbb{R}^m$$

and we want to choose inputs over time $u : [0,t] \to \mathbb{R}^m$ to minimize some cost

$$V(x,u) = \underbrace{\ell(t,x(t))}_{\text{'final' cost}} + \sum_{k=0}^{t-1} \underbrace{L(k,x(k),u(k))}_{\text{'running' cost}}$$

History Lesson: Richard Bellman developed the key insight for 'recursion' in optimal control in 1957. That is, the optimal control at time t depends only on $x(t)$ (i.e. not on previous states).

This key insight leads naturally to working backward from final time to determine optimal *policy*.

Let $V_k^*(x(k))$ denote the lowest (i.e. optimal) cost achievable from state $x(k)$ at time k,

$$V_k^*(x(k)) = \min_{u(k)\in\mathbb{R}^m} \big(L(k,x(k),u(k)) + V_{k+1}^*(x(k+1))\big)$$

where

$$x(k+1) = f(x(k),u(k))$$

depends on $u(k)$.

This is referred to as the **Bellman Equation** and it enables us (in principal) to determine optimal control inputs by solving a sequence of NLP in backward time!

[0]These notes are partially based on those of Sam Burden (UW), Stephen Boyd (Stanford), and Claire Tomlin (UC Berkeley).

6.2 Linear Quadratic Regulator DT

Consider
$$x(t+1) = Ax(t) + Bu(t),\ x(0) = x_0$$
and
$$J(u) = x_N^T Q_f x_N + \sum_{k=0}^{N-1} \big(x(k)^T Q x(k) + u(k)^T R u(k)\big)$$
where $u = (u(0), \dots, u(N-1))$ and $Q = Q^T \geq 0$, $Q_f = Q_f^T \geq 0$, $R = R^T > 0$ are the given state cost, final state cost, and input cost matrices.

- N is the time horizon
- first term measures state deviation
- second term measures input size or actuator authority
- last term measures final state deviation
- Q, R set relative weights of state deviation and input usage
- $R > 0$ means any (non-zero) input adds cost to J

LQR Problem: find u^* that minimizes $J(u)$.

Q: how does this compare to other problems you may be familiar with?

6.2.0.1 Comparison to Least Norm

Consider the least norm input problem where you must determine the least norm input that steers x to $x_N = 0$:

- no cost attached to $x_0, \dots, x_{N-1}$
- x_N must be exactly zero

We can approximate this problem in the above framework by letting $R = I$, $Q = 0$ and $Q_f >> I$ (e.g., $Q_f = 10^8 I$).

6.2.0.2 Multi-Objective Interpretation

Common form for Q and R:
$$R = \rho I,\ Q = Q_f = C^T C$$
where $C \in \mathbb{R}^{p \times n}$ (p size of output), $\rho \in \mathbb{R}$ with $\rho > 0$. Fix $x(0) = x_0$.

The cost is then
$$J(u) = \underbrace{\sum_{k=0}^{N} \|y(k)\|^2}_{\text{output cost: } J_o(u)} + \rho \underbrace{\sum_{k=0}^{N-1} \|u(k)\|^2}_{\text{input cost: } J_i(u)}$$

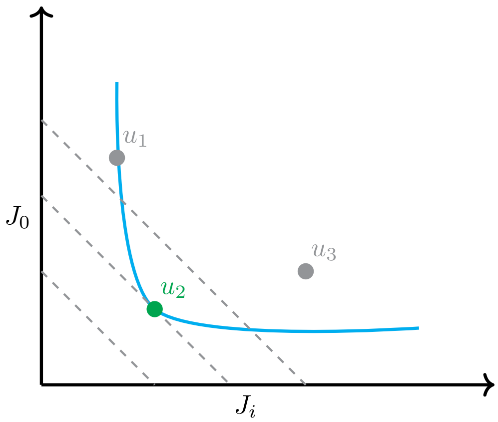
where

$$y = Cx$$

Here, $\sqrt{\rho}$ gives the relative weighting of output norm and input norm.

The input and output costs $J_i(u)$ and $J_o(u)$ respectively are competing objectives. We want both of them to be *small*.

Then the LQR cost is $J_o(u) + \rho J_i(u)$.



- Above the cyan curve shows (J_i, J_o) achieved by some u
- below the same cure shows (J_i, J_o) not achieved by any u
- Three samples inputs u_1, u_2, u_3 are shown
- u_3 is worse than u_2 on both parameters J_i, J_o
- u_1 is better than u_2 in J_i but worse in J_o

Where $J = J_o + \rho J_i =$ constant corresponds to lines with slope $-\rho$ on the (J_i, J_o) plot.

LQR optimal input is at boundary of shaded region, just touching the line of smallest possible J. Hence, u_2 in the plot is optimal for ρ. By varying ρ from 0 to $+\infty$, can sweep optimal tradeoff curve.

6.2.1 LQR as least squares

LQR can be formulated and solved as a least squares problem. Let $X = (x_0, \ldots, x_N)$ where $x_k = x(k)$. The state is a linear function of x_0 and $u = (u_0, \ldots, u_{N-1})$:

$$\begin{bmatrix} x_0 \\ \vdots \\ x_N \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & \cdots & \cdots & \cdots \\ B & 0 & \cdots & \cdots \\ AB & B & 0 & \cdots \\ \vdots & \vdots & \ddots & \vdots \\ A^{N-1}B & A^{N-2}B & \cdots & B \end{bmatrix}}_{G} \underbrace{\begin{bmatrix} u_0 \\ \vdots \\ u_{N-1} \end{bmatrix}}_{u} + \underbrace{\begin{bmatrix} I \\ A \\ \vdots \\ A^N \end{bmatrix}}_{H} x_0$$

Then we can express $X = Gu + Hx_0$ where $G \in \mathbb{R}^{Nn \times Nm}$ and $H \in \mathbb{R}^{Nn \times m}$

Thus the LQR cost is

$$J(u) = \left\| \mathrm{diag}(Q^{1/2}, \ldots, Q^{1/2}, Q_f^{1/2})(Gu + Hx_0) \right\|^2 + \left\| \mathrm{diag}(R^{1/2}, \ldots, R^{1/2} u \right\|^2$$

This is simply a big least squares problem! The solution method requires forming and solving a Least squares problem with size $N(n+m) \times Nm$. Using a naïve method (e.g., QR factorization), cost is $O(N^2 nm^2)$.

