Stochastic Adaptive Control (02421)

Lecture 12

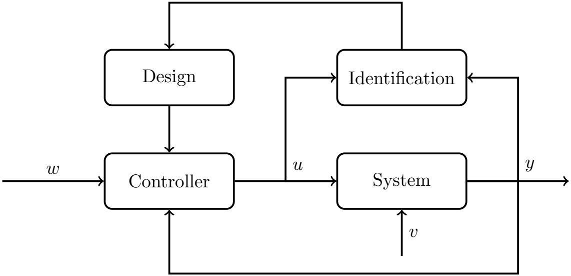
Tobias K. S. Ritschel

Assistant Professor

Section for Dynamical Systems, DTU Compute

Lecture Plan

1. ARX models
2. ARX prediction + control
3. ARX estimation
4. ARX model validation + adaptive control
5. ARMAX control
6. ARMAX estimation + adaptive control
7. Systems and control theory
8. Stochastic systems + Kalman filtering
9. SS estimation (recursive) + control
10. SS control
11. SS estimation (batch)
12. SS estimation (recursive)
13. SS nonlinear control



Today's Agenda

- Advanced topics in Kalman filtering
- Nonlinear Kalman filtering

Advanced topics in Kalman filtering

Kalman Filter: The Standard System

The Kalman filter is derived for the system

$$x_{t+1} = Ax_t + Bu_t + v_t \quad v_t \sim N(0, R_1), \tag{1}$$

$$y_t = Cx_t + Du_t + e_t \quad e_t \sim N(0, R_2), \tag{2}$$

$$\mathrm{Cov}(v_t, e_t) = 0, \quad e_t, v_t \text{ white} \perp x_s \quad s \leq t \tag{3}$$

The noises are independent of the state history and each other

1. $x_0 \sim N(\hat{x}_0, P_0)$
2. $v_t \sim N(0, R_1)$, white
3. $e_t \sim N(0, R_2)$, white
4. $\mathrm{Cov}(v_t, e_t) = 0$
5. $v_t, e_t \perp x_s, \quad s \leq t$

Kalman Filter: Deviations from the Standard Assumptions

We will now consider systems where one of these assumptions do not apply

1. Non-zero mean process disturbances
2. Non-zero mean output disturbances
3. Colored (non-white) process noise
4. Colored (non-white) output noise
5. Noise correlated with the state
6. Correlated noises

Uncertain Offset in the Process

System with stochastic process offset

$$x_{t+1} = Ax_t + Bu_t + Gd_t + v_t, \tag{4}$$
$$y_t = Cx_t + Du_t + e_t \tag{5}$$

Process offset

$$d_{t+1} = d_t + w_t \tag{6}$$

System in standard form

$$\begin{bmatrix} x_{t+1} \\ d_{t+1} \end{bmatrix} = \begin{bmatrix} A & G \\ 0 & I \end{bmatrix} \begin{bmatrix} x_t \\ d_t \end{bmatrix} + \begin{bmatrix} B \\ 0 \end{bmatrix} u_t + \begin{bmatrix} v_t \\ w_t \end{bmatrix}, \tag{7}$$
$$y_t = \begin{bmatrix} C & 0 \end{bmatrix} \begin{bmatrix} x_t \\ d_t \end{bmatrix} + Du_t + e_t \tag{8}$$

Uncertain Offset in the Output

System with stochastic measurement offset

$$x_{t+1} = Ax_t + Bu_t + v_t, \tag{9}$$

$$y_t = Cx_t + Du_t + Hd_t + e_t \tag{10}$$

Output offset

$$d_{t+1} = d_t + w_t \tag{11}$$

System in standard form

$$\begin{bmatrix} x_{t+1} \\ d_{t+1} \end{bmatrix} = \begin{bmatrix} A & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} x_t \\ d_t \end{bmatrix} + \begin{bmatrix} B \\ 0 \end{bmatrix} u_t + \begin{bmatrix} v_t \\ w_t \end{bmatrix}, \tag{12}$$

$$y_t = \begin{bmatrix} C & H \end{bmatrix} \begin{bmatrix} x_t \\ d_t \end{bmatrix} + Du_t + e_t \tag{13}$$

