Cambridge University Engineering Dept. Third year

Module 3F2: Systems and Control

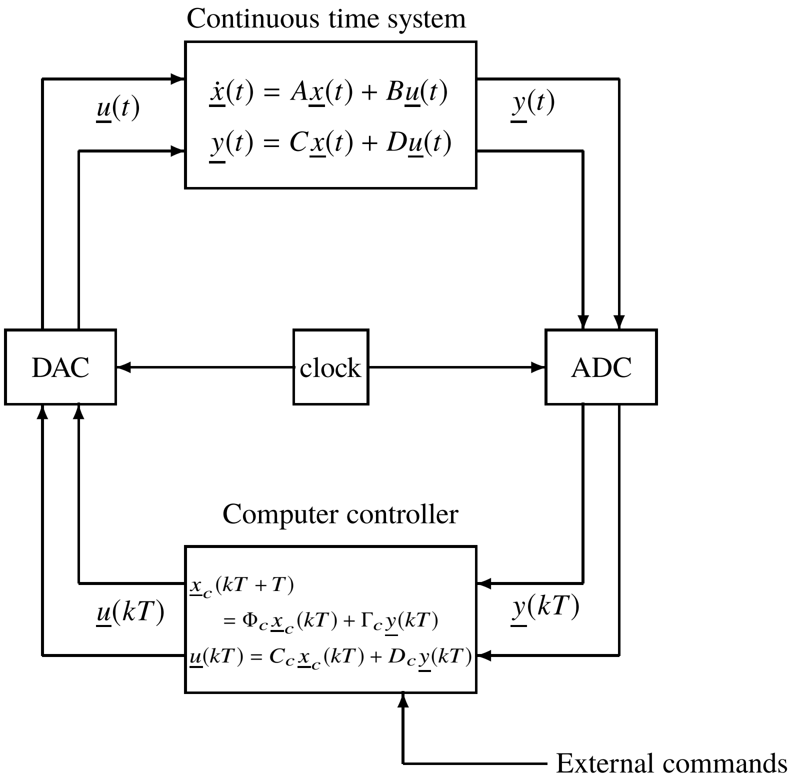
LECTURE NOTES 3: OBSERVABILITY & OBSERVERS

Contents

Contents

K. Glover Revised by G. Vinnicombe
Jan 2002 February 2018

1 Sampled Data Control System



The sampled data system satisfies:

$$\underline{\dot{x}}(t) = A\underline{x}(t) + B\underline{u}(t), \quad \text{with } \underline{u}(t) = \underline{u}(kT), \text{ for } kT \le t < (k+1)T.$$

Apply result from Handout 1, section 4.5 (Convolution integral):

$$\begin{aligned}
\underline{x}((k+1)T) &= \underbrace{e^{AT}}_{\Phi}\ \underline{x}(kT) + \int_{kT}^{(k+1)T} e^{A((k+1)T-\tau)} B d\tau \underline{u}(kT) \\
&= \Phi \underline{x}(kT) + \Gamma \underline{u}(kT) \\
\text{where} & \\
\Gamma &= \int_0^T e^{A\tau'} d\tau' B \ = A^{-1}\left(e^{AT} - I\right) B, \text{ if } \det(A) \neq 0. \\
\underline{y}(kT) &= C\underline{x}(kT) + D\underline{u}(kT)
\end{aligned}$$

This gives the standard state-space model in discrete time. Entirely analogous results can be obtained for the discrete time case as in the continuous time case:

- Solution of vector difference equations,
- Discrete-time convolution,
- $z-$transform for frequency response caculations etc,
- Notions of controllability and observability — coming next.

2 Solving Linear Equations

For convenience we will repeat some results and definitions from linear algebra.

