Lecture 25: Observability and Constructibility

7.4.1 Causal (Forward) Observer

A more practical observer law should give the current state \( x(t) \) according to the constructibility definition, e.g., by restricting the tail to the interval \([t_1, t]\). Let \( t_2 = t \) in the expression of the optimal state:

\[
(x(t))_{opt} = \Phi(t, t_1)(x_{t_1})_{opt} + x_{n_{[t_1, t]}}(t)
\]

\[
= \Phi(t, t_1)N^{-1}(t_1, t) \int_{t_1}^{t} \Phi^*(\tau, t_1)C^*(\tau) \left( y_1(\tau) + v(\tau) \right) d\tau + x_{n_{[t_1, t]}}(t) \quad .
\] (7.25)

\[
= \Phi(t, t_1)N^{-1}(t_1, t) \Phi^*(t, t_1) \int_{t_1}^{t} \Phi^*(\tau, t)C^*(\tau) \left( y_1(\tau) + v(\tau) \right) d\tau + x_{n_{[t_1, t]}}(t)
\]

Definition: Forward Observability Grammian

The forward observability Grammian is defined as:

\[
V(t_1, t) := \left[ \Phi(t, t_1)N^{-1}(t_1, t)\Phi^*(t, t_1) \right]^{-1}
\]

\[
= \int_{t_1}^{t} \Phi^*(\tau, t)C^*(\tau)C(\tau)\Phi(\tau, t) d\tau \quad .
\] (7.26)

The optimal observer law for the current state is thus given by:

\[
(x(t))_{opt} = V^{-1}(t_1, t) \int_{t_1}^{t} \Phi^*(\tau, t)C^*(\tau) \left( y_1(\tau) + v(\tau) \right) d\tau + \mathcal{L}u_{[t_1, t]} \quad .
\] (7.27)

The above integral defines a causal system, and the last term represents the forced component of the state produced by the input \( u \) restricted to the interval \([t_1, t]\).

Notes:

1. The \( L_2 \) optimal observer in (7.27), which removes errors caused by noise orthogonal to \( \mathcal{R}(\mathcal{L}) \), is sometimes called a deterministic Kalman filter.

2. In the absence of noise, the estimate \((x(t))_{opt}\) does not change if it is based on an observation of the tail \( y_1(\cdot) \) over the interval \([t_1, t_2] \supseteq [t_1, t]\), and we have an analogy to the principle of
optimality. If noise is present, this is no longer the case; lengthening the interval may improve or deteriorate the estimate.

3. $\mathcal{L}_2$-optimal observers have memory (because of the integral in (7.27)) unlike the optimal state feedback control law, which is memoryless. In fact, the integral in (7.27) is the controllability operator of the system: $\dot{z}(t) = A^*(t)z(t) + C^*(t)(y_1(t) + v(t))$

Block diagram of causal optimal observer

Recursion Interpretation of Observer Law

The observer law of (7.27) is has an equivalent simpler form as given in the theorem below.
Theorem:

Suppose there exists on \([t_1, t]\) a solution to the differential Riccati equation:

\[
\frac{d}{dt} V^{-1}(t_1, t) = A(t)V^{-1}(t_1, t) + V^{-1}(t_1, t)A^*(t) - V^{-1}(t_1, t)C^*(t)C(t)V^{-1}(t_1, t),
\]

\[V^{-1}(t_1, t_1) = \infty\]  \(7.28\)

Then, the “recursive” version of the \(L_2\)-optimal observer law is given by:

\[
\hat{x}(t) = A(t)\hat{x}(t) + Bu(t) + V^{-1}(t_1, t)C^*(t)[y_1(\tau) + v(\tau) - C(t)\hat{x}(t)].
\]

\[\hat{x}(t_1) = 0\]  \(7.29\)

The block diagram is shown below.

Note that by applying the optimal linear quadratic state feedback gain on the above state estimate, we obtain the so-called certainty equivalence controller.

