Lecture 3(a) ENEE 660 Fall 2010

1. Linear Inhomogeneous Equations

The basic model of linear system theory is the state evolution equation with inputs (controls) and outputs (sensor data) given in the form:

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

where A is $n \times n$, B is $n \times m$ and C is $p \times n$. Additionally, the direct feed-through of the input to the output is given by the term $D(t)\,u(t)$ where D is $p \times m$.
Let $\Phi(t, t_0)$ be the transition matrix associated to A.
Suppose the initial state $x(t_0) = x_0$.
Let $z(t) = \Phi(t_0, t)\,x(t)$. Then,

$$\dot{z}(t) = \frac{d\Phi}{dt}(t_0, t)\; x(t) + \Phi(t_0, t)\frac{dx}{dt}$$
$$= \left(\frac{d}{dt}\Phi^{-1}(t, t_0)\right) x(t) + \Phi(t_0, t)\frac{dx}{dt}$$
$$= -\Phi^{-1}(t, t_0)\frac{d\Phi}{dt}(t, t_0)\,\Phi^{-1}(t, t_0)\,x(t)$$
$$+ \Phi(t_0, t)\left(A(t)\,x(t) + B(t)\,u(t)\right)$$

$$= -\Phi^{-1}(t,t_0)A(t)\Phi(t,t_0)\Phi^{-1}(t,t_0)\cdot X(t)$$

$$+ \Phi(t_0,t)A(t)\cdot X(t) + \Phi(t_0,t)B(t)u(t)$$

$$= \Phi(t_0,t)B(t)u(t).$$

Integrating both sides, we obtain,

$$Z(t) = Z(t_0) + \int_{t_0}^{t} \Phi(t_0,\sigma)B(\sigma)u(\sigma)\,d\sigma$$

But $Z(t_0) = \Phi(t_0,t_0)X(t_0)$

$$= X(t_0).$$

Hence,

$$X(t) = \Phi(t,t_0)Z(t)$$

$$= \Phi(t,t_0)X(t_0) + \Phi(t,t_0)\int_{t_0}^{t}\Phi(t_0,\sigma)B(\sigma)u(\sigma)\,d\sigma$$

$$= \Phi(t,t_0)X(t_0) + \int_{t_0}^{t}\Phi(t,t_0)\Phi(t_0,\sigma)B(\sigma)u(\sigma)\,d\sigma$$

$$= \Phi(t,t_0)X(t_0) + \int_{t_0}^{t}\Phi(t,\sigma)B(\sigma)u(\sigma)\,d\sigma$$

This is the variation of constants formula.

From the output equation, we obtain,

$$\boxed{\begin{aligned} y(t) &= C(t)\,\Phi(t,t_0)\,X(t_0) \\ &\quad + \int_{t_0}^{t} C(t)\,\Phi(t,\sigma)\,B(\sigma)\,u(\sigma)\,d\sigma \\ &\quad + D(t)\,u(t). \end{aligned}}$$

This equation derived from the variation of Constants formula, shows the relationship between the input and the output vectors. Only inputs $u(\sigma)$ for $\sigma \leq t$ affect $u(t)$. This is the essence of causality. The term $C(t)\,\Phi(t,t_0)\,X(t_0)$ is referred to as the drift term.

A fundamental problem is to determine if state x_1 is reachable at time t_1, from the initial state x_0 at time t_0, i.e we seek to know if the equation

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

is solvable for a control function $u(\cdot)$ on the time interval $[t_0, t_1]$. This equation can be recast in the form:

$$x_0 - \Phi(t_0, t_1) x_1 = -\int_{t_0}^{t} \Phi(t_0, \sigma) B(\sigma) u(\sigma) d\sigma$$

$$= L(u)$$

This is a linear inhomogeneous equation for the map L from a function space into $\mathbb{R}^n$. To solve this problem we need a few additional concepts from linear algebra.

2. The concept of adjoint of a linear map

Recall that a vector space V over the reals $\mathbb{R}$ is an <u>inner product space</u>, if it is endowed with an inner-product

$$\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$$

$$(v, \tilde{v}) \mapsto \langle v, \tilde{v} \rangle$$

Satisfying the axioms:

(i) $\langle v, \tilde{v} \rangle = \langle \tilde{v}, v \rangle$ (symmetry)

(ii) $\langle v, ax + by \rangle = a\langle v, x\rangle + b\langle v, y\rangle$

where $a, b \in \mathbb{R}$ and $v, x, y \in V$. (bilinearity) and (iii) $\langle v, v \rangle \geq 0$ and $\langle v, v \rangle = 0$ iff $v = 0$ (positive definiteness).