Definition 2.1 *Let A be an $m \times n$ matrix then,*

(a) the set of all $\underline{x} \neq \underline{0}$ such that $A\underline{x} = \underline{0}$ is called the **Null Space** *of A ($\mathrm{null}(A)$). This is sometimes referred to as the* **Kernel** *of A.*

(b) the set of all $\underline{y}$ such that $\underline{y} = A\underline{x}$ for some $\underline{x}$ is called the **Range Space** *of A (or the range of A, $\mathrm{range}(A)$); This is sometimes referred to as the* **Column Space** *or* **Image** *of A.*

(c) A is said to have full row *rank if* $\mathrm{range}(A) = \mathbb{R}^m$ *(i.e. $\underline{z}^T A \neq \underline{0}$ for all $\underline{z} \neq \underline{0}$);*

(d) A is said to have full column *rank if* $\mathrm{null}(A) = \emptyset$ *(i.e. $A\underline{x} \neq \underline{0}$ for all $\underline{x} \neq \underline{0}$.)*

Theorem 2.2 *For any matrix A the row rank and the column rank are equal, and denoted* $\mathrm{rank}(A)$.

Given an $m \times n$ matrix A and an $m \times 1$ vector $\underline{b}$, consider the equation:

$$A\underline{x} = \underline{b},$$

in the unknown $\underline{x}$ in $\mathbb{R}^n$. Two natural questions are:

(a) Does there exist a solution, $\underline{x}$?

(b) If so, is it unique?

Fact 2.3 *For the case $m = n$:*

(a) If $\det(A) \neq 0$ *then for any $\underline{b}$ there exists a solution, $\underline{x}$, such that $A\underline{x} = \underline{b}$, and this solution is unique (Indeed it is given by $\underline{x} = A^{-1}\underline{b}$).*

(b) If $\det(A) = 0$ *then there exists $\underline{x} \neq \underline{0}$ such that $A\underline{x} = \underline{0}$.*

Fact 2.4 *For any $m \times n$ matrix, M,*

$$M^T M\underline{x} = \underline{0} \quad \Leftrightarrow \quad M\underline{x} = \underline{0}.$$

Fact 2.5 *For the case $m \leq n$,*

(a) If $\det(AA^T) \neq 0$ *then* $\underline{x} = A^T \left(AA^T\right)^{-1} \underline{b}$, *solves* $A\underline{x} = \underline{b}$ *for any* $\underline{b}$.

(b) If $\det(AA^T) = 0$ *then there exists a* $\underline{b} \neq \underline{0}$ *such that* $\underline{b} \perp A\underline{x}$ *(i.e.* $\underline{b}^T A\underline{x} = \underline{0}$*) for all* $\underline{x}$.

For the case $m \geq n$,

(c) If $\det(A^T A) \neq 0$ *then there may not be a solution to* $A\underline{x} = \underline{b}$, *but if there is then it is unique.*

(d) If $\det(A^T A) = 0$ *then there exists* $\underline{x} \neq \underline{0}$ *such that* $A\underline{x} = \underline{0}$.

For hand calculations it is generally easiest to use the following observations:

(a) If you can find a set of n rows of A such that the determinant of the $n \times n$ submatrix given by these rows is nonzero, then A has full column rank.

(b) If you can find a nonzero vector, $\underline{x}$, such that $A\underline{x} = \underline{0}$ then clearly A does not have full column rank.

(c) If you can find a set of m columns of A such that the determinant of the $m \times m$ submatrix given by these columns is nonzero, then A has full row rank.

(d) If you can find a nonzero vector, $\underline{z}$, such that $\underline{z}^T A = \underline{0}$ then clearly A does not have full row rank.

3 Observability

A system:

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

is called **observable** if we can deduce the state, $\underline{x}(t)$, from measurements of $\underline{u}(\tau)$ and $\underline{y}(\tau)$ over some time interval.

