MATH4406 (Control Theory)
Part 4: State Space Description and Control of Linear Systems

Prepared by Yoni Nazarathy, Last Updated: August 30, 2012

WARNING: This is NOT the final copy for this year.

1 About

This unit (composed of 9 lecture hours) is the center of the course. It is all about linear input–state–output systems of the form

$$\begin{array}{rcl} \dot{x}(t) & = & Ax(t)+Bu(t) \\ y(t) & = & Cx(t)+Du(t) \end{array} \qquad \text{or} \qquad \begin{array}{rcl} x(n+1) & = & Ax(n)+Bu(n) \\ y(n) & = & Cx(n)+Du(n) \end{array}. \tag{1}$$

The focus is mostly on the continuous time version $\big(u(t),x(t),y(t)\big)$, yet key results for the discrete time version, $\big(u(n),x(n),y(n)\big)$, are summarised,

Here $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{p\times n}$ and $D \in \mathbb{R}^{p\times m}$. So the dimension of the state (x) is n, the dimension of the input (u) is m and the dimension of the output (y) is p. The previous two units were about $m=1$ and $p=1$ (SISO systems). Note that in such systems the state may (and typically is) multi-dimensional - yet we did not explicitly consider and discuss the state.

We refer to the systems in (1) as the continuous (A,B,C,D) system and the discrete (A,B,C,D) system respectively.

Note that *some* of the results we present naturally extend to the time dependent cases, where the matrices A,B,C and D are allowed to depend on time. Nevertheless, we mostly focus on the time-independent case as presented in (1), except for the first section which deals with general (not necessarily linear) finite dynamical systems. I.e. ordinary differntial equation (ODE) systems.

2 ODE Systems

Basic Definitions

For $x \in \mathbb{R}^n$ and $f:\mathbb{R}^{n+1}\to\mathbb{R}^n$, we call the system of equations,

$$\dot{x} = f(t,x), \tag{2}$$

a *system of n first-order ordinary differential equations.* Let $\mathcal{D} \subset \mathbb{R}^{n+1}$ be an open, nonempty and connected set such that $f(\cdot)$ is continuous in $\mathcal{D}$. We call $\mathcal{D}$ a *domain* of

(2). A *solution* of (2) in the domain $\mathcal{D}$ is, some $\phi : \mathbb{R}^n \to \mathbb{R}^n$, defined on an interval $\mathcal{J} = (a, b)$ such that,

$$\big(t, \phi_1(t), \ldots, \phi_n(t)\big) \in \mathcal{D}, \ \forall t \in \mathcal{J},$$

and $\phi(\cdot)$ is continuously differentiable on $\mathcal{J}$ with,

$$\dot{\phi}(t) = f\big(t, \phi_1(t), \ldots, \phi_n(t)\big), \ \forall t \in \mathcal{J}.$$

The system is called an *initial value problem* if it is specified as follows:

$$\dot{x} = f(t, x), \qquad x(t_0) = x_0. \tag{3}$$

A solution of (3) in the domain $\mathcal{D}$ is as defined above, a-long with the requriment that,

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

Initial value problems may be equivalently expressed by the *integral equation*:

$$\phi(t) = x_0 + \int_{t_0}^{t} f\big(s, \phi(s)\big) ds$$

Linear ODE Systems

When $f(t, x)$ in the above is of the form $A(t)x + g(t)$ the system is a *linear system*. If $g(t) \equiv 0$ then the system is called *homogeneous*, otherwise it is *nonhomogeneous*. If $A(t) \equiv A$ then the system is called *time-independent* (or sometime *autonomous*).

That is, the *linear, time-independent, homogeneous* system is,

$$\dot{x} = Ax, \quad x(0) = x_0, \tag{4}$$

with $A \in \mathbb{R}^{n \times n}$. We shall typically refer to the system (4) in short as the *autonomous system*.

We shall also focus on the system,

$$\dot{x} = Ax + g(t), \quad x(0) = x_0, \tag{5}$$

and shall typically refer to the system (5) in short as the *nonhomogeneous system*.

Higher Order ODEs

Higher order ODEs are not harder than the ODE systems presented above. This can be shown for general ODEs, yet for simplicity we focus on the linear, constant coefficient (time-independent) case. Consider *linear, autonomous, homogeneous ordinary differential equation of order n*:

$$y^{(n)}(t) + a_{n-1}y^{(n-1)}(t) + \ldots + a_1 y^{(1)}(t) + a_0 y(t) = 0. \tag{6}$$

Set now,

$$y(\cdot) = x_1(\cdot),\ \ y^{(1)}(\cdot) = x_2(\cdot),\ \ \ldots\ \ y^{(n-1)}(\cdot) = x_n(\cdot),$$

and consider the autonomous system,

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \vdots \\ \dot{x}_{n-1}(t) \\ \dot{x}_n(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & & & 0 \\ 0 & 0 & 1 & 0 & & \\ & & & \ddots & & \\ & & & & \ddots & \\ 0 & & & & & 1 \\ -a_0 & -a_1 & \cdots & \cdots & -a_{n-1} & -1 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \\ \vdots \\ \vdots \\ x_n(t) \end{bmatrix}.$$

Then it is clear that solutions of the $n'th$ dimensional system also satisfy the (6). Note that the above matrix is called a *companion matrix* associated with $(a_0, \ldots, a_{n-1})$.

