Stochastic Adaptive Control (02421)

Lecture 4

Tobias K. S. Ritschel

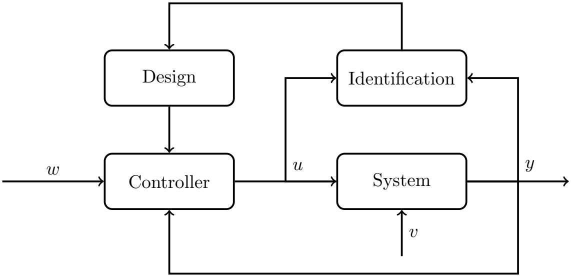
Section for Dynamical Systems

Department of Applied Mathematics and Computer Science

Technical University of Denmark

Lecture Plan

1. System theory
2. Stochastics
3. State estimation 1
4. **State estimation 2**
5. Optimal control 1
6. System identification 1 + adaptive control 1
7. External models + prediction
8. Optimal control 2
9. Optimal control 3
10. System identification 2
11. System identification 3 + model validation
12. System identification 4 + adaptive control 2
13. Adaptive control 3



Today's Agenda

- Follow-up from last lecture
- Advanced topics in Kalman filtering
- Nonlinear Kalman filtering
- Prediction
- Generalized predictive control (GPC)
- Examples

Follow-up From Last Time: Question 1

Find the stationary distribution of

$$x_{k+1} = \begin{bmatrix} 2/5 & 0 \\ -3/5 & 1/5 \end{bmatrix} x_k + v_k, \qquad v_k \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}\right), \tag{1}$$

$$y_k = \begin{bmatrix} 1 & 1 \end{bmatrix} x_k + e_k, \qquad e_k \sim N(0, 2) \tag{2}$$

Linear system and Gaussian noise ⇒ Gaussian distributions

$$\mathbb{E}[x_{k+1}] = A\mathbb{E}[x_k] + \mathbb{E}[v_k], \tag{3}$$

$$\mathrm{Cov}(x_{k+1}) = A\,\mathrm{Cov}(x_k)A^T + \mathrm{Cov}(v_k), \tag{4}$$

$$\mathbb{E}[y_k] = C\mathbb{E}[x_k] + \mathbb{E}[e_k], \tag{5}$$

$$\mathrm{Cov}(y_k) = C\,\mathrm{Cov}(x_k)C^T + \mathrm{Cov}(e_k) \tag{6}$$

Stationary distributions

$$x_\infty \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1.1905 & -0.3106 \\ -0.3106 & 2.6074 \end{bmatrix}\right), \tag{7}$$

$$y_\infty \sim N\left(0, 5.1768\right) \tag{8}$$

Follow-up from last time

Questions?

Advanced topics in Kalman filtering

Kalman Filter: The Standard System

The Kalman filter is derived for the system

$$x_{t+1} = Ax_t + Bu_t + v_t \quad v_t \sim N(0, R_1), \tag{9}$$

$$y_t = Cx_t + Du_t + e_t \quad e_t \sim N(0, R_2), \tag{10}$$

$$\mathrm{Cov}(v_t, e_t) = 0, \quad e_t, v_t \text{ white} \perp x_s \quad s \leq t \tag{11}$$

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

$$x_{t+1} = Ax_t + Bu_t + Gd_t + v_t, \tag{12}$$

$$y_t = Cx_t + Du_t + e_t \tag{13}$$

Process offset

$$d_{t+1} = d_t + w_t \tag{14}$$

System in standard form

$$\begin{bmatrix} x_{t+1} \\ d_{t+1} \end{bmatrix} = \begin{bmatrix} A & G \\ 0 & I \end{bmatrix} \begin{bmatrix} x_t \\ d_t \end{bmatrix} + \begin{bmatrix} B \\ 0 \end{bmatrix} u_t + \begin{bmatrix} v_t \\ w_t \end{bmatrix}, \tag{15}$$

$$y_t = \begin{bmatrix} C & 0 \end{bmatrix} \begin{bmatrix} x_t \\ d_t \end{bmatrix} + Du_t + e_t \tag{16}$$