Now consider differentiating $\underline{y}(t)$ to give

$$
\underbrace{\begin{bmatrix} \underline{y}(t) \\ \dot{\underline{y}}(t) \\ \ddot{\underline{y}}(t) \\ \vdots \\ \underline{y}^{(n-1)}(t) \end{bmatrix}}_{\text{known}} = \underbrace{\begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{bmatrix}}_{Q} \underbrace{\underline{x}(t)}_{?} + \underbrace{\begin{bmatrix} \underline{0} \\ CB\underline{u}(t) \\ CAB\underline{u}(t) + CB\dot{\underline{u}}(t) \\ \vdots \\ CA^{n-2}B\underline{u} + \cdots + CB\underline{u}^{(n-2)} \end{bmatrix}}_{\text{known}}
$$

We can solve the above equation uniquely for $\underline{x}(t)$ if and only if rank $Q = n$. Hence, defining the **observability matrix**

$$
Q = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix}
$$

we obtain the **Observability test:**

$$
\boxed{\text{The system is observable if and only if rank } Q = n}
$$

If a system is *not* observable, there will exist a vector $\underline{x}_o \neq 0$ for which $Q\underline{x}_o = 0$. This is called an **unobservable state**, for the following reason.

$$
\begin{aligned}
Q\underline{x}_o = \underline{0} \quad &\Rightarrow \quad CA^k\underline{x}_o = \underline{0} \text{ for } k = 0, \cdots n-1 \\
&\Rightarrow \quad CA^n\underline{x}_o = C\left(-\alpha_1 A^{n-1}\cdots - \alpha_{n-1}A - \alpha_n I\right)\underline{x}_o \text{ by Cayley-Hamilton Theorem} \\
&= \quad \underline{0} \\
&\Rightarrow \quad CA^k\underline{x}_o = \underline{0} \text{ for all } k \geq 0 \text{ by repeated use of Cayley-Hamilton theorem.} \\
&\Rightarrow \quad Ce^{At}\underline{x}_o = \underline{0} \text{ for all } t \text{ by the power series expansion of } e^{At}
\end{aligned}
$$

Conversely, $Ce^{At}\underline{x}_o = 0$ for all t implies $\frac{d^n}{dt^n}Ce^{At}\underline{x}_o = CA^ne^{At}\underline{x}_o = \underline{0}$ and so $Q\underline{x}_o = \underline{0}$.
Hence $Ce^{At}\underline{x}_0 = \underline{0}$ for all $t \iff Q\underline{x}_0 = \underline{0}$.

Recall that

$$
\underline{y}(t) = \underbrace{Ce^{At}\underline{x}(0)}_{\text{initial condition response}} \underbrace{+ D\underline{u}(t) + \int_0^t Ce^{A(t-\tau)}B\underline{u}(\tau)\,d\tau}_{\text{input response}}
$$

and so if two initial states $\underline{x}_1 \neq \underline{x}_2$ give the same outputs then $\underline{0} = \underline{y}_2 - \underline{y}_1 = Ce^{At}(\underline{x}_2 - \underline{x}_1)$. In this case, $\underline{x}_o = \underline{x}_1 - \underline{x}_2$ is an unobservable state.

3.1 Effect of Initial Condition on Output

Now consider the difference between two initial condition responses:

$$
\underline{y}_o(t) = Ce^{At}\underline{x}_o \text{ and } \underline{y}(t) = Ce^{At}\left(\underline{x}_o + \underline{d}\right) \quad \text{so} \quad \underline{y}(t) - \underline{y}_o(t) = Ce^{At}\underline{d}
$$

Can $\left(\underline{y}(t) - \underline{y}_o(t)\right)$ be small in spite of $\underline{d}$ being large? Measure the size of $\left(\underline{y}(t) - \underline{y}_o(t)\right)$ over the time interval $0 < t < t_1$ by

$$
\begin{aligned}
\int_0^{t_1} \left\|\underline{y}(t) - \underline{y}_o(t)\right\|^2 dt &= \int_0^{t_1} \left(\underline{y}(t) - \underline{y}_o(t)\right)^T \left(\underline{y}(t) - \underline{y}_o(t)\right) dt \\
&= \int_0^{t_1} \underline{d}^T\, e^{A^T t} C^T C e^{At}\, \underline{d}\, dt = \underline{d}^T\, W_o(t_1)\, \underline{d} \text{ where } W_o(t_1) = \int_0^{t_1} e^{A^T t} C^T C e^{At}\, dt
\end{aligned}
$$

Clearly this difference must be ≥ 0 so $W_o(t_1)$ is a positive semi-definite matrix. The system will be observable if $\underline{d}^T W_o(t_1)\underline{d} > 0$ for all $\underline{d} \neq \underline{0}$, i.e. if $W_o(t_1)$ is a positive definite matrix.