Coloured Noise

White noise (discrete-time):
In discrete time, a white noise signal ϵ_t has zero mean, finite variance, and it is uncorrelated in time: $\epsilon_t \perp \epsilon_s$ for $s \neq t$

If the noise w_t of a system is colored (non-white), it can be described as a system of white noises (η_t, ξ_t)

$$z_{t+1} = A_w z_t + \eta_t, \tag{14}$$

$$w_t = C_w z_t + \xi_t \tag{15}$$

Coloured Process Noise

System with colored process noise

$$x_{t+1} = Ax_t + Bu_t + v_t, \tag{16}$$

$$y_t = Cx_t + Du_t + e_t \tag{17}$$

Process noise

$$z_{t+1} = A_v z_t + \eta_t, \tag{18}$$

$$v_t = C_v z_t + \xi_t \tag{19}$$

System in standard form

$$\begin{bmatrix} x_{t+1} \\ z_{t+1} \end{bmatrix} = \begin{bmatrix} A & C_v \\ 0 & A_v \end{bmatrix} \begin{bmatrix} x_t \\ z_t \end{bmatrix} + \begin{bmatrix} B \\ 0 \end{bmatrix} u_t + \begin{bmatrix} \xi_t \\ \eta_t \end{bmatrix}, \tag{20}$$

$$y_t = \begin{bmatrix} C & 0 \end{bmatrix} \begin{bmatrix} x_t \\ z_t \end{bmatrix} + Du_t + e_t \tag{21}$$

Coloured Output Noise

System with colored measurement noise

$$x_{t+1} = Ax_t + Bu_t + v_t, \tag{22}$$

$$y_t = Cx_t + Du_t + e_t \tag{23}$$

Measurement noise

$$z_{t+1} = A_e z_t + \eta_t, \tag{24}$$

$$e_t = C_e z_t + \xi_t \tag{25}$$

System in standard form

$$\begin{bmatrix} x_{t+1} \\ z_{t+1} \end{bmatrix} = \begin{bmatrix} A & 0 \\ 0 & A_e \end{bmatrix} \begin{bmatrix} x_t \\ z_t \end{bmatrix} + \begin{bmatrix} B \\ 0 \end{bmatrix} u_t + \begin{bmatrix} v_t \\ \eta_t \end{bmatrix}, \tag{26}$$

$$y_t = \begin{bmatrix} C & C_e \end{bmatrix} \begin{bmatrix} x_t \\ z_t \end{bmatrix} + Du_t + \xi_t \tag{27}$$

Correlation with State History

System

$$x_{t+1} = Ax_t + Bu_t + Gw_t, \tag{28}$$
$$y_t = Cx_t + Du_t + \eta_t \tag{29}$$

w_t and η_t are correlated with the states

$$w_t = Hx_t + v_t, \qquad v_t \sim N(0, R_1), \qquad x_t \perp v_t, \tag{30}$$
$$\eta_t = Fx_t + e_t, \qquad e_t \sim N(0, R_2), \qquad x_t \perp e_t \tag{31}$$

System in standard form

$$x_{t+1} = (A + GH)x_t + Bu_t + Gv_t, \tag{32}$$
$$y_t = (C + F)x_t + Du_t + e_t \tag{33}$$

Kalman filter: Correlated noise

System with correlated process and measurement noise

$$x_{t+1} = Ax_t + Bu_t + v_t, \quad v_t \sim N(0, R_1), \tag{34}$$

$$y_t = Cx_t + Du_t + e_t, \quad e_t \sim N(0, R_2), \tag{35}$$

$$\text{Cov}(v_t, e_t) = R_{12}, \quad e_t, v_t \text{ white} \perp x_s, \quad s \le t \tag{36}$$

