

State-Space Control Fundamentals (00860721)

lecture no. 11

Leonid Mirkin

The Stephen B. Klein Faculty of Aerospace Engineering
Technion—IIT



1/84

Outline

Sampled-data control systems and discretization

Structural properties of discrete state-space realizations

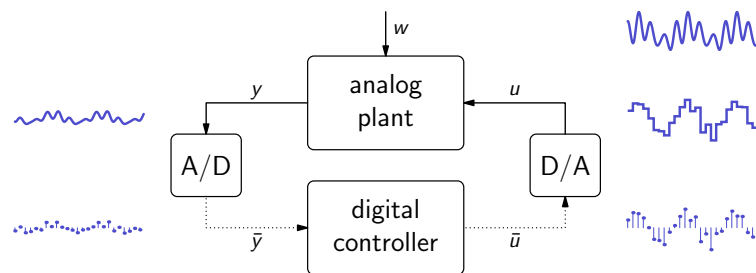
Discrete-time state-feedback

Finite-horizon LQR

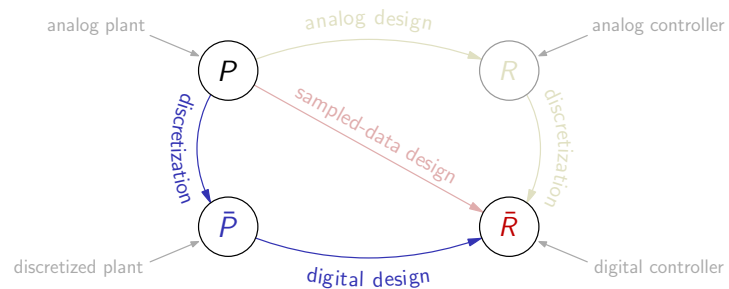
Discrete state-space realizations *ab initio* (for curious)

2/84

Sampled-data control

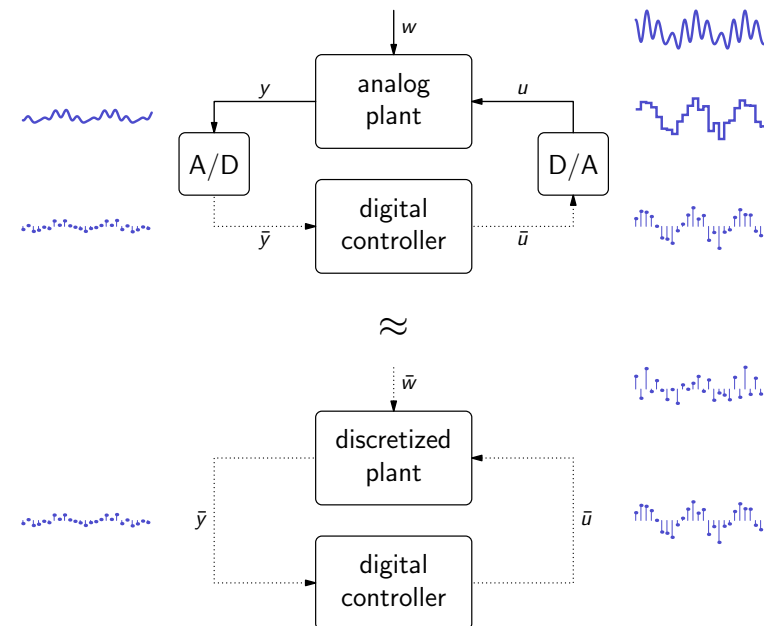


Approaches:



3/84

Problem discretization



4/84

Discretized plant

Given

- analog plant P
- A/D converter (sampler) $\mathcal{S}$ typically, the ideal sampler
- D/A converter (hold) $\mathcal{H}$ typically, the zero-order hold

find a discrete system $\bar{P}$ such that

$$\bar{P} = \mathcal{S}P\mathcal{H}.$$

This is not an approximation, it is the exact model of the dynamics $\bar{u} \mapsto \bar{y}$. Nevertheless, this leads to an approximate problem, for

- w is not necessarily of the form $w = \mathcal{H}\bar{w}$
hence, we have exact dynamics $u \mapsto y$, but not $w \mapsto y$
- a “good” behavior of $\bar{y} = \mathcal{S}y$ does not necessarily implies that of y

5/84

Discretization with ZOH

Let

$$P : \begin{cases} \dot{x}(t) = Ax(t) + Bu(t), & x(0) = x_0 \\ y(t) = Cx(t) \end{cases}$$

Its solution starting any $t_0 \geq 0$ is

$$x(t) = e^{A(t-t_0)}x(t_0) + \int_{t_0}^t e^{A(t-s)}Bu(s)ds.$$

Assume now that

$$u(t) = \bar{u}[i], \quad \forall t \in (ih, (i+1)h]$$

for some $h > 0$ (discretization interval). In this case

$$x(ih + \tau) = e^{A\tau}x(ih) + \int_{ih}^{ih+\tau} e^{A(ih+\tau-s)}dsB\bar{u}[i] = e^{A\tau}x(ih) + \int_0^\tau e^{A s}dsB\bar{u}[i]$$

for all $\tau \in [0, h]$.

6/84

Discretization with ZOH / Matlab: c2d(A,B,h)

Denoting $\bar{x} = \mathcal{S}x$, meaning $\bar{x}[i] = x(ih)$, we end up with the evolution

$$\bar{x}[i+1] = e^{Ah}\bar{x}[i] + \int_0^h e^{As}dsB\bar{u}[i] = \bar{A}\bar{x}[i] + \bar{B}\bar{u}[i],$$

where

$$[\bar{A} \ \bar{B}] := [e^{Ah} \ \int_0^h e^{As}dsB] = [I \ 0] \exp\left(\begin{bmatrix} A & B \\ 0 & 0 \end{bmatrix} h\right)$$

(the last relation follows by Van Loan's formulae, check it). Now,

$$y(t) = Cx(t) \implies y(ih) = Cx(ih) \implies \bar{y}[i] = C\bar{x}[i]$$

and we end up with

$$\bar{P} : \begin{cases} \bar{x}[i+1] = \bar{A}\bar{x}[i] + \bar{B}\bar{u}[i], & \bar{x}[0] = x_0 \\ \bar{y}[i] = C\bar{x}[i] \end{cases}$$

which is the discretized plant (aka its stroboscopic discretization).

7/84

Discretization with ZOH: two-tank example

With

$$\begin{bmatrix} A & B \end{bmatrix} = \begin{bmatrix} -1 & 1 & | & 1 \\ 1 & -2 & | & 0 \end{bmatrix},$$

we have

$$\bar{A} = \frac{1}{10} \begin{bmatrix} (5 - \sqrt{5})e^{\lambda_1 h} + (5 + \sqrt{5})e^{\lambda_2 h} & -2\sqrt{5}(e^{\lambda_1 h} - e^{\lambda_2 h}) \\ -2\sqrt{5}(e^{\lambda_1 h} - e^{\lambda_2 h}) & (5 + \sqrt{5})e^{\lambda_1 h} + (5 - \sqrt{5})e^{\lambda_2 h} \end{bmatrix}$$

and

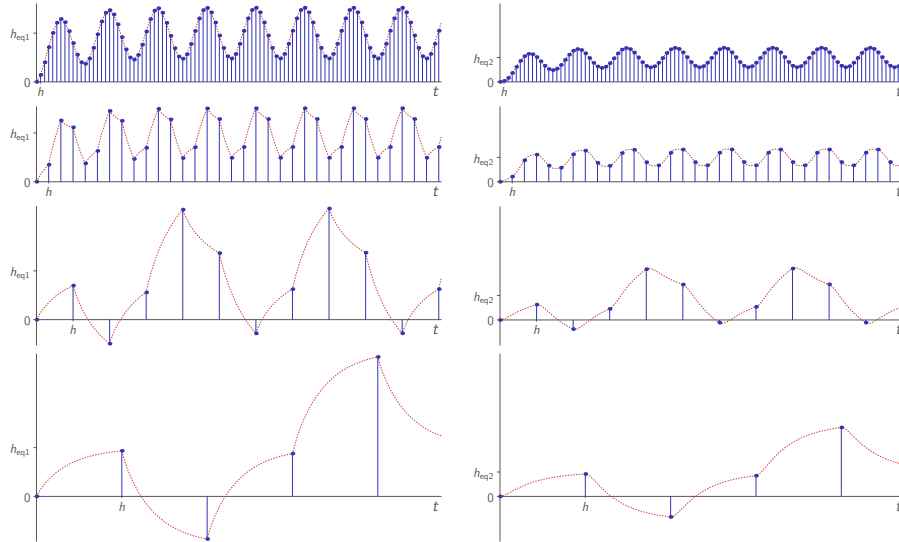
$$\bar{B} = \frac{1}{10} \begin{bmatrix} 20 - 2(5 - 2\sqrt{5})e^{\lambda_1 h} - 2(5 + 2\sqrt{5})e^{\lambda_2 h} \\ 10 - (5 - 3\sqrt{5})e^{\lambda_1 h} - (5 + 3\sqrt{5})e^{\lambda_2 h} \end{bmatrix}$$

where $\lambda_1 = -(3 + \sqrt{5})/2$ and $\lambda_2 = -(3 - \sqrt{5})/2$ are the eigenvalues of A .