6.2.2 Dynamic Programming Solution

- givens an efficient, recursive method to solve LQR least-squares problem; cost is $O(Nn^3)$
- (but in fact, a less naive approach to solve the LQR least-squares problem will have the same complexity)
- DP is a useful and important idea on its own (it is applied in a number of domains including search, RL, MDPs, etc.)

Definition 6.1 *For $t = 0, \ldots, N$ define the* **value function** $V_t : \mathbb{R}^n \to \mathbb{R}$ *by*

$$V_t(z) = \min_{u_t, \ldots, u_{N-1}} \sum_{k=t}^{N-1} (x_k^T Q x_k + u_k^T R u_k) + x_N^T Q_f x_N$$

subject to

$$x_t = z, \ x_{k+1} = Ax_k + Bu_k, \ k = t, \ldots, N$$

- $V_t(z)$ gives the **minimum LQR cost-to-go**, starting from state z at time t
- $V_0(x_0)$ is minimum LQR cost (from state x_0 at time 0

We will see that

- V_t is quadratic, i.e. $V_t(z) = z^T P_t z$, where $P_t = P_t^T \geq 0$ (this should look super familiar, like a Lyapunov function and hence we can tie this back to what we initially learned at the beginning of the quarter)
- P_t can be found recursively, working backward from $t = N$
- the LQR optimal u is easily expressed in terms of P_t

The **cost to go** with no time left is just the final state cost

$$V_N(z) = z^T Q_f z$$

Thus, we have that $P_N = Q_f$

6.2.2.1 Dynamic Programming Principle

Q: Suppose we know $V_{t+1}(z)$. What is the optimal choice for u_t?

choice of u_t impacts current cost incurred (through $u_t^T R u_t$) and where we end up, i.e. x_{t+1} (hene, the min-cost-to-go from x_{t+1}

Definition 6.2 *The dynamic programming principle is*

$$V_t(z) = \min_w (z^T Q z + w^T R w + V_{t+1}(Az + Bw))$$

where

- $z^T Q z + w^T R w$ *is cost incurred at time* t *if* $u_t = w$
- $V_{t+1}(Az + Bu)$ *is min cost-to-go from where you land at* $t+1$

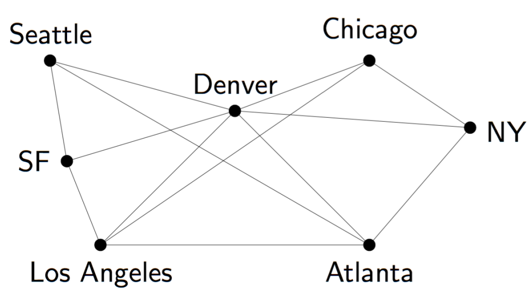
This follows from the fact that we can minimize in any order:

$$\min_{w_1,\ldots,w_k} f(w_1,\ldots,w_k) = \min_{w_1} \underbrace{\left(\min_{w_2,\ldots,w_k} f(w_1,\ldots,w_k) \right)}_{\text{a fn of } w_1}$$

in words: min cost-to-go from where you are = min over (current cost incurred + min cost-to-go from where you land)

Example: path optimization

- edges show possible flights and each has some cost (weight)
- e.g., want to find min cost route or path from Seattle to Atlanta



Q: what is the DP solution in this context?

- $V(i)$ is min cost from airport i to ATL, over all possible paths
- to find min cost from city i to ATL: minimize sum of flight cost plus min cost to ATL from where you land, over all flights out of city i (gives optimal flight out of city i on way to ATL)
- if we can find $V(i)$ for each i, we can find min cost path from any city to ATL
- **DP principle**: $V(i) = \min_j (c_{ji} + V(j))$, where c_{ji} is cost of flight from i to j, and minimum is over all possible flights out of i

6.2.2.2 HJ equation for LQR

$$V_t(z) = z^T Q z + \min_w (w^T R w + V_{t+1}(Az + Bw))$$

- called DP, Bellman, or Hamilton-Jacobi equation
- gives V_t recursively in terms of V_{t+1}
- any minimizing w gives optimal u_t:

$$u_t^* = \arg\min_w (w^T R w + V_{t+1}(Az + Bw))$$

Let us assume that $V_{t+1}(z) = z^T P_{t+1} z$, with $P_{t+1} = P_{t+1}^T \geq 0$. We will show that V_t has the same form!

By DP,

$$V_t(z) = z^T Q z + \min_w \left(w^T R w + (Az + Bw)^T P_{t+1} (Az + Bw) \right)$$

Here comes the NLP! we can solve by setting the derivative wrt w to zero:

$$2w^T R + 2(Az + Bw)^T P_{t+1} B = 0$$

Hence the optimal input is

$$w^* = -(R + B^T P_{t+1} B)^{-1} B^T P_{t+1} A z$$

and after a bunch of algebra (check your self!) we get

$$\begin{aligned} V_t(z) &= z^T Q z + (w^*)^T R w^* + (Az + Bw^*)^T P_{t+1} (Az + Bw^*) \\ &= z^T (Q + A^T P_{t+1} A - A^T P_{t+1} B (R + B^T P_{t+1} B)^{-1} B^T P_{t+1} A) z \\ &= z_t^P z \end{aligned}$$

where

$$P_t = Q + A^T P_{t+1} A - A^T P_{t+1} B (R + B^T P_{t+1} B)^{-1} B^T P_{t+1} A$$

This should look very familiar (discrete time Lyapunov?). Hence, it is easy to show that $P_t^T = P_t \geq 0$ (do this your self!)

6.2.3 Summary of LQR via DP

step 1 set $P_N = Q_f$

step 2 for $t = N, \ldots, 1$, do

$$P_{t-1} = Q + A^T P_t A - A^T P_t B (R + B^T P_{t+1} B)^{-1} B^T P_t A$$

step 3 for $t = 0, \ldots, N-1$, define

$$K_t = -(R + B^T P_{t+1} B)^{-1} B^T P_{t+1} A$$

step 4 for $t = 0, \ldots, N-1$, optima u is given by

$$u_t^* = K_t x_t$$

What else do we notice?

- optimal u is a linear function of the state (called linear state feedback)
- recursion for min cost-to-go runs backward in time

AA/EE547: Wi 2020

Lecture 11: Linear Quadratic Regulator: Continuous Time

Lecturer: L.J. Ratliff

Disclaimer: *These notes have not been subjected to the usual scrutiny reserved for formal publications, meaning you should take your own notes in class and review the provided references as opposed to taking these notes as your sole resource. I provide the lecture notes to you as a courtesy; it is not required that I do this. They may be distributed outside this class only with the permission of the Instructor.*

References. Chapter 10/20/21 **[JH]**

Note: these lecture notes follow the lectures of S. Boyd at Stanford pretty closely.