Uncertain Offset in the Output

System with stochastic measurement offset

$$x_{t+1} = Ax_t + Bu_t + v_t, \tag{17}$$

$$y_t = Cx_t + Du_t + Hd_t + e_t \tag{18}$$

Output offset

$$d_{t+1} = d_t + w_t \tag{19}$$

System in standard form

$$\begin{bmatrix} x_{t+1} \\ d_{t+1} \end{bmatrix} = \begin{bmatrix} A & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} x_t \\ d_t \end{bmatrix} + \begin{bmatrix} B \\ 0 \end{bmatrix} u_t + \begin{bmatrix} v_t \\ w_t \end{bmatrix}, \tag{20}$$

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

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)

$$z_{t+1} = A_w z_t + \eta_t, \tag{22}$$

$$w_t = C_w z_t + \xi_t \tag{23}$$

Coloured Process Noise

System with colored process noise

$$x_{t+1} = Ax_t + Bu_t + v_t, \tag{24}$$
$$y_t = Cx_t + Du_t + e_t \tag{25}$$

Process noise

$$z_{t+1} = A_w z_t + \eta_t, \tag{26}$$
$$v_t = C_v z_t + \xi_t \tag{27}$$

System in standard form

$$\begin{bmatrix} x_{t+1} \\ z_{t+1} \end{bmatrix} = \begin{bmatrix} A & C_v \\ 0 & A_w \end{bmatrix} \begin{bmatrix} x_t \\ z_t \end{bmatrix} + \begin{bmatrix} B \\ 0 \end{bmatrix} u_t + \begin{bmatrix} \xi_t \\ \eta_t \end{bmatrix}, \tag{28}$$
$$y_t = \begin{bmatrix} C & 0 \end{bmatrix} \begin{bmatrix} x_t \\ z_t \end{bmatrix} + Du_t + e_t \tag{29}$$

Coloured Output Noise

System with colored measurement noise

$$x_{t+1} = Ax_t + Bu_t + v_t, \tag{30}$$

$$y_t = Cx_t + Du_t + e_t \tag{31}$$

Measurement noise

$$z_{t+1} = A_w z_t + \eta_t, \tag{32}$$

$$e_t = C_e z_t + \xi_t \tag{33}$$

System in standard form

$$\begin{bmatrix} x_{t+1} \\ z_{t+1} \end{bmatrix} = \begin{bmatrix} A & 0 \\ 0 & A_w \end{bmatrix} \begin{bmatrix} x_t \\ z_t \end{bmatrix} + \begin{bmatrix} B \\ 0 \end{bmatrix} u_t + \begin{bmatrix} v_t \\ \eta_t \end{bmatrix}, \tag{34}$$

$$y_t = \begin{bmatrix} C & C_e \end{bmatrix} \begin{bmatrix} x_t \\ z_t \end{bmatrix} + Du_t + \xi_t \tag{35}$$

Correlation with State History

System

$$x_{t+1} = Ax_t + Bu_t + Gw_t, \tag{36}$$

$$y_t = Cx_t + Du_t + \eta_t \tag{37}$$

w_t and η_t are correlated with the states

$$w_t = Hx_t + v_t, \qquad v_t \sim N(0, R_1), \qquad x_t \perp v_t, \tag{38}$$

$$\eta_t = Fx_t + e_t, \qquad e_t \sim N(0, R_2), \qquad x_t \perp e_t \tag{39}$$

System in standard form

$$x_{t+1} = (A + GH)x_t + Bu_t + Gv_t, \tag{40}$$

$$y_t = (C + F)x_t + Du_t + e_t \tag{41}$$

Kalman filter: Correlated noise

System with correlated process and measurement noise

$$x_{t+1} = Ax_t + Bu_t + v_t, \quad v_t \sim N(0, R_1), \tag{42}$$

$$y_t = Cx_t + Du_t + e_t, \quad e_t \sim N(0, R_2), \tag{43}$$

$$\mathrm{Cov}(v_t, e_t) = R_{12}, \quad e_t, v_t \text{ white} \perp x_s, \quad s \leq t \tag{44}$$