8/84

Discretization with ZOH: two-tank example (contd)

Responses to $x[0] = -h_{eq}$ and $u = \mathcal{H}\bar{u}$, where $\bar{u}[i] = 0.5 \sin(0.5\pi ih)$, under various sampling intervals h :



9/84

Discretization with ZOH: properties

- $\bar{A}$ is always nonsingular
- $\lambda \in \text{spec}(A) \implies e^{\lambda h} \in \text{spec}(\bar{A})$
 - $\lambda \in \{s \in \mathbb{C} \mid \text{Re } s < 0\} \iff e^{\lambda h} \in \{s \in \mathbb{C} \mid |s| < 1\}$
 - $\lambda \in \{s \in \mathbb{C} \mid \text{Re } s = 0\} \iff e^{\lambda h} \in \{s \in \mathbb{C} \mid |s| = 1\}$
 - $\lambda \in \{s \in \mathbb{C} \mid \text{Re } s > 0\} \iff e^{\lambda h} \in \{s \in \mathbb{C} \mid |s| > 1\}$
- A is Hurwitz iff $\bar{A}$ is Schur stable
- sampling is termed **pathological** with respect to $\lambda \in \text{spec}(A)$ if there is $\mu \in \text{spec}(A)$ such that $\mu \neq \lambda$, whereas $e^{\mu h} = e^{\lambda h}$
e.g. $\text{spec}(\begin{bmatrix} 0 & \pi \\ -\pi & 0 \end{bmatrix}) = \{\pm j\pi\}$ and $e^{\lambda_1 h} = e^{\lambda_2 h} = -1$ for every odd $h \in \mathbb{Z}$
- if a bounded $\bar{u}$ produces a bounded $\bar{x}$ in $\bar{x}[i+1] = \bar{A}\bar{x}[i] + \bar{B}\bar{u}[i]$, then $u = \mathcal{H}\bar{u}$ produces a bounded x in $\dot{x}(t) = Ax(t) + Bu(t)$ for all $h > 0$ follows by

$$x(ih + \tau) = e^{A\tau} \bar{x}[i] + \int_0^\tau e^{A(\tau-s)} B \bar{u}[i] ds \quad : \quad \text{[plot of } x(t) \text{ over one sample interval]}$$

and the boundedness of $e^{A\tau}$ over any finite τ , so x cannot escape in the intersample

10/84

Discrete-time state equation

A causal finite-dimensional LTI system $G : u \mapsto y$ can be described by

$$\begin{cases} x[t+1] = Ax[t] + Bu[t], & x[0] = x_0 \\ y[t] = Cx[t] + Du[t] \end{cases}$$

where x is its state. The solution of the state equation (for $x_0 = 0$) is

$$x[t] = \sum_{s=0}^{t-1} A^{t-1-s} Bu[s] \implies y[t] = \sum_{s=0}^{t-1} CA^{t-1-s} Bu[s] + Du[t]$$

The latter is the convolution sum (δ is the unit pulse and $\mathbb{1}$ is the unit step)

$$y[t] = \sum_{s=-\infty}^{\infty} g[t-s]u[s], \quad \text{where } g[t] = D\delta[t] + CA^{t-1}B\mathbb{1}[t-1]$$

Thus

- we just replace integrals with sums, exponents with powers, et cetera.

11/84

Funniest thing about discrete-time systems...



Vincent: But you know what the funniest thing about **discrete-time systems** is?

Jules: What?

Vincent: It's the little differences. I mean, they got the same **properties** over there that they got here, but it's just... just their's a little different.

Jules: Examples?

12/84

Transfer functions

Naturally,

$$\begin{cases} x[t+1] = Ax[t] + Bu[t] \\ y[t] = Cx[t] + Du[t] \end{cases} \xrightarrow{\text{z-transform}} \begin{cases} zX(z) = AX(z) + BU(z) \\ Y(z) = CX(z) + DU(z) \end{cases}$$

and

$$G(z) = D + C(zI - A)^{-1}B =: \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array} \right].$$

Like in the s-domain,

- similarity transformations do not affect transfer functions
- canonical realizations remain the same
- $G(z)$ is **rational** (i.e. a ratio between two finite-order polynomials)
- **poles** of $G(z)$ are among the eigenvalues of A , i.e. in $\text{spec}(A)$
- **zeros** of $G(z)$ are among points where $\begin{bmatrix} A - zI & B \\ C & D \end{bmatrix}$ loses its rank

All system interconnection formulae remain the same as well...

13/84

Response with nonzero initial conditions

If $x_0 \neq 0$, then

$$y[t] = CA^t x_0 + \sum_{s=0}^{t-1} CA^{t-1-s} Bu[s] + Du[t]$$

It may be useful to visualize this relation in the form

$$\begin{bmatrix} y[0] \\ y[1] \\ y[2] \\ \vdots \end{bmatrix} = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \end{bmatrix} x_0 + \begin{bmatrix} D & 0 & 0 & \cdots \\ CB & D & 0 & \cdots \\ CAB & CB & D & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} u[0] \\ u[1] \\ u[2] \\ \vdots \end{bmatrix}$$

especially when considering the system over a finite time interval (horizon). This representation can be treated as a standard matrix relation, facilitating certain numerical optimizations, like those used in MPC.

14/84

Outline

Sampled-data control systems and discretization

Structural properties of discrete state-space realizations

Discrete-time state-feedback

Finite-horizon LQR

Discrete state-space realizations *ab initio* (for curious)

15/84

Reachability and controllability

Consider the state equation

$$x[t+1] = Ax[t] + Bu[t], \quad x[0] = x_0$$

for $A \in \mathbb{R}^{n \times n}$ and $B \in \mathbb{R}^{n \times m}$.

Definition 11.1 (reachability from the origin)

A state x_f is said to be **reachable** if, given $x_0 = 0$, there exist a time $t_f \in \mathbb{N}$ and a control signal $u : \mathbb{Z}_{0..t_f-1} \rightarrow \mathbb{R}^m$ such that $x[t_f] = x_f$. If every state is reachable, we say that the system or (A, B) is (completely) reachable.

Definition 11.2 (controllability to the origin)

An initial state x_0 is said to be **controllable** if there exist a time $t_f \in \mathbb{N}$ and a control signal $u : \mathbb{Z}_{0..t_f-1} \rightarrow \mathbb{R}^m$ such that $x[t_f] = 0$. If every initial state is controllable, we say that the system or (A, B) is (completely) controllable.

16/84

Reachability: the upper bound on t_f

Lemma 11.3

If $x_f \in \mathbb{R}^n$ is reachable in $t_f > n$ steps, then it is reachable in n steps as well.

Proof: By Cayley–Hamilton, $A^{n+j} = \sum_{i=0}^{n-1} \alpha_{j,i} A^{n-1-i}$ for some $\alpha_{j,i}$. Thus,

$$\begin{aligned} x[n+t] &= \sum_{s=0}^{n+t-1} A^{n+t-1-s} Bu[s] = \sum_{s=t}^{n+t-1} A^{n+t-1-s} Bu[s] + \sum_{s=0}^{t-1} A^{n+t-1-s} Bu[s] \\ &= \sum_{i=0}^{n-1} A^{n-1-i} Bu[i+t] + \sum_{s=0}^{t-1} \sum_{i=0}^{n-1} \alpha_{t-1-s,i} A^{n-1-i} Bu[s] \\ &= \sum_{i=0}^{n-1} A^{n-1-i} B \left(\underbrace{u[i+t] + \sum_{s=0}^{t-1} \alpha_{t-1-s,i} u[s]}_{\tilde{u}[i]} \right) = \sum_{s=0}^{n-1} A^{n-1-s} B \tilde{u}[s] \end{aligned}$$

So if $x[n+t] = x_f$, then the very same x_f is reachable in n steps by $\tilde{u}$. $\square$

17/84

Reachability and the controllability matrix

The matrix

$$M_c := [B \ AB \ \dots \ A^{n-1}B] \in \mathbb{R}^{n \times mn}$$

is known as the **controllability matrix** associated with (A, B) .

Proposition 11.4

(A, B) is reachable iff $\text{rank } M_c = n$ (i.e. M_c has full row rank).

Proof: We know that

$$x[n] = \sum_{i=0}^{n-1} A^{n-1-i} Bu[i] = \underbrace{[B \ AB \ \dots \ A^{n-1}B]}_{M_c} \underbrace{\begin{bmatrix} u[n-1] \\ \vdots \\ u[1] \\ u[0] \end{bmatrix}}_{U_n}.$$

if: $\text{rank } M_c = n \implies x_f = M_c U_n$ is solvable in U_n for all $x_f \in \mathbb{R}^n$.

only if: $\text{rank } M_c < n \implies \exists \eta \neq 0$ such that $\eta' M_c = 0 \implies \eta' x_f = 0$ for all $x_f = M_c U_n$ (i.e. reachable) $\implies x_f = \eta$ is not reachable.