1 CT LQR Problem

Consider the CT system

$$\dot{x} = Ax + Bu, \; x(0) = x_0$$

The problem is to choose $u : [0, T] \to \mathbb{R}^m$ so as to minimize

$$J = \int_0^T (x(\tau)^T Q x(\tau) + u(\tau)^T R u(\tau)) \, d\tau + x(T)^T Q_f x(T)$$

We have a similar set up:

- T is horizon
- $Q = Q^T \geq 0$, $Q_f = Q_f^T \geq 0$, $R = R^T > 0$ are the state cost, final state cost and input cost matrices

This is now an *infinite dimensional* problem since $u : [0, T] \to \mathbb{R}^m$ is the variable.

2 DP solution

As in the DT case, we define a **value function** $V_t : \mathbb{R}^n \to \mathbb{R}$ by

$$V_t(z) = \min_u \int_t^T \left(x(\tau)^T Q x(\tau) + u(\tau)^T R u(\tau) \right) \, d\tau + x(T)^T Q_f x(T)$$

subject to $x(t) = z$, $\dot{x} = Ax + Bu$

Some notes:

- The minimization is taken over all possible signals $u : [t, T] \to \mathbb{R}^m$.
- $V_t(z)$ gives the minimum LQR cost-to-go, starting from state z at time t
- $V_T(z) = z^T Q_f z^T$

Note. V_t is quadratic, i.e., $V_t(z) = z^T P_t z$, where $P_t = P_t^T \geq 0$

Moreover, like the DT case: P_t can be found from a differential equation running backward in time from $t = T$. And, the LQR optimalu is easily expressed in terms of P_t.

AA/EE547: Wi 2020

Lecture 10: Linear Quadratic Regulator: Discrete Time

Lecturer: L.J. Ratliff

Disclaimer: *These notes have not been subjected to the usual scrutiny reserved for formal publications, meaning you should take your own notes in class and review the provided references as opposed to taking these notes as your sole resource. I provide the lecture notes to you as a courtesy; it is not required that I do this. They may be distributed outside this class only with the permission of the Instructor.*

References. Chapter 10/20/21 **[JH]**

Note: these lecture notes follow the lectures of S. Boyd at Stanford pretty closely.

1 Overview

Last time we started to discuss the LQR problem and feedback invariants. We started in CT but we will switch to DT because I think its more intuitive from an optimization perspective. Then we will come back to CT.

Goal: Learn the basics of dynamic programming via discrete time linear quadratic regulator (LQR)

In previous lectures, we have discussed the Lyapunov equation. We will see how the Lyapunov equation can be combined with non-linear programming (optimization with a nonlinear cost) to solve optimal control problems.

An optimal control problem is the problem of finding a control input to a system that minimizes a given objective which typically is a cost on the state and control such as minimum energy control, state tracking, minimum energy output.

We will start with discrete time in order to give some intuition for dynamic programming.

2 Bellman's Principle

Suppose we have a DT difference equation

$$x^+ = f(x,u),\ x\in\mathbb{R}^n,\ u\in\mathbb{R}^m$$

and we want to choose inputs over time $u:[0,t]\to\mathbb{R}^m$ to minimize some cost

$$V(x,u) = \underbrace{\ell(t,x(t))}_{\text{'final' cost}} + \sum_{k=0}^{t-1}\underbrace{L(k,x(k),u(k))}_{\text{'running' cost}}$$

History Lesson: Richard Bellman developed the key insight for 'recursion' in optimal control in 1957. That is, the optimal control at time t depends only on $x(t)$ (i.e. not on previous states).

This key insight leads naturally to working backward from final time to determine optimal *policy*.

Let $V_t(x(t)) = \ell(t, x(t))$, i.e., the final cost. Let $V_k^*(x(k))$ denote the lowest (i.e. optimal) cost achievable from state $x(k)$ at time k,

$$V_k^*(x(k)) = \min_{u(k)\in\mathbb{R}^m} \big(L(k, x(k), u(k)) + V_{k+1}^*(x(k+1))\big)$$

where

$$x(k+1) = f(x(k), u(k))$$

depends on $u(k)$.

This is referred to as the **Bellman Equation** and it enables us (in principal) to determine optimal control inputs by solving a sequence of nonlinear programs in backward time!

3 Linear Quadratic Regulator DT

Consider

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

and

$$J(u) = x_N^\top Q_f x_N + \sum_{k=0}^{N-1} \big(x(k)^\top Q x(k) + u(k)^\top R u(k)\big)$$

where $u = (u(0), \ldots, u(N-1))$ and $Q = Q^\top \geq 0$, $Q_f = Q_f^\top \geq 0$, $R = R^\top > 0$ are the given state cost, final state cost, and input cost matrices.

- N is the time horizon
- first term measures state deviation
- second term measures input size or actuator authority
- last term measures final state deviation
- Q, R set relative weights of state deviation and input usage
- $R > 0$ means any (non-zero) input adds cost to J

LQR Problem: find u^* that minimizes $J(u)$ subject to the state dynamics.

Q: how does this compare to other problems you may be familiar with?

Comparison to Least Norm. Consider the least norm input problem where you must determine the least norm input that steers x to $x_N = 0$:

- no cost attached to $x_0, \ldots, x_{N-1}$
- x_N must be exactly zero

We can approximate this problem in the above framework by letting $R = I$, $Q = 0$ and $Q_f >> I$ (e.g., $Q_f = 10^8 I$).

Multi-Objective Interpretation. Common form for Q and R:

$$R = \rho I,\ Q = Q_f = C^\top C$$

where $C \in \mathbb{R}^{p\times n}$ (p size of output), $\rho \in \mathbb{R}$ with $\rho > 0$. Fix $x(0) = x_0$.

