Outline

- Observability and Detectability
  - Illustrative Example
- Stability of State Estimator
- Obtaining Covariances from Data
- Conclusions
- Additional Reading
Observability Property

- Consider the linear system \((A, C)\) with \(n\) measurements \(Y(n - 1)\)

\[
  \begin{align*}
  x(k + 1) &= Ax(k) \\
  y(k) &= Cx(k) \\
  Y(n - 1) &= \{y(0), y(1), \ldots, y(n - 1)\}
  \end{align*}
\]

in which \(A \in \mathbb{R}^{n \times n}\) and \(C \in \mathbb{R}^{p \times n}\)

**Definition (Observability)**

\((A, C)\) is *observable* if these \(n\) measurements *uniquely* determine the system's initial state \(x(0)\).

- Observability is a property of the deterministic model equations

---

Observability Matrix

- For the \(n\) measurements, the system model gives

\[
\begin{bmatrix}
  y(0) \\
  y(1) \\
  \vdots \\
  y(n - 1)
\end{bmatrix}
= \begin{bmatrix}
  C \\
  CA \\
  \vdots \\
  CA^{n-1}
\end{bmatrix}
\begin{bmatrix}
  x(0)
\end{bmatrix}
\]

in which \(O \in \mathbb{R}^{np \times n}\) is the Observability Matrix

- \((A, C)\) is *observable* if and only if \(\text{rank}(O) = n\)
For the linear system
\[
\begin{bmatrix}
    x_{k+1} \\
    y_k
\end{bmatrix} =
\begin{bmatrix}
    A \\
    C
\end{bmatrix}
\begin{bmatrix}
    x_k \\
    y_k
\end{bmatrix}
\]

Find a similarity transformation \( T \):
\[
\begin{bmatrix}
    z_1 \\
    z_2
\end{bmatrix} =
\begin{bmatrix}
    T_1 \\
    T_2
\end{bmatrix}
\begin{bmatrix}
    x
\end{bmatrix} \Rightarrow \tilde{A} = TAT^{-1}, \tilde{C} = CT^{-1}
\]

So the transformed system has the following canonical form
\[
\begin{bmatrix}
    z_1 \\
    z_2
\end{bmatrix}_{k+1} =
\begin{bmatrix}
    \tilde{A}_{11} & 0 \\
    \tilde{A}_{21} & \tilde{A}_{22}
\end{bmatrix}
\begin{bmatrix}
    z_1 \\
    z_2
\end{bmatrix}_k
\]
\[
y_k =
\begin{bmatrix}
    \tilde{C}_1 & 0 \\
    \tilde{C}
\end{bmatrix}
\begin{bmatrix}
    z_1 \\
    z_2
\end{bmatrix}_k
\]

where \((\tilde{A}_{11}, \tilde{C}_1)\) is observable

In this structure
- \(z_1\) are the observable modes
- \(z_2\) are the unobservable modes
  - however, the system is still detectable if \(\lambda(\tilde{A}_{22}) \leq 1\)
Detectability Property

Definition (Detectability)
A linear system is detectable when all the unobservable modes are stable.

- This property is important for partially observable systems.
- An observable system is also detectable.

The property of detectability is important for control because one may successfully design a control system for an unobservable but detectable system so as to estimate and control the unstable modes.

Advanced Process Control.

Illustrative Example: CSTR Reactor (Ray, 1981)

- Reaction: \( A \rightarrow B \rightarrow C \)
- Controlled variables: \( C_A, C_B \)
- Manipulated variables: \( C_{Af}, C_{Bf} \)
- State variables: dimensionless concentrations - \( x_1(C_A), x_2(C_B) \)
- Input variables: dimensionless feeds - \( u_1(C_{Af}), u_2(C_{Bf}) \)

\[
\begin{align*}
\dot{x}_1 &= \begin{bmatrix} -A_1 & 0 \\ A_2 & -A_3 \end{bmatrix} x_1 + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u_1 \\
\dot{x}_2 &= \begin{bmatrix} 0 & 1 \end{bmatrix} u_2 \\
y &= C x
\end{align*}
\]

in which \( A_1, A_2, A_3 \geq 0 \) are constants.
Solution for 2 Cases

- Only $x_1$ is measured
  - $C = [1 \ 0]$
  - $O = \begin{bmatrix} 1 & 0 \\ -A_1 & 0 \end{bmatrix} \Rightarrow \text{rank}(O) = 1$
  - Unobservable system!
  - But it is still detectable: $\lambda(A) < 0$ (continuous system)

- Only $x_2$ is measured
  - $C = [0 \ 1]$
  - $O = \begin{bmatrix} 0 & 1 \\ A_2 & -A_3 \end{bmatrix} \Rightarrow \text{rank}(O) = 2$
  - Completely observable system!