Measurement update

$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + \kappa_t(y_t - C\hat{x}_{t|t-1} - Du_t), \tag{45}$$

$$P_{t|t} = (1 - \kappa_t C)P_{t|t-1}, \tag{46}$$

$$\kappa_t = P_{t|t-1}C^T(CP_{t|t-1}C^T + R_2)^{-1} \tag{47}$$

Time update

$$\hat{x}_{t+1|t} = A\hat{x}_{t|t} + Bu_t + M(y_t - C\hat{x}_{t|t} - Du_t), \tag{48}$$

$$= (A - MC)\hat{x}_{t|t} + (B - MD)u_t + My_t, \tag{49}$$

$$P_{t+1|t} = (A - MC)P_{t|t}(A - MC)^T + R_1 - MR_{12}, \quad M = R_{12}R_2^{-1} \tag{50}$$

Derivation of Kalman filter: Correlated noise

Original system

$$
\begin{aligned}
x_{t+1} &= Ax_t + Bu_t + v_t + M(y_t - y_t) && (51)\\
&= Ax_t + Bu_t + v_t + M(y_t - Cx_t - Du_t - e_t) && (52)\\
&= (A - MC)x_t + (B - MD)u_t + My_t + \tilde{v}_t, \tilde{v}_t = v_t - Me_t && (53)
\end{aligned}
$$

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

$$
\begin{aligned}
\tilde{R}_{12} &= \mathbb{E}[\tilde{v}_t e_t^T] = \mathbb{E}[(v_t - Me_t)e_t^T] = R_{12} - MR_2 = 0 && (54)\\
&\Rightarrow M = R_{12}R_2^{-1}, && (55)\\
\tilde{R}_1 &= \mathbb{E}[\tilde{v}_t \tilde{v}_t^T] = R_1 - R_{12}R_2^{-1}R_{12}^T && (56)
\end{aligned}
$$

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

$$
\begin{aligned}
\hat{x}_{t+1|t} &= (A - MC)\hat{x}_{t|t} + (B - MD)u_t + My_t, && (57)\\
P_{t+1|t} &= (A - MC)P_{t|t}(A - MC)^T + \tilde{R}_1 && (58)
\end{aligned}
$$

Predictive Kalman filter: Correlated noise

System with correlated process and measurement noise

$$x_{t+1} = Ax_t + Bu_t + v_t, \quad v_t \in N(0, R_1), \tag{59}$$

$$y_t = Cx_t + Du_t + e_t, \quad e_t \in N(0, R_2), \tag{60}$$

$$\text{Cov}(v_t, e_t) = R_{12}, \quad e_t, v_t \text{ white} \perp x_s, \quad s \leq t \tag{61}$$

Predictive Kalman filter

$$\hat{x}_{t+1|t} = A\hat{x}_{t|t-1} + Bu_t + K_t(y_t - C\hat{x}_{t|t-1} - Du_t), \tag{62}$$

$$P_{t+1|t} = AP_{t|t-1}A^T + R_1 - K_t(AP_{t|t-1}C^T + R_{12})^T, \tag{63}$$

$$K_t = (AP_{t|t-1}C^T + R_{12})(CP_{t|t-1}C^T + R_2)^{-1} \tag{64}$$

Relationship between gains

$$K_t = (A - MC)\kappa_t + M \tag{65}$$

Correlated Noise - Derivation (Predictive)

Conditional distribution

$$\begin{bmatrix} x_{t+1} \\ y_t \end{bmatrix} |Y_{t-1} \sim \mathbb{E}\left(\begin{bmatrix} A\hat{x}_{t|t-1} + Bu_t \\ C\hat{x}_{t|t-1} + Du_t \end{bmatrix}, \begin{bmatrix} AP_{t|t-1}A^T + R_1 & AP_{t|t-1}C^T + R_{12} \\ CP_{t|t-1}A^T + R_{12}^T & CP_{t|t-1}C^T + R_2 \end{bmatrix} \right) \tag{66}$$