Measurement update

$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + \kappa_t(y_t - C\hat{x}_{t|t-1} - Du_t), \tag{37}$$

$$P_{t|t} = (1 - \kappa_t C)P_{t|t-1}, \tag{38}$$

$$\kappa_t = P_{t|t-1}C^T(CP_{t|t-1}C^T + R_2)^{-1} \tag{39}$$

Time update

$$\hat{x}_{t+1|t} = A\hat{x}_{t|t} + Bu_t + M(y_t - C\hat{x}_{t|t} - Du_t), \tag{40}$$

$$= (A - MC)\hat{x}_{t|t} + (B - MD)u_t + My_t, \tag{41}$$

$$P_{t+1|t} = (A - MC)P_{t|t}(A - MC)^T + R_1 - MR_{12}, \quad M = R_{12}R_2^{-1} \tag{42}$$

Derivation of Kalman filter: Correlated noise

Original system

$$x_{t+1} = Ax_t + Bu_t + v_t + M(y_t - y_t) \quad (43)$$
$$= Ax_t + Bu_t + v_t + M(y_t - Cx_t - Du_t - e_t) \quad (44)$$
$$= (A - MC)x_t + (B - MD)u_t + My_t + \tilde{v}_t, \tilde{v}_t = v_t - Me_t \quad (45)$$

Define the covariance of $\tilde{v}_t$ and e_t to be zero

$$\tilde{R}_{12} = \mathbb{E}[\tilde{v}_t e_t^T] = \mathbb{E}[(v_t - Me_t)e_t^T] = R_{12} - MR_2 = 0 \quad (46)$$
$$\Rightarrow M = R_{12}R_2^{-1}, \quad (47)$$
$$\tilde{R}_1 = \mathbb{E}[\tilde{v}_t\tilde{v}_t^T] = R_1 - R_{12}R_2^{-1}R_{12}^T \quad (48)$$

Kalman filter for systems with correlated noise: Original data update with new time update

$$\hat{x}_{t+1|t} = (A - MC)\hat{x}_{t|t} + (B - MD)u_t + My_t, \quad (49)$$
$$P_{t+1|t} = (A - MC)P_{t|t}(A - MC)^T + \tilde{R}_1 \quad (50)$$

Predictive Kalman filter: Correlated noise

System with correlated process and measurement noise

$$x_{t+1} = Ax_t + Bu_t + v_t, \quad v_t \in N(0, R_1), \tag{51}$$

$$y_t = Cx_t + Du_t + e_t, \quad e_t \in N(0, R_2), \tag{52}$$

$$\mathrm{Cov}(v_t, e_t) = R_{12}, \quad e_t, v_t \text{ white} \perp x_s, \quad s \leq t \tag{53}$$

Predictive Kalman filter

$$\hat{x}_{t+1|t} = A\hat{x}_{t|t-1} + Bu_t + K_t(y_t - C\hat{x}_{t|t-1} - Du_t), \tag{54}$$

$$P_{t+1|t} = AP_{t|t-1}A^T + R_1 - K_t(AP_{t|t-1}C^T + R_{12})^T, \tag{55}$$

$$K_t = (AP_{t|t-1}C^T + R_{12})(CP_{t|t-1}C^T + R_2)^{-1} \tag{56}$$

Relationship between gains

$$K_t = (A - MC)\kappa_t + M \tag{57}$$

Correlated Noise - Derivation (Predictive)

Conditional distribution

$$\begin{bmatrix} x_{t+1} \\ y_t \end{bmatrix} |Y_{t-1} \sim \mathbb{E}\left(\begin{bmatrix} A\hat{x}_{t|t-1} + Bu_t \\ C\hat{x}_{t|t-1} + Du_t \end{bmatrix}, \begin{bmatrix} AP_{t|t-1}A^T + R_1 & AP_{t|t-1}C^T + R_{12} \\ CP_{t|t-1}A^T + R_{12}^T & CP_{t|t-1}C^T + R_2 \end{bmatrix} \right) \tag{58}$$