The proof follows as reachability in n steps is equivalent to reachability. $\square$

18/84

Reachability from an arbitrary initial condition

Proposition 11.5

If (A, B) is reachable, then every x_f is reachable from all initial conditions x_0 .

Proof: Let x_f and x_0 be arbitrary and define $x_1 := x_f - A^n x_0 \in \mathbb{R}^n$. Because (A, B) is reachable, there is $u[t] = u_1[t]$, $t \in \mathbb{Z}_{0..n-1}$, such that

$$x_1 = x_f - A^n x_0 = \sum_{s=0}^{n-1} A^{n-1-s} B u_1[s].$$

But this implies that x_f is reachable from x_0 :

$$x_f = A^n x_0 + \sum_{s=0}^{n-1} A^{n-1-s} B u_1[s].$$

The result follows by the arbitrariness of x_f and x_0 . $\square$

19/84

Controllability and the controllability matrix

A corollary of the last lemma is that

– reachability $\implies$ controllability.

The other direction, reachability $\Leftarrow$ controllability, is *not necessarily* true:

Example: Let

$$[A \mid B] = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \implies \text{rank } M_c = \text{rank} \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} = 1,$$

meaning this (A, B) is not reachable. But the dynamics

$$\begin{cases} x_1[t+1] = x_1[t] + x_2[t] + u[t], & x_1[0] = x_{0,1} \\ x_2[t+1] = 0, & x_2[0] = x_{0,2} \end{cases}$$

can be controlled to the origin in one step via $u[0] = -x_{0,1} - x_{0,2}$, meaning this (A, B) is controllable.

20/84

Controllability and the controllability matrix (contd)

Proposition 11.6

(A, B) is controllable iff $\text{rank } M_c = \text{rank} \begin{bmatrix} A^n & M_c \end{bmatrix}$.

Proof: An initial condition $x_0 \neq 0$ is controllable iff

$$0 = A^n x_0 + \sum_{i=0}^{n-1} A^{n-1-i} B u[i] \iff M_c U_n = -A^n x_0$$

is solvable in $U_n \iff \text{rank } M_c = \text{rank} \begin{bmatrix} -A^n x_0 & M_c \end{bmatrix}$. Because

$$\begin{bmatrix} -A^n x_0 & M_c \end{bmatrix} = \begin{bmatrix} A^n & M_c \end{bmatrix} \begin{bmatrix} -x_0 & 0 \\ 0 & I_{mn} \end{bmatrix} \text{ and } \text{rank} \begin{bmatrix} -x_0 & 0 \\ 0 & I_{mn} \end{bmatrix} = mn + 1,$$

$\text{rank } M_c \leq \text{rank} \begin{bmatrix} -A^n x_0 & M_c \end{bmatrix} \leq \text{rank} \begin{bmatrix} A^n & M_c \end{bmatrix}$ holds for all x_0 , whence sufficiency follows. If $\text{rank } M_c < \text{rank} \begin{bmatrix} A^n & M_c \end{bmatrix}$, then there is $\eta \neq 0$ such that $\text{rank } M_c < \text{rank} \begin{bmatrix} A^n \eta & M_c \end{bmatrix}$. Hence, $x_0 = -\eta$ is not controllable and we have necessity. $\square$

21/84

Controllability and the controllability matrix (contd)

Example: Return to $\begin{bmatrix} A & B \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$. In this case

$$M_c = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \text{ and } \begin{bmatrix} A^n & M_c \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \left(\text{as } A^2 = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \right)$$

and

$$\text{rank} \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} = \text{rank} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} = 1,$$

as expected.

If $\det A \neq 0$, then $\text{rank} \begin{bmatrix} A^n & M_c \end{bmatrix} = n$ irrespective of M_c , so that

$$\text{rank } M_c = \text{rank} \begin{bmatrix} A^n & M_c \end{bmatrix} \iff \text{rank } M_c = n$$

and **reachability** $\iff$ **controllability**.

22/84

Reachability and controllability criteria

Theorem 11.7

The following statements are equivalent:

1. (A, B) is reachable
2. $M_c := \begin{bmatrix} B & AB & \dots & A^{n-1}B \end{bmatrix} \in \mathbb{R}^{n \times (nm)}$ has full row rank
3. $W_c[t] > 0$ for all $t \geq n$, where $W_c[t] := \sum_{s=0}^{t-1} A^s B B' (A^s)' \in \mathbb{R}^{n \times n}$
4. $\begin{bmatrix} A - \lambda I & B \end{bmatrix} \in \mathbb{C}^{n \times (n+m)}$ has full row rank for all $\lambda \in \mathbb{C}$ **PBH test**

The following statements are equivalent as well:

5. (A, B) is controllable
6. $\text{rank } M_c = \text{rank} \begin{bmatrix} A^n & M_c \end{bmatrix}$
7. $\text{rank} \begin{bmatrix} A - \lambda I & B \end{bmatrix} = n$ for all $\lambda \in \mathbb{C} \setminus \{0\}$

Remark: It is readily seen that $W_c[n] = M_c M_c'$, which immediately yields that 2. $\iff$ 3.

Remark: If A is Schur, i.e. $\text{spec}(A) \in \mathbb{D}_1 := \{z \in \mathbb{C} \mid |z| < 1\}$, then $W_c := \lim_{t \rightarrow \infty} W_c[t]$ exists and solves the discrete Lyapunov $W_c - A W_c A' - B B' = 0$ (**controllability Gramian**).

23/84

Observability

Consider the state equation with zero input

$$\begin{cases} x[t+1] = A x[t], & x[0] = x_0 \\ y[t] = C x[t] \end{cases}$$

for $A \in \mathbb{R}^{n \times n}$ and $C \in \mathbb{R}^{p \times n}$.

Definition 11.8 (observability)

An initial state x_0 is said to be **observable** if there exists a time $t_f \in \mathbb{N}$ such that x_0 can be reconstructed from measurements $y[t]$, $t \in \mathbb{Z}_{0..t_f-1}$. If every initial state is observable, we say that the system or (C, A) is observable.

24/84

Observability: the upper bound on t_f

Lemma 11.9

If $x_0 \in \mathbb{R}^n$ is observable in $t_f > n$ steps, then it is observable in n steps too.

Proof: By Cayley–Hamilton, $A^{n+k} = \sum_{i=0}^{n-1} \beta_{k,i} A^i$ for some $\beta_{k,i}$. Thus,

$$y[n+k] = CA^{n+k}x_0 = C \sum_{i=0}^{n-1} \beta_{k,i} A^i x_0 = \sum_{i=0}^{n-1} \beta_{k,i} CA^i x_0 = \sum_{i=0}^{n-1} \beta_{k,i} y[i]$$

for every x_0 . Hence, measurements $y[t]$ for $t \geq n$ are linear combinations of measurements $y[t]$ for $t \in \mathbb{Z}_{0..n-1}$ and contain no new information. $\square$

25/84

Observability and the observability matrix

The matrix

$$M_o := \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} \in \mathbb{R}^{pn \times n}$$

is known as the **observability matrix** associated with (C, A) .

Theorem 11.10

(C, A) is observable iff $\text{rank } M_o = n$ (i.e. M_o has full column rank).

Proof: Let $Y_n := [y'[0] \ y'[1] \ \cdots \ y'[n-1]]'$. The first n measurements satisfy $Y_n = M_o x_0$.

if: $\text{rank } M_o = n \implies Y_n = M_o x_0 \in \mathbb{R}^n$ has a unique x_0 satisfying it.

only if: $\text{rank } M_o < n \implies \exists \eta \neq 0$ such that $M_o \eta = 0 \implies y \equiv 0$ under both $x_0 = 0$ and $x_0 = \eta$ and we cannot tell whether measurements are caused by η or by 0. Hence, $x_0 = \eta$ is not observable. $\square$

26/84

Observability criteria

Theorem 11.11

The following statements are equivalent:

1. (C, A) is observable
2. $\text{rank } M_o = n$

3. $W_o[k] > 0$ for all $k \geq n$, where $W_o[k] := \sum_{i=0}^{k-1} (A^i)' C' C A^i \in \mathbb{R}^{n \times n}$

4. $\begin{bmatrix} A - \lambda I \\ C \end{bmatrix}$ has full column rank for all $\lambda \in \mathbb{C}$ **PBH test**

5. (A', C') is reachable

Remark: It is readily seen that $W_o[n] = M_o' M_o$, which immediately yields that 2. $\iff$ 3.