### 7.4.2 Observer Example and Interpretation for an Integrator System

Consider the linear time-invariant scalar integrator system used to integrate an accelerometer's signal:
\[
\dot{x}(t) = u(t) \\
y(t) = Cx(t) \\
z(t) = Cx(t) + v(t)
\] (7.30)

**Optimal Observer/Filtering Problem:**

We wish to find optimal estimates \( \hat{x}(t), \hat{y}(t) \) of the state of the integrator \( x(t) \), and the velocity \( y(t) \) from measurements of \( u(t) \) and \( y(t) + v(t) \).

Let's derive the optimal observer from a knowledge of the form of the solution rather than by using the pseudoinverse formula. Here, the tail \( y_1 \), defined on \([t_1, t]\) (where \( t \) is fixed,) is constrained to be a step signal \( q(t - t_1) \), which is the zero-input response of an integrator. Thus, \( y_1 \in \mathcal{R}\{El\} = \text{span}\{q(t - t_1)\} \), or

\[
y_1(t) = \xi q(t - t_1)
\] (7.31)

The observability equation is:

\[
(Elx_1)(t) = y_1(t) = C\Phi(t, t_1)x_1 = Cx_1q(t - t_1).
\] (7.32)

Determination of the optimal tail estimate \( \hat{y}_1 \) boils down to finding the constant \( \xi \), in (7.31), and the corresponding optimal initial state \( \hat{x}_1 \).

Parameter \( \xi \) can be found from the knowledge that the filtering error \( y_1 + v - \hat{y}_1 \) must be orthogonal to \( \mathcal{R}\{El\} \). Thus,

\[
\langle y_1 + v - \hat{y}_1, \hat{y}_1 \rangle = 0 \\
\int_{t_1}^{t} \left[ y_1(\tau) + v(\tau) - \hat{y}_1(\tau) \right]^T \hat{y}_1(\tau) d\tau = 0
\] (7.33)

which, because the signal are real and scalar valued, reduces to

\[
\int_{t_1}^{t} \left[ y_1(\tau) + v(\tau) - \xi \right] \xi d\tau = 0
\]

\[\Leftrightarrow \int_{t_1}^{t} \xi d\tau = \int_{t_1}^{t} \left[ y_1(\tau) + v(\tau) \right] d\tau
\] (7.34)

\[\Leftrightarrow \xi = \frac{1}{t - t_1} \int_{t_1}^{t} \left[ y_1(\tau) + v(\tau) \right] d\tau\]
which is the time average of the noisy tail. Hence the optimal tail estimate for \( t > t_i \) is given by:

\[
\hat{y}_i(t) = \frac{1}{t - t_i} \int_{t_i}^{t} [y_i(\tau) + v(\tau)] d\tau
\]  

(7.35)

and the corresponding optimal state estimate is

\[
\hat{x}(t) = \frac{1}{C(t - t_i)} \int_{t_i}^{t} [y_i(\tau) + v(\tau)] d\tau.
\]  

(7.36)

The state estimate error is

\[
x(t) - \hat{x}(t) = \frac{1}{C(t - t_i)} \int_{t_i}^{t} v(\tau) d\tau.
\]  

(7.37)

Notes:

1. In many stochastic problems the state estimate error approaches a constant value as \( t \to \infty \). Clearly, if the deterministic noise in our example has a nonzero DC component, the error will ramp up with time.

2. The state estimate error will tend to some steady-state value regardless (assuming the noise is zero-mean) if the original system is stable or not (our example here is an unstable integrator.)

### 7.5 Generalized Pseudoinverse

Recall that if the linear operator \( \mathcal{A} \) on Hilbert space is onto, then its Moore-Penrose pseudoinverse is defined as \( \mathcal{A}^{\oplus} := \mathcal{A}^* \left( \mathcal{A} A^* \right)^{-1} \), and the solution \( u_{opt} = \mathcal{A}^{\oplus} v \) to \( v = \mathcal{A} u \) has minimum norm.

We can generalize the pseudoinverse by giving a separate definition when \( \mathcal{A} \) is one-to-one:

\[
\mathcal{A}^{\ominus} := \left( A^* A \right)^{-1} A^*.
\]  

(7.38)

The pseudoinverse solution in this case \( u_{opt} = \mathcal{A}^{\ominus} v \) minimizes \( \| v - \mathcal{A} u_{opt} \| \) by knocking off the orthogonal component of the error in \( \mathcal{N}\left( A^* \right) \).