Also,

$$
\begin{aligned}
\underline{d} \text{ in Null Space of } W_o(t_1) \quad &\Leftrightarrow \quad W_o(t_1)\underline{d} = \underline{0} \Leftrightarrow \underline{d}^T W_o(t_1)\underline{d} = \underline{0} \Leftrightarrow Ce^{At}\underline{d} = \underline{0} \text{ for all } t < t_1 \\
&\Leftrightarrow \quad \underline{d} \text{ is an unobservable state.} \\
\Rightarrow \text{Null Space of } W_o(t_1) \quad &= \quad \text{Null Space of } Q.
\end{aligned}
$$

Example

$$\underline{\dot{x}} = \begin{bmatrix} -1 & 0 \\ 0 & -2 \end{bmatrix} \underline{x}, \; y = \begin{bmatrix} 1 & 1 \end{bmatrix} \underline{x} \;\; \Rightarrow C e^{At} = \begin{bmatrix} e^{-t} & e^{-2t} \end{bmatrix}$$

$$W_o(t_1) \;\; = \;\; \int_0^{t_1} \begin{bmatrix} e^{-2t} & e^{-3t} \\ e^{-3t} & e^{-4t} \end{bmatrix} dt = \begin{bmatrix} \frac{1}{2}\left(1 - e^{-2t_1}\right) & \frac{1}{3}\left(1 - e^{-3t_1}\right) \\ \frac{1}{3}\left(1 - e^{-3t_1}\right) & \frac{1}{4}\left(1 - e^{-4t_1}\right) \end{bmatrix} \xrightarrow[\text{as } t_1 \to \infty]{} \begin{bmatrix} \frac{1}{2} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{4} \end{bmatrix}$$

3.2 Change of State Coordinates when System is not Observable

If (A, C) is not observable then we can make a change of state coordinates to isolate the unobservable states as follows.

If the rank $Q = r < n$ then there exists a nonsingular $n \times n$ matrix T and a $pn \times r$ matrix $\tilde{Q}_1$ of rank r, such that *(Recall QR factorization)*

$$Q = \begin{bmatrix} \tilde{Q}_1 & 0 \end{bmatrix} T$$

Now change the state coordinates to $\underline{\tilde{x}} = T\underline{x}$:

$$\underline{\dot{\tilde{x}}} = \underbrace{TAT^{-1}}_{\tilde{A}} \underline{\tilde{x}} + \underbrace{TB}_{\tilde{B}} \underline{u}, \quad \underline{y} = \underbrace{CT^{-1}}_{\tilde{C}} \underline{\tilde{x}}.$$

Theorem 3.1 *In these coordinates if we partition the state,* $\underline{\tilde{x}} = \begin{bmatrix} \underline{\tilde{x}}_1 \\ \underline{\tilde{x}}_2 \end{bmatrix}$ *with* $\underline{\tilde{x}}_1$ *of dimension r, and compatibly partition:*

$$\tilde{A} = \begin{bmatrix} \tilde{A}_{11} & \tilde{A}_{12} \\ \tilde{A}_{21} & \tilde{A}_{22} \end{bmatrix}; \;\; \tilde{B} = \begin{bmatrix} \tilde{B}_1 \\ \tilde{B}_2 \end{bmatrix}; \;\; \tilde{C} = \begin{bmatrix} \tilde{C}_1 & \tilde{C}_2 \end{bmatrix}$$

then

$$\tilde{C}_2 = 0, \quad \tilde{A}_{12} = 0, \quad \textit{and } (\tilde{A}_{11}, \tilde{C}_1) \textit{ is observable}$$

Proof: Firstly $\tilde{C}\tilde{A}^k = CT^{-1}TA^kT^{-1} = CA^kT^{-1}$ and so the observability matrix in the transformed coordinates is given by

$$\tilde{Q} = \begin{bmatrix} \tilde{C} \\ \tilde{C}\tilde{A} \\ \vdots \\ \tilde{C}\tilde{A}^{n-1} \end{bmatrix} = \begin{bmatrix} CT^{-1} \\ CAT^{-1} \\ \vdots \\ CA^{n-1}T^{-1} \end{bmatrix} = QT^{-1} = \begin{bmatrix} \tilde{Q}_1 & 0 \end{bmatrix}$$