Remark: If A is Schur, i.e. $\text{spec}(A) \in \mathbb{D}_1 := \{z \in \mathbb{C} \mid |z| < 1\}$, then $W_o := \lim_{t \rightarrow \infty} W_o[t]$ exists and solves the discrete Lyapunov $W_o - A' W_o A - C' C = 0$ (**observability Gramian**).

27/84

Consequences of PBH

Exactly like in the continuous-time case, eigenvalues of A at which

- PBH for (A, B) fails are called **unreachable modes**
- PBH for (C, A) fails are called **unobservable modes**

Also,

- every unreachable mode of (A, B) is an eigenvalue of $A + BK$ for all K
- every unobservable mode of (C, A) is an eigenvalue of $A + LC$ for all L

and

- pair (A, B) is stabilizable iff $\text{rank} \begin{bmatrix} A - \lambda I & B \end{bmatrix} = n$ for all $\lambda \in \mathbb{C} \setminus \mathbb{D}_1$
- pair (C, A) is detectable iff $\text{rank} \begin{bmatrix} A - \lambda I \\ C \end{bmatrix} = n$ for all $\lambda \in \mathbb{C} \setminus \mathbb{D}_1$

with natural definitions of stabilizability and detectability.

Moreover,

- realization (A, B, C, D) is minimal iff it is reachable and observable

28/84

*ability and similarity transformations

Similarity transformations do not affect reachability, observability, et cetera.

If $(\tilde{A}, \tilde{B}, \tilde{C}) := (TAT^{-1}, TB, CT^{-1})$, then

- $\tilde{M}_c = TM_c$ not similar
- $\tilde{W}_c[t] = TW_c[t]T'$ for all $t \in \mathbb{Z}_+$ not similar
- $[\tilde{A} - \lambda I \quad \tilde{B}] = T[A - \lambda I \quad B] \begin{bmatrix} T^{-1} & 0 \\ 0 & I \end{bmatrix}$ for all $\lambda \in \mathbb{C}$
- $\tilde{A} + \tilde{B}\tilde{K} = T(A + BK)T^{-1}$ if $\tilde{K} = FT^{-1}$ similar
- $\tilde{M}_o = M_o T^{-1}$ not similar
- $\tilde{W}_o[t] = (T^{-1})'W_o[t]T^{-1}$ for all $t \in \mathbb{Z}_+$ not similar
- $\begin{bmatrix} \tilde{A} - \lambda I \\ \tilde{C} \end{bmatrix} = \begin{bmatrix} T & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} A - \lambda I \\ C \end{bmatrix} T^{-1}$ for all $\lambda \in \mathbb{C}$
- $\tilde{A} + \tilde{L}\tilde{C} = T(A + LC)T^{-1}$ if $\tilde{L} = TL$ similar

29/84

Outline

Sampled-data control systems and discretization

Structural properties of discrete state-space realizations

Discrete-time state-feedback

Finite-horizon LQR

Discrete state-space realizations *ab initio* (for curious)

30/84

State feedback in discrete time

If $x[t+1] = Ax[t] + Bu[t]$, then the state-feedback control law is still

$$u = Kx + v$$

with the closed-loop dynamics $x[t+1] = (A + BK)x[t]$. We may design K

- to assign eigenvalues of $A_K := A + BK$ to desired locations
- to satisfy some quadratic Lyapunov function $x'Px$
- via minimizing a quadratic cost function
- via solving an H_∞ problem

These approaches are quite straightforward extensions of the corresponding continuous-time results, normally with only slight deviations...

31/84

Discrete pole placement

Ackermann's formula, including its derivation, remains unchanged. Namely, if $m = 1$, then

$$K = - \begin{bmatrix} 0 & \cdots & 0 & 1 \end{bmatrix} M_c^{-1} \chi_{cl}(A),$$

where $\chi_{cl}(z)$ is the required closed-loop characteristic polynomial. The only technical difference from the continuous-time case is that

- $\chi_{cl}(z)$ is supposed to be Schur (roots are in $\mathbb{D}_1$), rather than Hurwitz.

One special case has no continuous-time counterpart though. If $\chi_{cl}(z) = z^n$, for which $K = - \begin{bmatrix} 0 & \cdots & 0 & 1 \end{bmatrix} M_c^{-1} A^n$, then A_K is nilpotent, i.e.

- there is $k \in \mathbb{Z}_{1..n}$ such that $A_K^i = 0$ for all $i \geq k$.

This, in turn, implies that

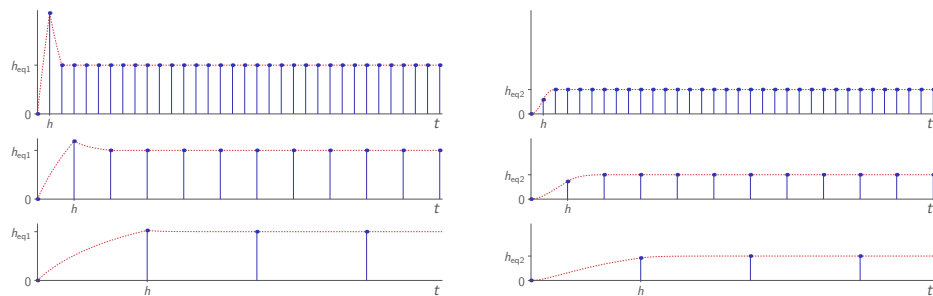
$$x[t] = A_K^t x_0 = 0 \quad \forall t \geq k$$

i.e. transients end after a finite number, upperbounded by the system order, of steps for every initial condition. Such systems are known as **deadbeat**.

32/84

Two-tank example: deadbeat control

Responses to $x[0] = -h_{eq}$ under various sampling intervals h ,



converge to their equilibria in at most 2 discrete steps.

Beware, deadbeat control tends to

- produce large control signals the fastest possible modes
- result in systems sensitive to modeling errors
- have poor intersample analog behavior in some cases

33/84

Stability of unforced motion (Lyapunov stability)

Let

$$x[t+1] = f(x[t]).$$

Its equilibrium $x_{eq} \in \mathbb{R}^n$, satisfying $x_{eq} = f(x_{eq})$, is said to be

- **stable** if $\forall \epsilon > 0$ there is $\delta = \delta(\epsilon) > 0$ such that

$$\|x[0] - x_{eq}\| < \delta \implies \|x[t] - x_{eq}\| < \epsilon, \quad \forall t \in \mathbb{R}_+$$

- **asymptotically stable** if stable and $\exists \delta > 0$ such that

$$\|x[0] - x_{eq}\| < \delta \implies \lim_{t \rightarrow \infty} \|x[t] - x_{eq}\| = 0$$

The **region of attraction** and **global stability** notions are the very same as in the continuous-time case. The linear system $x[t+1] = Ax[t]$ is

- stable iff $\text{spec}(A) \subset \{z \in \mathbb{C} \mid |z| \leq 1\}$ and every unit circle eigenvalue is not defective (i.e. its algebraic and geometric multiplicities are equal)
- asymptotically stable iff A is Schur, i.e. $\text{spec}(A) \subset \{z \in \mathbb{C} \mid |z| < 1\}$

34/84

Lyapunov's direct method in discrete time

Consider the system $x[t+1] = f(x[t])$ with an equilibrium at $x_{eq} = 0$. Let $\mathbb{X} \subset \mathbb{R}^n$ be so that $0 \in \mathbb{X}$. A *continuous* function $V: \mathbb{X} \rightarrow \mathbb{R}$ is a Lyapunov function candidate if

- $V(x) > 0$ for all $x(t) \in \mathbb{X} \setminus \{0\}$ and $V(0) = 0$.

Its forward difference along trajectories of the system

$$\Delta V(x[t]) := V(x[t+1]) - V(x[t]) = V(f(x[t])) - V(x[t]).$$

The system $x[t+1] = f(x[t])$ is

- stable if $\Delta V(x[t]) \leq 0$ for all $x[t] \in \mathbb{X}$
 x converges to the largest invariant set s.t. $\Delta V(x) \equiv 0$ LaSalle's invariance principle
- asymptotically stable if $\Delta V(x[t]) < 0$ for all $x[t] \in \mathbb{X} \setminus \{0\}$

Global versions require, in addition,

- g $\mathbb{X} = \mathbb{R}^n$,
- g radial unboundedness of V , i.e. that $V(x) \rightarrow \infty$ whenever $\|x\| \rightarrow \infty$.

35/84

Quadratic Lyapunov functions for linear systems

A popular choice in the linear case, $x[t+1] = Ax[t]$, is

$$V(x) = x'Px$$

for some $P = P' > 0$, which is a design parameter. In this case

$$\begin{aligned} \Delta V(x[t]) &= x'[t+1]Px[t+1] - x'[t]Px[t] = x'[t]A'PAx[t] - x'[t]Px[t] \\ &= x'[t](A'PA - P)x[t] \end{aligned}$$

Hence, by Lyapunov's direct method the system is

- asymptotically stable if there is $P > 0$ such that $A'PA - P < 0$.

