The problem statement

- Question: Can we design continuous-time Kalman filters?
- **GOAL:** Develop an estimator $\hat{x}(t)$ that is a linear function of measurements $z(\tau)$, $0 \le \tau \le t$, and minimizes the function:
$$\mathbb{E}\left[(x(t) - \hat{x}(t))^T (x(t) - \hat{x}(t))\right].$$
- **ASSUMED:** Dynamics of linear system are (ignore $u(t)$ for now):
$$\begin{aligned}\dot{x}(t) &= Ax(t) + B_w w(t)\\ z(t) &= Cx(t) + v(t),\end{aligned}$$
where noises are uncorrelated and white and $\hat{x}(0)$ and $\Sigma_{\tilde{x}}(0)$ are given.
  - Specifically, $\mathbb{E}[w(t)] = \mathbb{E}[v(t)] = 0$, $\mathbb{E}[w(t_1)w(t_2)^T] = S_w\delta(t_1 - t_2)$, $\mathbb{E}[v(t_1)v(t_2)^T] = S_v\delta(t_1 - t_2)$, $\mathbb{E}[w(t_1)v(t_2)] = 0$, $S_w > 0$, $S_v > 0$,
- **APPROACH:** As with some analysis in weeks 2 and 3 of Course #1, analyze discrete-time case as $\Delta t \to 0$.

Conversion: The Kalman gain $L(t)$

- To convert the Kalman gain, we start with:
$$L_k = \Sigma_{\tilde{x},k}^{-} C^T \left[C\Sigma_{\tilde{x},k}^{-} C^T + \Sigma_{\tilde{v}}\right]^{-1}.$$
- Then, anticipating a future step in the development, we compute:
$$\frac{L_k}{\Delta t} = \frac{\Sigma_{\tilde{x},k}^{-} C^T}{\Delta t}\left[C\Sigma_{\tilde{x},k}^{-} C^T + \Sigma_{\tilde{v}}\right]^{-1} = \Sigma_{\tilde{x},k}^{-} C^T \left[C\Sigma_{\tilde{x},k}^{-} C^T \Delta t + \Sigma_{\tilde{v}}\Delta t\right]^{-1}.$$
- As $\Delta t \to 0$, if $\Sigma_{\tilde{x},k}^{-}$ is finite,
$$\left.\begin{aligned} &C\Sigma_{\tilde{x},k}^{-} C^T \Delta t \to 0\\ &\Sigma_{\tilde{v}}\Delta t \to S_v\\ &\Sigma_{\tilde{x},k}^{-} \to \Sigma_{\tilde{x}}(t)\end{aligned}\right\}\text{ or, } \frac{L_k}{\Delta t} \to L(t) = \Sigma_{\tilde{x}}(t) C^T S_v^{-1},$$
where we recall from Lesson 1.3.7 that $\Sigma_{\tilde{v}} \approx S_v/\Delta t$ as $\Delta t \to 0$.

Conversion: The estimation-error covariance $\Sigma_{\tilde{x}}(t)$

- We examine the prediction-error covariance time update:
$$\begin{aligned}\Sigma_{\tilde{x},k+1}^{-} &= A_d \Sigma_{\tilde{x},k}^{+} A_d^T + \Sigma_{\widetilde{w}}\\ &\approx (I + A\Delta t)\Sigma_{\tilde{x},k}^{+}(I + A\Delta t)^T + B_w S_w B_w^T \Delta t\\ &= \Sigma_{\tilde{x},k}^{+} + \Delta t\left[A\Sigma_{\tilde{x},k}^{+} + \Sigma_{\tilde{x},k}^{+} A^T + B_w S_w B_w^T\right] + \mathcal{O}(\Delta t^2).\end{aligned}$$
- In this analysis, we have approximated
$$\begin{aligned}A_d &= e^{A\Delta t}\\ &\approx I + A\Delta t,\end{aligned}$$
like we did in Lesson 1.2.5.
- We also approximated $\Sigma_{\widetilde{w}} \approx B_w S_w B_w^T \Delta t$ as we did in Lesson 1.3.6.
- Both of these approximations are valid only when $\Delta t \to 0$.