Predictive distribution

$$x_{t+1}|Y_t \sim N(\hat{x}_{t+1|t}, P_{t+1|t}) \tag{67}$$

Use projection theorem

$$\hat{x}_{t+1|t} = A\hat{x}_{t|t-1} + Bu_t + K_t(y_t - C\hat{x}_{t|t-1} - Du_t), \tag{68}$$

$$P_{t+1|t} = AP_{t|t-1}A^T + R_1 - K_t(AP_{t|t-1}C^T + R_{12})^T, \tag{69}$$

$$K_t = (AP_{t|t-1}C^T + R_{12})(CP_{t|t-1}C^T + R_2)^{-1} \tag{70}$$

Correlated Noise - A Special Case

Perfect correlation between process noise and measurement noise

$$v_t = Ge_t, \qquad R_1 = GR_2G^T, \qquad R_{12} = GR_2 \tag{71}$$

Stationary predictive Kalman filter

$$P_\infty = 0, \qquad K_\infty = G \tag{72}$$

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

$$P_{t+1|t} = R_1 - K_tR_{12}^T, \qquad K_t = R_{12}R_2^{-1} \tag{73}$$

Substitute R_1 and R_{12}

$$P_{t+1|t} = GR_2G^T - K_t(GR_2)^T = 0, \qquad K_t = GR_2R_2^{-1} = G \tag{74}$$

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

$$x_{t+1} = f(x_t, u_t, v_t), \qquad v_t \sim N(m_v, R_v), \tag{75}$$

$$y_t = g(x_t, u_t, e_t), \qquad e_t \sim N(m_e, R_e) \tag{76}$$

Linearize right-hand side functions

$$f(x_t, u_t, v_t) \approx f(x_t^*, u_t^*, v_t^*) + \frac{\partial f}{\partial x}(x_t - x_t^*) + \frac{\partial f}{\partial u}(u_t - u_t^*) + \frac{\partial f}{\partial v}(v_t - v_t^*), \tag{77}$$

$$g(x_t, u_t, v_t) \approx g(x_t^*, u_t^*, e_t^*) + \frac{\partial g}{\partial x}(x_t - x_t^*) + \frac{\partial g}{\partial u}(u_t - u_t^*) + \frac{\partial g}{\partial v}(v_t - v_t^*) \tag{78}$$

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

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial x}(x_t^*, u_t^*, v_t^*) \tag{79}$$

Linearization

Deviation variables

$$X_t = x_t - x_t^*, \qquad Y_t = y_t - y_t^*, \qquad y_t^* = g(x_t^*, u_t^*, e_t^*), \tag{80}$$

$$U_t = u_t - u_t^*, \qquad V_t = v_t - v_t^*, \qquad E_t = e_t - e_t^* \tag{81}$$

Linearized system

$$X_{t+1} = x_{t+1} - x_{t+1}^* = f(x_t, u_t, v_t) - f(x_t^*, u_t^*, v_t^*) \tag{82}$$

$$= A_t X_t + B_t U_t + G_t V_t, \tag{83}$$

$$Y_t = y_t - y_t^* = g(x_t, u_t, e_t) - g(x_t^*, u_t^*, e_t^*) \tag{84}$$

$$= C_t X_t + D_t U_t + F_t E_t \tag{85}$$

System matrices

$$A_t = \frac{\partial f}{\partial x}(x_t^*, u_t^*, v_t^*), \; B_t = \frac{\partial f}{\partial u}(x_t^*, u_t^*, v_t^*), \; G_t = \frac{\partial f}{\partial v}(x_t^*, u_t^*, v_t^*), \tag{86}$$

$$C_t = \frac{\partial g}{\partial x}(x_t^*, u_t^*, e_t^*), \; D_t = \frac{\partial g}{\partial u}(x_t^*, u_t^*, e_t^*), \; H_t = \frac{\partial g}{\partial e}(x_t^*, u_t^*, e_t^*) \tag{87}$$

Linearized Kalman filter

Measurement update

$$\hat{Y}_{t|t-1} = C_t \hat{X}_{t|t-1} + D_t U_t, \tag{88}$$

$$\hat{X}_{t|t} = \hat{X}_{t|t-1} + \kappa_t (Y_t - \hat{Y}_{t|t-1}), \tag{89}$$

$$\kappa_t = P_{t|t-1} C_t^T (C_t P_{t|t-1} C_t^T + R_e)^{-1}, \tag{90}$$

$$P_{t|t} = P_{t|t-1} - \kappa_t C_t^T P_{t|t-1} \tag{91}$$

Time update

$$\hat{X}_{t+1|t} = A_t \hat{X}_{t|t} + B_t U_t, \tag{92}$$

$$P_{t+1|t} = A_t P_{t|t} A_t^T + R_v \tag{93}$$

Extended Kalman filter: Measurement update

Linearization point

$$x_t^* = \hat{x}_{t|t-1}, \qquad u_t^* = u_t, \qquad e_t^* = m_e \tag{94}$$

Deviation variables

$$\hat{X}_{t|t-1} = 0, \qquad U_t = 0, \qquad V_t = v_t - m_v, \qquad E_t = e_t - m_e, \tag{95}$$

