Stochastic Adaptive Control (02421)

Lecture 10

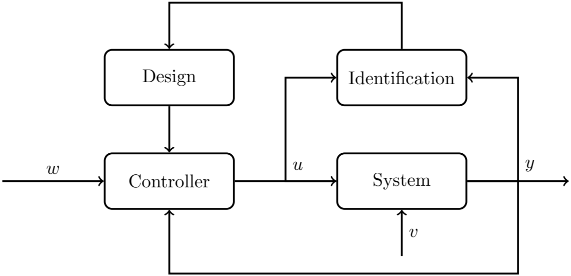
Tobias K. S. Ritschel

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

- Linear quadratic regulator
- Linear quadratic-Gaussian regulator

Linear quadratic regulator

Optimal Control - Quadratic cost functions

System

$$
\begin{aligned}
x_{k+1} &= Ax_k + Bu_k + v_k, \quad & x_0 \sim N(m_0, P_0), \quad & v_k \sim N(0, R_1), \quad & \text{(1a)} \\
y_k &= Cx_k + e_k, & e_k \sim N(0, R_2), \quad & v_k \perp e_k \perp x_k & \text{(1b)}
\end{aligned}
$$

Deviation from reference and control usage

$$
J = \mathbb{E}\left[\sum_{k=0}^{N-1} (y_k - w_k)^T Q_y (y_k - w_k) + u_k^T Q_u u_k \mid \mathcal{F}\right] \tag{2}
$$

Deviation from reference and initial control

$$
J = \mathbb{E}\left[\sum_{k=0}^{N-1} (y_k - w_k)^T Q_y (y_k - w_k) + (u_k - u_0)^T Q_u (u_k - u_0) \mid \mathcal{F}\right] \tag{3}
$$

Deviation from reference and control rate-of-movement

$$
J = \mathbb{E}\left[\sum_{k=0}^{N-1} (y_k - w_k)^T Q_y (y_k - w_k) + (u_k - u_{k-1})^T Q_u (u_k - u_{k-1}) \mid \mathcal{F}\right] \tag{4}
$$

Optimal Control - Linear Quadratic Regulator

Important: Specify which data/information that is available

Assume perfect state information ($y_k = x_k$)

$$J = \mathbb{E}\left[x_N^T Q_0 x_N + \sum_{k=0}^{N-1}\left(x_k^T Q_1 x_k + u_k^T Q_2 u_k\right)\right] \tag{5}$$

Split the equation at time t

$$J = \mathbb{E}\left[\sum_{k=0}^{t-1}\left(x_k^T Q_1 x_k + u_k^T Q_2 u_k\right)\right] \tag{6}$$

$$+ \mathbb{E}\left[x_N^T Q_0 x_N + \sum_{k=t}^{N-1}\left(x_k^T Q_1 x_k + u_k^T Q_2 u_k\right)\right] \tag{7}$$

The first term is independent of $u_t, \ldots, u_{N-1}$

Optimal Control - minimum and definition

Assume that $l(x,u)$ has a unique minimum with respect to u for all x, and let $u^0(x)$ denote the value of u where this minimum is attained. Then,

$$\min_{u(x)} \mathbb{E}\left[l(x,u)\right] = \mathbb{E}\left[l(x,u^0(x))\right] = \mathbb{E}\left[\min_{u} \; l(x,u)\right] \tag{8}$$

Apply result

$$\min_{u_t,\ldots,u_{N-1}} \mathbb{E}\left[x_N^T Q_0 x_N + \sum_{k=t}^{N-1}\left(x_k^T Q_1 x_k + u_k^T Q_2 u_k\right)\right] = \mathbb{E}\Big[V_t(x_t)\Big] \tag{9}$$

where

$$V_t(x_t) = \min_{u_t,\ldots,u_{N-1}} \mathbb{E}\left[x_N^T Q_0 x_N + \sum_{k=t}^{N-1}\left(x_k^T Q_1 x_k + u_k^T Q_2 u_k\right) \mid x_t\right] \tag{10}$$