Properties of Solutions: Existence and Uniqueness

Example 2.1 *Take $\dot{x} = x^{1/3}$, $x(0) = 0$, then there are at least two solutions:*

$$\phi(t) = 0, \quad \textit{and} \quad \phi(t) = \left(\frac{2}{3}\,t\right)^{3/2}.$$

Example 2.2 *Take $\dot{x} = ax$, $x(0) = 0$, then there is a unique solution,*

$$\phi(t) = e^{at}.$$

So when are we guaranteed to have unique solutions? We now present a series of general results about ODEs. We do not cover the proofs.

Theorem 2.3 (Cauchy-Peano Existance Theorem) *Let $f(\cdot)$ be continuous in a domain $\mathcal{D}$, then for any $(t_0, x_0) \in \mathcal{D}$, the initial value problem has a solution defined on $[t_0, t_0 + c)$ for some $c > 0$.*

Theorem 2.4 *Let $f(\cdot)$ be continuous in a domain $\mathcal{D}$. If for every compact set $\mathcal{K} \subset \mathcal{D}$,*

$$||(f(t,x) - f(t,y)|| \leq L_{\mathcal{K}}||x - y||,$$

for all $(t,x), (t,y) \in \mathcal{K}$ where $L_{\mathcal{K}} > 0$ is a constant (allowed to depend on $\mathcal{K}$) then the intitial value problem has at most one solution on any interval $[t_0, t_0 + c)$, $c > 0$.

Note that if $f(\cdot)$ is continuously differentiable on $\mathcal{D}$ then the above *local Lipschitz condition* is automatically satisfied.

Theorem 2.5 *Let $f(\cdot)$ be continuous in some domain $\mathcal{D} = \mathcal{J} \times \mathbb{R}^n$ and assume that $f(\cdot)$ satisfies the local Lipschitz condition. Then for any $(t_0, x_0) \in \mathcal{J} \times \mathbb{R}^n$, the initial value problem has a unique solution that exists on the entire interval $\mathcal{J}$.*

Consider now the system,

$$\dot{x} = A(t)x + g(t),$$

Theorem 2.6 *Let $A(t)$ and $g(t)$ be continuous on some open interval $\mathcal{J}$, then for any $t_0 \in \mathcal{J}$ and $x_0 \in \mathbb{R}^n$ the above system has a unique solution satisfying $x(t_0) = x_0$. The solution exists on the entire interval $\mathcal{J}$ and is continuous in (t, t_0, x_0).*

This theorem is an application of the previous theorems (and others), where the essence is to verify the Lipschitz condition. Denote, $f(t, x) = A(t)x + g(t)$ then:

$$||f(t,x) - f(t,y)||_1 = ||A(t)(x-y)||_1 \le ||A(t)||_1 ||x-y||_1 \le \Big(\sum_{i=1}^{n} \max_{1 \le j \le n} |a_{ij}(t)| \Big) ||x-y||_1.$$

Picard Iterations

Given an initial value problem, a *Picard iteration sequence* is a sequence of functions on $[t_0, t_0 + c]$ constructed as follows:

$$\begin{aligned} \phi_0(t) &= x_0 \\ \phi_{m+1}(t) &= x_0 + \int_{t_0}^{t} f\big(s, \phi_m(s)\big) ds, \quad m = 0, 1, 2, 3, \ldots \end{aligned}$$

Theorem 2.7 *If $f(\cdot)$ satisfies the local Lipschitz condition on some compact set $\mathcal{K} \subset \mathcal{D}$ then the successive approximations ϕ_m, $m = 0, 1, 2, \ldots$ exist on $[t_0, t_0 + c]$, are continuous there and converge uniformly as $m \to \infty$ to the unique solution $\phi(\cdot)$. I.e. for every $\epsilon > 0$ there exists N such that for all $t \in [t_0, t_0 + c]$,*

$$||\phi(t) - \phi_m(t)|| < \epsilon,$$

whenever $m > N$.

3 e^{At} Through Picard Iterations

To do: Touch up to match what was done on board during class.

It is useful to briefly consider the non-homogeneous and time-dependent (non-autonomous) system:

$$\dot{x} = A(t)x + g(t), \quad x(t_0) = x_0.$$

As can be seen by successive Picard iterations, the solution of the above system is:

$$\phi(t, t_0, x_0) = \Phi(t, t_0)x_0 + \int_{t_0}^{t} \Phi(t, s)g(s)ds,$$

where the *state transition* matrix $\Phi(t, t_0)$ is defined as follows:

$$\Phi(t, t_0) = I + \int_{t_0}^{t} A(s_1)ds_1 + \int_{t_0}^{t} A(s_1) \int_{t_0}^{s_1} A(s_2)ds_2\, ds_1 + \int_{t_0}^{t} A(s_1) \int_{t_0}^{s_1} A(s_2) \int_{t_0}^{s_2} A(s_3)ds_3\, ds_2\, ds_1$$

$$\ldots\ldots\ldots + \int_{t_0}^{t} A(s_1) \int_{t_0}^{s_1} A(s_2) \ldots\ldots \int_{t_0}^{s_{m-1}} A(s_m)ds_m\, ds_{m-1} \ldots ds_1 + \ldots\ldots$$

The above expression is called the is the *Peano-Baker series.* Note that, $\Phi(t, t) = I$. When differentiating the Peano-Baker series with respect to t it is evident that,

$$\dot{\Phi}(t, t_0) = A(t)\Phi(t, t_0).$$

In the time-independent case of $A(t) = A$, the m'th term in the Peano-Baker series reduces to:

$$A^m \int_{t_0}^{t} \int_{t_0}^{s_1} \int_{t_0}^{s_2} \cdots \int_{t_0}^{s_{m-1}} 1ds_m \ldots ds_1 = \frac{(t - t_0)^m}{m!} A^m.$$

Hence in this case, the state transition matrix reduces to the form,

$$\Phi(t, t_0) = \sum_{k=1}^{\infty} \frac{(t - t_0)^k}{k!} A^k$$

Theorem 3.1 *Let $A \in \mathbb{R}^{n \times n}$ (or $A \in \mathbb{C}^{n \times n}$). Denote,*

$$S_m(t) = \sum_{k=1}^{m} \frac{t^k}{k!} A^k.$$

Then each element of the matrix $S_m(t)$ converges absolutely and uniformly on an finite $t \in \mathbb{R}$ interval containing 0, as $m \to \infty$.