Hence

$$\tilde{Q} \begin{bmatrix} 0 \\ I_{n-r} \end{bmatrix} = \begin{bmatrix} \tilde{Q}_1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ I_{n-r} \end{bmatrix} = 0$$

From which it follows that

$$\tilde{C}\tilde{A}^k \begin{bmatrix} 0 \\ I_{n-r} \end{bmatrix} = 0 \text{ for all } k.$$

In particular, $\tilde{C}_2 = 0$. Furthermore

$$\tilde{Q}\tilde{A} \begin{bmatrix} 0 \\ I_{n-r} \end{bmatrix} = \begin{bmatrix} \tilde{C}\tilde{A} \\ \tilde{C}\tilde{A}^2 \\ \vdots \\ \tilde{C}\tilde{A}^n \end{bmatrix} \begin{bmatrix} 0 \\ I_{n-r} \end{bmatrix} = 0$$

But

$$\tilde{Q}\tilde{A} \begin{bmatrix} 0 \\ I_{n-r} \end{bmatrix} = \begin{bmatrix} \tilde{Q}_1 & 0 \end{bmatrix} \begin{bmatrix} \tilde{A}_{11} & \tilde{A}_{12} \\ \tilde{A}_{21} & \tilde{A}_{22} \end{bmatrix} \begin{bmatrix} 0 \\ I_{n-r} \end{bmatrix} = \tilde{Q}_1\tilde{A}_{12}$$

which implies that $\tilde{A}_{12} = 0$ since $\tilde{Q}_1$ is full column rank.

Hence in these state coordinates we have,

$$\dot{\tilde{\underline{x}}}_1 = \tilde{A}_{11}\tilde{\underline{x}}_1 + \tilde{B}_1\underline{u}, \quad \underline{y} = \tilde{C}_1\tilde{\underline{x}}_1$$

and the input/output response (i.e. the transfer function) depends only on $\tilde{\underline{x}}_1$ and the states $\tilde{\underline{x}}_2$ are all unobservable.

3.2.1 A subspace interpretation

As before, we start by factorising Q as $Q = \begin{bmatrix} \tilde{Q}_1 & 0 \end{bmatrix} T$.

Now put $T^{-1} = [X\ Y]$.

Y in $\mathbb{R}^{n \times r}$ is a basis for null(Q), which we shall call $\bar{O}$, the unobservable subspace. (i.e. whenever $a = Yb$, $Qa = 0$)

and X complements Y

(i.e. range$[X\ Y] = \mathbb{R}^n$ and, whenever $a_1 = Yb_1$ and $a_2 = Xb_2$, then $a_1^T a_2 = 0$.)

Note that $A\bar{O} \subseteq \bar{O}$ and $\bar{O} \subseteq$ null(C).

Since $AT^{-1} = T^{-1}\hat{A}$, we have

$$A[X\ Y] = [X\ Y] \begin{bmatrix} \hat{A}_{11} & \hat{A}_{12} \\ \hat{A}_{21} & \hat{A}_{22} \end{bmatrix}$$

or

$$AY = [X\ Y] \begin{bmatrix} \hat{A}_{12} \\ \hat{A}_{22} \end{bmatrix} = X\hat{A}_{12} + Y\hat{A}_{22}$$

and so $\hat{A}_{12} = 0$.