Optimal Control - Dynamic Programming

Repeat to obtain the Bellman equation

$$V_t(x_t) = \min_{u_t} \ \mathbb{E}\left[x_t^T Q_1 x_t + u_t^T Q_2 u_t + V_{t+1}(x_{t+1}) \mid x_t\right] \tag{11a}$$

$$= \min_{u_t} \ x_t^T Q_1 x_t + u_t^T Q_2 u_t + \mathbb{E}\left[V_{t+1}(x_{t+1}) \mid x_t\right] \tag{11b}$$

End-point condition $(t = N)$

$$V_N(x_N) = \min_{u_N} \ \mathbb{E}\left[x_N^T Q_0 x_N \mid x_N\right] = x_N^T Q_0 x_N \tag{12}$$

The solution to this end value problem is a quadratic function

$$V_t(x_t) = x_t^T S_t x_t + s_t \tag{13}$$

S_t is non-negative definite

Optimal Control - Dynamic Programming

It is true for $t = N$

$$V_N(x_N) = x_N^T Q_0 x_N \tag{14}$$

Proof by induction: Assume that it holds for $t+1$ and show that it holds for t

By assumption

$$V_{t+1}(x_{t+1}) = x_{t+1}^T S_{t+1} x_{t+1} + s_{t+1} \tag{15}$$

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

$$\mathbb{E}\left[V_{t+1}(x_{t+1}) \mid x_t\right] = \left(Ax_t + Bu_t\right)^T S_{t+1} \left(Ax_t + Bu_t\right) \tag{16}$$

$$+ \operatorname{Tr}(S_{t+1}R_1) + s_{t+1} \tag{17}$$

Optimal Control - Dynamic Programming

Insert result from previous slide

$$V_t(x_t) = \min_{u_t} \; x_t^T Q_1 x_t + u_t^T Q_2 u_t + (Ax_t + Bu_t)^T S_{t+1}(Ax_t + Bu_t) \tag{18}$$

$$+ \operatorname{Tr}(S_{t+1}R_1) + s_{t+1} \tag{19}$$

Minimum

$$u_t = -L_t x_t \tag{20}$$

Control gain

$$L_t = (Q_2 + B^T S_{t+1} B)^{-1} B^T S_{t+1} A \tag{21}$$

Collect terms

$$V_t(x_t) = x_t^T (A^T S_{t+1} A + Q_1 - L_t^T (Q_2 + B^T S_{t+1} B) L_t) x_t \tag{22}$$

$$+ \operatorname{Tr}(S_{t+1}R_1) + s_{t+1} \tag{23}$$

Optimal Control - Dynamic Programming

$V_t(x_t)$ is quadratic

$$S_t = A^T S_{t+1} A + Q_1 - L_t^T (Q_2 + B^T S_{t+1} B) L_t \tag{24a}$$

$$s_t = \mathrm{Tr}(S_{t+1} R_1) + s_{t+1} \tag{24b}$$

We still need to show that S_t is non-negative definite

Rearrange terms

$$S_t = (A - BL_t)^T S_{t+1} (A - BL_t) + L_t^T Q_2 L_t + Q_1 \tag{25}$$

If S_{t+1} is non-negative definite, then S_t is also non-negative definite (due to the properties of Q_1 and Q_2)