Predictive distribution

$$x_{t+1}|Y_t \sim N(\hat{x}_{t+1|t}, P_{t+1|t}) \tag{59}$$

Use projection theorem

$$\hat{x}_{t+1|t} = A\hat{x}_{t|t-1} + Bu_t + K_t(y_t - C\hat{x}_{t|t-1} - Du_t), \tag{60}$$

$$P_{t+1|t} = AP_{t|t-1}A^T + R_1 - K_t(AP_{t|t-1}C^T + R_{12})^T, \tag{61}$$

$$K_t = (AP_{t|t-1}C^T + R_{12})(CP_{t|t-1}C^T + R_2)^{-1} \tag{62}$$

Correlated Noise - A Special Case

Perfect correlation between process noise and measurement noise

$$v_t = Ge_t, \qquad R_1 = GR_2G^T, \qquad R_{12} = GR_2 \tag{63}$$

Stationary predictive Kalman filter

$$P_\infty = 0, \qquad K_\infty = G \tag{64}$$

Proof: Predictive variance and gain if $P_{t|t-1} = 0$

$$P_{t+1|t} = R_1 - K_t R_{12}^T, \qquad K_t = R_{12} R_2^{-1} \tag{65}$$

Substitute R_1 and R_{12}

$$P_{t+1|t} = GR_2G^T - K_t(GR_2)^T = 0, \qquad K_t = GR_2R_2^{-1} = G \tag{66}$$

This proves that $P_\infty = 0$ and $K_\infty = G$ is a solution

If $(A - GC, R_1)$ is reachable and $R_2 \succ 0$, it is the only solution (from the theory of Riccati equations)

Nonlinear Kalman filtering

Linearization

Nonlinear system

$$x_{t+1} = f(x_t, u_t, v_t), \qquad v_t \sim N(m_v, R_v), \tag{67}$$

$$y_t = g(x_t, u_t, e_t), \qquad e_t \sim N(m_e, R_e) \tag{68}$$

Linearize right-hand side functions

$$f(x_t, u_t, v_t) \approx f(x_t^*, u_t^*, v_t^*) + \frac{\partial f}{\partial x}(x_t - x_t^*) + \frac{\partial f}{\partial u}(u_t - u_t^*) + \frac{\partial f}{\partial v}(v_t - v_t^*), \tag{69}$$

$$g(x_t, u_t, v_t) \approx g(x_t^*, u_t^*, e_t^*) + \frac{\partial g}{\partial x}(x_t - x_t^*) + \frac{\partial g}{\partial u}(u_t - u_t^*) + \frac{\partial g}{\partial v}(v_t - v_t^*) \tag{70}$$

The Jacobian matrices are evaluated in the linearization point, e.g.,

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial x}(x_t^*, u_t^*, v_t^*) \tag{71}$$

Linearization

Deviation variables

$$X_t = x_t - x_t^*, \qquad Y_t = y_t - y_t^*, \qquad y_t^* = g(x_t^*, u_t^*, e_t^*), \tag{72}$$

$$U_t = u_t - u_t^*, \qquad V_t = v_t - v_t^*, \qquad E_t = e_t - e_t^* \tag{73}$$

Linearized system

$$X_{t+1} = x_{t+1} - x_{t+1}^* = f(x_t, u_t, v_t) - f(x_t^*, u_t^*, v_t^*) \tag{74}$$

$$= A_t X_t + B_t U_t + G_t V_t, \tag{75}$$

$$Y_t = y_t - y_t^* = g(x_t, u_t, e_t) - g(x_t^*, u_t^*, e_t^*) \tag{76}$$

$$= C_t X_t + D_t U_t + F_t E_t \tag{77}$$

System matrices

$$A_t = \frac{\partial f}{\partial x}(x_t^*, u_t^*, v_t^*),\ B_t = \frac{\partial f}{\partial u}(x_t^*, u_t^*, v_t^*),\ G_t = \frac{\partial f}{\partial v}(x_t^*, u_t^*, v_t^*), \tag{78}$$

$$C_t = \frac{\partial g}{\partial x}(x_t^*, u_t^*, e_t^*),\ D_t = \frac{\partial g}{\partial u}(x_t^*, u_t^*, e_t^*),\ H_t = \frac{\partial g}{\partial e}(x_t^*, u_t^*, e_t^*) \tag{79}$$