$$Y_t - \hat{Y}_{t|t-1} = y_t - \hat{y}_{t|t-1}, \tag{96}$$

$$\hat{Y}_{t|t-1} = \hat{y}_{t|t-1} - y_t^* = 0 \quad \Rightarrow \quad \hat{y}_{t|t-1} = y_t^* = g(x_t^*, u_t^*, e_t^*) \tag{97}$$

Measurement update

$$\hat{y}_{t|t-1} = g(\hat{x}_{t|t-1}, u_t, m_e), \tag{98}$$

$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + \kappa_t(y_t - \hat{y}_{t|t-1}), \tag{99}$$

$$\kappa_t = P_{t|t-1} C_t^T (C_t P_{t|t-1} C_t^T + R_e)^{-1}, \tag{100}$$

$$P_{t|t} = P_{t|t-1} - \kappa_t C_t^T P_{t|t-1} \tag{101}$$

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

$$x_t^* = \hat{x}_{t|t}, \qquad u_t^* = u_t, \qquad v_t^* = m_v, \qquad e_t^* = m_e \tag{102}$$

Deviation variables

$$\hat{X}_{t|t} = 0, \qquad U_t = 0, \qquad V_t = v_t - m_v, \qquad E_t = e_t - m_e, \tag{103}$$

$$\hat{X}_{t+1|t} = \hat{x}_{t+1|t} - x_{t+1}^* = \hat{x}_{t+1|t} - f(x_t^*, u_t^*, v_t^*) = 0 \tag{104}$$

Time update

$$\hat{x}_{t+1|t} = f(\hat{x}_{t|t}, u_t, m_v), \tag{105}$$

$$P_{t+1|t} = A_t P_{t|t} A_t^T + R_v \tag{106}$$

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

Prediction

The prediction problem: State prediction

System

$$x_{t+1} = Ax_t + Bu_t + Gv_t, \qquad v_t \sim N(0, R_1), \tag{107}$$
$$y_t = Cx_t + Du_t + Fe_t, \qquad e_t \sim N(0, R_2) \tag{108}$$

State predictions

$$x_1 = Ax_0 + Bu_0 + Gv_0, \tag{109}$$
$$x_2 = Ax_1 + Bu_1 + Gv_1 = A(Ax_0 + Bu_0 + Gv_0) + Bu_1 + Gv_1 \tag{110}$$
$$= A^2 x_0 + \begin{bmatrix} AB & B \end{bmatrix} \begin{bmatrix} u_0 \\ u_1 \end{bmatrix} + \begin{bmatrix} AG & G \end{bmatrix} \begin{bmatrix} v_0 \\ v_1 \end{bmatrix}, \tag{111}$$
$$x_3 = Ax_2 + Bu_2 + Gv_2 \tag{112}$$
$$= A\left(A^2 x_0 + \begin{bmatrix} AB & B \end{bmatrix} \begin{bmatrix} u_0 \\ u_1 \end{bmatrix} + \begin{bmatrix} AG & G \end{bmatrix} \begin{bmatrix} v_0 \\ v_1 \end{bmatrix} \right) + Bu_2 + Gv_2 \tag{113}$$
$$= A^3 x_0 + \begin{bmatrix} A^2 B & AB & B \end{bmatrix} \begin{bmatrix} u_0 \\ u_1 \\ u_2 \end{bmatrix} + \begin{bmatrix} A^2 G & AG & G \end{bmatrix} \begin{bmatrix} v_0 \\ v_1 \\ v_2 \end{bmatrix} \tag{114}$$