For $v = \mathbb{R}^n$, we can define

$$\langle \beta, \tilde{\beta} \rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} \beta_i \, q_{ij} \, \tilde{\beta}_j = \beta^T Q \tilde{\beta}$$

where $Q = Q^T$ is a positive definite matrix. If $Q = I$ the identity matrix, we call this inner product <u>Euclidean</u>.

More generally, on an abstract vector V of dimension n, let $S_v = \{v_1, v_2, \cdots, v_n\}$ be a <u>basis</u>. For each $v \in V$, let β denote the corresponding coordinate vector $\in \mathbb{R}^n$. We can define an inner product on V by the coordinate expression:

$$\langle v, \tilde{v} \rangle = \sum_{i=1}^{n} \beta_i \, \tilde{\beta}_i = \beta^T \tilde{\beta}$$

Observe that $v = v_i$ has the coordinate vector $(0, 0, \ldots, 1, 0, \ldots, 0)^T = e_i$ a standard basis vector in $\mathbb{R}^n$.

$\uparrow$ i^{th} place

It follows that,

$$\langle v_i, v_j \rangle = e_i^T e_j = \delta_j^i$$

Thus the basis vectors in S_V are <u>mutually orthogonal</u> and of length given by $\|v_i\| = \|e_i\| = 1$. We call S_V an <u>orthonormal basis</u>.

<u>Remark</u> For an inner product space $(V, \langle\cdot,\cdot\rangle_V)$, a general basis $S_V = \{v_1, v_2, \cdots, v_n\}$ does not have the orthogonal property above. But there is an algorithm, known as the Gram-Schmidt process, that converts it into $\hat{S}_V = \{\hat{v}_1, \hat{v}_2, \cdots, \hat{v}_n\}$ a new basis whose vectors constitute an orthonormal set:

$$\hat{v}_1 = \frac{v_1}{\|v_1\|} = \frac{v_1}{\sqrt{\langle v_1, v_1\rangle}}$$

$$\hat{v}_k = \frac{\left(v_k - \sum_{i=1}^{k-1} \langle v_k, \hat{v}_i\rangle \hat{v}_i\right)}{\left\| v_k - \sum_{i=1}^{k-1} \langle v_k, \hat{v}_i\rangle \hat{v}_i \right\|}$$

$k = 2, 3, \cdots, n.$ $\square$

We now define the adjoint of a linear map between inner product spaces.

Let $(V, \langle\cdot,\cdot\rangle_V)$ and $(W, \langle\cdot,\cdot\rangle_W)$ be two inner product spaces, and let

$$A : V \to W$$

be a linear map. We say that a map $A^* : W \to V$ is <u>the</u> adjoint of A if

$$\langle Av, w \rangle_W = \langle v, A^* w \rangle_V$$

for all $v \in V$ and $w \in W$.

Remark A^* is unique. To see this, suppose maps A_1^* and A_2^* both satisfy the above adjoint relation for given A. Then

$$\langle v, A_1^* w \rangle_V - \langle v, A_2^* w \rangle_V$$

$$= \langle Av, w \rangle_W - \langle Av, w \rangle_W$$

$$= 0 \qquad \forall\, v \in V,\ w \in W.$$

Thus $\langle v, (A_1^* - A_2^*) w \rangle_V = 0 \quad \forall\ v \in V, w \in W.$

letting $v = (A_1^* - A_2^*) w$ and using the positive definiteness property of $\langle \cdot, \cdot \rangle_V$ we conclude that

$$(A_1^* - A_2^*) w = 0. \qquad \forall\, w \in W$$

Hence $A_1^* = A_2^*$

If V and W are finite dimensional, then we can compute the matrix representation of A^* from the matrix representation of A with respect to the fixed bases $S_V = \{v_1, v_2, \cdots, v_n\}$ of V and $S_W = \{w_1, w_2, \cdots, w_m\}$ of W. Assume that these are orthonormal basis. If not, we can use the Gram-Schmidt process to obtain orthonormal bases.

Let $\tilde{A} = [a_{ij}]$ be the $m \times n$ matrix representation of $A : V \to W$. Thus,

$$A v_j = \sum_{i=1}^{m} a_{ij} w_i$$

$$\begin{aligned} \langle A v_j, w_\ell \rangle_W &= \Big\langle \sum_{i=1}^{m} a_{ij} w_i , w_\ell \Big\rangle_W \\ &= \sum_{i=1}^{m} a_{ij} \langle w_i, w_\ell \rangle_W \\ &= \sum_{i=1}^{m} a_{ij} \delta_\ell^i \qquad \text{(orthonormality of } S_W) \\ &= a_{\ell j} \end{aligned}$$