Example Remarks

\[ A \xrightarrow{k_1} B \xrightarrow{k_2} C \quad r_1 = k_1 C_A \quad r_2 = k_2 C_B \]

- Physical Reasons
  - $C_B$ depends on both $C_A$ and $C_B$
  - $C_A$ is independent of $C_B$

- Consequences
  - By measuring $x_2(C_B)$ and knowing $u$, $x_1(C_A)$ can be determined
  - By measuring $x_1(C_A)$ and knowing $u$, $x_2(C_B)$ can take any value
Deterministic Stability of State Estimator

Definition (Asymptotic Stability of the State Estimator)

The estimator is **asymptotically stable** in sense of an observer if the estimator is able to “recover” from the incorrect initial value of state as data with no measurement noise are collected.

For example:

- assume an incorrect initial estimate
- the estimator converges (asymptotically) to the correct value

State estimation: probabilistic optimality versus stability

*Kalman filtering was first publicly presented (to somewhat more than polite applause) on April 1, 1959. But please note: Kalman filtering is not a triumph of applied probability: the theory has only a slight inheritance from probability theory while it has become an important pillar of system theory.*

R. Kalman, 1994

As Kalman has often stressed the major contribution of his work is not perhaps the actual filter algorithm, elegant and useful as it no doubt is, but the proof that under certain technical conditions called “controllability” and “observability,” the optimum filter is “stable” in the sense that the effects of initial errors and round-off and other computational errors will die out asymptotically.

T. Kailath, 1974
State Estimation: Optimality does not Ensure Stability

- For the linear system
  \[
  x(k + 1) = Ax(k) \\
  y(k) = Cx(k)
  \]

- Consider the case when \( A = I, C = 0 \)
  - Optimal estimate is \( \hat{x}(k) = x(0) \) (for a chosen initial condition)
  - Estimator does not converge to the true state \( x(0) \)
    - unless we have luckily chosen \( x(0) = x(0) \)
  - Unobservable (and undetectable) system

Cost Convergence and Stability Lemmas

**Lemma (Convergence of estimator cost)**

*Given noise-free measurements \( Y(T) \), the optimal estimator cost \( \Phi^0(Y(T)) \) converges as \( T \) increases, regardless of the system observability.*

**Lemma (Estimator stability - convergence to the true state)**

*For \((A, C)\) observable and \( Q, R > 0 \) (positive definite), the optimal linear state estimator is asymptotically stable

\[
\hat{x}(T) \to x(T) \quad \text{as} \quad T \to \infty
\]
Obtaining $Q$ and $R$ from Data

Model discretized with $t_k = k\Delta t$:

$$x_{k+1} = f(x_k, u_k) + g(x_k, u_k)w_k$$

$$\begin{bmatrix} y_1^k \\ y_2^k \end{bmatrix} = \begin{bmatrix} 1 & \cdots & 0 \\ 0 & \cdots & 1 \end{bmatrix} x_k + \begin{bmatrix} v_1^k \\ v_2^k \end{bmatrix}$$

- Measurements are only $X^D, X^B$ at the discretization times
- Noise $w_k$ affects all the states
- Noise $v_k$ corrupts the measurements

Motivation for Using Autocovariances

Idea of Autocovariances

- The state noise $w_k$ gets propagated in time
- The measurement noise $v_k$ appears only at the sampling times and is not propagated in time
- Taking autocovariances of data at different time lags gives covariances of $w_k$ and $v_k$

Let $w_k, v_k$ have zero means and covariances $Q$ and $R$
Mathematical Formulation of the ALS

Linear State-Space Model:

\[
\begin{align*}
x_{k+1} &= Ax_k + Gw_k & \quad w_k &\sim \mathcal{N}(0, Q) \\
y_k &= Cx_k + v_k & \quad v_k &\sim \mathcal{N}(0, R)
\end{align*}
\]

- Model \((A, C, G)\) known from the linearization, finite set of measurements: \(\{y_0, \ldots, y_k\}\) given.
- Only unknowns are noises \(w_k\) and \(v_k\).