Conversion of $\Sigma_{\tilde{x}}(t)$, continued

- Now, we recognize that $\Sigma^+_{\tilde{x},k} = \Sigma^-_{\tilde{x},k} - L_k C \Sigma^-_{\tilde{x},k}$. Inserting:

$$\Sigma^-_{\tilde{x},k+1} = \Sigma^-_{\tilde{x},k} - L_k C \Sigma^-_{\tilde{x},k} + \Delta t \Big[A\Sigma^-_{\tilde{x},k} - AL_k C \Sigma^-_{\tilde{x},k} + \Sigma^-_{\tilde{x},k} A^T - L_k C \Sigma^-_{\tilde{x},k} A^T + B_w S_w B_w^T \Big] + \mathcal{O}(\Delta t^2).$$

- Rearrange...

$$\frac{\Sigma^-_{\tilde{x},k+1} - \Sigma^-_{\tilde{x},k}}{\Delta t} = \frac{-L_k}{\Delta t} C \Sigma^-_{\tilde{x},k} + A\Sigma^-_{\tilde{x},k} + \Sigma^-_{\tilde{x},k} A^T + B_w S_w B_w^T - AL_k C \Sigma^-_{\tilde{x},k} - L_k C \Sigma^-_{\tilde{x},k} A^T + \mathcal{O}(\Delta t).$$

- As $\Delta t \to 0$, $\text{LHS} \to \dot{\Sigma}_{\tilde{x}}(t)$, $\frac{L_k}{\Delta t} \to L(t)$, $\Sigma^-_{\tilde{x},k} \to \Sigma_{\tilde{x}}(t)$ and $AL_k C \Sigma^-_{\tilde{x},k} \to 0$.

Conversion of $\Sigma_{\tilde{x}}(t)$, final form

- The resulting expression is:

$$\dot{\Sigma}_{\tilde{x}}(t) = \overbrace{\underbrace{A\Sigma_{\tilde{x}}(t) + \Sigma_{\tilde{x}}(t)A^T + B_w S_w B_w^T}_{\text{Lyapunov for propagation}} - \underbrace{\Sigma_{\tilde{x}}(t) C^T S_v^{-1} C \Sigma_{\tilde{x}}(t)}_{\geq 0}}^{\text{Riccati for measurement update}}.$$

- This is a continuous-time *differential Riccati equation* for error covariance.
- The equation includes the following impact on $\dot{\Sigma}_{\tilde{x}}(t)$:
  - $A\Sigma_{\tilde{x}}(t) + \Sigma_{\tilde{x}}(t)A^T$ ➠ Homogeneous part.
  - $B_w S_w B_w^T$ ➠ Increase due to process noise.
  - $\Sigma_{\tilde{x}}(t) C^T S_v^{-1} C \Sigma_{\tilde{x}}(t)$ ➠ Decrease due to measurements.

Conversion: State estimate

- Recall the state-prediction and state-estimate steps:
  1. $\hat{x}_k^- = A_d \hat{x}_{k-1}^+$.
  2. $\hat{x}_k^+ = \hat{x}_k^- + L_k(y_k - C\hat{x}_k^-)$.
- Substitute (1) into (2):

$$\hat{x}_k^+ = A_d \hat{x}_{k-1}^+ + L_k(z_k - CA_d \hat{x}_{k-1}^+).$$

- Substitute $A_d \approx I + A\Delta t$:

$$\hat{x}_k^+ = (I + A\Delta t)\hat{x}_{k-1}^+ + L_k\left(z_k - C(I + A\Delta t)\hat{x}_{k-1}\right).$$