This condition is known as the **discrete-time Lyapunov inequality** on P . If $A'PA - P < 0$, then $\exists Q = Q' > 0$ such that

$$P = A'PA + Q,$$

which is a discrete matrix Lyapunov equation.

36/84

Quadratic Lyapunov functions: necessity

Like in the continuous-time case, it may be sufficient to look for quadratic Lyapunov functions V for LTI systems.

Lemma 11.12

The system $x[t+1] = Ax[t]$ is asymptotically stable *iff* there is a quadratic Lyapunov function of the form $V = x'Px$ for some $P = P' > 0$ for it.

Proof.

if: By Lyapunov's direct method.

only if: If A is Schur, then $P = A'PA + Q$ is solvable for every Q and $P > 0$ whenever $Q > 0$. Thus, $V = x'Px$ is a quadratic Lyapunov function and $\Delta V = -x'Qx < 0$ for all $x \neq 0$. □

Lemma 11.13 (alternative formulation)

A is Schur *iff* there is $P = P' > 0$ such that $A'PA - P < 0$.

37/84

State-feedback design via quadratic Lyapunov functions

Designing stabilizing state feedback $u = Kx$ is thus equivalent to finding K and $P > 0$ such that

$$(A' + K'B')P(A + BK) - P < 0$$

This is not an LMI in the design variables P and K because of the presence of bilinear terms $A'PBK$ and $K'B'PA'$ and the cubic term $K'B'PBK$. And neither is

$$(MA' + N'B')M^{-1}(AM + BN) - M < 0,$$

where $M := P^{-1} > 0$ and $N := KP^{-1}$ ("the trick" in continuous time). We need another transformation to end up with the LMI

$$\begin{bmatrix} -M & MA' + N'B' \\ AM + BN & -M \end{bmatrix} < 0$$

cf. the Schur complement of its (2,2) block, also includes $M > 0$ as its part. Once solved, $K = NM^{-1}$. Voilà!

38/84

The LQR problem

Given

$$x[t+1] = Ax[t] + Bu[t], \quad x[0] = x_0$$

with $A \in \mathbb{R}^{n \times n}$ and $B \in \mathbb{R}^{n \times m}$, find u that

- stabilizes the systems ($\lim_{t \rightarrow \infty} x[t] = 0$),
- minimizes

$$\mathcal{J} = \sum_{t=0}^{\infty} \begin{bmatrix} x'[t] & u'[t] \end{bmatrix} \begin{bmatrix} Q & S \\ S' & R \end{bmatrix} \begin{bmatrix} x[t] \\ u[t] \end{bmatrix}$$

for given $Q = Q' \in \mathbb{R}^{n \times n}$, $R = R' \in \mathbb{R}^{m \times m}$, and $S \in \mathbb{R}^{n \times m}$.

Assumptions:

$\mathcal{A}_1$: (A, B) is stabilizable

$\mathcal{A}_2$: $\begin{bmatrix} Q & S \\ S' & R \end{bmatrix} \geq 0$ no need for $R > 0$

$\mathcal{A}_3$: $\begin{bmatrix} A - e^{-j\theta}I & B \\ Q & S \\ S' & R \end{bmatrix}$ has full column rank for all $\theta \in [-\pi, \pi]$

39/84

DARE / Matlab: `idare(A,B,Q,R,S)` or `idare(A,B,Q,R)`

Let $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$, $Q = Q' \in \mathbb{R}^{n \times n}$, $R = R' \in \mathbb{R}^{m \times m}$, and $S \in \mathbb{R}^{n \times m}$ be given. The matrix equation

$$\bar{X} = A'\bar{X}A + Q - (S + A'\bar{X}B)(R + B'\bar{X}B)^{-1}(S' + B'\bar{X}A)$$

is called the **discrete-time algebraic Riccati equation** (DARE). Its solution $\bar{X} \in \mathbb{R}^{n \times n}$ is said to be stabilizing if $\det(R + B'\bar{X}B) \neq 0$ and the matrix

$$A_K := A + BK, \quad \text{where } K = -(R + B'\bar{X}B)^{-1}(S' + B'\bar{X}A),$$

is Schur, i.e. having eigenvalues in $\mathbb{D}_1 := \{z \in \mathbb{C} \mid |z| < 1\}$. For LQR we are interested in DAREs with

$$\begin{bmatrix} Q & S \\ S' & R \end{bmatrix} \geq 0$$

which requires $Q \geq 0$ and $R \geq 0$, but not $R > 0$.

40/84

DARE: example

Let $A = a$, $B = 1$, $Q = q \geq 0$, $S = 0$, and $R = r \geq 0$. The DARE reads

$$\bar{X} = a^2 \bar{X} + q - \frac{a^2 \bar{X}^2}{r + \bar{X}} = \frac{ra^2 \bar{X}}{r + \bar{X}} + q$$

1. If $r = 0$, then the DARE becomes $\bar{X} = q$ and $r + \bar{X} \neq 0 \iff q \neq 0$. In the latter case, $A_K = 0$ is Schur.
2. If $r > 0$, then the equation reads

$$\bar{X}^2 - (r(a^2 - 1) + q)\bar{X} - qr = 0 \quad \text{and} \quad \bar{X} \neq -r.$$

The first equation is solved by

$$\bar{X} = \frac{r(a^2 - 1) + q}{2} \pm \sqrt{\frac{(r(a^2 - 1) + q)^2}{4} + qr}$$

with $\bar{X} \geq 0$ for the “+” solution ($\bar{X} = 0$ iff $q = 0$ and $|a| \leq 1$ then).

41/84

DARE: example (contd)

2. Now,

$$r + \bar{X} = \frac{r(a^2 + 1) + q}{2} \pm \sqrt{\frac{(r(a^2 - 1) + q)^2}{4} + qr}$$

is zero only for its “−” part for $a = 0$. Furthermore,

$$A_K = a - \frac{a\bar{X}}{r + \bar{X}} = \frac{r(a^2 + 1) + q}{2ar} \mp \frac{\sqrt{(r(a^2 + 1) + q)^2 - (2ar)^2}}{2ar}$$

and it is Schur for the “+” solution of $\bar{X}$. Hence, the stabilizing

$$\bar{X} = \frac{r(a^2 - 1) + q}{2} + \sqrt{\frac{(r(a^2 - 1) + q)^2}{4} + qr}$$

and the resulting

$$A_K = \frac{a^2 + 1 + q/r}{2a} - \frac{\sqrt{(a^2 + 1 + q/r)^2 - 4a^2}}{2a} \xrightarrow{r \rightarrow 0} 0$$

42/84

DARE: solvability

Theorem 11.14

Let $\begin{bmatrix} Q & S \\ S' & R \end{bmatrix} \geq 0$. The DARE

$$\bar{X} = A'\bar{X}A + Q - (S + A'\bar{X}B)(R + B'\bar{X}B)^{-1}(S' + B'\bar{X}A)$$

has a stabilizing solution $\bar{X} \geq 0$ iff

1. (A, B) is stabilizable,
2. $\begin{bmatrix} A - \lambda I & B \\ Q & S \\ S' & R \end{bmatrix}$ has full column rank for all $\lambda \in \{z \in \mathbb{C} \mid |z| = 1\}$.

Moreover, $\bar{X} > 0$ iff the second condition is replaced with

- 2'. $\begin{bmatrix} A - \lambda I & B \\ Q & S \\ S' & R \end{bmatrix}$ has full column rank for all $\lambda \in \{z \in \mathbb{C} \mid |z| \leq 1\}$.

In any case, the stabilizing solution is unique.

43/84

Discrete LQR: solution

Theorem 11.15

If $\mathcal{A}_{1-3}$ hold, then the unique stabilizing controller that minimizes

$$\mathcal{J} = \sum_{t=0}^{\infty} \begin{bmatrix} x'[t] & u'[t] \end{bmatrix} \begin{bmatrix} Q & S \\ S' & R \end{bmatrix} \begin{bmatrix} x[t] \\ u[t] \end{bmatrix}$$

is

$$u[t] = -(R + B'\bar{X}B)^{-1}(S' + B'\bar{X}A)x[t],$$

where $\bar{X} = \bar{X}' \geq 0$ is the stabilizing solution to the DARE

$$\bar{X} = A'\bar{X}A + Q - (S + A'\bar{X}B)(R + B'\bar{X}B)^{-1}(S' + B'\bar{X}A)$$

The optimal cost is $\mathcal{J}_{\text{opt}} = x_0'\bar{X}x_0$ then.

Remark: Like in the continuous-time case, the optimal control law is the state feedback with $K = -(R + B'\bar{X}B)^{-1}(S' + B'\bar{X}A)$. The closed-loop “A” matrix $A + BK$ is Schur because $\bar{X}$ is the stabilizing solution.

