Lecture 2(a) ENEE 660 Fall 2010

<u>Functions of linear maps (matrices)</u>

1. Given a scalar polynomial $q(\lambda) = q_k \lambda^k + q_{k-1} \lambda^{k-1} + \cdots + q_1 \lambda + q_0$ we define the associated <u>matrix-valued function</u> (here polynomial) of a <u>matrix argument</u> by

$$q(A) = \sum_{j=0}^{k} q_j A^j$$

where $A^0 = I$ the identity matrix. Since powers are involved, A is required to be square.

Immediately, we verify that if $\lambda \in$ spectrum(A) then $q(\lambda) \in$ spectrum of $q(A)$.

<u>proof</u>: Suppose λ, x s.t.

$$Ax = \lambda x.$$

Then $A^j x = \lambda^j x$ for $j = 0, 1, 2, \ldots$

Hence $q(A)x = q(\lambda)x$ ▨

In fact <u>one can say more</u>. For an $n \times n$ matrix A, if spectrum $(A) = \{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ then spectrum $(q(A)) = \{p(\lambda_1), p(\lambda_2), \ldots, p(\lambda_n)\}$

This brings us to the question of how would one define $f(A)$ for a square matrix A, for a 'general' function f. One approach is to consider the following:

suppose,

$$f(\lambda) = \sum_{k=0}^{\infty} a_k \lambda^k \qquad \lambda \in \mathbb{C}$$

with coefficients $a_k \in \mathbb{C}$, the complex numbers, is a convergent power series expansion for $\lambda \in D_\rho$, a disk of radius $\rho > 0$ centered at 0 in the complex plane. Then we define,

$$(*) \qquad f(A) = \sum_{k=0}^{\infty} a_k A^k .$$

This definition makes sense provided we limit A to lie in a suitable set of matrices allowing convergence of the series $(*)$ just as we limit $\lambda \in D_\rho$. We postpone details of this for later. We will also discuss later, alternatives to the definition $(*)$.

One particular matrix-valued function is of great importance in this course, namely the exponential. Recall that the scalar exponential is defined by the expansion

$$e^{x} = \sum_{k=0}^{\infty} \frac{x^k}{k!}.$$

This is a convergent series expansion. For the matrix exponential we simply define,

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

<u>Remark</u> Whereas for a scalar x, y

$$e^{x} e^{y} = e^{y} e^{x} = e^{x+y}$$

(prove this), the analogous result does not hold for matrix arguments, i.e. for matrices A, B, in general,

$$e^{A} e^{B} \neq e^{B} e^{A}.$$

This only holds if $AB = BA$.

<u>2.</u> The importance of the <u>matrix exponential</u> derives from the following fact

$$(**) \qquad \frac{d}{dt} e^{tA} = A e^{tA} = e^{tA} A \quad \forall t \in \mathbb{R}$$

We "prove" this as follows:

$$\frac{d}{dt} e^{tA} = \frac{d}{dt} \sum_{k=0}^{\infty} \frac{(tA)^k}{k!}$$

$$= \sum_{k=0}^{\infty} \frac{d}{dt} \frac{(tA)^k}{k!}$$

(assuming $\frac{d}{dt}$ can be done term-by-term)

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

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

$$= A \sum_{k=0}^{\infty} \frac{t^\ell A^\ell}{\ell!} = \left(\sum_{\ell=0}^{\infty} \frac{t^\ell A^\ell}{\ell!}\right) A$$

$$= A e^{tA} = e^{tA} A.$$

This property of the exponential has

an immediate consequence.

<u>Theorem</u>: Suppose

(†) $$\dot{x} = Ax$$

is a linear time invariant (LTI) differential equation (i.e. A is a constant matrix). Then <u>a solution</u> to (†) with initial condition

$$x(0) = x_0$$

is given by

$$x(t) = e^{tA} x_0 .$$

<u>Proof</u>: (a) $$e^{0A} x_0 = \sum_{k=0}^{\infty} \frac{0^k A^k}{k!} x_0 = I x_0 = x_0$$

(b) $$\dot{x}(t) = \frac{d}{dt} e^{tA} x_0$$

$$= e^{tA} A x_0 = A e^{tA} x_0$$

$$= \cancel{e^{tA}} A x(t). \quad \blacksquare$$

We have <u>existence</u> of a solution.