Optimal Control - Dynamic Programming

Original equation and control gain

$$S_t = A^T S_{t+1} A + Q_1 - L_t^T (Q_2 + B^T S_{t+1} B) L_t, \tag{26}$$

$$L_t = (Q_2 + B^T S_{t+1} B)^{-1} B^T S_{t+1} A \tag{27}$$

Intermediate results

$$L_t^T (Q_2 + B^T S_{t+1} B) L_t = L_t^T B^T S_{t+1} A, \tag{28}$$

$$L_t^T = A^T S_{t+1} B (Q_2 + B^T S_{t+1} B)^{-1}, \tag{29}$$

$$L_t^T Q_2 L_t + L_t^T B^T S_{t+1} B L_t - A^T S_{t+1} B L_t = 0, \tag{30}$$

$$S_t = A^T S_{t+1} A + Q_1 - L_t^T B^T S_{t+1} A + L_t^T Q_2 L_t \tag{31}$$

$$+ L_t^T B^T S_{t+1} B L_t - A^T S_{t+1} B L_t \tag{32}$$

Final result

$$S_t = (A - B L_t)^T S_{t+1} (A - B L_t) + L_t^T Q_2 L_t + Q_1 \tag{33}$$

Optimal Control - Dynamic Programming

Optimal control law

$$u_t = -L_t x_t \tag{34}$$

Optimal control gain

$$L_t = (Q_2 + B^T S_{t+1} B)^{-1} B^T S_{t+1} A \tag{35}$$

The matrix S_t is

$$S_t = (A - BL_t)^T S_{t+1} (A - BL_t) + L_t^T Q_2 L_t + Q_1 \tag{36}$$

End condition

$$S_N = Q_0 \tag{37}$$

How does L_t depend on x_t, how would you implement the controller in practice, and how could you simplify the implementation?

Think about it for yourself for <u>one minute</u> and then discuss with the person next to you for <u>two minutes</u>.

LQR - Linear Quadratic Regulator w. complete state info

Finite-horizon LQR

$$J_t = \mathbb{E}\left[\sum_{k=t}^{t+N} \begin{bmatrix} x_k^T & u_k^T \end{bmatrix} \begin{bmatrix} Q_1 & Q_{12} \\ Q_{12}^T & Q_2 \end{bmatrix} \begin{bmatrix} x_k \\ u_k \end{bmatrix}\right], \qquad x_t \sim N(m_0, P_0) \tag{38}$$

$$x_{k+1} = Ax_k + Bu_k + v_k, \qquad v_k \sim N(0, R_1) \tag{39}$$

Optimal control law

$$\boxed{u_t = -L_t x_t = -(B^T S_{t+1} B + Q_2)^{-1}(B^T S_{t+1} A + Q_{12}^T)x_t} \tag{40}$$

Optimal state weight

$$S_t = A^T S_{t+1} A + Q_1 - A^T S_{t+1} B(B^T S_{t+1} B + Q_2)^{-1} B^T S_{t+1} A \tag{41}$$

$$S_{t+N+1} = 0 \tag{42}$$

LQR - Closed-loop analysis (complete state info)

System

$$x_{t+1} = Ax_t + Bu_t + v_t, \tag{43}$$

$$y_t = Cx_t + Du_t + e_t \tag{44}$$

Closed-loop description ($u_t = -L_t x_t$)

$$x_{t+1} = (A - BL_t)x_t + v_t = A_{cl}x_t + v_t, \tag{45}$$

$$y_t = (C - DL_t)x_t = C_{cl}x_t + e_t \tag{46}$$

State mean/variance

$$\mathbb{E}[x_t] = A_{cl}\mathbb{E}[x_{t-1}], \quad \mathbb{E}[x_0] = m_0, \tag{47}$$

$$\mathrm{Cov}(x_t) = A_{cl}\,\mathrm{Cov}(x_{t-1})A_{cl}^T + R_1, \quad \mathrm{Cov}(x_0) = P_0 \tag{48}$$

Output mean/variance

$$\mathbb{E}[y_t] = C_{cl}\mathbb{E}[x_t], \tag{49}$$

$$\mathrm{Cov}(y_t) = C_{cl}\,\mathrm{Cov}(x_t)C_{cl}^T + R_2 \tag{50}$$

LQR - stationary control

Infinite horizon LQR ($N = \infty$) is a stationary controller

Discrete algebraic Ricatti equation (DARE)

$$\boxed{\begin{aligned} S_\infty &= A^T S_\infty A + Q_1 - A^T S_\infty B(B^T S_\infty B + Q_2)^{-1} B^T S_\infty A, \\ L_\infty &= -(B^T S_\infty B + Q_2)^{-1}(B^T S_\infty A + Q_{12}) \end{aligned}} \tag{51, 52}$$

This applicable iff (A, B) is at least stabilizable (controllable, reachable)

If (A, Q_1) is observable, then the DARE has a unique positive semi-definite solution, and $A - BL$ is asymptotically stable