The cost is then

$$J(u) = \underbrace{\sum_{k=0}^{N} \|y(k)\|^2}_{\text{output cost: } J_o(u)} + \rho \underbrace{\sum_{k=0}^{N-1} \|u(k)\|^2}_{\text{input cost: } J_i(u)}$$

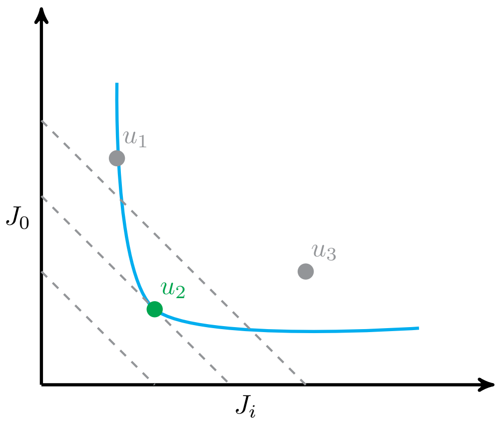
where

$$y = Cx$$

Here, $\sqrt{\rho}$ gives the relative weighting of output norm and input norm.

The input and output costs $J_i(u)$ and $J_o(u)$ respectively are competing objectives. We want both of them to be *small*.

Then the LQR cost is $J_o(u) + \rho J_i(u)$.



- Above the cyan curve shows (J_i, J_o) achieved by some u
- below the same cure shows (J_i, J_o) not achieved by any u
- Three samples inputs u_1, u_2, u_3 are shown
- u_3 is worse than u_2 on both parameters J_i, J_o
- u_1 is better than u_2 in J_i but worse in J_o

Where $J = J_o + \rho J_i =$ constant corresponds to lines with slope $-\rho$ on the (J_i, J_o) plot.

LQR optimal input is at boundary of the region above/to right of the cyan curve, just touching the line of smallest possible J. Hence, u_2 in the plot is optimal for ρ. By varying ρ from 0 to $+\infty$, can sweep optimal tradeoff curve.

3.1 LQR as least squares

LQR can be formulated and solved as a least squares problem. Let $X = (x_0, \ldots, x_N)$ where $x_k = x(k)$. The state is a linear function of x_0 and $u = (u_0, \ldots, u_{N-1})$:

$$\begin{bmatrix} x_0 \\ \vdots \\ x_N \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & \cdots & \cdots & \cdots \\ B & 0 & \cdots & \cdots \\ AB & B & 0 & \cdots \\ \vdots & \vdots & \ddots & \vdots \\ A^{N-1}B & A^{N-2}B & \cdots & B \end{bmatrix}}_{G} \underbrace{\begin{bmatrix} u_0 \\ \vdots \\ u_{N-1} \end{bmatrix}}_{u} + \underbrace{\begin{bmatrix} I \\ A \\ \vdots \\ A^N \end{bmatrix}}_{H} x_0$$

Then we can express $X = Gu + Hx_0$ where $G \in \mathbb{R}^{Nn \times Nm}$ and $H \in \mathbb{R}^{Nn \times m}$

Thus the LQR cost is

$$J(u) = \left\| \operatorname{diag}(Q^{1/2}, \ldots, Q^{1/2}, Q_f^{1/2})(Gu + Hx_0) \right\|^2 + \left\| \operatorname{diag}(R^{1/2}, \ldots, R^{1/2})u \right\|^2$$

This is simply a big least squares problem! The solution method requires forming and solving a Least squares problem with size $N(n+m) \times Nm$. Using a naïve method (e.g., QR factorization), cost is $O(N^2nm^2)$.

3.2 Dynamic Programming Solution

Given an efficient, recursive method to solve LQR least-squares problem the computational cost is $O(Nn^3)$. However, there is in fact a less naive approach to solve the LQR least-squares problem will have the same complexity: dynamic programming. DP is a useful and important idea on its own (it is applied in a number of domains including search, RL, MDPs, etc.)

Definition 1. For $t = 0, \ldots, N$ define the **value function** $V_t : \mathbb{R}^n \to \mathbb{R}$ by

$$V_t(z) = \min_{u_t,\ldots,u_{N-1}} \sum_{k=t}^{N-1} (x_k^\top Q x_k + u_k^\top R u_k) + x_N^\top Q_f x_N$$

subject to

$$x_t = z,\ x_{k+1} = Ax_k + Bu_k,\ k = t, \ldots, N$$

- $V_t(z)$ gives the **minimum LQR cost-to-go**, starting from state z at time t
- $V_0(x_0)$ is minimum LQR cost (from state x_0 at time 0

We will see that

- V_t is quadratic, i.e. $V_t(z) = z^\top P_t z$, where $P_t = P_t^\top \geq 0$ (this should look super familiar, like a Lyapunov function and hence we can tie this back to what we initially learned at the beginning of the quarter)
- P_t can be found recursively, working backward from $t = N$
- the LQR optimal u is easily expressed in terms of P_t

The **cost to go** with no time left is just the final state cost

$$V_N(z) = z^\top Q_f z$$

Thus, we have that $P_N = Q_f$

3.2.1 Dynamic Programming Principle

Q: Suppose we know $V_{t+1}(z)$. What is the optimal choice for u_t?

choice of u_t impacts current cost incurred (through $u_t^\top R u_t$) and where we end up, i.e. x_{t+1} (here, the min-cost-to-go from x_{t+1}

Definition 2. The dynamic programming principle is

$$V_t(z) = \min_w (z^\top Q z + w^\top R w + V_{t+1}(Az + Bw))$$

where

- $z^\top Q z + w^\top R w$ is cost incurred at time t if $u_t = w$
- $V_{t+1}(Az + Bu)$ is min cost-to-go from where you land at $t + 1$

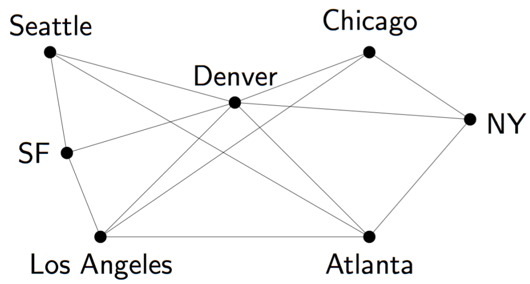
This follows from the fact that we can minimize in any order:

$$\min_{w_1,\ldots,w_k} f(w_1, \ldots, w_k) = \min_{w_1} \underbrace{\left(\min_{w_2,\ldots,w_k} f(w_1, \ldots, w_k)\right)}_{\text{a fn of } w_1}$$

in words: min cost-to-go from where you are = min over (current cost incurred + min cost-to-go from where you land)

Example: path optimization

- edges show possible flights and each has some cost (weight)
- e.g., want to find min cost route or path from Seattle to Atlanta



Q: what is the DP solution in this context?