We can thus define the *matrix exponential* matrix for any $t \in \mathbb{R}$ as,

$$e^{At} = \sum_{k=1}^{\infty} \frac{t^k}{k!} A^k.$$

Thus for the linear autonomous system, we have,

$$\Phi(t, t_0) = e^{A(t-t_0)}. \tag{7}$$

4 Basic Descriptions of the (A, B, C, D) Systems

We first consider the time-varying system,

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

with $x(t_0) = x_0$. In this case,

$$y(t) = C(t)\Phi(t, t_0)x_0 + C(t)\int_{t_0}^{t} \Phi(t, s)B(s)u(s)ds + D(t)u(t).$$

Specializing to the (time-invariant), (A, B, C, D) system we get from (7) that,

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

For the discrete time (A, B, C, D) system with $x(n_0) = x_0$ we get by similar development:

$$y(n) = CA^{n-n_0}x_0 + C\sum_{j=n_0}^{n-1} A^{n-(j+1)}Bu(j) + Du(n).$$

Exercise 4.1 *Verify using the above (discrete and continuous time) descriptions that if $\mathcal{O}(\cdot)$ is taken as the input-output mapping, then,*

$$\mathcal{O}\Big(\alpha_1 u_1 + \alpha_2 u_2\Big) = \alpha_1 \mathcal{O}\Big(u_1\Big) + \alpha_2 \mathcal{O}\Big(u_2\Big).$$

The definitions of memoryless, causality and time-invariance follow for these MIMO systems as they did for the SISO systems of unit 2.

The development (or treatment) of the *impulse response* and *transfer function* follows for MIMO systems in a similar ways to SISO systems. First consider the operator/system/input-output mapping $\mathcal{O}$:

$$y(\cdot) = \mathcal{O}\big(u(\cdot)\big).$$

We assume it admits an *integral representation*,

$$y(t) = \mathcal{O}\big(u(\cdot)\big)(t) = \int_{-\infty}^{\infty} h(t-\tau)u(\tau)d\tau = (h * u)(t),$$

with $h(t) \in \mathbb{R}^{p\times n}$ being the *impulse response matrix*. Note that for inputs $u(t)$ that have coordinates 0 except for the j'th coordinate, $u_j(t)$, the i'th component of the output has the form,

$$y_i(t) = \int_{-\infty}^{\infty} h_{ij}(t-\tau)u_j(\tau)d\tau,$$

as a SISO system with impulse response $h_{ij}(t)$.

The system is *causal* if and only if $h(t) = 0_{p \times n}$ for $t < 0$ and thus for inputs with positive support,

$$y(t) = \int_0^t h(t-\tau)u(\tau)d\tau.$$

The relation of convolutions and Laplace transforms exploited in Unit 2, carries over easily to the non-scalar version here. If the Laplace transform, $H(s)$ of $h(t)$ exists then,

$$\hat{y}(s) = H(s)\hat{u}(s).$$

A matrix Laplace transform such as this is simply a Laplace transform of each of the elements. In this case, $H(s)$ is the *transfer function matrix*.

We can further get the following useful representations:

$$h(t) = \mathbf{1}_{p \times p}(t)\Big(Ce^{At}B + D\delta_{m \times m}(t)\Big),$$

where we use a diagonal matrix of m delta-functions, $\delta_{m \times m}(t)$. Our treatment here is again not rigorous.

Of **independent interest**, note that in the SISO case, C is $1 \times n$ (single output), B is $n \times 1$ (single input) and D is a scalar. In this case if $det(A) \neq 0$, the *step-response* for $t > 0$ is,

$$s(t) = (\mathbf{1} * u)(t) = D + C\int_0^t e^{A\tau}d\tau = D + CA^{-1}(e^{At} - I)B = D - CA^{-1}B + B'e^{A't}A'^{-1}C'.$$

In the study of matrix-analytic methods in applied probability, the above is essentially the form of the so-called *Matrix Exponential (ME)* probability distribution. In that case, it is generally an open problem to characterize the $(A_{n\times n}, B_{n\times 1}, C_{1\times n}, D_{1\times 1})$ systems for which $s(t)$ is monotonic.

The Resolvent

The Laplace transform $H(s)$ takes on a very specific form for (A, B, C, D) systems.

Consider first the system autonomous system $\dot{x} = Ax$ with $x(0) = x_0$. In this case,

$$s\hat{x}(s) - x_0 = A\hat{x}(s),$$

and thus for s that are not eigenvalues of A,

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

Hence the Laplace transform of e^{At} is $(sI - A)^{-1}$. This is called the *resolvent* of the system.