This is the only solution. Formally,

Theorem Suppose $x(t)$ and $y(t)$ both satisfy a LTI differential equation

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

$$x(0) = x_0$$

Then $x(t) \equiv y(t) \quad \forall\, t$

Proof: Let us denote

$$\|v\| = \left(\sum_{i=1}^{n} v_i^2\right)^{1/2}$$

the Euclidean norm. Observe that, for any component v_i of v,

$$v_i \le |v_i| \le \|v\|.$$

By hypothesis,

$$z(t) = x(t) - y(t)$$

satisfies,

$$z(0) = 0$$

and
$$\dot{z}(t) = \dot{x}(t) - \dot{y}(t)$$
$$= A\,x(t) - A\,y(t)$$
$$= A\,z(t)$$

Then, letting
$$\rho(t) = \| z(t) \| = \left(z^T(t)\, z(t) \right)^{1/2}$$
where the superscript T denotes transpose, we have

$$\frac{d}{dt}\rho^2 = \frac{d}{dt}\left(z^T(t)\, z(t) \right)$$

$$= \dot{z}^T(t)\, z(t) + z^T(t)\, \dot{z}(t)$$

$$= z^T(t)\, A^T z(t) + z^T(t)\, A\, z(t)$$

$$= 2\, z^T A\, z$$

$$= 2 \sum_{i=1}^{n} \sum_{j=1}^{n} z_i\, a_{ij}\, z_j$$

$$\leq 2 \sum_{i=1}^{n} \sum_{j=1}^{n} \|z\|\, |a_{ij}|\, \|z\|$$

$$\leq 2 \|z\|^2 n^2 \max_{i,j} |a_{ij}| \,.$$

Let $\eta = 2n^2 \max_{i,j} |a_{ij}|$. Then we have

$$\frac{d}{dt}\rho^2 \leq \rho^2 \eta \,.$$

Let $\theta(t) = \rho^2(t) \exp(-t\eta)$, we see that,

$$\frac{d}{dt}\theta(t) = \left(\frac{d}{dt}\rho^2 - \eta\rho^2\right)e^{-t\eta}$$

$$\leq 0 \qquad (\text{exponential} > 0)$$

$$\Rightarrow \quad \theta(t) \leq \theta(0).$$

But $\theta(0) = \rho^2(0) = 0$ (by hypothesis)
and

$\theta(t) \geq 0$ (by definition)

Hence $\theta(t) \equiv 0 \quad \forall t$

$$\Rightarrow \quad z(t) \equiv 0$$

$$\Rightarrow \quad x(t) \equiv y(t) \qquad \square$$

Remark It we replace the given LTI differential equation by a linear time-varying differential equation of the form

$$\dot{x}(t) = A(t)\, x(t), \qquad x(0) = x_0$$

then, the conclusions of the above

theorem can be extended to this LTV setting as well, provided $A(t)$ is continuous in t on interval of interest, η is replaced by $\eta(t) = \max_{\substack{i,j \\ \in \{1,2,\dots,n\}}} |a_{ij}(t)|$ and the integrating factor $e^{-t\eta}$ is replaced by $\exp\left(-\int_0^t \eta(\sigma)d\sigma\right)$ in the proof.

Thus we get uniqueness of solutions to time-varying linear systems of ordinary differential equations in a straightforward way. To have existence, we need to replace the e^{tA} for LTI system by a suitable (family of) linear maps (matrices). What might this be? We answer this question in a constructive way in the next subsection.

3. Solving Linear Time Varying Systems

Consider the system

$$\dot{x}(t) = A(t)x(t) \quad ; \quad x(t_0) = x_0$$

The clock time at which the system is initialized is t_0 and is crucial to keep track of since $A(\cdot)$ is time-varying.

Integrating both sides of the differential equation we get,

$$x(t) - x(t_0) = \int_{t_0}^{t} A(\sigma)\, x(\sigma)\, d\sigma$$

or

$$x(t) = x_0 + \int_{t_0}^{t} A(\sigma)\, x(\sigma)\, d\sigma,$$

with time horizon of interest

specified as $t \in [t_0, t_1]$. We are interested in solving the above integral equation for <u>arbitrary initial conditions</u>. Suppose we denote by $\phi_k(t, t_0)$ <u>the</u> solution for initial condition

$x_0 = e_k$ = standard basis vector in $\mathbb{R}^n$ with 1 in k^{th} place and 0 elsewhere. Then

$$\phi_k(t, t_0) = e_k + \int_{t_0}^{t} A(\sigma)\, \phi_k(\sigma, t_0)\, d\sigma.$$

By arranging such solutions in

$$\Phi(t, t_0) = \left[\phi_1(t,t_0), \phi_2(t,t_0), \ldots, \phi_n(t,t_0)\right]$$

we obtain matrix, called <u>transition matrix</u> of the LTV system, also referred to in earlier literature as the <u>matrizant</u> and it has to satisfy

$$\Phi(t, t_0) = I + \int_{t_0}^{t} A(\sigma)\Phi(\sigma, t_0)\,d\sigma$$

where $I = [e_1, e_2, \ldots, e_n]$ is the identity matrix.