Notes:

1. \( \mathcal{A}^{\ominus} \) for \( \mathcal{A} \) one-to-one is a left inverse, i.e., \( \mathcal{A}^{\ominus} A = \left( A^* A \right)^{-1} A^* A = I \).

2. If \( \mathcal{A}^{-1} \) exists, then \( \mathcal{A}^{-1} = \mathcal{A}^{\ominus} \).
This is the pseudoinverse (of $\mathcal{L}$) that we used to find the optimal observer, assuming the system was C.O.

### 7.6 Duality and Observability of LTI Systems

#### 7.6.1 Duality

For a given LTV system $S := (A(\cdot), B(\cdot), C(\cdot), D(\cdot))$, the controllability operator $\mathcal{L} : L_2^\infty \rightarrow \mathbb{R}^n$ and the adjoint of the observability operator $\mathcal{L}^* : L_2^\infty \rightarrow \mathbb{R}^n$ look very similar:

\[
(\mathcal{L} u)(t_1) := \int_{t_0}^{t_1} \Phi(t_1, \tau) B(\tau) u(\tau) d\tau
\]  
(7.39)

\[
(\mathcal{L}^* y)(t_1) := \int_{t_0}^{t_1} \Phi^*(\tau, t_1) C^*(\tau) y(\tau) d\tau
\]  
(7.40)

except that $C^*(\tau)$ takes the role of $B(\tau)$, and $\Phi^*(\tau, t_1)$ is substituted for $\Phi(t_1, \tau)$.

**Question:**

Does there exists a dual system $(A_d(\cdot), B_d(\cdot), C_d(\cdot), D_d(\cdot))$ such that the adjoint of its observability operator is equal to the controllability operator of $(A(\cdot), B(\cdot), C(\cdot), D(\cdot))$, i.e., $\mathcal{L}^* = \mathcal{L}$?

The answer is yes, and this system is given by:

\[
S_d := (A_d(\cdot), B_d(\cdot), C_d(\cdot), D_d(\cdot)) := (-A^*(\cdot), C^*(\cdot), B^*(\cdot), D^*(\cdot))
\]  
(7.41)

and this dual system is assumed to be running backwards in time.

#### 7.6.2 LTI Observability by Duality

Consider the continuous-time LTI system $S := (A, B, C, D)$. Let the observability matrix be defined as:

\[
\mathcal{O} := \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix}
\]  
(7.42)

We have the following result on observability of $S$:
Theorem:

(a) The set of observable states of $S$ is given by: $\mathcal{N}\{\mathcal{E}\}^\perp = \mathcal{R}\{\mathcal{O}\}^\perp$

(b) $S$ is C.O. $\iff$ rank $\{\mathcal{O}\} = n$.

Proof:

Apply the controllability results to the dual system $S_d := (-A^\ast, B^\ast, C^\ast, D^\ast)$ ...

7.6.3 Decomposition into Observable and Unobservable Parts

Suppose the LTI system $S := (A, B, C, D)$ is not completely observable.

Theorem:

There exists a similarity transformation $T$, $\tilde{x} = Tx$ such that the system:

\[
\dot{\tilde{x}}(t) = \begin{bmatrix} TAT^{-1} & \frac{T}{B} \\ \frac{B}{A} & \frac{C}{A} \end{bmatrix} \tilde{x}(t) + \begin{bmatrix} \frac{T}{B} \\ \frac{B}{A} \end{bmatrix} u(t)
\]

\[
y(t) = \begin{bmatrix} \frac{C}{A} \end{bmatrix} \tilde{x}(t) + Du(t)
\]

is separated into its observable and unobservable parts, i.e.,

\[
\tilde{A} = \begin{bmatrix} \tilde{A}_{11} & \theta_{q\times(n-q)} \\ \tilde{A}_{12} & \tilde{A}_{22} \end{bmatrix}, \quad \tilde{C} = \begin{bmatrix} \tilde{C}_1 & \theta_{p\times(n-q)} \end{bmatrix}, \quad \tilde{A}_{11} \in \mathbb{R}^{q\times q}, \tilde{A}_{12} \in \mathbb{R}^{(n-q)\times q}, \tilde{A}_{22} \in \mathbb{R}^{(n-q)\times(n-q)}, \tilde{C}_1 \in \mathbb{R}^{p\times q}
\]