Linearized Kalman filter

Measurement update

$$\hat{Y}_{t|t-1} = C_t \hat{X}_{t|t-1} + D_t U_t, \tag{80}$$

$$\hat{X}_{t|t} = \hat{X}_{t|t-1} + \kappa_t (Y_t - \hat{Y}_{t|t-1}), \tag{81}$$

$$\kappa_t = P_{t|t-1} C_t^T (C_t P_{t|t-1} C_t^T + R_e)^{-1}, \tag{82}$$

$$P_{t|t} = P_{t|t-1} - \kappa_t C_t^T P_{t|t-1} \tag{83}$$

Time update

$$\hat{X}_{t+1|t} = A_t \hat{X}_{t|t} + B_t U_t, \tag{84}$$

$$P_{t+1|t} = A_t P_{t|t} A_t^T + R_v \tag{85}$$

Extended Kalman filter: Measurement update

Linearization point

$$x_t^* = \hat{x}_{t|t-1}, \qquad u_t^* = u_t, \qquad e_t^* = m_e \tag{86}$$

Deviation variables

$$\hat{X}_{t|t-1} = 0, \qquad U_t = 0, \qquad V_t = v_t - m_v, \qquad E_t = e_t - m_e, \tag{87}$$

$$Y_t - \hat{Y}_{t|t-1} = y_t - \hat{y}_{t|t-1}, \tag{88}$$

$$\hat{Y}_{t|t-1} = \hat{y}_{t|t-1} - y_t^* = 0 \quad \Rightarrow \quad \hat{y}_{t|t-1} = y_t^* = g(x_t^*, u_t^*, e_t^*) \tag{89}$$

Measurement update

$$\hat{y}_{t|t-1} = g(\hat{x}_{t|t-1}, u_t, m_e), \tag{90}$$

$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + \kappa_t (y_t - \hat{y}_{t|t-1}), \tag{91}$$

$$\kappa_t = P_{t|t-1} C_t^T (C_t P_{t|t-1} C_t^T + R_e)^{-1}, \tag{92}$$

$$P_{t|t} = P_{t|t-1} - \kappa_t C_t^T P_{t|t-1} \tag{93}$$

The Jacobian matrix C_t is evaluated in $\hat{x}_{t|t-1}$, u_t, and m_e

Extended Kalman filter: Time update

Linearization point

$$x_t^* = \hat{x}_{t|t}, \qquad u_t^* = u_t, \qquad v_t^* = m_v, \qquad e_t^* = m_e \tag{94}$$

Deviation variables

$$\hat{X}_{t|t} = 0, \qquad U_t = 0, \qquad V_t = v_t - m_v, \qquad E_t = e_t - m_e, \tag{95}$$

$$\hat{X}_{t+1|t} = \hat{x}_{t+1|t} - x_{t+1}^* = \hat{x}_{t+1|t} - f(x_t^*, u_t^*, v_t^*) = 0 \tag{96}$$

Time update

$$\boxed{\begin{aligned} \hat{x}_{t+1|t} &= f(\hat{x}_{t|t}, u_t, m_v), &(97)\\ P_{t+1|t} &= A_t P_{t|t} A_t^T + R_v &(98) \end{aligned}}$$

The Jacobian matrix A_t is evaluated in $\hat{x}_{t|t}$, u_t, and m_v

In general, nonlinear Kalman filters are not optimal

Questions?