The prediction problem: State and output predictions

N-step state and output prediction

$$x_N = A^N x_0 + \begin{bmatrix} A^{N-1}B & A^{N-2}B & \cdots & B \end{bmatrix} \begin{bmatrix} u_0 \\ u_1 \\ \vdots \\ u_{N-1} \end{bmatrix} \tag{115}$$

$$+ \begin{bmatrix} A^{N-1}G & A^{N-2}G & \cdots & G \end{bmatrix} \begin{bmatrix} v_0 \\ v_1 \\ \vdots \\ v_{N-1} \end{bmatrix}, \tag{116}$$

$$y_N = CA^N x_0 + \begin{bmatrix} CA^{N-1}B & CA^{N-2}B & \cdots & CB & D \end{bmatrix} \begin{bmatrix} u_0 \\ u_1 \\ \vdots \\ u_{N-1} \\ u_N \end{bmatrix} \tag{117}$$

$$+ \begin{bmatrix} CA^{N-1}G & CA^{N-2}G & \cdots & CG & 0 \end{bmatrix} \begin{bmatrix} v_0 \\ v_1 \\ \vdots \\ v_{N-1} \\ v_N \end{bmatrix} + Fe_N \tag{118}$$

The prediction problem: N-step predictions

N-step state and output predictions

$$\begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_N \end{bmatrix} = \begin{bmatrix} I \\ A \\ \vdots \\ A^N \end{bmatrix} x_0 + \begin{bmatrix} 0 & & & \\ B & 0 & & \\ \vdots & \ddots & \ddots & \\ A^{N-1}B & \cdots & B & 0 \end{bmatrix} \begin{bmatrix} u_0 \\ u_1 \\ \vdots \\ u_N \end{bmatrix} + \begin{bmatrix} 0 & & & \\ G & 0 & & \\ \vdots & \ddots & \ddots & \\ A^{N-1}G & \cdots & G & 0 \end{bmatrix} \begin{bmatrix} v_0 \\ v_1 \\ \vdots \\ v_N \end{bmatrix}, \tag{119}$$

$$\begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_N \end{bmatrix} = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^N \end{bmatrix} x_0 + \begin{bmatrix} D & & & \\ CB & D & & \\ \vdots & \ddots & \ddots & \\ CA^{N-1}B & \cdots & CB & D \end{bmatrix} \begin{bmatrix} u_0 \\ u_1 \\ \vdots \\ u_N \end{bmatrix} + \begin{bmatrix} 0 & & & \\ G & 0 & & \\ \vdots & \ddots & \ddots & \\ CA^{N-1}G & \cdots & CG & 0 \end{bmatrix} \begin{bmatrix} v_0 \\ v_1 \\ \vdots \\ v_N \end{bmatrix} \tag{120}$$

$$+ \begin{bmatrix} F & & & \\ & F & & \\ & & \ddots & \\ & & & F \end{bmatrix} \begin{bmatrix} e_0 \\ e_1 \\ \vdots \\ e_N \end{bmatrix} \tag{121}$$

The prediction problem: Compact notation

Compact notation

$$X_N = \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_N \end{bmatrix}, \quad Y_N = \begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_N \end{bmatrix}, \quad U_N = \begin{bmatrix} u_0 \\ u_1 \\ \vdots \\ u_N \end{bmatrix}, \quad V_N = \begin{bmatrix} v_0 \\ v_1 \\ \vdots \\ v_N \end{bmatrix}, \quad E_N = \begin{bmatrix} e_0 \\ e_1 \\ \vdots \\ e_N \end{bmatrix}, \tag{122}$$

$$\Phi_{xx}^N = \begin{bmatrix} I \\ A \\ \vdots \\ A^N \end{bmatrix}, \quad \Gamma_{xu}^N = \begin{bmatrix} 0 & & & \\ B & 0 & & \\ \vdots & \ddots & \ddots & \\ A^{N-1}B & \cdots & B & 0 \end{bmatrix}, \quad \Gamma_{xv}^N = \begin{bmatrix} 0 & & & \\ G & 0 & & \\ \vdots & \ddots & \ddots & \\ A^{N-1}G & \cdots & G & 0 \end{bmatrix}, \tag{123}$$