(7.43)

Proof:

Use duality: Let $A_d := A^\ast$, $B_d := C^\ast$, find a similarity transformation that decomposes $(A_d, B_d)$ into its controllable and uncontrollable parts, and get $(\tilde{C}, \tilde{A})$ by taking the transpose of $(\tilde{A}_d, \tilde{B}_d)$ ...

7.7 Kalman's Decomposition Theorem

Let $(A, B)$ be not completely controllable and $(C, A)$ be not completely observable, then there exists a similarity transformation $Q$ such that the system $(A, B, C)$ can be separated into controllable and observable parts as follows:
\[
\tilde{A} = Q^{-1}AQ = \begin{bmatrix}
\tilde{A}_{11} & 0 & \tilde{A}_{13} & 0 \\
0 & \tilde{A}_{22} & \tilde{A}_{23} & \tilde{A}_{24} \\
0 & 0 & \tilde{A}_{33} & 0 \\
0 & 0 & \tilde{A}_{43} & \tilde{A}_{44}
\end{bmatrix},
\]

(7.45)

\[
\tilde{B} = Q^{-1}B = \begin{bmatrix}
\tilde{B}_1 \\
\tilde{B}_2 \\
0
\end{bmatrix}, \quad \tilde{C} = CQ = \begin{bmatrix}
\tilde{C}_1 & 0 & \tilde{C}_3 & 0
\end{bmatrix}
\]

where \((A_C, B_C), \quad A_C := \begin{bmatrix}
\tilde{A}_{11} & 0 \\
0 & \tilde{A}_{22}
\end{bmatrix}, \quad B_C = \begin{bmatrix}
\tilde{B}_1 \\
\tilde{B}_2
\end{bmatrix}\) is C.C., and

\((C_O, A_O), \quad A_O := \begin{bmatrix}
\tilde{A}_{11} & \tilde{A}_{13} \\
0 & \tilde{A}_{33}
\end{bmatrix}, \quad C_O = \begin{bmatrix}
\tilde{C}_1 & \tilde{C}_3
\end{bmatrix}\) is C.O.

Proof:

First apply the decomposition into controllable and uncontrollable parts, and then apply the decomposition into observable and unobservable parts.

7.8 Observable Canonical Form Realization

We now derive a state-space model for the system below called observable canonical form realization. The idea is to write the transfer function \(H(s)\) as a rational function of \(s^{-1}\):

\[
H(s) = \frac{b_M s^{M-N} + b_{M-1}s^{M-N-1} + \cdots + b_0 s^{-N}}{1 + a_{N-1}s^{-1} + \cdots + a_0 s^{-N}}.
\]

(7.46)

Assume without loss of generality that \(M = N\) (if it is not equal, then just set \(b_N = b_{N-1} = \cdots = b_{M+1} = 0\) ) so that

\[
H(s) = \frac{b_N s^{-1} + \cdots + b_0 s^{-N}}{1 + a_{N-1}s^{-1} + \cdots + a_0 s^{-N}}.
\]

(7.47)

Then, the input-output relationship between \(U(s)\) and \(Y(s)\) can be written as

\[
Y(s) = b_N U(s) + (b_{N-1}U(s) - a_{N-1}Y(s))s^{-1} + (b_{N-2}U(s) - a_{N-2}Y(s))s^{-2} + \cdots + (b_0 U(s) - a_0 Y(s))s^{-N}
\]

(7.48)
The interpretation is that the output is a linear combination of successive integrals of linear combinations of itself and the input. The block diagram for the observable canonical form consists of a chain of \( N \) integrators with summing junctions at the input of each integrator.