The same computations can be carried out for the (A, B, C, D) system to get,

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

The Similarity Transform and Equivalent Representations

Given $P \in \mathbb{R}^{n\times n}$, with $det(P) \neq 0$, we can change the coordinates of the state-space based on the similarity transform,

$$P\tilde{x} = x.$$

The resulting system is,

$$\Big(P^{-1}AP, \;\; PB, \;\; CP, D\Big). \tag{8}$$

Both systems have the same external representations (i.e. same impulse response/transfer function) and are thus called *equivalent systems.*

Sampling a Continuous Time System

Consider a continuous time (A, B, C, D) system that is sampled at time intervals of T. In this case, the discrete time system,

$$\Big(e^{AT}, \;\; \int_0^T e^{A\tau} d\tau B, \;\; C, \;\; D\Big)$$

agrees with the continuous time one at the sampling points.

5 Computation of e^{At} - Understanding "Diagonalization"

Simple Evaluation Examples

Let,

$$A = \begin{bmatrix} 0 & 0 \\ \gamma & 0 \end{bmatrix},$$

then,

$$e^{At} = I + tA + \frac{t^2}{2}A^2 + \ldots = I + \begin{bmatrix} 0 & 0 \\ t\gamma & 0 \end{bmatrix} + 0_{2\times 2} = \begin{bmatrix} 1 & 0 \\ t\gamma & 1 \end{bmatrix}$$

Let,

$$A = \mathrm{diag}(\gamma_1, \ldots, \gamma_n),$$

i.e. the diagonal matrix with diagonal elements $\gamma_1, \ldots, \gamma_n$, then:

$$e^{At} = \mathrm{diag}\Big(\sum_{k=1}^{\infty} \frac{t^k}{k!}\gamma_1^k, \ldots\ldots\ldots, \sum_{k=1}^{\infty} \frac{t^k}{k!}\gamma_n^k\Big) = \mathrm{diag}\Big(e^{\gamma_1 t}, \ldots\ldots, e^{\gamma_n t}\Big).$$

Simple Diagonalization

To be updated

When is a Matrix Diagonalizable?

To be updated.

Non-Diagonalizable systems and Jordan Form

To be updated.

The Cayley-Hamilton Theorem Method

Theorem 5.1 *Every square matrix satisfies its characteristic polynomial equation.*

Proof To be updated $\square$

A consequence of of the Cayley-Hamilton Theorem with respect to e^{At} is that there exist scalar functions,

$$\alpha_1(t), \ldots, \alpha_{n-1}(t),$$

such that,

$$e^{At} = \sum_{i=0}^{n-1} \alpha_i(t) A^i.$$

The Laplace Transform of the State Transition Matrix

Since,

$$\mathcal{L}\big(e^{At}\big) = (sI - A)^{-1},$$

we can get the i, j's element of e^{At} by inverting the corresponding element of the resolvent. It is useful to look at an example:

Example 5.2 *Consider,*

$$A = \begin{bmatrix} -1 & 3 \\ 0 & 1 \end{bmatrix}.$$

Then,

$$(sI - A)^{-1} = \begin{bmatrix} s+1 & -3 \\ 0 & s-1 \end{bmatrix}^{-1} = \frac{1}{(s+1)(s-1)} \begin{bmatrix} s-1 & 3 \\ 0 & s+1 \end{bmatrix} = \begin{bmatrix} \frac{1}{s+1} & \frac{3/2}{s-1} + \frac{3/2}{s+1} \\ 0 & \frac{1}{s-1} \end{bmatrix}$$

So,

$$e^{At} = \begin{bmatrix} e^{-t} & \frac{3}{2}(e^t - e^{-t}) \\ 0 & e^t \end{bmatrix}.$$

6 Stability Based on Modes (w-out Lyapounov)

Some Stability Definitions

We now discuss the system,

$$\dot{x}(t) = Ax(t), \quad , x(0) = x_0, \tag{9}$$

and see when it is stable. As a brief introduction, consider the more general system,

$$\dot{x}(t) = f\big(x(t)\big), \quad x(0) = x_0.$$

and assume it has a unique solution $\phi(t)$ and that there is an *equilibrium point* at the origin. I.e,

$$f(0) = 0.$$

This obviously holds for the linear case (9). Alternativly, for the general system with $f(\cdot)$, if there is an equilibrium point $\tilde{x}_0$, one may make a change of coordinates so that in the new system 0 is an equilibrium point.

We say the equilibrium point, $x = 0$ is *stable* if for every $\epsilon > 0$ there exists a $\delta > 0$ (that may depend on ϵ) such that,

$$||\phi(t)|| < \epsilon, \ \forall t \geq 0,$$

whenever, $||x_0|| < \delta$.

Note that in this definition, ϵ and δ are not necessarily quantifies that are "considered small" (this is different than the normal use of ϵ, δ in say the definition of the limit of a function at a point).

A stronger notion of stability is *asymptotic stability*. We say that an equilibrium at $x = 0$ is said to be *asymptotically stable* if it is stable and further there is an $\eta > 0$ such that,

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

whenever $||x_0|| < \eta$.

An equilibrium point is *unstable* if it is not stable.

Exercise 6.1 *Spell out the definition of an unstable system. I.e." there exists and $\epsilon > 0$ such that ... ".*

Modes

We recall the system, (9) and denote,

$$det(sI - A) = \prod_{i=1}^{\sigma} (s - \lambda_i)^{n_i},$$

where $\lambda_1, \dots, \lambda_\sigma$ are distinct eigenvalues, each with a multiplicity of n_i (i.e. $\sum_{i=1}^{\sigma} n_i = n$).