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

How do we solve the above integral equation for the transition matrix? We state

<u>Theorem</u>: (Transition Matrix)
Let $A(\cdot)$ be a square matrix whose elements are continuous functions of time on the interval $t_0 \le t \le t_1$. Define the sequence of matrix-valued functions defined recursively by:

$$S_0(t, t_0) = I$$

$$S_k(t, t_0) = I + \int_{t_0}^{t} A(\sigma)\, S_{k-1}(\sigma, t_0)\, d\sigma$$

$$k = 1, 2, \ldots$$

(This is a sequence of partial sums of a matrix series.) Then the sequence converges uniformly on the given interval $[t_0, t_1]$. Moreover the limit, denoted by $\Phi(t, t_0)$ satisfies the <u>matrix</u> differential equation

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

and <u>the</u> solution to the LTV system

$$\dot{x}(t) = A(t)\, x(t)$$
$$x(t_0) = x_0$$

is

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

Remark The theorem as stated can be found R.W. Brockett, Finite Dimensional Linear Systems, John Wiley and Sons Inc., New York, 1970 and with minor variations, in many other places. Our proof follows Brockett.

Proof. For a series of scalar function of $t \in [t_0, t_1]$ of the form $x_1(t) + x_2(t) + x_3(t) + \cdots$, we say the series converges absolutely and uniformly on the interval $[t_0, t_1]$ if there exists a sequence of positive constants c_i such that

$$|x_i(t)| \le c_i \qquad \forall\, t \in [t_0, t_1]$$

And series $c_1 + c_2 + c_3 + \cdots$ converges. Converges of the latter is simply convergence of the sequence $\mu_1, \mu_2, \mu_3, \cdots$ of partial sums

$$\mu_k = \sum_{i=1}^{k} c_i \qquad k = 1, 2, \cdots$$

$$\text{Let} \quad \eta(t) = \max_{i,j=1,2,\ldots,n} |a_{ij}(t)|$$

$$\text{Let} \quad \gamma(t) = \int_{t_0}^{t} \eta(\sigma)\, d\sigma .$$

For any $n \times n$ matrices A, B, the product $C = AB = [c_{ij}]$ satisfies,

$$|c_{ij}| = \left| \sum_{k=1}^{n} a_{ik} b_{kj} \right|$$

$$\le \sum_{k=1}^{n} |a_{ik}|\, |b_{kj}|$$

$$\leq \sum \max_{\substack{\ell m \\ =1,2,\ldots,n}} |a_{\ell m}| \max_{\substack{r,s \\ =1,2,\ldots,n}} |b_{rs}|$$

$$= n \max_{\ell, m} |a_{\ell m}| \max_{r,s} |b_{rs}|$$

It follows that

$$\left(S_k(t,t_0) - S_{k-1}(t,t_0)\right)_{ij}$$

$$= \left(\int_{t_0}^{t} A(\sigma_1) \int_{t_0}^{\sigma_1} A(\sigma_2) \cdots \int_{t_0}^{\sigma_{k-1}} A(\sigma_k)\, d\sigma_k\, d\sigma_{k-1} \cdots d\sigma_1\right)_{ij}$$

$$\leq \int_{t_0}^{t}\int_{t_0}^{\sigma_1} \cdots \int_{t_0}^{\sigma_{k-1}} n^{k-1} \eta(\sigma_1)\eta(\sigma_2) \cdots \eta(\sigma_k)\, d\sigma_k \cdots d\sigma_1$$

$$= n^{k-1} \frac{\gamma^k(t)}{k!}$$

Thus, the (i,j)th element of each term in the series

$$M_0(t,t_0) + \sum_{k=1}^{\infty} M_k(t,t_0) - M_{k-1}(t,t_0)$$

is smaller the corresponding term in the series

$$1 + \gamma(t) + n\frac{\gamma^2(t)}{2!} + n^2\frac{\gamma^3(t)}{3!} + \cdots$$

But the latter converges for all n to

$$1 - \frac{1}{n} + \frac{e^{n\gamma(t)}}{n} + \cdots$$

So each element of the matrix series must converge too.

To show that the limit Φ actually satisfies the differential equation, observe that term-by-term differentiation is permitted (because of uniform convergence) and

$$\frac{d}{dt}\Phi(t,t_0) = \frac{d}{dt}\Big(I + \int_{t_0}^{t} A(\sigma_1)d\sigma_1 + \int_{t_0}^{t} A(\sigma_1)\int_{t_0}^{\sigma_1} A(\sigma_2)d\sigma_2 d\sigma_1 + \cdots\Big)$$

$$= 0 + A(t) + A(t)\int_{t_0}^{t} A(\sigma_2)d\sigma_1 + \cdots$$

$$= A(t)\left[\Phi(t,t_0)\right]$$

The expression

$$\Phi(t,t_0) = I + \int_{t_0}^{t} A(\sigma_1)\,d\sigma_1$$

$$+ \int_{t_0}^{t} A(\sigma_1) \int_{t_0}^{\sigma_1} A(\sigma_2)\,d\sigma_2\,d\sigma_1$$

$$+ \dots$$

is known as the Peano – Baker series.

Remark special case :

suppose $A(t) = a(t)\,\bar{A}$

where $\bar{A}$ is a constant matrix. Then

$$\Phi(t,t_0) = \exp\left(A \int_{t_0}^{t} a(\sigma)\,d\sigma\right)$$