- $V(i)$ is min cost from airport i to ATL, over all possible paths
- to find min cost from city i to ATL: minimize sum of flight cost plus min cost to ATL from where you land, over all flights out of city i (gives optimal flight out of city i on way to ATL)
- if we can find $V(i)$ for each i, we can find min cost path from any city to ATL
- **DP principle**: $V(i) = \min_j(c_{ji} + V(j))$, where c_{ji} is cost of flight from i to j, and minimum is over all possible flights out of i

Hamilton-Jacobo equation for LQR.

Theorem 3. The optimal cost-to-go and the optimal control at time t are given by

$$V_k^*(z) = z^\top P_k z,\ u_k^* = -K_k z$$

where

$$K_k = (R + B^\top P_{k+1} B)^{-1} B^\top P_{k+1} A$$

and

$$P_t = Q + A^\top P_{t+1} A - A^\top P_{t+1} B (R + B^\top P_{t+1} B)^{-1} B^\top P_{t+1} A$$

$$V_t(z) = z^\top Q z + \min_w (w^\top R w + V_{t+1}(Az + Bw))$$

- called DP, Bellman, or Hamilton-Jacobi equation
- gives V_t recursively in terms of V_{t+1}
- any minimizing w gives optimal u_t:

$$u_t^* = \arg\min_w (w^\top R w + V_{t+1}(Az + Bw))$$

Let us assume that $V_{t+1}(z) = z^\top P_{t+1} z$, with $P_{t+1} = P_{t+1}^\top \geq 0$. We will show that V_t has the same form!

By DP,

$$V_t(z) = z^\top Q z + \min_w \left(w^\top R w + (Az + Bw)^\top P_{t+1} (Az + Bw)\right)$$

We can solve by setting the derivative wrt w to zero:

$$2w^\top R + 2(Az + Bw)^\top P_{t+1} B = 0$$

Hence the optimal input is

$$w^* = -(R + B^\top P_{t+1} B)^{-1} B^\top P_{t+1} A z$$

and after a bunch of algebra (check your self!) we get

$$\begin{aligned} V_t(z) &= z^\top Q z + (w^*)^\top R w^* + (Az + Bw^*)^\top P_{t+1}(Az + Bw^*) \\ &= z^\top (Q + A^\top P_{t+1} A - A^\top P_{t+1} B(R + B^\top P_{t+1} B)^{-1} B^\top P_{t+1} A) z \\ &= z^\top P_t z \end{aligned}$$

where

$$P_t = Q + A^\top P_{t+1} A - A^\top P_{t+1} B(R + B^\top P_{t+1} B)^{-1} B^\top P_{t+1} A$$

This should look very familiar (discrete time Lyapunov?). Hence, it is easy to show that $P_t^\top = P_t \geq 0$ (do this your self!)

3.3 Summary of LQR via DP

step 1 set $P_N = Q_f$

step 2 for $t = N, \ldots, 1$, do

$$P_{t-1} = Q + A^\top P_t A - A^\top P_t B(R + B^\top P_{t+1} B)^{-1} B^\top P_t A$$

step 3 for $t = 0, \ldots, N-1$, define

$$K_t = -(R + B^\top P_{t+1} B)^{-1} B^\top P_{t+1} A$$

step 4 for $t = 0, \ldots, N-1$, optima u is given by

$$u_t^* = K_t x_t$$

What else do we notice?

- optimal u is a linear function of the state (called linear state feedback)
- recursion for min cost-to-go runs backward in time

4 Riccati and Steady-State DT LQR

Above is a summary of a procedure for how to solve the DT LQR via DP. The following are some additional comments:

- another name for the recursion for P_t is the Riccati recursion
- usually P_t rapidly converges as t decreases below N (remember we are going backward in time)
- the limit or steady-state (ss) value P_{ss} satisfies the following Riccati equation:

$$P_{ss} = Q + A^\top P_{ss} A - A^\top P_{ss} B(R + B^\top P_{ss} B)^{-1} B^\top P_{ss} A$$

  i.e. the DT algebraic Riccati equation (ARE)
- for t not close to the horizon N, LQR optimal input is approximately a linear constant state feedback

$$u_t = K_{ss} x_t, \;\; K_{ss} = -(R + B^\top P_{ss} B)^{-1} B^\top P_{ss} A$$

5 LQR via Constrained Non-Linear Programming

Matrix inversion identities:

1. $(I + \tilde{A}\tilde{B})^{-1} = I - \tilde{A}(I + \tilde{A}\tilde{B})^{-1}\tilde{B}$
2. $\tilde{A}(I + \tilde{A}\tilde{B})^{-1} = (I + \tilde{B}\tilde{A})^{-1}\tilde{B}$
3. $(I + \tilde{A}\tilde{C}^{-1}\tilde{B})^{-1} = I - \tilde{A}(\tilde{C} + \tilde{B}\tilde{A})^{-1}\tilde{B}$

We saw the following matrix inversion lemma last quarter:

Lemma 4.

$$(\tilde{A} + \tilde{B}\tilde{C})^{-1} = \tilde{A}^{-1} - \tilde{A}^{-1}\tilde{B}(I + \tilde{C}\tilde{A}^{-1}\tilde{B})^{-1}\tilde{C}\tilde{A}^{-1}$$

This allows us to convert the Riccati recursion into a different form:

$$\begin{aligned}
P_{t-1} &= Q + A^\top P_t A - A^\top P_t B(R + B^\top P_t B)^{-1}B^\top P_t A \\
&= Q + A^\top P_t(I - B(R + B^\top P_t B)^{-1}B^\top P_t)A \quad \text{[factor]} \\
&= Q + A^\top P_t(I - B\left((I + B^\top P_t B R^{-1})R\right)^{-1} B^\top P_t)A \quad \text{[pull } R \text{ out]} \\
&= Q + A^\top P_t(I - BR^{-1}(I + B^\top P_t B R^{-1})^{-1}B^\top P_t)A \quad \text{[apply inverse]} \\
&= Q + A^\top P_t(I + BR^{-1}B^\top P_t)^{-1}A \quad \text{[apply woodbury]} \\
&= Q + A^\top(I + P_t B R^{-1}B^\top)^{-1}P_t A \quad \text{[apply 2 above]}
\end{aligned}$$

which can be converted to the following symmetric form

$$P_{t-1} = Q + A^\top P_t^{1/2}(I + P_t^{1/2}BR^{-1}B^\top P_t^{1/2})^{-1}P_t^{1/2}A$$

where we applied woodbury with

$$\tilde{A} = I, \tilde{C} = B^\top P_t,\ \tilde{B} = BR^{-1}$$

5.1 Recall Some Basics of Non-Linear Programming

Recall the constrained NLP's we looked at before the midterm:

$$\begin{aligned}
&\min_x \ F(x) \\
&\text{s.t. } Gx = h
\end{aligned}$$

where $F : \mathbb{R}^n \to \mathbb{R}$ and $G \in \mathbb{R}^{m\times n}$.