- So, rearranging and taking the limit as $\Delta t \to 0$:

$$\frac{\hat{x}_k^+ - \hat{x}_{k-1}^+}{\Delta t} = A\hat{x}_{k-1}^+ + \frac{L_k}{\Delta t}\left(z_k - C\hat{x}_{k-1}^+ - CA\Delta t\,\hat{x}_{k-1}^+\right)$$

$$\dot{\hat{x}}(t) = A\hat{x}(t) + L(t)\left[z(t) - C\hat{x}(t)\right].$$

Kalman–Bucy filter equations

- Finally, adding back in the effect of $u(t)$, we have:

$$\begin{aligned}\dot{\hat{x}}(t) &= A\hat{x}(t) + Bu(t) + L(t)\left[z(t) - C\hat{x}(t) - Du(t)\right]\\ L(t) &= \Sigma_{\tilde{x}}(t)C^T S_v^{-1}\\ \dot{\Sigma}_{\tilde{x}}(t) &= A\Sigma_{\tilde{x}}(t) + \Sigma_{\tilde{x}}(t)A^T + B_w S_w B_w^T - \Sigma_{\tilde{x}}(t)C^T S_v^{-1} C\Sigma_{\tilde{x}}(t).\end{aligned}$$

- This is called a "Kalman–Bucy" filter, or a continuous-time Kalman filter/estimator.
- Often very difficult to implement due to the challenges of evaluating the differential Riccati equation in real time.
- But, with a determined attitude, it can be done using operational-amplifier ("analog computer") circuits.

Steady-state in continuous time

- A steady-state solution is more practical to implement than a time-varying solution for LTI continuous-time systems.
- We let $\dot{\Sigma}_{\tilde{x}}(t) \to 0$, giving the continuous-time algebraic Riccati equation (ARE):

$$A\Sigma_{\tilde{x}} + \Sigma_{\tilde{x}}A^T + B_w S_w B_w^T - \Sigma_{\tilde{x}}C^T S_v^{-1} C\Sigma_{\tilde{x}} = 0,$$

  where we also have $L = \Sigma_{\tilde{x}}C^T S_v^{-1}$.
- Same as time-varying case, but with constant $\Sigma_{\tilde{x}}$ and L. Solve with: `lqe.m`
- The tradeoff between sensor and process noise is now explicit: $L = \Sigma_{\tilde{x}}C^T S_v^{-1}$.
  - If we are uncertain about the estimate (i.e., $\Sigma_{\tilde{x}}$ is large), then the innovation $(z(t) - C\hat{x}(t))$ is weighted heavily (big L).
  - If S_v is small (measurements are accurate) then new measurements are heavily weighted (big L).
- We can think of $\Sigma_{\tilde{x}}C_y^T S_v^{-1}$ as analogous to a signal-to-noise ratio (SNR).

Error dynamics of the steady-state filter

- We are interested in finding the estimation-error dynamics.
- We start with the plant dynamics:

$$\begin{aligned}\dot{x} &= Ax + B_u u + B_w w\\ z &= Cx + v\end{aligned}$$

  where $w \sim \mathcal{N}(0, S_w)$, $B_w S_w B_w^T > 0$, $v \sim \mathcal{N}(0, S_v)$, $S_v > 0$; w, v white and independent, $[A, C]$ observable. (Strong form of assumptions.)
- Steady-state Kalman filter: $\dot{\hat{x}} = A\hat{x} + B_u u + L(z - C\hat{x})$.
- This gives estimation error dynamics: $\tilde{x} = x - \hat{x}$

$$\begin{aligned}\dot{\tilde{x}} = \dot{x} - \dot{\hat{x}} &= Ax + B_w w - [(A - LC)\hat{x} + L(Cx + v)]\\ &= (A - LC)\tilde{x} + B_w w - Lv.\end{aligned}$$

- Filter stability governed by eigenvalues of $(A - LC)$.

Frequency-domain interpretation

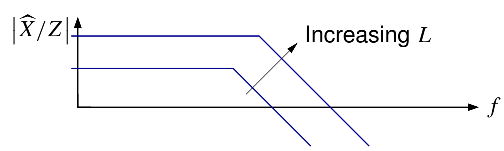
- This equation makes explicit a tradeoff between:
  - Speed of error decay (L big, $\text{eig}(A - LC)$ far in LHP).
  - Susceptibility of error to corruption by sensor noise (L small so Lv small).
- Kalman filter selects the *optimal balance* between these two goals.
- Can view in the frequency domain: Have $\dot{\hat{x}} = A\hat{x} + L(z - C\hat{x})$
- Laplace transform both sides... (assume scalar state for simplicity):

$$\frac{\widehat{X}(s)}{Z(s)} = \frac{L}{s - A + LC} = \frac{\frac{L}{LC-A}}{\frac{s}{LC-A} + 1}.$$

- Pole at $s = -(LC - A)$ and dc-gain of $L/(LC - A)$.
- This is the transfer function of the filter applied to the measurements to form the estimate $\hat{x}$ (a low-pass filter!).