LQR - complete/incomplete state information

More general form of the Bellman equation ($\mathcal{F}_t \in \{x_t, Y_t, Y_{t-1}\}$)

$$V_t(\mathcal{F}_t) = \min_{u_t,\ldots,u_{t+N}} \mathbb{E}\left[\sum_{k=t}^{t+N} I_k(x_k, u_k) \mid \mathcal{F}_t\right] \tag{53}$$

$$= \min_{u_t} \mathbb{E}\left[I_t(x_t, u_t) + V_{t+1}(\mathcal{F}_{t+1}) \mid \mathcal{F}_t\right] \tag{54}$$

Using the same derivation, the LQR control law becomes

$$u_t = -L_t \mathbb{E}\left[x_t \mid \mathcal{F}_t\right], \tag{55}$$

$$L_t = (B^T S_{t+1} B + Q_2)^{-1}(B^T S_{t+1} A + Q_{12}), \tag{56}$$

$$S_t = A^T S_{t+1} A + Q_1 - L_t^T (B^T S_{t+1} B + Q_2) L_t, \tag{57}$$

$$S_{t+N+1} = 0 \tag{58}$$

Control law for incomplete state information

$$u_t = -L_t \mathbb{E}[x_t \mid Y_t] = -L_t \hat{x}_{t|t}, \tag{59}$$

$$u_t = -L_t \mathbb{E}[x_t \mid Y_{t-1}] = -L_t \hat{x}_{t|t-1} \tag{60}$$

Linear quadratic-Gaussian regulator

Optimal linear quadratic Gaussian observer-based controller

We have discussed both controllers and observers/state estimation

1. LQR: Optimal state control based on perfect state and system knowledge
2. Kalman filter: Optimal state estimation based on perfect system knowledge

When full state knowledge is not possible, we combine the controller with an observer

The optimal observer-based controller is the linear quadratic Gaussian controller (LQG)

$$\min_{u_t,\ldots,u_{t+N}} \mathbb{E}\left[\sum_{k=t}^{t+N} \begin{bmatrix} x_k \\ u_k \end{bmatrix}^T \begin{bmatrix} Q_1 & Q_{12} \\ Q_{12} & Q_2 \end{bmatrix} \begin{bmatrix} x_k \\ u_k \end{bmatrix} \mid \mathcal{F}\right], \tag{61}$$

$$x_{k+1} = Ax_k + Bu_k + v_k, \quad v_k \sim N(0, R_1), \tag{62}$$

$$y_k = Cx_k + e_k, \quad e_k \sim N(0, R_2), \mathrm{Cov}(v_k, e_k) = R_{12} \tag{63}$$

The controller and observer can be designed independently (the separation principle)

LQG - Duality and Stationarity

Control vs. observation - two sides of the same coin

Consider quadratic optimal control (LQ) and quadratic optimal observers (Kalman filter)

Optimal gains

$$L_t^T = (A^T S_{t+1} B + Q_{12})(B^T S_{t+1} B + Q_2)^{-1}, \tag{64}$$

$$K_t = (A P_t C^T + R_{12})(C P_t C^T + R_2)^{-1} \tag{65}$$

Riccati equations

$$S_t = A^T S_{t+1} A + Q_1 - L_t^T (B^T S_{t+1} B + Q_2) L_t, \quad S_{N+1} = 0, \tag{66}$$

$$P_{t+1} = A P_t A^T + R_1 - K_t (C P_t C^T + R_2) K_t^T, \quad P_0 \text{ is given} \tag{67}$$

Algebraic Riccati Equations (Stationary case)

$$S = A^T S A + Q_1 - (A^T S B + Q_{12})(B^T S B + Q_2)^{-1}(B^T S A + Q_{12}^T), \tag{68}$$

$$P = A P A^T + R_1 - (A P C^T + R_{12})(C P C^T + R_2)^{-1}(C P A^T + R_{12}^T) \tag{69}$$

Sketch of proof of separation principle

Separation principle:
Independently designed optimal controller and observer design is optimal

System

$$x_{t+1} = Ax_t + Bu_t + v_t, \tag{70}$$
$$y_t = Cx_t + e_t \tag{71}$$

Closed-loop system (predictive Kalman filter)

$$\begin{bmatrix} x_{t+1} \\ \hat{x}_{t+1|t} \end{bmatrix} = \begin{bmatrix} A & -BL_t \\ K_tC & A - K_tC - BL_t \end{bmatrix} \begin{bmatrix} x_t \\ \hat{x}_{t|t-1} \end{bmatrix} + \begin{bmatrix} I & 0 \\ 0 & K_t \end{bmatrix} \begin{bmatrix} v_t \\ e_t \end{bmatrix} \tag{72}$$