$$\Phi_{yx}^N = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^N \end{bmatrix}, \quad \Gamma_{yu}^N = \begin{bmatrix} D & & & \\ CB & D & & \\ \vdots & \ddots & \ddots & \\ CA^{N-1}B & \cdots & CB & D \end{bmatrix}, \quad \Gamma_{yv}^N = \begin{bmatrix} 0 & & & \\ CG & 0 & & \\ \vdots & \ddots & \ddots & \\ CA^{N-1}G & \cdots & CG & 0 \end{bmatrix}, \tag{124}$$

$$\Gamma_{ye} = \begin{bmatrix} F & & & \\ & F & & \\ & & \ddots & \\ & & & F \end{bmatrix}, R_V^N = \begin{bmatrix} R_1 & & & \\ & R_1 & & \\ & & \ddots & \\ & & & R_1 \end{bmatrix}, R_E^N = \begin{bmatrix} R_2 & & & \\ & R_2 & & \\ & & \ddots & \\ & & & R_2 \end{bmatrix}, \tag{125}$$

$$X_N = \Phi_{xx}^N x_0 + \Gamma_{xu}^N U_N + \Gamma_{xv}^N V_N, \qquad V_N \sim N(0, R_V^N), \tag{126}$$

$$Y_N = \Phi_{yx}^N x_0 + \Gamma_{yu}^N U_N + \Gamma_{yv}^N V_N + \Gamma_{ye}^N E_N, \qquad E_N \sim N(0, R_E^N) \tag{127}$$

General predictive control

Optimal control - General predictive control

Optimal control problem

$$\min_{U_N} \quad J_N, \qquad J_N = \mathbb{E}[(Y_N - W_N)^T Q_y (Y_N - W_N) + U_N^T Q_u U_N] \tag{128}$$

Write out the objective function

$$\begin{aligned} J_N &= \mathbb{E}[(Y_N - W_N)^T Q_y (Y_N - W_N) + U_N^T Q_u U_N] &(129)\\ &= \mathbb{E}[Y_N^T Q_y Y_N - W_N^T Q_y Y_N - Y_N^T Q_y W_N - W_N^T Q y W_N + U_N^T Q_u U_N] &(130)\\ &= \mathbb{E}[Y_N^T Q_y Y_N - 2 W_N^T Q_y Y_N - W_N^T Q y W_N + U_N^T Q_u U_N] &(131)\\ &= \mathbb{E}[Y_N^T Q_y Y_N] - 2 W_N^T Q_y \mathbb{E}[Y_N] - W_N^T Q y W_N + U_N^T Q_u U_N &(132) \end{aligned}$$

Expectation of quadratic form

$$\mathbb{E}[Y_N^T Q_y Y_N] = \mathbb{E}[Y_N]^T Q_y \mathbb{E}[Y_N] + \mathrm{Tr}(A \, \mathrm{Cov}(Y_N)) \tag{133}$$

*We omit the superscript N on the matrices for brevity of notation.

Output expectation and variance

Expectation

$$\mathbb{E}[Y_N] = \Phi_{yx}\mathbb{E}[x_0] + \Gamma_{yu}U_N + \Gamma_{yv}\mathbb{E}[V_N] + \Gamma_{ye}\mathbb{E}[E_N] \tag{134}$$

$$= \Phi_{yx}\mathbb{E}[x_0] + \Gamma_{yu}U_N \tag{135}$$

Deviation from expectation

$$Y_N - \mathbb{E}[Y_N] = \Phi_{yx}(x_0 - \mathbb{E}[x_0]) + \Gamma_{yv}V_N + \Gamma_{ye}E_N \tag{136}$$

Covariance

$$\mathrm{Cov}(Y_N) = \mathbb{E}[(Y_N - \mathbb{E}[Y_N])(Y_N - \mathbb{E}[Y_N])^T] \tag{137}$$

$$= \Phi_{yx}\mathbb{E}[(x_0 - \mathbb{E}[x_0])(x_0 - \mathbb{E}[x_0])^T]\Phi_{yx}^T \tag{138}$$

$$+ \Gamma_{yv}\mathbb{E}[V_N V_N^T]\Gamma_{yv}^T + \Gamma_{ye}\mathbb{E}[E_N E_N^T]\Gamma_{ye}^T + \cdots \tag{139}$$

$$= \Phi_{yx}P_0\Phi_{yx}^T + \Gamma_{yv}R_V\Gamma_{yv}^T + \Gamma_{ye}R_E\Gamma_{ye}^T \tag{140}$$