44/84

Proof

As $K = -(R + B' \bar{X} B)^{-1}(S' + B' \bar{X} A) \implies S' + B' \bar{X} A = -(R + B' \bar{X} B)K$, we have that

$$\bar{X} = A' \bar{X} A + Q - (S + A' \bar{X} B)(R + B' \bar{X} B)^{-1}(S' + B' \bar{X} A)$$

$\Downarrow$

$$Q = \bar{X} - A' \bar{X} A + K'(R + B' \bar{X} B)K$$

and, therefore, that

$$\begin{bmatrix} Q & S \\ S' & R \end{bmatrix} = \begin{bmatrix} \bar{X} & 0 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} A' \\ B' \end{bmatrix} \bar{X} \begin{bmatrix} A & B \end{bmatrix} + \begin{bmatrix} -K' \\ I \end{bmatrix} (R + B' \bar{X} B) \begin{bmatrix} -K & I \end{bmatrix}$$

This is a discrete counterpart of a key relation in the continuous-time case and it plays a similar role in the discrete proof.

45/84

Proof (contd)

The quadratic form in the cost $\mathcal{J}$ is

$$\begin{aligned} & \begin{bmatrix} x'[t] & u'[t] \end{bmatrix} \begin{bmatrix} Q & S \\ S' & R \end{bmatrix} \begin{bmatrix} x[t] \\ u[t] \end{bmatrix} \\ &= \begin{bmatrix} x'[t] & u'[t] \end{bmatrix} \begin{bmatrix} -K' \\ I \end{bmatrix} (R + B' \bar{X} B) \begin{bmatrix} -K & I \end{bmatrix} \begin{bmatrix} x[t] \\ u[t] \end{bmatrix} \\ &\quad + \begin{bmatrix} x'[t] & u'[t] \end{bmatrix} \left(\begin{bmatrix} \bar{X} & 0 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} A' \\ B' \end{bmatrix} \bar{X} \begin{bmatrix} A & B \end{bmatrix} \right) \begin{bmatrix} x[t] \\ u[t] \end{bmatrix} \\ &= (u[t] - Kx[t])'(R + B' \bar{X} B)(u[t] - Kx[t]) \\ &\quad + x'[t] \bar{X} x[t] - x'[t+1] \bar{X} x[t+1] \\ &= \|(R + B' \bar{X} B)^{1/2}(u[t] - Kx[t])\|^2 + x'[t] \bar{X} x[t] - x'[t+1] \bar{X} x[t+1] \end{aligned}$$

where $M^{1/2} > 0$ is the matrix square root (Lecture 1, Slide 29, Property 10) of $M > 0$.

46/84

Proof (contd)

Thus,

$$\begin{aligned} \mathcal{J} &= \sum_{t=0}^{\infty} \left(\|(R + B' \bar{X} B)^{1/2}(u[t] - Kx[t])\|^2 \right. \\ &\quad \left. + x'[t] \bar{X} x[t] - x'[t+1] \bar{X} x[t+1] \right) \\ &= \sum_{t=0}^{\infty} \|(R + B' \bar{X} B)^{1/2}(u[t] - Kx[t])\|^2 + x'_0 \bar{X} x_0 - \lim_{t \rightarrow \infty} x'[t] \bar{X} x[t] \end{aligned}$$

If u is stabilizing, then $\lim_{t \rightarrow \infty} x'[t] \bar{X} x[t] = 0$. Because $(R + B' \bar{X} B)^{1/2} > 0$,

$$\mathcal{J} = \sum_{t=0}^{\infty} \|(R + B' \bar{X} B)^{1/2}(u[t] - Kx[t])\|^2 + x'_0 \bar{X} x_0 \geq x'_0 \bar{X} x_0$$

whenever the system is asymptotically stable and the equality is attained iff $u = Kx$, which is stabilizing ($\bar{X}$ is the stabilizing solution). $\square$

47/84

Suboptimal solutions

Corollary 11.16

Let $\mathcal{A}_{1-3}$ hold and the cost function

$$\mathcal{J} = \sum_{t=0}^{\infty} \begin{bmatrix} x'[t] & u'[t] \end{bmatrix} \begin{bmatrix} Q & S \\ S' & R \end{bmatrix} \begin{bmatrix} x[t] \\ u[t] \end{bmatrix}.$$

The control law

$$u[t] = -(R + B' \bar{X} B)^{-1}(S' + B' \bar{X} A)x[t] + (R + B' \bar{X} B)^{-1/2}v[t],$$

where $\bar{X} = \bar{X}' \geq 0$ is the stabilizing solution to the DARE

$$\bar{X} = A' \bar{X} A + Q - (S + A' \bar{X} B)(R + B' \bar{X} B)^{-1}(S' + B' \bar{X} A)$$

renders

$$\mathcal{J} = x'_0 \bar{X} x_0 + \sum_{t=0}^{\infty} \|v[t]\|^2 = x'_0 \bar{X} x_0 + \|v\|_2^2.$$

48/84

What if $R = 0$ (and $S = 0$)?

In this case

$$K = -(B' \bar{X} B)^{-1} B' \bar{X} A \implies A + BK = (I - B(B' \bar{X} B)^{-1} B' \bar{X}) A.$$

and

$$B' \bar{X} (A + BK) = B' \bar{X} A - B' \bar{X} B (B' \bar{X} B)^{-1} B' \bar{X} A = 0.$$

To have $B' \bar{X} B$ nonsingular, $B' \bar{X} \in \mathbb{R}^{m \times n}$ must have full row rank. Indeed, if there is $0 \neq \eta \in \mathbb{R}^m$ such that $\eta' B' \bar{X} = 0$, then $\eta' B' \bar{X} B = 0$ as well, which contradicts the nonsingularity of $B' \bar{X} B$. So there are m linearly independent vectors $\eta_i \in \mathbb{R}^m$ (e.g. $\eta_i = \bar{X} B e_i$, where e_i is the i th standard basis of $\mathbb{R}^m$), such that $\eta_i' (A + BK) = 0$. Therefore,

- $A + BK$ has at least m eigenvalues at the origin (deadbeat modes).

49/84

Discrete-time return difference equality

Using the key relation, it can be shown that

$$(I + L'(z^{-1}))(R + B' \bar{X} B)(I + L(z)) = \begin{bmatrix} P'_x(z^{-1}) & I \end{bmatrix} \begin{bmatrix} Q & S \\ S' & R \end{bmatrix} \begin{bmatrix} P_x(z) \\ I \end{bmatrix}$$

where $L(z) = -K(zI - A)^{-1}B$ and $P_x(z) = (zI - A)^{-1}B$. If $m = 1$, $R = r$, and $S = 0$, then under $z = e^{j\theta}$

$$|1 + L(e^{j\theta})|^2 = \frac{r}{r + B' \bar{X} B} + \frac{[P_x(e^{j\theta})]' Q P_x(e^{j\theta})}{r + B' \bar{X} B} \geq \underbrace{\frac{r}{r + B' \bar{X} B}}_{< 1 \text{ unless } \bar{X} B = 0}$$

and we can no longer guarantee that $|1 + L(e^{j\theta})| \geq 1$ for all θ (in fact, it is not). As a result, unless $\bar{X} B = 0$ (implying $K = 0$ if $S = 0$)

- there are frequencies θ where the sensitivity magnitude $|S(e^{j\theta})| > 1$
- continuous-time LQR margins ($\mu_g = \infty / \mu_{ph} \geq 60^\circ$) cannot be attained perhaps expectably, discrete-time strictly proper systems cannot have $\mu_g = \infty$

50/84

Outline

Sampled-data control systems and discretization

Structural properties of discrete state-space realizations

Discrete-time state-feedback

Finite-horizon LQR

Discrete state-space realizations *ab initio* (for curious)

51/84

Yet another LQR problem

Consider

$$x[t+1] = Ax[t] + Bu[t], \quad x[0] = x_0$$

for given $A \in \mathbb{R}^{n \times n}$ and $B \in \mathbb{R}^{n \times m}$. Suppose it

- operates over the interval $\mathbb{Z}_{0..\eta}$ for a given $\eta \in \mathbb{N}$ (aka horizon).

The goal is to find u that

- minimizes

$$\mathcal{J} = x'[\eta] Q_\eta x[\eta] + \sum_{t=0}^{\eta-1} \begin{bmatrix} x'[t] & u'[t] \end{bmatrix} \begin{bmatrix} Q & S \\ S' & R \end{bmatrix} \begin{bmatrix} x[t] \\ u[t] \end{bmatrix}.$$

for conventional Q , R , and S and some $Q_\eta = Q'_\eta \geq 0$.