L_t and K_t are LQR and Kalman filter gains

Sketch of proof of separation principle

The LQR control law is the same for complete and partial state information

Consequently, we only need to prove that the Kalman filter is optimal for the LQR control law

System estimation error $(\tilde{x}_{t|t-1} = x_t - \hat{x}_{t|t-1})$

$$\begin{bmatrix} x_{t+1} \\ \tilde{x}_{t+1|t} \end{bmatrix} = \begin{bmatrix} A - BL_t & BL_t \\ 0 & A - K_t C \end{bmatrix} \begin{bmatrix} x_{t+1} \\ \tilde{x}_{t|t-1} \end{bmatrix} + \begin{bmatrix} I & 0 \\ I & -K_t \end{bmatrix} \begin{bmatrix} v_t \\ e_t \end{bmatrix} \tag{73}$$

The estimation error is independent of the control gain and true state

The system matrix is triangular: Its eigenvalues only depend on the eigenvalues of $A - BL_t$ and $A - K_t C$

Linear-quadratic Gaussian regulator

How would you implement an LQG controller in practice?

Think about it for yourself for one minute and then discuss with the person next to you for one minute.

Closed loop LQG - Predictive

Closed-loop system: LQG controller based on predictive Kalman filter

$$\begin{bmatrix} x_{t+1} \\ \tilde{x}_{t+1|t} \end{bmatrix} = \begin{bmatrix} A - BL_t & BL_t \\ 0 & A - K_t C \end{bmatrix} \begin{bmatrix} x_t \\ \tilde{x}_{t|t-1} \end{bmatrix} + \begin{bmatrix} I & 0 \\ I & -K_t \end{bmatrix} \begin{bmatrix} v_t \\ e_t \end{bmatrix} \tag{74}$$

$$= A_{cl} \begin{bmatrix} x_{t+1} \\ \tilde{x}_{t|t-1} \end{bmatrix} + G \begin{bmatrix} v_t \\ e_t \end{bmatrix} \tag{75}$$

Closed-loop mean and covariance

$$m_{t+1} = A_{cl} m_t \rightarrow 0 \qquad \text{(iff asym. stable)} \tag{76}$$

$$\Sigma_{t+1} = A_{cl} \Sigma_t A_{cl}^T + G \bar{R}_1 G^T \rightarrow \begin{bmatrix} P_x & P_\infty \\ P_\infty & P_\infty \end{bmatrix} \qquad \text{(iff asym. stable)} \tag{77}$$

$$\bar{R}_1 = \mathrm{diag}(R_v, R_e) \tag{78}$$

Stationary covariance (Ricatti equation for the predictive Kalman filter)

$$P_\infty = A P_\infty A^T + R_1 - K_\infty (C P_\infty C^T + R_2) K_\infty^T \tag{79}$$

Closed loop LQG - Predictive

Closed-loop system: LQG controller based on predictive Kalman filter

$$\begin{bmatrix} x_{t+1} \\ \tilde{x}_{t+1|t} \end{bmatrix} = \begin{bmatrix} A - BL_t & BL_t \\ 0 & A - K_t C \end{bmatrix} \begin{bmatrix} x_t \\ \tilde{x}_{t|t-1} \end{bmatrix} + \begin{bmatrix} I & 0 \\ I & -K_t \end{bmatrix} \begin{bmatrix} v_t \\ e_t \end{bmatrix} \tag{80}$$

$$= A_{cl} \begin{bmatrix} x_{t+1} \\ \tilde{x}_{t|t-1} \end{bmatrix} + G \begin{bmatrix} v_t \\ e_t \end{bmatrix} \tag{81}$$