Important observation: $\mathrm{Cov}(Y_N)$ is independent of U_N

Optimal control - General predictive control

Disregard terms that are independent of U_N

$$\begin{aligned} J_N &= \mathbb{E}[Y_N^T Q_y Y_N] - 2W_N^T Q_y \mathbb{E}[Y_N] - W_N^T Q y W_N + U_N^T Q_u U_N && (141)\\ &= \mathbb{E}[Y_N]^T Q_y \mathbb{E}[Y_N] + \mathrm{Tr}(A\,\mathrm{Cov}(Y_N)) && (142)\\ &- 2W_N^T Q_y \mathbb{E}[Y_N] - W_N^T Q y W_N + U_N^T Q_u U_N && (143)\\ &= \mathbb{E}[Y]^T Q_y \mathbb{E}[Y_N] - 2W_N^T Q_y \mathbb{E}[Y_N] + U_N^T Q_u U_N + \text{const.} && (144) \end{aligned}$$

Insert expectation

$$\begin{aligned} J_N &= (\Phi_{yx}\mathbb{E}[x_0] + \Gamma_{yu}U_N)^T Q_y (\Phi_{yx}\mathbb{E}[x_0] + \Gamma_{yu}U_N) && (145)\\ &- 2W_N^T Q_y (\Phi_{yx}\mathbb{E}[x_0] + \Gamma_{yu}U_N) + U_N^T Q_u U_N + \text{const.} && (146)\\ &= U_N^T \Gamma_{yu}^T Q_y \Gamma_{yu} U_N + 2\mathbb{E}[x_0]^T \Phi_{yx}^T Q_y \Gamma_{yu} U_N && (147)\\ &+ \mathbb{E}[x_0]^T \Phi_{yx}^T Q_y \Phi_{yx} \mathbb{E}[x_0] - 2W_N^T Q_y \Gamma_{yu} U_N + U_N^T Q_u U_N + \text{const.} && (148)\\ &= U_N^T \left(\Gamma_{yu}^T Q_y \Gamma_{yu} + Q_u\right) U_N + 2\left(\Phi_{yx}\mathbb{E}[x_0] - W_N\right)^T Q_y \Gamma_{yu} U_N + \text{const.} && (149) \end{aligned}$$

Optimal control - General predictive control

Gradient of objective function

$$\nabla J_N = 2\left(\Gamma_{yu}^T Q_y \Gamma_{yu} + Q_u\right) U_N + 2\left(\left(\Phi_{yx}\mathbb{E}[x_0] - W_N\right)^T Q_y \Gamma_{yu}\right)^T \quad (150)$$

$$= 2\left(\Gamma_{yu}^T Q_y \Gamma_{yu} + Q_u\right) U_N + 2\,\Gamma_{yu}^T Q_y \left(\Phi_{yx}\mathbb{E}[x_0] - W_N\right) \quad (151)$$

The gradient is zero for the optimal solution

$$\nabla J_N = 0 \quad (152)$$

Optimal solution

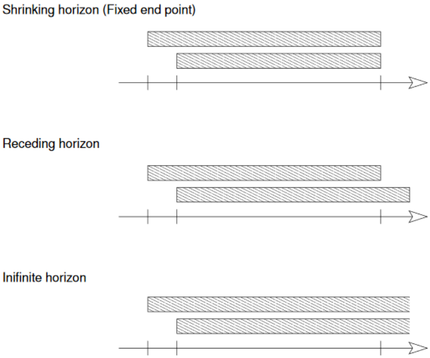
$$\boxed{U_N = -\left(\Gamma_{yu}^T Q_y \Gamma_{yu} + Q_u\right)^{-1} \Gamma_{yu}^T Q_y \left(\Phi_{yx}\mathbb{E}[x_0] - W_N\right)} \quad (153)$$

General predictive control: Prediction and control horizons

Closed-loop predictive control

$$u_t = \begin{bmatrix} I & 0 & \cdots & 0 \end{bmatrix} U_N \tag{154}$$

U_N is updated at every time instance



Questions

Questions?