Unusual:

- finite horizon $\implies$ no stability requirements
- terminal penalty $x'[\eta] Q_\eta x[\eta]$

52/84

Difference Riccati equation

For every $t \in \mathbb{Z}_{0..\eta-1}$ define the sequence of matrices, running back in time,

$$\bar{X}[t] = A' \bar{X}[t+1] A + Q - (S + A' \bar{X}[t+1] B) \times (R + B' \bar{X}[t+1] B)^{-1} (S' + B' \bar{X}[t+1] A), \quad \bar{X}[\eta] = \bar{X}_\eta$$

If $\mathcal{A}_{1-3}$ hold and $\bar{X}_\eta \geq 0$, then

$$\lim_{t \rightarrow -\infty} \bar{X}[t] = \bar{X}_\infty$$

where $\bar{X}_\infty$ is the stabilizing solution to the DARE

$$\bar{X}_\infty = A' \bar{X}_\infty A + Q - (S + A' \bar{X}_\infty B)(R + B' \bar{X}_\infty B)^{-1} (S' + B' \bar{X}_\infty A).$$

Moreover,

- if horizon is very long, then $\bar{X}[t] \approx \bar{X}_\infty$ for almost all t
- $\bar{X}_\eta = \bar{X}_\infty \implies \bar{X}[t] = \bar{X}_\infty$ for all t

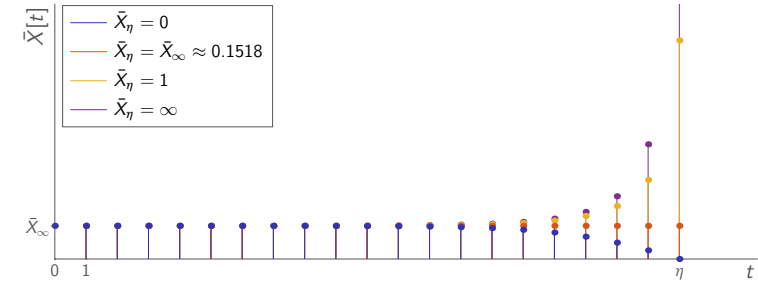
53/84

Example

If $A = a$, $B = 1$, $Q = q \geq 0$, $S = 0$, and $R = r \geq 0$, then the DDRE reads

$$\bar{X}[t] = q + \frac{ra^2 \bar{X}[t+1]}{r + \bar{X}[t+1]}.$$

For $a = r = 1$, $q = 0.02$, and the horizon $\eta = 20$, different $\bar{X}_\eta$ yield



54/84

Solution

Theorem 11.17

If $R + B' Q_\eta B > 0$, then the unique controller minimizing

$$\mathcal{J} = x'[\eta] Q_\eta x[\eta] + \sum_{t=0}^{\eta-1} [x'[t] \quad u'[t]] \begin{bmatrix} Q & S \\ S' & R \end{bmatrix} \begin{bmatrix} x[t] \\ u[t] \end{bmatrix}$$

is

$$u[t] = -(R + B' \bar{X}[t+1] B)^{-1} (S' + B' \bar{X}[t+1] A) x[t],$$

where $\bar{X}[t] = \bar{X}'[t] \geq 0$ is the solution to the difference ARE

$$\bar{X}[t] = A' \bar{X}[t+1] A + Q - (S + A' \bar{X}[t+1] B) \times (R + B' \bar{X}[t+1] B)^{-1} (S' + B' \bar{X}[t+1] A), \quad \bar{X}[\eta] = Q_\eta$$

The optimal cost is $\mathcal{J}_{opt} = x'_0 \bar{X}[0] x_0$ then.

55/84

Proof

The relation

$$\begin{bmatrix} Q & S \\ S' & R \end{bmatrix} = \begin{bmatrix} \bar{X} & 0 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} A' \\ B' \end{bmatrix} \bar{X} \begin{bmatrix} A & B \end{bmatrix} + \begin{bmatrix} -K' \\ I \end{bmatrix} (R + B' \bar{X} B) \begin{bmatrix} -K & I \end{bmatrix}$$

for the DARE solution $\bar{X}$ has a counterpart for DDRE, in the form

$$\begin{bmatrix} Q & S \\ S' & R \end{bmatrix} = \begin{bmatrix} \bar{X}[t] & 0 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} A' \\ B' \end{bmatrix} \bar{X}[t+1] \begin{bmatrix} A & B \end{bmatrix} + \begin{bmatrix} -K'[t] \\ I \end{bmatrix} (R + B' \bar{X}[t+1] B) \begin{bmatrix} -K[t] & I \end{bmatrix}$$

where

$$K[t] := -(R + B' \bar{X}[t+1] B)^{-1} (S' + B' \bar{X}[t+1] A).$$

It holds for all $t \in \mathbb{Z}_{0..\eta-1}$.

56/84

Proof (contd)

The quadratic form in the cost $\mathcal{J}$ is

$$\begin{aligned}
 & \begin{bmatrix} x'[t] & u'[t] \end{bmatrix} \begin{bmatrix} Q & S \\ S' & R \end{bmatrix} \begin{bmatrix} x[t] \\ u[t] \end{bmatrix} \\
 &= \begin{bmatrix} x'[t] & u'[t] \end{bmatrix} \begin{bmatrix} -K'[t] \\ I \end{bmatrix} (R + B' \bar{X}[t+1]B) \begin{bmatrix} -K[t] & I \end{bmatrix} \begin{bmatrix} x[t] \\ u[t] \end{bmatrix} \\
 &\quad + \begin{bmatrix} x'[t] & u'[t] \end{bmatrix} \left(\begin{bmatrix} \bar{X}[t] & 0 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} A' \\ B' \end{bmatrix} \bar{X}[t+1] \begin{bmatrix} A & B \end{bmatrix} \right) \begin{bmatrix} x[t] \\ u[t] \end{bmatrix} \\
 &= (u[t] - K[t]x[t])'(R + B' \bar{X}[t+1]B)(u[t] - K[t]x[t]) \\
 &\quad + x'[t] \bar{X}[t] x[t] - x'[t+1] \bar{X}[t+1] x[t+1] \\
 &= \|(R + B' \bar{X}[t+1]B)^{1/2}(u[t] - K[t]x[t])\|^2 \\
 &\quad + x'[t] \bar{X}[t] x[t] - x'[t+1] \bar{X}[t+1] x[t+1]
 \end{aligned}$$

57/84

Proof (contd)

Thus,

$$\begin{aligned}
 \mathcal{J} &= x'[\eta] Q_{\eta} x[\eta] + \sum_{t=0}^{\eta-1} \left(\|(R + B' \bar{X}[t+1]B)^{1/2}(u[t] - K[t]x[t])\|^2 \right. \\
 &\quad \left. + x'[t] \bar{X}[t] x[t] - x'[t+1] \bar{X}[t+1] x[t+1] \right) \\
 &= x'[\eta] Q_{\eta} x[\eta] + x'_0 \bar{X}[0] x_0 - x'[\eta] \bar{X}[\eta] x[\eta] \\
 &\quad + \sum_{t=0}^{\eta-1} \|(R + B' \bar{X}[t+1]B)^{1/2}(u[t] - K[t]x[t])\|^2 \\
 &= x'_0 \bar{X}[0] x_0 + \sum_{t=0}^{\eta-1} \|(R + B' \bar{X}[t+1]B)^{1/2}(u[t] - K[t]x[t])\|^2 \geq x'_0 \bar{X}[0] x_0
 \end{aligned}$$

Therefore, the unique control law minimizing $\mathcal{J}$ is $u[t] = K[t]x[t]$ and the optimal cost is the lower bound. $\square$

58/84

Example: DC motor with rigid mechanical load

Consider an LTI system G having the transfer function

$$G(s) = \frac{1}{s(s+1)},$$

which can be thought of as a voltage controlled DC motor with normalized time constant. Its companion form

$$\begin{cases} \dot{x}(t) = \begin{bmatrix} 0 & 1 \\ 0 & -1 \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(t) \\ y(t) = \begin{bmatrix} 1 & 0 \end{bmatrix} x(t) \end{cases}$$

with the state $x = \begin{bmatrix} y \\ \dot{y} \end{bmatrix}$. Note that the state feedback

$$u = Kx = \begin{bmatrix} k_1 & k_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = k_1 y + k_2 \dot{y}$$

is a PD controller in this case.

59/84

Example: requirements and LQR cost function

We would like to have

- fast convergence of y $\implies$ penalize $\|y\|_2$
- smooth convergence of y $\implies$ penalize $\|\dot{y}\|_2$
- small control signal u $\implies$ penalize $\|u\|_2$

Cost

$$\mathcal{J} = \int_0^{\infty} (\lambda y^2(t) + (1-\lambda) \dot{y}^2(t) + r u^2(t)) dt$$

for some weights $\lambda \in [0, 1]$ and $r > 0$ ($\lambda \uparrow 1$ encourages faster convergence, $\lambda \downarrow 0$ encourages smoother convergence, while the increase of r discourages large control energy).