We can show that,

$$e^{At} = \sum_{i=1}^{\sigma} \sum_{k=1}^{n_i - 1} A_{ik} t^k e^{\lambda_i t},$$

with the matrices A_{ik}, $i = 1, \dots, \sigma$ and for $k = 0, \dots, n_i - 1$,

$$A_{ik} = \frac{1}{k!(n_i - 1 - k)!} \lim_{s \to \lambda_i} \frac{d^{n_i - 1 - k}}{ds^{n_i - 1 - k}} (s - \lambda_i)^{n_i} (sI - A)^{-1}$$

We call each of the terms,

$$A_{ik} t^k e^{\lambda_i t},$$

a *mode* of the system. The system with state space dynamics $A \in \mathbb{R}^{n \times n}$ has exactly n modes.

The above representation can be derived by partial fraction expansion of the resolvent:

$$(sI - A)^{-1} = \sum_{i=1}^{\sigma} \sum_{k=1}^{n_i - 1} (k! A_{ik})(s - \lambda_i)^{-(k+1)}.$$

When all eigenvalues of A are distinct ($\sigma = n$) we get,

$$e^{At} = \sum_{i=1}^{n} A_i e^{\lambda_i t}.$$

In this case, $A_i = v_i \tilde{v}_i$ where v_i and $\tilde{v}_i'$ are the right and left eigenvectors of A corresponding to λ_i.

Stability Based on Modes

The stability of (9) is determined by the modes as follows:

1. It is *asymptotically stable* if and only if all eigenvalues of A have negative real part.

2. It is *stable* if and only if all eigenvalues are non-positive (i.e. some may be 0) and for eigenvalues, λ_j, with $Re\,\lambda_j = 0$ and multiplicity n_j,

$$\lim_{s \to \lambda_j} \frac{d^{n_i - 1 - k}}{ds^{n_i - 1 - k}} (s - \lambda_i)^{n_i} (sI - A)^{-1} = 0, \quad k = 1, \dots, n_j - 1.$$

3. The system is *unstable* if and only if (2) is not true.

7 Controllability and Observability

In this section we briefly introduce the two *regularity conditions*: *controllability* and *observability* (and the similar terms *reachability* and *constructibility*). The key results are briefly summarized with anticipation of the next two sections. A deeper look into controllability, observability and related forms follows later on.

With respect to controllability we consider both the discrete time system:

$$x(n+1) = Ax(n) + Bu(n),$$

and the continuous time system:

$$\dot{x}(t) = Ax(t) + Bu(t).$$

With respect observability we need to also consider the,

$$y(\tau) = Cx(\tau) + D(\tau),$$

part, where τ is either a discrete or continuous time index.

Controllability in Brief

A state $x_d \in \mathbb{R}^n$ is said to be *reachable* (synonymous with *controllable-from-the-origin*) if there exists an input $u(\cdot)$ that transfers $x(t)$ from the zero state to x_d in some finite time. A state $x_s \in \mathbb{R}^n$ is said to be *controllable* if there exists an input that transfers the state from x_s to the zero state in some finite time. These definitions are applicable to both discrete and continuous time systems.

As is evident later on, while reachability always implies controllability, controllability implies reachability only when the state transition matrix, $\Phi(\cdot)$ is nonsingular. This is always true for continuous time systems but for discrete time systems requires that A be non-singular. We will mostly ignore discrete time systems with singular A and thus treat reachability of a state and controllability of a state as essentially synonymous terms.

The set of all reachable/controllable states is called the *reachable / controllable subspace* of the system (it will be evident that this set is a linear sub-space of $\mathbb{R}^n$).

We say the whole system is *reachable / controllable* if any state is *reachable / controllable*, i.e. if the reachable / controllable subspace is $\mathbb{R}^n$. In this case we may also say that the pair (A, B) is *reachable / controllable.*

A key structure in the development is the matrix $\text{con}_k(A, B)$, defined for positive integer k as follows:

$$\text{con}_k(A, B) = \left[B, AB, A^2B, \ldots, A^{k-1}B\right] \in \mathbb{R}^{n \times mk}.$$

The following are important properties of $\text{con}_k(A, B)$:

Lemma 7.1 *For $k \geq n$, range$\Big(con_k(A,B)\Big) =$ range$\Big(con_n(A,B)\Big)$. For $k < n$, range$\Big(con_k(A,B)\Big) \subset$ range$\Big(con_n(A,B)\Big)$.*

Proof The statement for $k < n$ is obvious as adding columns to a matrix can only increase the dimension of its range.

Now the Cayley-Hamilton theorem states that,

$$A^n = -\frac{\alpha_{n-1}}{\alpha_n}A^{n-1} - \ldots - \frac{\alpha_1}{\alpha_n}A - \frac{\alpha_0}{\alpha_n}I,$$

where α_i are the coefficients of the characteristic polynomial of A with $\alpha_n \neq 0$. Alternatively,

$$A^n B = -\frac{\alpha_0}{\alpha_n}B - \frac{\alpha_1}{\alpha_n}AB - \ldots - \frac{\alpha_{n-1}}{\alpha_n}A^{n-1}B.$$

So the additional m columns in $\mathrm{con}_{n+1}(A,B)$ are linear combinations of the columns of $\mathrm{con}_n(A,B)$. Further the additional m columns in $\mathrm{con}_{n+2}(A,B)$ (that are not in $\mathrm{con}_n(A,B)$) are,

$$AA^n B = -\frac{\alpha_0}{\alpha_n}AB - \frac{\alpha_1}{\alpha_n}A^2B + \ldots + -\frac{\alpha_{n-2}}{\alpha_n}A^{n-1}B. - \frac{\alpha_{n-1}}{\alpha_n}A^nB.$$

and these are linear combinations of columns of $\mathrm{con}_{n+1}(A,B)$. Continuing by induction the result is proved. $\square$