The Lagrangian is given by

$$L(x, \lambda) = F(x) + \lambda^\top(h - Gx)$$

If x is optimal, then

$$D_x L = DF(x) - G^\top\lambda = 0,\ D_\lambda L = h - Gx = 0$$

Interesting connection to the LA we saw last quarter: the first equation implies that $DF(x) = G^\top\lambda$ for some λ. This is true iff $DF(x)^\top \in \mathcal{R}(G^\top)$ iff $DF(x)^\top \perp \mathcal{N}(G)$ (behold! the many applications of FROL—Finite Rank Operator Lemma).

Recall Descent Directions. Suppose x is current, feasible point (i.e. $Gx = h$). Consider a small step in the direction v, to $x + \gamma v$ (γ small, positive step size).

Q: When is $x + \gamma v$ better than x?

A: We need the following:

1. $x + \gamma v$ to be feasible:
$$G(x + \gamma v) = h + \gamma G v = h \implies Gv = 0$$
Hence, $v \in \mathcal{N}(G)$ is called a **feasible direction**.
2. $x + \gamma v$ to have a smaller objective than x:
$$F(x + \gamma v) \simeq F(x) + \gamma DF(x)v < F(x)$$
so we need $DF(x)v < 0$ (i.e. **descent direction**)
Note that if $DF(x)v > 0$, $-v$ is a descent direction, so we need only $DF(x)v \neq 0$.

x is not optimal if there exists a feasible descent direction.

if x is optimal, every feasible direction satisfies $DF(x)v = 0$. Hence,

$$\begin{aligned} Gv = 0 \implies DF(x)v = 0 \quad &\iff \quad \mathcal{N}(G) \subset \mathcal{N}(DF(x)) \\ &\iff \quad \mathcal{R}(G^\top) \supset \mathcal{R}(DF(x)^\top) \\ &\iff \quad DF(x)^\top \in \mathcal{R}(G^\top) \\ &\iff \quad DF(x)^\top = G^\top \lambda \ \text{ for some } \lambda \in \mathbb{R}^n \\ &\iff \quad DF(x)^\top \perp \mathcal{N}(G) \end{aligned}$$

5.2 LQR as Constrained Minimization Problem

Consider the following

$$\begin{aligned} \min_u \ \ & J = \frac{1}{2} \sum_{t=0}^{N-1} \left(x_t^\top Q x_t + u_t^\top R u_t \right) + \frac{1}{2} x_N^\top Q_f x_N \\ \text{s.t.} \ \ & x_{t+1} = A x_t + B u_t, \ t = 0, \ldots, N-1 \end{aligned}$$

- variables are $u_0, \ldots, u_{N-1}$ and $x_1, \ldots, x_N$ (x's can actually be specified in terms of u's via the constraint as we have seen before).
- x_0 is given
- the objective is quadratic (i.e. it is convex)

First step is to introduce Lagrange multipliers.

Q: How many do we need?

A: $\lambda_i \in \mathbb{R}^n$ for $i = 1, \ldots, N$.

Next step is to form the Lagrangian:

$$L = J + \sum_{t=0}^{N-1} \lambda_{t+1}^\top (A x_t + B u_t - x_{t+1})$$

5.2.1 Optimality Conditions

Q: What are the optimality conditions?

A: First, x_0 is given and we know that

$$x_{t+1} = Ax_t + Bu_t,\ t = 0, \ldots, N-1$$

We can take the derivative of L wrt u_t to get

$$D_{u_t}L = Ru_t + B^\top \lambda_{t+1} = 0,\ t = 0, \ldots, N-1$$

Hence,

$$u_t = -R^{-1}B^\top \lambda_{t+1}$$

Similarly, we can write first order conditions for x_t:

$$D_{x_t}L = Qx_t + A^\top \lambda_{t+1} - \lambda_t = 0,\ t = 0, \ldots, N-1$$

Hence

$$\lambda_t = A^\top \lambda_{t+1} + Qx_t$$

Finally,

$$D_{x_N}L = Q_f x_N - \lambda_N = 0$$

so that $\lambda_N = Q_f x_N$

This is a set of linear equations in the variables

$$u_0, \ldots, u_{N-1},\ x_1, \ldots, x_N,\ \lambda_1, \ldots, \lambda_N$$

5.2.2 Co-State Equations

Q: What do you notice about these equations?

A: we have a forward equation for x and a backward equation for λ...

Optimality conditions break into two parts:

$$x_{t+1} = Ax_t + Bu_t,\ x_0$$

this recursion for state x runs forward in time, with initial condition

$$\lambda_t = A^\top \lambda_{t+1} + Qx_t,\ \lambda_N = Q_f x_N$$

this recursion for **co-state** λ runs backward in time, with final condition

As mentioned before, the recursion for λ is sometimes called the **adjoint system**.

5.2.3 Solution via Riccati Recursion

We will see that $\lambda_t = P_t x_t$ where P_t is the min-cost-to-go matrix defined by the Riccati recursion we saw last time. This implies that the Riccati recursion gives a useful way to solve this set of linear equations.

Q: How do we show this?

A: via induction like last time.

First, it holds for $t = N$, since $P_N = Q_f$ and $\lambda_N = Q_f x_N$

Now, suppose it holds for $t+1$, that is, $\lambda_{t+1} = P_{t+1}x_{t+1}$. Let's show it holds for t. Using the state equation $x_{t+1} = Ax_t + Bu_t$ and $u_t = -R^{-1}B^\top \lambda_t$ from our optimality conditions, we get that

$$\lambda_{t+1} = P_{t+1}(Ax_t + Bu_t) = P_{t+1}(Ax_t - BR^{-1}B^\top \lambda_{t+1})$$

so that

$$\lambda_{t+1} = (I + P_{t+1} B R^{-1} B^\top)^{-1} P_{t+1} A x_t$$

We can then use the co-state equation $\lambda_t = A^\top \lambda_{t+1} + Q x_t$ to get that

$$\lambda_t = A^\top (I + +P_{t+1} B R^{-1} B^\top)^{-1} P_{t+1} A x_t + Q x_t = P_t x_t$$

since by the Riccati recursion we know that

$$P_t = Q + A^\top (I + P_{t+1} B R^{-1} B^\top)^{-1} P_{t+1} A.$$

This completes the induction argument so that $\lambda_t = P_t x_t$.