Closed-loop input and output mean and covariance

$$u_t = -L_t \hat{x}_{t|t-1} \tag{82}$$

$$= -L_t(x_t - \tilde{x}_{t|t-1}) \sim N\left(\begin{bmatrix} -L_t & L_t \end{bmatrix} m_t, \begin{bmatrix} -L_t & L_t \end{bmatrix} \Sigma_t \begin{bmatrix} -L_t^T \\ L_t^T \end{bmatrix} \right), \tag{83}$$

$$y_t = Cx_t \sim N\left(\begin{bmatrix} C & 0 \end{bmatrix} m_t, \begin{bmatrix} C & 0 \end{bmatrix} \Sigma_t \begin{bmatrix} C \\ 0 \end{bmatrix} \right) = N(Cm_{x,t}, CP_{x,t}C^T) \tag{84}$$

Stationary: $\tilde{x} \sim N(0, P_\infty)$

Closed loop LQG - Ordinary

Closed-loop system: LQG controller based on ordinary Kalman filter

$$\begin{bmatrix} x_{t+1} \\ \tilde{x}_{t+1|t+1} \end{bmatrix} = \begin{bmatrix} A - BL_t & BL_t \\ 0 & A - \kappa_t CA \end{bmatrix} \begin{bmatrix} x_t \\ \tilde{x}_{t|t} \end{bmatrix} + \begin{bmatrix} I & 0 \\ I - \kappa C & -\kappa_t \end{bmatrix} \begin{bmatrix} v_t \\ e_t \end{bmatrix} \tag{85}$$

$$= A_{cl} \begin{bmatrix} x_t \\ \tilde{x}_{t|t} \end{bmatrix} + G \begin{bmatrix} v_t \\ e_t \end{bmatrix} \tag{86}$$

Closed-loop mean and covariance

$$m_{t+1} = A_{cl} m_t \to 0 \quad \text{(iff asym. stable)} \tag{87}$$

$$\Sigma_{t+1} = A_{cl} \Sigma_t A_{cl}^T + G \bar{R}_1 G^T \to \begin{bmatrix} P_x & \bar{P}_\infty \\ \bar{P}_\infty & \bar{P}_\infty \end{bmatrix} \quad \text{(iff asym. stable)} \tag{88}$$

$$\bar{R}_1 = \mathrm{diag}(R_v, R_e) \tag{89}$$

Stationary covariance (Riccati equation for the ordinary Kalman filter)

$$\bar{P}_\infty = (I - \kappa_\infty C)(A \bar{P}_\infty A^T + R_1)(I - \kappa_\infty C)^T + \kappa_\infty R_2 \kappa_\infty \tag{90}$$

Closed loop LQG - Ordinary

Closed-loop system: LQG controller based on ordinary Kalman filter

$$\begin{bmatrix} x_{t+1} \\ \tilde{x}_{t+1|t+1} \end{bmatrix} = \begin{bmatrix} A - BL_t & BL_t \\ 0 & A - \kappa_t CA \end{bmatrix} \begin{bmatrix} x_t \\ \tilde{x}_{t|t} \end{bmatrix} + \begin{bmatrix} I & 0 \\ I - \kappa C & -\kappa_t \end{bmatrix} \begin{bmatrix} v_t \\ e_t \end{bmatrix} \quad (91)$$

$$= A_{cl} \begin{bmatrix} x_t \\ \tilde{x}_{t|t} \end{bmatrix} + G \begin{bmatrix} v_t \\ e_t \end{bmatrix} \quad (92)$$

Closed-loop input and output mean and covariance

$$u_t = -L_t \hat{x}_{t|t} \quad (93)$$

$$= -L_t(x_t - \tilde{x}_{t|t}) \sim N\left(\begin{bmatrix} -L_t & L_t \end{bmatrix} m_t, \begin{bmatrix} -L_t & L_t \end{bmatrix} \Sigma_t \begin{bmatrix} -L_t^T \\ L_t^T \end{bmatrix} \right), \quad (94)$$

$$y_t = Cx_t \sim N\left(\begin{bmatrix} C & 0 \end{bmatrix} m_t, \begin{bmatrix} C & 0 \end{bmatrix} \Sigma_t \begin{bmatrix} C \\ 0 \end{bmatrix} \right) = N(Cm_{x,t}, CP_{x,t}C^T) \quad (95)$$

Stationary: $\tilde{x} \in N(0, \bar{P}_\infty)$.

Questions?