- \(y_k = Cx_k + v_k\)
- \(y_{k+1} = CAx_k + CGw_k + v_{k+1}\)
- \(y_{k+2} = CA^2x_k + CAGw_k + CGw_{k+1} + v_{k+2}\)

The Autocovariance Least-Squares (ALS) Problem

Skipping a lot of algebra, we can write:

**Autocovariance Least Squares**

\[
\Phi = \min_{Q, R} \left\| \mathcal{A}_N \begin{bmatrix} (Q)_s \\ (R)_s \end{bmatrix} - \hat{b} \right\|^2
\]

- A least-squares problem in a vector of unknowns, \(Q, R\)
- Form \(\mathcal{A}_N\) from known system matrices
- \(\hat{b}\) is a vector containing the estimated correlations from data

\[
\hat{b} = \frac{1}{T} \sum_{k=1}^{T} \begin{bmatrix} y_ky_k^T \\ \vdots \\ y_{k+N-1}y_{k+N-1}^T \end{bmatrix}_s
\]
What about this idea?

Our new proposal!

- Choose a suboptimal state estimator gain $L$ and apply state estimation to $\{y_k\}$ to obtain preliminary $\{\hat{x}_k\}$.
- Obtain estimates of $w_k$ and $v_k$ from

$$G \hat{w}_k = \hat{x}_{k+1} - A \hat{x}_k$$
$$\hat{v}_k = y_k - C \hat{x}_k$$

- Obtain estimates of $Q$ and $R$ from sample variances!

$$\hat{Q} = \frac{1}{T} \sum_{k=1}^{T} \hat{w}_k \hat{w}_k^T$$
$$\hat{R} = \frac{1}{T} \sum_{k=1}^{T} \hat{v}_k \hat{v}_k^T$$

The bad news . . .

Unfortunately an estimate is not the same as the true noise

$$G \hat{Q} G^T = AL(CS C^T + R) L^T A^T \neq GQG^T$$
$$\hat{R} = CSC^T + R \neq R$$

in which $S$ satisfies the Lyapunov equation

$$S = (A - ALC) S (A - ALC)^T + GQG^T + ALRL^T A^T$$
Maybe they are close?

- Example:

\[
A = \begin{bmatrix}
0.04 & 0.99 & 0.15 \\
0.31 & 0.16 & 0.48 \\
0.02 & 0.18 & 0.74 \\
\end{bmatrix}
\]
\[
C = \begin{bmatrix}
1.30 \\
0.50 \\
0.05 \\
\end{bmatrix}
\]
\[
G = \begin{bmatrix}
0.51 & 0.00 \\
0.00 & 0.92 \\
0.30 & 0.00 \\
\end{bmatrix}
\]

- True values:

\[
Q = \begin{bmatrix}
0.10 & 0.00 \\
0.00 & 0.10 \\
\end{bmatrix}, \quad R = 0.10
\]

- Using \(\hat{w}_k, \hat{v}_k\) samples:

\[
\hat{Q} = \begin{bmatrix}
0.07 & 0.04 \\
0.04 & 0.02 \\
\end{bmatrix}, \quad \hat{R} = 0.32
\]

- Using ALS:

\[
\hat{Q}_{als} = \begin{bmatrix}
0.08 & 0.00 \\
0.00 & 0.13 \\
\end{bmatrix}, \quad \hat{R}_{als} = 0.08
\]

Comparison of Results
Still more bad news

If you were lucky and somehow guessed (or estimated) the optimal $L$ for processing the data...

**Optimal Pre-filtering of the measurements**

Still incorrect $\hat{Q}$, $\hat{R}$

\[
G\hat{Q}G^T = AL(CP^-C^T + R)L^T A^T \neq GQG^T
\]

\[
\hat{R} = CP^-C^T + R \neq R
\]

in which $P^-$ satisfies the filtering Riccati equation

\[
P^- = GQG^T + AP^-A^T - AP^-C^T(CP^-C^T + R)^{-1}CP^-A^T
\]

**Conclusions**

Today we have learned...

- Concepts of observability and detectability of linear systems
  - Illustrated through chemical reactor example
- Introduction to State Estimator Stability
  - Stability versus optimality
  - Cost convergence and estimator stability lemmas
- Obtaining Covariances from Data
  - Separating effects of $Q$ and $R$ in measurement $y$
  - Autocovariance Least-Squares (ALS) technique to estimate $Q$ and $R$
Additional Reading