5.2.4 Check for consistency across the two methods

$$\begin{aligned}
u_t &= -R^{-1} B^\top \lambda_{t+1} \\
&= -R^{-1} B^\top (I + P_{t+1} B R^{-1} B^\top)^{-1} P_{t+1} A x_t \\
&= -R^{-1} (I + B^\top P_{t+1} B R^{-1})^{-1} B^\top P_{t+1} A x_t \\
&= -\left((I + B^\top P_{t+1} B R^{-1}) R\right)^{-1} B^\top P_{t+1} A x_t \\
&= -(R + B^\top P_{t+1} B)^{-1} B^\top P_{t+1} A x_t
\end{aligned}$$

2.1 DP Approach

Start with $x(t) = z$. Let us take $u(t) = w \in \mathbb{R}^m$, a constant, over the time interval $[t, t+h]$, where $h > 0$ is small.

Just as with the usual DP approach, we want to approximate the value at time t forward as the *instantaneous cost* incurred where we are plus the cost to go over the remainder of the horizon.

We will do this in an approximate manner...

First, the cost incurred over $[t, t+h]$ is

$$\int_t^{t+h} \left(x(\tau)^T Q x(\tau) + u(\tau)^T R u(\tau)\right) \, d\tau \simeq h(z^T Q + w^T R w)$$

and we end up at $x(t+h) \simeq z + h(Az + Bw)$

Second, the min-cost-to-go from where we end-up is approximately given by

$$\begin{aligned} V_{t+h}(z + h(Az + Bw)) &= (z + h(Az + Bw))^T P_{t+h}(z + h(Az + Bw)) \\ &\simeq (z + h(Az + Bw))^T (P_t + h\dot{P}_t)(z + h(Az + Bw)) \\ &\simeq z^T P_t z + h\left((Az + Bw)^T P_t z + z^T \dot{P}_t (Az + Bw) + z^T \dot{P}_t z\right) \end{aligned}$$

(of course, dropping h^2 and h.o.t's)

Adding them up, the cost incurred plus min-cost-to-go is approximately given by

$$z^T P_t z + h\left(z^T Q z + w^T R w + (Az + Bw)^T P_t z + z^T P_t (Az + Bw) + z^T \dot{P}_t z\right)$$

Now, minimize over w to get (approximately) optimal w:

$$2hw^T R + 2Hz^T P_t B = 0$$

so that

$$w^* = -R^{-1} B^T P_t z$$

(boy does that look similar!)

Thus optimal u is time-varying linear state feedback is given by

$$u^*(t) = K_t x(t), \;\; K_t = -R^{-1} B^T P_t$$

Ok, so this gives us the optimal control in terms of P_t. But, what is P_t?

2.2 (Approx.) Hamilton-Jacobi Equation

Let's substitute w^* into the (approx.) HJ equation

$$z^T P_t z \simeq z^T P_t z + h\left(z^T Q z + (w^*)^T R w^* + (Az + Bw^*)^T P_t z + z^T P_t (Az + Bw^*) + z^T \dot{P}_t z\right)$$

After you plug things in and simplify we get

$$-\dot{P}_t = A^T P_t + P_t A - P_t B R^{-1} B^T P_t Q$$

which is the Riccati differential equation for the LQR problem.

We can solve it (numerically) using the final condition $P_T = Q_f$.

Summary:

step 1 solve Riccati differential equation

$$-\dot{P}_t = A^T P_t + P_t A - P_t B R^{-1} B^T P_t Q,\ P_T = Q_f$$

backward in time

step 2 optimal u is $u^*(t) = K_t x(t)$, $K_t = -R^{-1} B^T P_t$

3 Some Additional comments

- DP method readily extends to time-varying A, B, Q, R, and tracking problem
- usually P_t rapidly converges as t decreases below T (same comment as DT)
- limit P_{ss} satisfies (cts-time) algebraic Riccati equation (ARE)

$$A^T P + PA - PBR^{-1} B^T P + Q = 0$$

  which is a quadratic matrix equation
- P_{ss} can be found by (numerically) integrating the Riccati differential equation, or by direct methods (you will do this in your project)
- for t not close to horizon T, LQR optimal input is approximately a linear, constant state feedback

$$u(t) = K_{ss} x(t),\ K_{ss} = -R^{-1} B^T P_{ss}$$

There is a similar co-state formulation for continuous time. Please see Lewis chapter 3 for this.

4 CT LQR via Lagrange

Constrained NLP:

$$\begin{aligned} \min_u \ & \frac{1}{2}\int_0^T x(\tau)^T Q x(\tau)\ d\tau \quad =: J + \frac{1}{2} x(T)^T Q_f x(T) \\ \text{s.t. } & \dot{x}(t) = Ax(t) + Bu(t),\ t \in [0, T] \end{aligned}$$

Introduce Lagrange multiplier (function) $\lambda : [0, T] \to \mathbb{R}^n$ and write

$$L = J + \int_0^T \lambda(\tau)^T (Ax(\tau) + Bu(\tau) - \dot{x}(\tau))\ d\tau$$

We can compute the derivatives (recall: you need distributions or test function to actually compute the derivatives here...)

wrt u:

$$D_{u(t)} L = Ru(t) + B\lambda^T(t) = 0 \implies u = -R^{-1} B^T \lambda(t)$$

wrt x: use the fact that (just basic fundamental theorem of calculus)

$$\int_0^T \lambda(\tau)^T \dot{x}(\tau) + \dot{\lambda}(\tau)^T x(\tau)\ d\tau = \lambda(T)^T x(T) - \lambda(0)^T x(0)$$

$$D_{x(t)} L = Qx(t) + A^T \lambda(t) + \dot{\lambda}(t) = 0 \implies \dot{\lambda}(t) = -A^T \lambda(t) - Qx(t)$$

wrt $x(T)$:

$$D_{x(T)} L = Q_f x(T) - \lambda(T) = 0 \implies \lambda(T) = Q_f x(T)$$

The above gives us optimality conditions just as in the DT case:

$$\dot{x} = Ax + Bu,\ x(0) = x_0,\quad \dot{\lambda} = -A^T\lambda - Qx,\ \lambda(T) = Q_f x(T),\ u(t) = -R^{-1} B^T \lambda(t)$$

4.1 Adjoint or Co-state system

Using $u(t) = -R^{-1}B^T\lambda(t)$, we can write a system of equations (two-point boundary value problem):

$$\frac{d}{dt}\begin{bmatrix} x(t) \\ \lambda(t) \end{bmatrix} = \underbrace{\begin{bmatrix} A & -BR^{-1}B^T \\ -Q & -A^T \end{bmatrix}}_{\text{Hamiltonian}} \begin{bmatrix} x(t) \\ \lambda(t) \end{bmatrix} \tag{1}$$

You can argue that $\lambda(t) = P_t x(t)$. That is, use the Riccati DE with final time condition $P_T = Q_f$ and the co-state equation to show that $\lambda(t) = P_t x(t)$ as above is a solution to the costate equation.