To see the source of the $\mathrm{con}_k(A,B)$ matrix, consider the discrete time system with k-step input sequence reversed in time:

$$\overline{u}_k = \big[u(k-1)', u(k-2)', \ldots, u(0)'\big]' \in \mathbb{R}^{km}.$$

Since the evolution of state is,

$$x(k) = A^n x(0) + \sum_{i=0}^{k-1} A^{k-(i+1)}Bu(i),$$

we have that with input $\overline{u}$ over time steps, $0, 1, \ldots, k-1$, the state at time k can be represented by:

$$x(k) = A^k x(0) + \mathrm{con}_k(A,B)\overline{u}_k. \tag{10}$$

Hence the $\mathrm{con}_k(A,B)$ matrix captures the propagation of state in discrete time systems. As we shall see, it is also used in continuous time systems.

The key condition for reachability is based on the so-called *controllability matrix*:

$$\mathrm{con}(A,B) := \mathrm{con}_n(A,B).$$

I.e. it is the matrix that can be used to examine the state propogation over inputs for a number of time steps equal to the dimension of the state of the system.

Theorem 7.2 *A discrete system is reachable if and only if rank$\Big($con$(A,B)\Big) = n$.*

Proof

It is possible to transfer from state x_s to state x_d in k steps if an only if there exists an input sequence, $\overline{u}$ such that

$$\text{con}_k(A,B)\,\overline{u}_k = x_d - A^k x_s.$$

That is for reachability, set $x_s = 0$ and the system is reachable if and only if there is an integer k, such that,

$$x_d \in \text{range}\Big(\text{con}_k(A,B)\Big).$$

Now if rank$\Big(\text{con}(A,B)\Big) = n$ then x_d can be reached in n steps and thus it is reachable.

Conversley if it is reachable, since x_d is arbitrary, there is a k for which,

$$\text{rank}\Big(\text{con}_k(A,B)\Big) = n.$$

But then by Lemma 7.1, the above must be true for $k = n$.

$\square$

For continuous time systems, $\text{con}_k(A,B)$ does not have the same direct meaning as in (10) yet plays a central role. Assume $x(0) = x_s$ and an input $\{u(t),\, t \in [0,T]\}$ is applied such that $x(T) = x_d$, then,

$$x_d = e^{AT}x_s + \int_0^T e^{A(T-\tau)}Bu(\tau)d\tau.$$

The reachability sub-space during time $[0,T]$ is then:

$$\mathcal{R}_T := \Big\{x \in \mathbb{R}^n : \exists\ \{u(t),\, t \in [0,T]\},\ \text{such that,}\ x = \int_0^T e^{A(T-\tau)}Bu(\tau)d\tau\Big\}.$$

Lemma 7.3 *For any $T > 0$,*

$$\mathcal{R}_T \subset \mathit{range}\Big(\mathit{con}(A,B)\Big).$$

Proof As we saw earlier, using the Cayley-Hamilton theorem, there exist scalars $\alpha_i(t)$, $i = 0,\ldots,n-1$, such that

$$e^{At} = \alpha_0(t)I + \alpha_1(t)A + \ldots + \alpha_{n-1}(t)A^{n-1}.$$

Thus,

$$x = \int_0^T e^{At}Bu(T-\tau)d\tau = \int_0^T \Big(\sum_{i=0}^{n-1}\alpha_i(\tau)A^i\Big)Bu(T-\tau)d\tau$$

$$= \sum_{i=0}^{n-1} A^i B \int_0^T \alpha_i(\tau) u(T-\tau) d\tau = \mathrm{con}(A,B) r,$$

where

$$r = \begin{bmatrix} \int_0^T \alpha_1(\tau) u(T-\tau) d\tau \\ \vdots \\ \vdots \\ \int_0^T \alpha_m(\tau) u(T-\tau) d\tau \end{bmatrix},$$

Hence $x(t) \in \mathrm{range}\Big(\mathrm{con}(A,B)\Big)$. $\square$

Theorem 7.4 *A continuous time system is reachable if and only if rank$\Big(con(A,B)\Big) = n$.*

Proof The system is reachable if and only if $\mathrm{rank}\big(\mathcal{R}_T\big) = n$ for some T. Suppose first that $\mathrm{rank}\Big(\mathrm{con}(A,B)\Big) \neq n$ then by the lemma above $\mathbb{R}^n \setminus \mathcal{R}_T \neq \emptyset$ hence the system is not reachable. Thus if the system is reachable, $\mathrm{rank}\Big(\mathrm{con}(A,B)\Big) = n$.

To show the other direction we denote for any T, the $n \times n$ matrix,

$$W_T = \int_0^T e^{-A\tau} B B' e^{-A'\tau} d\tau.$$

We first show W_T is non singular. Suppose there is a vector $a \in \mathbb{R}^n$ such that, $W_T\, a = 0$, then $a'\, W_T\, a = 0$ or,

$$\int_0^T a' e^{-A\tau} B B' e^{-A'\tau} a \, d\tau = 0.$$

The integrand is of the form $c(t)'c(t)$ where $c(t) = B' e^{-A't} a$. Thus the integrand is nonnegative and thus for the integral to vanish we must have for all $t \in [0,T]$,

$$a' e^{-At} B = 0.$$

Take now derivatives with respect to t at $t = 0$,

$$\begin{aligned} a'B &= 0, \\ a'AB &= 0, \\ a'A^2B &= 0, \\ &\vdots \\ a'A^{n-1}B &= 0. \end{aligned}$$

Thus a is orthogonal to all columns of $\text{con}(A, B)$ and thus $a = 0$. Hence thus W_T is non-singular.