Also $CT^{-1} = \hat{C}$, i.e.

$$C[X\ Y] = [\hat{C}_1\ 0]$$

4 Observers

4.1 Differentiating signals is a bad idea

Typically the state is not available for measurement,
but we can estimate $\underline{x}(t)$ from $\underline{y}$ and $\underline{u}$

In the section on observability we saw how to exactly deduce $\underline{x}(t)$ from

$$y, \dot{y}, .., y^{(n-1)}, u, \dot{u}, ..u^{(n-2)}$$

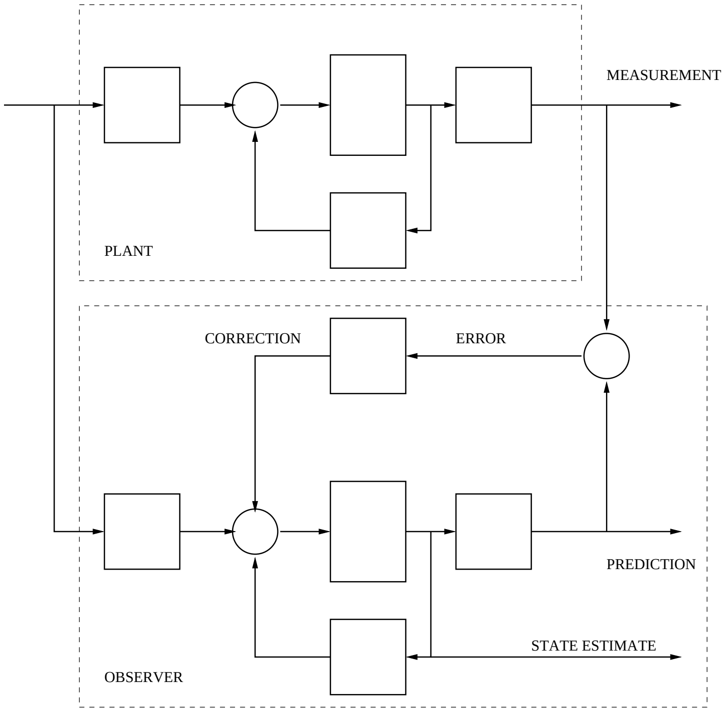
but differentiating signals has bad noise amplification problems:

$$\begin{array}{rcll} y(t) & = & \sin \omega t + \epsilon \sin \omega_n t & \text{S/N ratio } = 1/\epsilon \\ \dot{y}(t) & = & \omega \cos \omega t + \epsilon \omega_n \cos \omega_n t & \text{S/N ratio } = (\omega/\epsilon\omega_n) \\ \ddot{y}(t) & = & -\omega^2 \sin \omega t - \epsilon \omega_n^2 \sin \omega_n t & \text{S/N ratio } = \frac{1}{\epsilon}\left(\frac{\omega}{\omega_n}\right)^2 \end{array}$$

4.2 Observer structure

Instead we will use a *state observer* (Luenberger Observer) which contains a dynamic model of the system and whose state, $\underline{\hat{x}}(t)$, approaches $\underline{x}(t)$ as $t \to \infty$.

$$\left\{ \begin{array}{rcl} \underline{\dot{\hat{x}}} & = & A\underline{\hat{x}} + B\underline{u} + L(\underline{y} - \underline{\hat{y}}) \\ \underline{\hat{y}} & = & C\underline{\hat{x}} \end{array} \right.$$



Consider the error $\underline{e}(t) = \underline{x}(t) - \underline{\hat{x}}(t)$

$$\begin{aligned} \underline{\dot{e}} &= \underline{\dot{x}} - \dot{\hat{x}} = (A\underline{x} + B\underline{u}) - (A\underline{\hat{x}} + B\underline{u} + L(\underline{y} - \underline{\hat{y}})) \\ &= A(\underline{x} - \underline{\hat{x}}) - LC(\underline{x} - \underline{\hat{x}}) = (A - LC)\underline{e} \end{aligned}$$

$$\boxed{\underline{\dot{e}} = (A - LC)\underline{e}}$$

We want $e^{(A-LC)t} \to 0$ quickly as t increases.

This is achieved if the eigenvalues of $(A - LC)$ are large and negative, for example.

Can we assign the eigenvalues of $(A - LC)$ by choice of L?