4.2 Solving the Riccati DE via Hamiltonian

Consider

$$-\dot{P} = A^T P + PA - PBR^{-1}B^T P + Q$$

and recall (1). These two differential equations are related.

The fact that $\lambda(t) = P_t x(t)$ suggests that P should have the form $\lambda(t)x(t)^{-1}$ but of course this only makes sense if they are scalars.

So what do we do? How do we generalize this intuition?

Consider the matrix version of the Hamiltonian DE:

$$\frac{d}{dt}\begin{bmatrix} X(t) \\ Y(t) \end{bmatrix} = \underbrace{\begin{bmatrix} A & -BR^{-1}B^T \\ -Q & -A^T \end{bmatrix}}_{\text{Hamiltonian}} \begin{bmatrix} X(t) \\ Y(t) \end{bmatrix} \tag{2}$$

where $X, Y \in \mathbb{R}^{n\times n}$.

Then $Z(t) = Y(t)X(t)^{-1}$ satisfies the Riccati DE:

$$-\dot{Z} = A^T Z + ZA - ZBR^{-1}B^T Z + Q$$

Check your self using the facts that

$$\frac{d}{dt}(F(t)G(t)) = \dot{F}(t)G(t) + F(t)\dot{G}(t)$$
$$\frac{d}{dt}(F(t)^{-1}) = -F(t)^{-1}\dot{F}(t)F(t)^{-1} \quad \text{we saw this last quarter!!}$$

$$\begin{aligned}
\dot{Z} = \frac{d}{dt}YX^{-1} &= \dot{Y}X^{-1} - YX^{-1}\dot{X}X^{-1} \\
&= (-QX - AY)X^{-1} - YX^{-1}(AX - BR^{-1}B^T Y)X^{-1} \\
&= -Q - A^T Z - ZA + ZBR^{-1}B^T Z
\end{aligned}$$

Hence, we can solve the Riccati DE by solving a linear matrix (Hamiltonian) DE with final conditions $X(T) = I, Y(T) = Q_f$ and forming $P(t) = Y(t)X(t)^{-1}$ after.

4.3 Solving ARE via Hamiltonian

Recall the ARE (which gives you the steady state solution for the Riccati DE and also gives you the infinite horizon LQR solution—see hw 5):

$$Q + A^T P + PA - PBR^{-1}B^T P = 0$$

with $P \geq 0$.

$$\begin{bmatrix} A & -BR^{-1}B^T \\ -Q & -A^T \end{bmatrix} \begin{bmatrix} I \\ P \end{bmatrix} = \begin{bmatrix} A - BR^{-1}B^T P \\ -Q - A^T P \end{bmatrix} = \begin{bmatrix} A + BK \\ -Q - A^T P \end{bmatrix}$$

Aside:
Note $u^* = Kx = -BR^{-1}B^T Px$ and the closed loop dynamics look like

$$\dot{x} + Ax + Bu = (A + BK)x$$

It turns out that the closed loop system is stable when (Q, A) is observable and (A, B) is controllable (that is, finding an LQR solution will stabilize the system if these two conditions are met).

Let the eigenvalues of the closed loop system be

$$\lambda_1, \ldots, \lambda_n$$

And with the assumptions above, $\text{Re}(\lambda_i) < 0$.

Coming back to our algebra:

$$\begin{bmatrix} I & 0 \\ -P & I \end{bmatrix} \begin{bmatrix} A & -BR^{-1}B^T \\ -Q & -A^T \end{bmatrix} \begin{bmatrix} I & 0 \\ P & 0 \end{bmatrix} = \begin{bmatrix} A + BK & -BR^{-1}B^T \\ 0 & -(A + BK)^T \end{bmatrix} \tag{3}$$

where the zero comes from the ARE. Note

$$\begin{bmatrix} I & 0 \\ P & 0 \end{bmatrix}^{-1} = \begin{bmatrix} I & 0 \\ -P & I \end{bmatrix}$$

Now, looking at the diagonal of RHS, the eigs of the Hamiltonian are $\lambda_1, \ldots, \lambda_n$ and $-\lambda_1, \ldots, -\lambda_n$. Hence, closed-loop eigenvalues are the eigenvalues of H with negative real part (under assumptions above).

You can show that if $A + BK$ is diagonalizable, i.e.

$$M^{-1}(A + BK)M = \Lambda = \text{diag}(\lambda_1, \ldots, \lambda_n),$$

then

$$M^T(-A - BK)^T M^{-T} = -\Lambda$$

so that

$$\begin{bmatrix} M^{-1} & 0 \\ 0 & M^T \end{bmatrix} \begin{bmatrix} A + BK & -BR^{-1}B^T \\ 0 & -(A + BK)^T \end{bmatrix} \begin{bmatrix} M & 0 \\ 0 & M^{-T} \end{bmatrix} = \begin{bmatrix} \Lambda & -M^{-1}BR^{-1}B^T M^{-T} \\ 0 & -\Lambda \end{bmatrix}$$

Combine this with (3) to get

$$H \begin{bmatrix} M \\ PM \end{bmatrix} = \begin{bmatrix} M \\ PM \end{bmatrix} \Lambda$$

so that the n columns of

$$\begin{bmatrix} M \\ PM \end{bmatrix}$$

are the eigenvectors of H associated with the stable eigenvalues $\lambda_1, \ldots, \lambda_n$.

Thus, to solve the ARE, take the n stable eigs (there will be n stable and n unstable ones) and finding the eigenvectors associated with the n stables ones. Stack them up as columns in a matrix and partition that matrix as

$$\begin{bmatrix} v_1 & \cdots v_n \end{bmatrix} = \begin{bmatrix} X \\ Y \end{bmatrix} \in \mathbb{R}^{2n\times n}$$

Then $P = YX^{-1}$ is the unique PSD solution of the ARE.