Now for a given x_d, select any $T > 0$ and set,

$$u(t) = B' e^{-A't} W_T^{-1} e^{-AT} x_d.$$

Thus,

$$x(T) = \int_0^T e^{A(T-t)} B u(t) dt = \int_0^T e^{A(T-t)} B B' e^{-A't} W_T^{-1} e^{-AT} x_d dt$$

$$= e^{AT} \int_0^T e^{-At} B B' e^{-A't} dt \;\; W_T^{-1} e^{-AT} x_d = e^{AT} W_T W_T^{-1} e^{-AT} x_d = x_d.$$

$\square$

Note: The above proof actually shows that to reach x_d in T time units, an input that can be applied over $[0, T]$ is,

$$u(t) = B' e^{-A't} \Big(\int_0^T e^{-A\tau} B B' e^{-A'\tau} d\tau \Big)^{-1} e^{-AT} x_d.$$

Observability in Brief

A system is said to be *observable* if knowledge of the outputs and the inputs over some finite time interval is enough to determine the initial state $x(0)$. For a discrete time system this means that $x(0)$ can be uniquely identified based on $y(0), y(1), \ldots, y(N-1)$ and $u(0), \ldots, u(N-1)$ for some finite N. For continuous time systems it means that $x(0)$ can be uniquely identified by $\{y(t),\, t \in [0, T]\}$ and $\{u(t),\, t \in [0, T]\}$ for some finite T.

The development of observability criteria parallels that of controllability. For discrete time systems,

$$y(k) = C A^k x(0) + \sum_{i=0}^{k-1} C A^{k-(i+1)} B u(i) + D u(k).$$

or alternatively define,

$$\tilde{y}(k) = y(k) - \big(\sum_{i=0}^{k-1} C A^{k-(i+1)} B u(i) + D u(k) \big),$$

and,

$$\text{obs}_k(A, C) = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{k-1} \end{bmatrix} \in \mathbb{R}^{pk \times n}.$$

Then,

$$\text{obs}_k(A, C)\, x(0) = \begin{bmatrix} \tilde{y}(0) \\ \tilde{y}(1) \\ \vdots \\ \tilde{y}(k-1) \end{bmatrix}. \tag{11}$$

The system is thus observable in k time units if (11) has the same unique solution, $x(0)$ for any k.

We define the *observability matrix* as :

$$\text{obs}(A, C) := \text{obs}_n(A, C).$$

Theorem 7.5 *A discrete or continuous system* (A, B, C, D) *is observable if and only if,*

$$rank\Big(obs(A, C)\Big) = n.$$

Duality between Controllability and Observability

Consider the "usual" system,

$$\begin{array}{rcl} \dot{x}(t) & = & Ax(t) + Bu(t) \\ y(t) & = & Cx(t) + Du(t) \end{array}. \tag{12}$$

The *dual system* is,

$$\begin{array}{rcl} \dot{x}(t) & = & A'x(t) + C'u(t) \\ y(t) & = & B'x(t) + D'u(t) \end{array}. \tag{13}$$

Notice that in the dual system, the state dimension is still n, but the dimensions of the input and the output were switched: The new input dimension is p and the new output dimension is m. The same definition holds for discrete time systems.

Theorem 7.6 *The system* (12) *is controllable if and only if the dual system* (13) *is observable. Similarly the system* (12) *is observable if and only if the dual system* (13) *is controllable.*

Proof We have that,

$$\text{con}(A, B) = \text{obs}(A', B')', \quad \text{obs}(A, C)' = \text{con}(A', C').$$

□

8 Linear State Feedback with Full State Information

The *linear state feedback law* is:

$$u(t) = Fx(t) + r(t),$$

for some $F \in \mathbb{R}^{m\times n}$ and $r(t) \in \mathbb{R}^m$ some external input vector refereed to as the *reference*. We often assume in the analysis that $r(t) \equiv 0$.

So we get,

$$\begin{aligned} \dot{x}(t) &= (A + BF)x(t) + Br(t) \\ y(t) &= (C + DF)x(t) + Dr(t) \end{aligned}$$

Theorem 8.1 *Given $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{n\times m}$ there exists $F \in \mathbb{R}^{m\times n}$ such that the n eigenvalues of $A + BF$ are assigned to arbitrary, real or complex conjugate locations if and only if (A, B) is a controllable pair.*

Proof

To be updated.

$\square$

Choosing F Using Controller Form

To be updated.

9 Observers

We now show how to design a system based on the original system whose state is denoted by $\hat{x}$ and is designed so that $\hat{x}(t)$ is an estimate of the (generally unobservable) $x(t)$. This simple (yet very powerful idea) is called the *Luenberger observer*. The basic equation in the design of the "observer system" is this:

$$\dot{\hat{x}}(t) = A\hat{x}(t) + Bu(t) + K(y - \hat{y}),$$

where,

$$\hat{y}(t) = C\hat{x}(t) + Du(t).$$

Combining we have,

$$\begin{aligned} \dot{\hat{x}}(t) &= (A - KC)\hat{x}(t) + \begin{bmatrix} B - KD, K \end{bmatrix} \begin{bmatrix} u \\ y \end{bmatrix} \\ \hat{y} &= C\hat{x} + \begin{bmatrix} D, 0 \end{bmatrix} \begin{bmatrix} u \\ y \end{bmatrix} \end{aligned}$$

Thus the Luenberger observer system, associated with the system (A, B, C, D) is the system $(A - KC, \big[B - KD, K\big], C, \big[D, 0\big])$ whose input is $[u', y']$, i.e. the input of the original system together with the output of the original system.