Suppose (A, C) is **not** observable then in section 3.2 we found a change of coordinates, $\underline{\tilde{x}} = T\underline{x}$ such that,

$$\begin{bmatrix} \dot{\underline{\tilde{x}}}_1 \\ \dot{\underline{\tilde{x}}}_2 \end{bmatrix} = \begin{bmatrix} \tilde{A}_{11} & 0 \\ \tilde{A}_{21} & \tilde{A}_{22} \end{bmatrix} \begin{bmatrix} \underline{\tilde{x}}_1 \\ \underline{\tilde{x}}_2 \end{bmatrix} + \begin{bmatrix} \tilde{B}_1 \\ \tilde{B}_2 \end{bmatrix} \underline{u}, \quad \underline{y} = \begin{bmatrix} \tilde{C}_1 & 0 \end{bmatrix} \underline{\tilde{x}} + D\underline{u}$$

Hence

$$T(A - LC)T^{-1} = \tilde{A} - \tilde{L}\tilde{C} = \begin{bmatrix} \tilde{A}_{11} & 0 \\ \tilde{A}_{21} & \tilde{A}_{22} \end{bmatrix} - \begin{bmatrix} \tilde{L}_1 \\ \tilde{L}_2 \end{bmatrix} \begin{bmatrix} \tilde{C}_1 & 0 \end{bmatrix} = \begin{bmatrix} (\tilde{A}_{11} - \tilde{L}_1\tilde{C}_1) & 0 \\ (\tilde{A}_{21} - \tilde{L}_2\tilde{C}_1) & \tilde{A}_{22} \end{bmatrix},$$

and the eigenvalues of the observer,

$$\lambda_i(A - LC) = \lambda_i(\tilde{A} - \tilde{L}\tilde{C}) = \lambda_i(\tilde{A}_{11} - \tilde{L}_1\tilde{C}_1) \cup \lambda_i(\tilde{A}_{22}),$$

and $\lambda_i(\tilde{A}_{22})$ are not changed by $\tilde{L}$.

However *it can be shown that*

We can arbitrarily assign the eigenvalues of $(A - LC)$ by choice of L if and only if the system is observable.

- We can thus make the error, $\underline{e}(t) \to 0$ arbitrarily quickly.

- But high gains might imply very large transient errors and noisy estimates.

4.3 Tracking disturbances, ignoring noise

Imagine tracking aircraft by radar (1-D). Aircraft position z is affected by random turbulence. Take $\underline{x} = [z, \dot{z}]^T$:

$$\dot{\underline{x}}(t) = A\underline{x}(t) + Bd(t) = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \underline{x}(t) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} d(t)$$

The radar measurement is corrupted by noise:

$$y(t) = C\underline{x}(t) + n(t) = [1 \quad 0]\underline{x}(t) + n(t)$$

Observer: $\dot{\hat{\underline{x}}}(t) = A\hat{\underline{x}}(t) + L[y(t) - C\hat{\underline{x}}(t)]$ NB: $d(t)$ not known, so not used.

d large, n small: Believe the measurements. Use large L. *React quickly.*
d small, n large: Don't trust measurements, believe model. Use small L.
— *Smooth the measurements.*

4.4 Kalman Filter

Assume we have measurements of $\underline{u}(t)$ and $\underline{y}(t)$ and the model

$$\dot{\underline{x}} = A\underline{x} + B(\underline{u} + \underline{d})$$
$$\underline{y} = C\underline{x} + \underline{n}$$

What are the *smallest* $\underline{d}$ and $\underline{n}$, in terms of $(\int_0^\infty \underline{d}^T \underline{d}\, dt)^2 + (\int_0^\infty \underline{n}^T \underline{n}\, dt)^2$, which make the measurement consistent with the model, and what is the corresponding estimate of the state?

The solution is given by *Kalman Filter* theory, which gives an
optimal trade-off between tracking d and rejecting n. The solution is a Luenberger observer with $L = \Sigma C^T$ where $\Sigma > 0$ solves the quadratic matrix equation

$$A\Sigma + \Sigma A^T + BB^T - \Sigma C^T C\Sigma = 0$$

(if the system is observable, then it can be shown that such a solution exists, is unique, and that the resulting observer is stable).

Generalises to arbitrary disturbance/noise spectra. Very widely used *Navigation & guidance, Telecomms, Control, Finance, . . .*
Especially in discrete time. Software implementation *Matlab*: `kalman, dkalman, estim` etc.