The state variables are defined as the integrator outputs, which gives us

\[
\begin{align*}
    sX_1(s) &= -a_0 X_N(s) + (b_0 - a_0 b_N) U(s) \\
    sX_2(s) &= X_1(s) - a_1 X_N(s) + (b_1 - a_1 b_N) U(s) \\
    sX_3(s) &= X_2(s) - a_2 X_N(s) + (b_2 - a_2 b_N) U(s) \\
    &\vdots \\
    sX_N(s) &= X_{N-1}(s) - a_{N-1} X_N(s) + (b_{N-1} - a_{N-1} b_N) U(s)
\end{align*}
\]

The state equations on the right can be rewritten in matrix form:

\[
\begin{bmatrix}
    \dot{x}_1 \\
    \dot{x}_2 \\
    \vdots \\
    \dot{x}_{N-1} \\
    \dot{x}_N
\end{bmatrix} =
\begin{bmatrix}
    0 & 0 & \cdots & 0 & -a_0 \\
    1 & 0 & 0 & -a_1 \\
    \vdots & \vdots & \ddots & \vdots \\
    0 & 0 & 0 & -a_{N-2} \\
    0 & 0 & \cdots & 1 & -a_{N-1}
\end{bmatrix}
\begin{bmatrix}
    x_1 \\
    x_2 \\
    \vdots \\
    x_{N-1} \\
    x_N
\end{bmatrix} +
\begin{bmatrix}
    b_0 - a_0 b_N \\
    b_1 - a_1 b_N \\
    \vdots \\
    b_{N-2} - a_{N-2} b_N \\
    b_{N-1} - a_{N-1} b_N
\end{bmatrix} u.
\]

The output equation is
If \( H(s) \) is strictly proper, then again \( D = 0 \).

### 7.9 LTI State Estimation via Pole Placement: The Luenberger Observer

#### 7.9.1 Full-state Luenberger observer

As the name implies, the **full-state observer** produces an estimate of the entire state vector from measurement of the output and input signals. Consider the state-space system representing the plant whose state is not completely measured, and for which both the state and the output are corrupted by the deterministic noise signals \( w \) and \( v \), respectively:

\[
\dot{x} = Ax + Bu + w \\
y = Cx + v
\]  
(7.52)

The structure of an LTI full-state observer for this system is shown below.

Assume that we know the state-space matrices of the plant with perfect accuracy. The state-space system describing the dynamics of the observer is as follows:

\[
\dot{x} = A\hat{x} + L(y - \hat{y}) + Bu \\
\hat{y} = C\hat{x}
\]  
(7.53)

or,
\[
\dot{x} = (A - LC)\dot{x} + Ly + Bu \\
y = C\dot{x}
\]

(7.54)

The goal is to design the observer gain \( L \) such that the state estimate will track the state. This can be expressed in terms of the state-space system governing the evolution of the error \( e(t) := x(t) - \hat{x}(t) \).

A bit of algebra shows that this system is given by:

\[
\dot{e} = (A - LC)e + w + Lv.
\]

(7.55)

Therefore, it suffices to find a matrix \( L \) such that all the eigenvalues of \( A - LC \) (the poles of the observer) are in the open left half-plane to ensure that the error will tend to zero when the noises are zero. Note that this is true even for quickly varying inputs since (7.55) doesn't depend on the input signal. A fast observer is obtained with a "large" matrix gain \( L \). For systems with a single output, the technique of pole placement can be used to design \( L \) (a column vector) with the difference that the state-space system \((A, B, C, 0)\) should be expressed in an observable canonical form first.

### 7.9.2 Pole placement of the observer using the observable canonical realization

Assume that the state-space system is SISO. Using a similarity transformation, any state-space system can be transformed into the observable canonical form in which the \( A \) matrix is in transposed companion form. The state equation is then given by:

\[
\begin{bmatrix}
\dot{x}_1 \\
\dot{x}_2 \\
\vdots \\
\dot{x}_{N-1} \\
\dot{x}_N
\end{bmatrix} =
\begin{bmatrix}
0 & 0 & \cdots & 0 & -a_0 \\
1 & 0 & 0 & \cdots & -a_1 \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
0 & 0 & \cdots & 0 & -a_{N-2} \\
0 & 0 & \cdots & 1 & -a_{N-1}
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2 \\
\vdots \\
x_{N-1} \\
x_N
\end{bmatrix} +
\begin{bmatrix}
b_0 \\
b_1 \\
\vdots \\
b_{N-2} \\
b_{N-1}
\end{bmatrix} u .
\]