As opposed to the original system which typical has some physical manifestation, the observer is typically implemented in one way or another (often using digital computers). The *state* of the observer, $\hat{x}(t)$ is thus accessible by design and as we show now can yield a very good estimate of the actual (non-fully accessible) state, $x(t)$.

The error between the state and the estimate is $e(t) = x(t) - \hat{x}(t)$. Thus,

$$\begin{aligned}
\dot{e}(t) &= \dot{x}(t) - \dot{\hat{x}}(t) = (Ax(t) + Bu(t)) - \big(A\hat{x}(t) + Bu(t) + K(y(t) - \hat{y}(t)\big) \\
&= (Ax(t) + Bu(t)) - \big(A\hat{x}(t) + Bu(t) + K((Cx(t) + Du(t)) - (C\hat{x}(t) + Du(t))\big) \\
&= (A - KC)(x(t) - \hat{x}(t)) = (A - KC)\dot{e}(t).
\end{aligned}$$

Hence the estimation error associated with the Luenberger observer behaves like the autonomous system,

$$\dot{e}(t) = (A - KC)e(t).$$

If K is designed so that $(A - KC)$ has eigenvalues strictly in the LHP then $e(t) \to 0$ as $t \to \infty$ yielding an *asymptotic state estimator*. I.e. the estimation error would vanish as time progresses!!!! This is for any initial condition of both the system, $x(0)$ and the the observer $\hat{x}(t)$.

It turns out that the *observabillity* condition is exactly the condition that specifies when the autonomous system $(A - KC)$ can be shaped arbitrarlly:

Theorem 9.1 *There is a $K \in \mathbb{R}^{n \times p}$ so that eigenvalues of $A - KC$ are assigned to arbitrary locations if and only if the pair (A, C) is observable.*

Proof The eigenvalues of $(A - KC)' = A' - C'K'$ are arbitrarily assigned via K' if and only if the pair (A', C') is controllable (Theorem 8.1). This by duality (Theorem 7.6) occurs if and only if (A, C) is observable. □

10 Observer + State Feedback

Now that we know about state feedback and observers, we can combine them practically into a controlled system that has an observer for generating $\hat{x}(t)$ and then uses $\hat{x}(t)$ as input to a "state feedback" controller. This means that the input is,

$$u(t) = F\hat{x}(t) + r(t). \tag{14}$$

Remember that the observer follows,

$$\dot{\hat{x}}(t) = A\hat{x}(t) + Bu(t) + K\big(y(t) - \hat{y}(t)\big).$$

where,

$$\hat{y}(t) = C\hat{x}(t) + Du(t).$$

Combining the above with $y(t) = Cx(t) + Du(t)$ we get,

$$\dot{\hat{x}}(t) = (A - KC)\hat{x}(t) + KCx(t) + Bu(t).$$

Hence if we now combine (14) and look at the *compensated system* (original plant together with a state feedback law operating on an observer estimate), we get:

$$\begin{aligned}
\begin{bmatrix} \dot{x}(t) \\ \dot{\hat{x}}(t) \end{bmatrix} &= \begin{bmatrix} A & BF \\ KC & A - KC + BF \end{bmatrix} \begin{bmatrix} x(t) \\ \hat{x}(t) \end{bmatrix} + \begin{bmatrix} B \\ B \end{bmatrix} r(t) \\
y(t) &= \begin{bmatrix} C & DF \end{bmatrix} \begin{bmatrix} x(t) \\ \hat{x}(t) \end{bmatrix} + Dr(t)
\end{aligned}$$

Thus the compensated system is of state dimension $2n$ and has as state variables both the state variables of the system x and the observer "virtual"-state variables $\hat{x}$.

It is useful to apply the following similarity transform to the system:

$$P \begin{bmatrix} x \\ \hat{x} \end{bmatrix} = \begin{bmatrix} I & 0 \\ I & -I \end{bmatrix} \begin{bmatrix} x \\ \hat{x} \end{bmatrix} = \begin{bmatrix} x \\ e \end{bmatrix}$$

Hence as in (8), the resulting system is:

$$\begin{aligned}
\begin{bmatrix} \dot{x}(t) \\ \dot{e}(t) \end{bmatrix} &= \begin{bmatrix} A + BF & -BF \\ 0 & A - KC \end{bmatrix} \begin{bmatrix} x(t) \\ e(t) \end{bmatrix} + \begin{bmatrix} B \\ 0 \end{bmatrix} r(t) \\
y(t) &= \begin{bmatrix} C + DF & -DF \end{bmatrix} \begin{bmatrix} x(t) \\ e(t) \end{bmatrix} + Dr(t)
\end{aligned}$$

Exercise 10.1 *Show that the system above (with state (x,e)) is not fully controllable.*

The reason for not being fully controllable is that the state at the coordinates of corresponding to $e(t)$ should converge to 0, independently of $r(t)$.

Notice that,

$$det\Big(\begin{bmatrix} A + BF & -BF \\ 0 & A - KC \end{bmatrix} \Big) = det\big(sI - (A + BF)\big)\ det\big(sI - (A - KC)\big).$$

This implies that the behavior (fully governed by the eigenvalues) of the compensated system can be fully determined by selecting F and K separately!!!!

This is called the *separation principle* and it has far reaching implications: One may design the controller and the state estimator in separation and then combine. The dynamics of one will not affect the dynamics of the other.