4.5 Application to sensor fusion

Satellite, 1 axis of rotation: $J\ddot{\theta} = u + d$ (u = control torque, d = disturbance torque).

Two noisy sensors: Star sensor: $y_1 = \theta + n_\theta$, Rate gyro: $y_2 = \dot{\theta} + n_\omega$

Let $\underline{x} = [\theta, \dot{\theta}]^T$. State-space model:

$$
\begin{aligned}
\dot{\underline{x}} &= \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \underline{x} + \begin{bmatrix} 0 & 0 \\ 1/J & 1/J \end{bmatrix} \begin{bmatrix} u \\ d \end{bmatrix} \\
\underline{y} &= \begin{bmatrix} \theta \\ \dot{\theta} \end{bmatrix} + \begin{bmatrix} n_\theta \\ n_\omega \end{bmatrix} = I\underline{x} + \begin{bmatrix} n_\theta \\ n_\omega \end{bmatrix}
\end{aligned}
$$

Observable? Yes. ($C = I$, so rank $C = 2$, so rank$\begin{bmatrix} C \\ CA \end{bmatrix} = 2$.)

Observer:

$$
\begin{aligned}
\dot{\hat{\underline{x}}} &= A\hat{\underline{x}} + B\begin{bmatrix} u \\ 0 \end{bmatrix} + L(\underline{y} - C\hat{\underline{x}}) \qquad (d \text{ not known}) \\
&= (A - LC)\hat{\underline{x}} + \begin{bmatrix} 0 \\ 1/J \end{bmatrix} u + L\underline{y} \qquad \text{but } C = I \text{ so:} \\
&= \begin{bmatrix} -\ell_{11} & 1 - \ell_{12} \\ -\ell_{21} & -\ell_{22} \end{bmatrix} \hat{\underline{x}} + \begin{bmatrix} 0 \\ 1/J \end{bmatrix} u + \begin{bmatrix} \ell_{11} & \ell_{12} \\ \ell_{21} & \ell_{22} \end{bmatrix} \underline{y}
\end{aligned}
$$

Place both eigenvalues at -10 (say): Using trace$(A - LC) = \sum_i \lambda_i$ and det$(A - LC) = \prod_i \lambda_i$: $-\ell_{11} - \ell_{22} = -20$ and $\ell_{11}\ell_{22} + \ell_{21}(1 - \ell_{12}) = 100$. This leaves some design freedom.

$n_\theta \ll n_\omega$: Make $\ell_{11} \gg \ell_{12}$ and $\ell_{21} \gg \ell_{22}$.

$n_\theta \gg n_\omega$: Make $\ell_{11} \ll \ell_{12}$ and $\ell_{21} \ll \ell_{22}$.

Optimal trade-off: *Kalman Filter* again.

4.6 Application to sensor bias estimation

Satellite, as before: $J\ddot{\theta} = u$

Sensors: Star tracker measures angular position: $y_1 = \theta$
Rate gyro measures angular velocity with bias: $y_2 = \dot{\theta} + b_\omega$.

Augment state vector: $\underline{x} = [\theta, \dot{\theta}, b_\omega]^T$, and assume bias is constant: $\dot{b}_\omega = 0$.

State-space model:

$$\begin{aligned}
\dot{\underline{x}} &= \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \underline{x} + \begin{bmatrix} 0 \\ 1/J \\ 0 \end{bmatrix} u \\
\underline{y} &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix} \underline{x}
\end{aligned}$$

Is the state observable?

$$\begin{bmatrix} C \\ CA \\ CA^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

First 3 rows are linearly independent (Or: All three columns are linearly independent).

So rank = 3. Hence: **Observable**. So can use observer to estimate $\underline{x}$:

$$\dot{\hat{\underline{x}}} = A\hat{\underline{x}} + Bu + L(y - C\hat{\underline{x}})$$

$A - LC$ stable $\Rightarrow \hat{x}_3 \to b_\omega$ as $t \to \infty$. Rate of convergence depends on eigenvalues of $A - LC$.