Visualizing filter frequency response

- Increasing L ($\Sigma_{\tilde{x}}$ is large or S_v is small) pushes the filter magnitude response up and out.



- Eventually, the estimate would be too corrupted by the noise in the measurements; this is why the Kalman filter must choose an optimal L.
- Note that balancing the sensor-noise impact is done with respect to the process noise ($B_w w$), which is implicitly present in $\Sigma_{\tilde{x}}$.
- It turns out that the ratio S_w/S_v plays a key role in the selection of L, as an example will demonstrate.

Example of a time-varying filter

- EXAMPLE: System driven by noise:

$$\begin{aligned}\dot{x}(t) &= w(t)\\ z(t) &= x(t) + v(t),\end{aligned}$$

  where $\mathbb{E}[w] = \mathbb{E}[v] = 0$ and w and v independent; $\mathbb{E}[w(t)w(t+\tau)] = S_w\delta(\tau)$ and $\mathbb{E}[v(t)v(t+\tau)] = S_v\delta(\tau)$. Also, $\mathbb{E}[x(0)] = 0$, $\mathbb{E}[x(0)^2] = \Sigma_{\tilde{x}}(0)$.
- The optimal time-varying filter is:

$$\dot{\hat{x}}(t) = \frac{\Sigma_{\tilde{x}}(t)}{S_v}(z(t) - \hat{x}(t)), \qquad \text{and} \qquad \Sigma_{\tilde{x}}(t) = \sqrt{S_w S_v}\,\frac{1 + be^{-2\alpha t}}{1 - be^{-2\alpha t}},$$

  where $\alpha = \sqrt{S_w/S_v}$ and $b = (\Sigma_{\tilde{x}}(0) - \sqrt{S_w S_v})/(\Sigma_{\tilde{x}}(0) + \sqrt{S_w S_v})$.
- Convergence speed of $\Sigma_{\tilde{x}}(t)$ is determined by S_w/S_v via α.

Example of steady-state filter

- In steady-state, the Kalman-filter equations become:

$$\Sigma_{\tilde{x}}(t) \to \Sigma_{\tilde{x},\text{ss}} = \sqrt{S_w S_v}$$

$$L(t) \to L_{\text{ss}} = \frac{\Sigma_{\tilde{x},\text{ss}}}{S_v} = \sqrt{S_w / S_v}.$$

- Therefore, the steady-state filter implements:

$$\dot{\hat{x}}(t) = -\sqrt{\frac{S_w}{S_v}}\hat{x}(t) + \sqrt{\frac{S_w}{S_v}}z(t),$$

  and the closed-loop pole location is determined by S_w / S_v.
  - If S_w / S_v small, sensors are relatively noisy, state converges slowly.
  - If S_w / S_v large, sensors relatively clean, state converges quickly.

Summary

- You learned that we can develop Kalman-filter equations in continuous-time, resulting in the Kalman–Bucy filter.
- It is generally difficult to perform these calculations (e.g., solving the continuous-time differential Riccati equation) in real time although it can be done using operational-amplifier (“analog computer”) circuits.
- More often, a steady-state Kalman–Bucy filter is implemented, where the Kalman gain is computed as the solution to a continuous-time algebraic Riccati equation.
  - Since this gain is constant, it is easier to implement this solution using analog circuits.
- An example showed a frequency-domain interpretation of the Kalman–Bucy filter, and how it acts as an optimal low-pass filter on noisy measurements.
- The example also demonstrated the impact of the ratio S_w / S_v on the convergence speed of the filter and the product $S_w S_v$ on the confidence of the final answer.