(7.56)

and the output equation is:

\[
y = \begin{bmatrix}
0 & \cdots & 0 & 1
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2 \\
\vdots \\
x_{N-1} \\
x_N
\end{bmatrix}
\]

(7.57)

where the last column of \( \tilde{A} \) contains the coefficients of the characteristic polynomial of the system (=denominator of the transfer function):

\[
p(s) = s^N + a_{N-1}s^{N-1} + \cdots + a_1s + a_0
\]

In fact, it is easy to write down an observable canonical state-space realization directly from a transfer function \( G(s) \). Now write the observer gain matrix \( L \) as

L2S- 11/13
The A matrix of the error system \( A_e := A - LC \) is then computed to obtain

\[
A_e = \begin{bmatrix}
0 & 0 & \cdots & 0 & -a_0 \\
1 & 0 & 0 & \cdots & -a_1 \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
0 & \ddots & \ddots & \ddots & 0 \\
0 & 0 & \cdots & 1 & -a_{N-1}
\end{bmatrix} - \begin{bmatrix}
l_0 \\
l_1 \\
\vdots \\
l_{N-2} \\
l_{N-1}
\end{bmatrix} = \begin{bmatrix}
0 & 0 & \cdots & 0 & -(a_0 + l_0) \\
1 & 0 & 0 & \cdots & -(a_1 + l_1) \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
0 & \ddots & \ddots & \ddots & 0 \\
0 & 0 & \cdots & 1 & -(a_{N-1} + l_{N-1})
\end{bmatrix}
\]

(7.59)

This matrix is still in transposed companion form, hence the characteristic polynomial of the error system is:

\[
p_e(s) = s^N + (a_{N-1} + l_{N-1})s^{N-1} + \cdots + (a_1 + l_1)s + (a_0 + l_0)
\]

(7.60)

It is now clear how to compute the entries of the observer gain matrix \( L = \begin{bmatrix} l_0 \\ \vdots \\ l_{N-1} \\ l_N \end{bmatrix} \).

Pole placement procedure for observer

1. Select desired poles for the error system \( p_1, \ldots, p_N \)

2. Transform the state-space system \((A, B, C, D)\) into its observable canonical form if it is not in this form already using a similarity transformation. This can be done as follows.
   - Compute the eigenvalues of \( A \), form its characteristic polynomial
     \[
p(s) = s^N + a_{N-1}s^{N-1} + \cdots + a_1s + a_0,
\]
     and then write down its transposed companion form matrix \( \tilde{A} \). The similarity transformation matrix \( T \) for which \( x = T\tilde{x} \), where \( \tilde{x} \) is the state of the observable canonical realization, can be computed from the following identities. Let
\[ M = T^{-1} = \begin{bmatrix} M_1 \\ M_2 \\ \vdots \\ M_N \end{bmatrix}, \text{then } \tilde{A} = T^{-1}AT \iff \tilde{A}M = MA \text{ and } \tilde{C}M = C. \] From these we can compute the rows of \( M \) recursively:

- \( M_N = C \)
- \( M_{N-1} = M_N A + a_{N-1}M_N \)
- \( M_{N-2} = M_{N-1} A + a_{N-2}M_N \)
  \[ \vdots \]
- \( M_1 = M_2 A + a_0M_N \)

3. Form the desired characteristic polynomial of the error system with the set of desired poles:

\[ p_e(s) = s^N + c_{N-1}s^{N-1} + \cdots + c_1s + c_0 \]

4. Solve for the entries of the observer matrix \( L \) by identifying the desired \( p_e(s) \) with that of Equation (7.60) above: \( l_i = c_i - a_i, \quad i = 0, \ldots, n - 1 \)

5. The final observer gain matrix is computed as \( L_f = TL \)