Let $\tilde{A}^* = [a^*_{ij}]$ denote the matrix representation of A^*. Then,

$$A^* w_j = \sum_{i=1}^{n} a^*_{ij} v_i$$

$$\langle v_j, A^* w_\ell \rangle = \langle v_j, \sum_{i=1}^{n} a^*_{i\ell} v_i \rangle$$

$$= \sum_{i=1}^{n} a^*_{i\ell} \langle v_j, v_i \rangle$$

$$= \sum_{i=1}^{n} a^*_{i\ell} \, \delta^i_j \quad \text{(orthonormality of } S_V\text{)}$$

$$= a^*_{j\ell}$$

From the definition of adjoint

$$\langle A v_j, w_\ell \rangle_W = \langle v_j, A^* w_\ell \rangle$$

or equivalently

$$a_{\ell j} = a^*_{j\ell}$$

Thus the matrix representation

$$\tilde{A}^* = \tilde{A}^T$$

So, when we speak about adjoints of linear maps from a finite dimensional inner product space to another finite dimensional inner product space, then we are simply giving an abstraction of the concept of matrix transpose.

But, the finite dimensional setting is not the only thing of interest to us.

Example. Let $V = \{ f : [t_1, t_2] \to \mathbb{R}^m \mid f \text{ continuous} \}$. Consider an inner product on V defined by

$$\langle f, g \rangle = \int_{t_1}^{t_2} \sum_{i=1}^{m} f_i(t)\, g(t_i)\, dt$$

$$= \int_{t_1}^{t_2} f^T g \, dt .$$

Consider a linear map

$$L_M : V \to \mathbb{R}^n$$
$$f \mapsto \int_{t_1}^{t_2} M(\sigma) f(\sigma)\, d\sigma$$

where M is a specific $n \times m$ matrix-valued continuous function.

When $\mathbb{R}^n$ is given the usual Euclidean inner-product what is L_M^* ?

Claim: $L_M^* : \mathbb{R}^n \to V$ is given by
$$\eta \mapsto g(\cdot) = M^T(\cdot)\eta .$$

Proof

$$\langle L_M f, \eta \rangle_{\mathbb{R}^n} = \eta^T \int_{t_1}^{t_2} M(\sigma) f(\sigma) d\sigma$$

$$= \int_{t_1}^{t_2} \eta^T M(\sigma) f(\sigma) d\sigma$$

$$= \int_{t_1}^{t_2} (M^T(\sigma) \eta)^T f(\sigma) d\sigma$$

$$= \int_{t_1}^{t_2} f^T(\sigma) \; M^T(\sigma) \eta \; d\sigma$$

$$= \langle f, g \rangle_V$$

(where $g(\cdot) = M^T(\cdot)\eta$)

$$= \langle f, L_M^* \eta \rangle_V \qquad \boxtimes$$

We now state a key result about range spaces.

Theorem Suppose $(V, \langle \cdot, \cdot \rangle_V)$ and $(W, \langle \cdot, \cdot \rangle_W)$ are finite dimensional inner product spaces. Suppose $A : V \to W$ has the adjoint $A^* : W \to V$. Then

$$\mathcal{R}(A)^\perp = \operatorname{Ker}(A^*).$$

(Here $\mathcal{R}(A)^\perp = \{ p : \langle p, y \rangle_W = 0 \;\; \forall \; y \in \mathcal{R}(A) \}$).

__Proof__: ($\Rightarrow$) Let $y^* \in$ ~~▒~~ $\mathrm{Ker}(A^*)$, Hence $A^*y^* = 0$

Let y be an arbitrary element of $\mathcal{R}(A)$.
Then $y = Ax$ for some $x \in V$.
Then

$$\begin{aligned}
\langle y, y^* \rangle_W &= \langle Ax, y^* \rangle_W \\
&= \langle x, A^*y^* \rangle_V \qquad \text{(def}^{n}\text{ of adjoint)} \\
&= \langle x, 0 \rangle_V \\
&= 0
\end{aligned}$$

Thus $y \perp y^*$.
Thus we have shown $y^* \in \mathcal{R}(A)^\perp$
Thus

$$\mathrm{Ker}(A^*) \subseteq \mathcal{R}(A)^\perp$$

($\Leftarrow$) Assume $y^* \in \mathcal{R}(A)^\perp$. Then __for every__ $x \in V$, $\langle Ax, y^* \rangle_W = 0$.

$$\begin{aligned}
&\Rightarrow \quad \langle x, A^*y^* \rangle_V = 0 \qquad \forall\, x \in V \\
&\Rightarrow \quad A^*y^* = 0 \\
&\Rightarrow \quad y^* \in \mathrm{Ker}(A^*).
\end{aligned}$$

Thus $\mathcal{R}(A)^\perp \subseteq \mathrm{Ker}(A^*)$

Combining both parts $\mathcal{R}(A)^\perp = \mathrm{Ker}(A^*)$ ▨

<u>Question</u> Do we need the <u>finite dimensionality hypothesis</u> above?