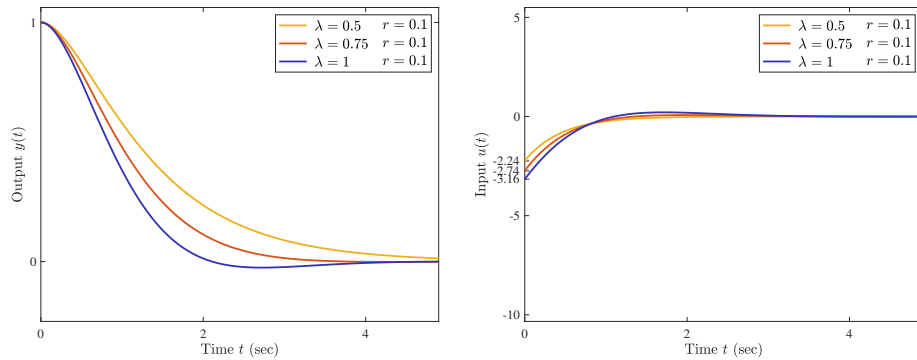
Taking into account that $y = x_1$ and $\dot{y} = x_2$, we get

$$\begin{bmatrix} Q & S \\ S' & R \end{bmatrix} = \begin{bmatrix} \lambda & 0 & 0 \\ 0 & 1-\lambda & 0 \\ 0 & 0 & r \end{bmatrix}.$$

60/84

Example: simulations of analog controller

With (large) $r = 0.1$:

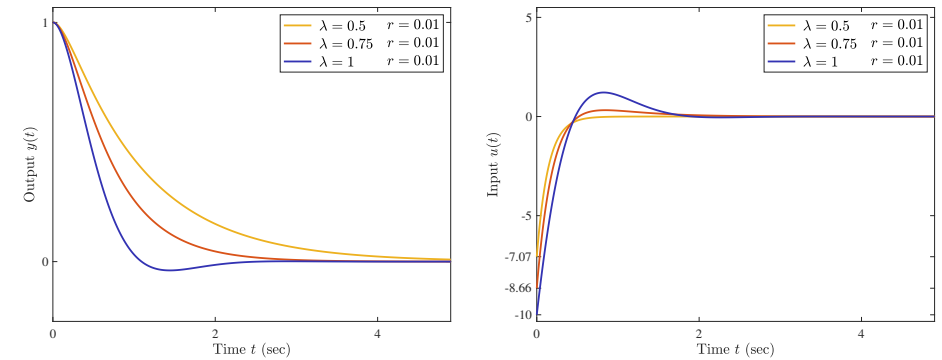


- under $\lambda = 1$ the response is faster, but overshoots (complex poles)
- as λ decreases, the response slows down, but smooths
- control effort decreases as λ decreases (always peaks at $-\sqrt{\lambda/r}$)

61/84

Example: simulations of analog controller (contd)

With (middle) $r = 0.01$:

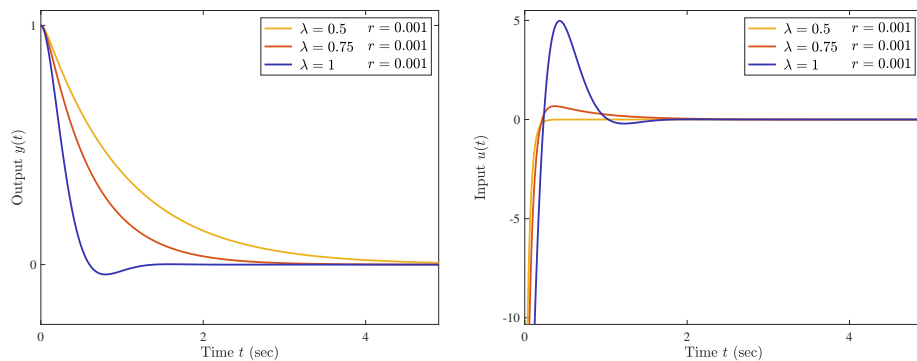


- under $\lambda = 1$ the response is faster, but overshoots (complex poles)
- as λ decreases, the response slows down, but smooths
- control effort decreases as λ decreases (always peaks at $-\sqrt{\lambda/r}$)

62/84

Example: simulations of analog controller (contd)

With (small) $r = 0.001$:



- under $\lambda = 1$ the response is faster, but overshoots (complex poles)
- as λ decreases, the response slows down, but smooths
- control effort decreases as λ decreases (always peaks at $-\sqrt{\lambda/r}$)

63/84

Example: discretization

The same DC motor has the discretized realization

$$\begin{cases} x[t+1] = \begin{bmatrix} 1 & 1 - e^{-h} \\ 0 & e^{-h} \end{bmatrix} x[t] + \begin{bmatrix} h - 1 + e^{-h} \\ 1 - e^{-h} \end{bmatrix} u[t] \\ y[t] = \begin{bmatrix} 1 & 0 \end{bmatrix} x[t] \end{cases}$$

and the state still has $x_1[i] = y(ih)$ and $x_2[i] = \dot{y}(ih)$. Hence, choosing

$$\begin{bmatrix} Q & S \\ S' & R \end{bmatrix} = \begin{bmatrix} \lambda & 0 & 0 \\ 0 & 1 - \lambda & 0 \\ 0 & 0 & r \end{bmatrix}$$

yields

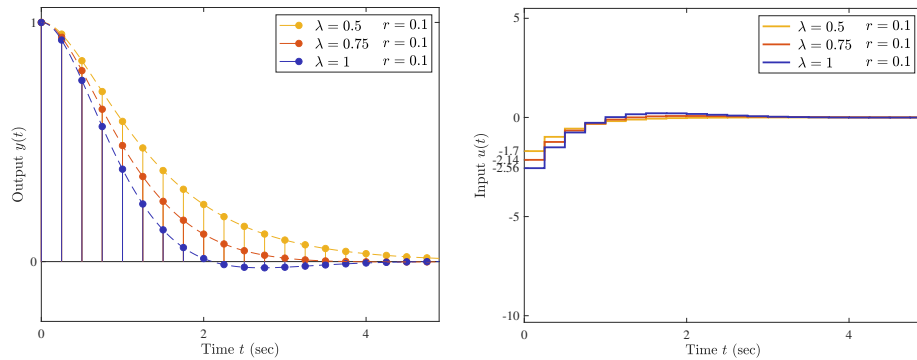
$$\mathcal{J} = \sum_{i=0}^{\infty} (\lambda y^2(ih) + (1 - \lambda) \dot{y}^2(ih) + r u^2[i])$$

which has the advantage of being physical.

64/84

Example: simulations of digital controller

With (large) $r = 0.1$:

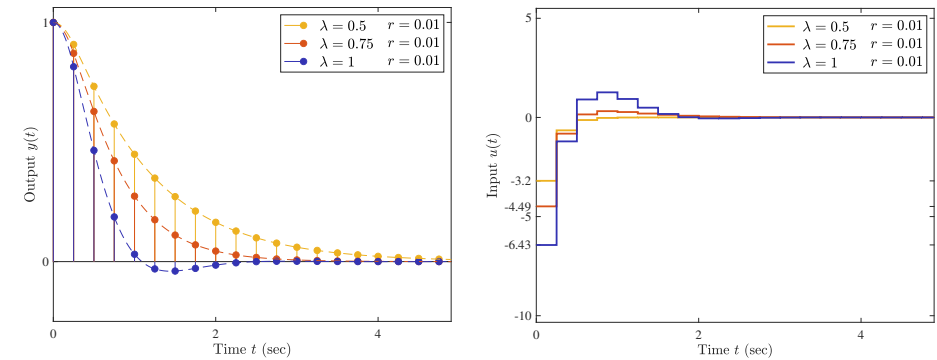


- under $\lambda = 1$ the response is faster, but overshoots (complex poles)
- as λ decreases, the response slows down, but smooths
- control effort decreases as λ decreases
- intersample behavior is as expected

65/84

Example: simulations of digital controller (contd)

With (middle) $r = 0.01$:

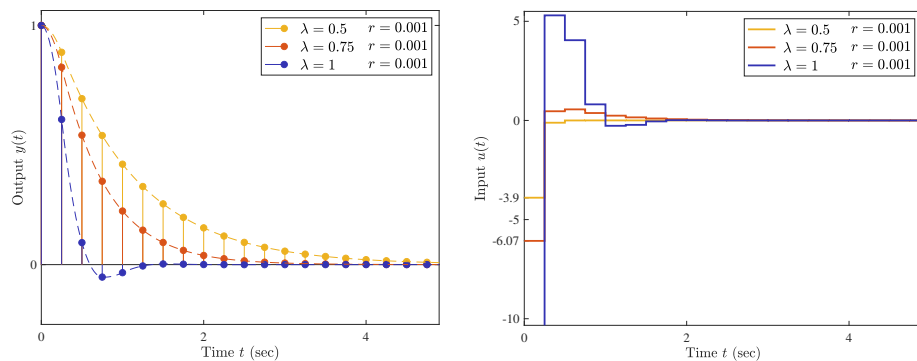


- under $\lambda = 1$ the response is faster, but overshoots (complex poles)
- as λ decreases, the response slows down, but smooths
- control effort decreases as λ decreases
- intersample behavior is as expected

66/84

Example: simulations of digital controller (contd)

With (small) $r = 0.001$:

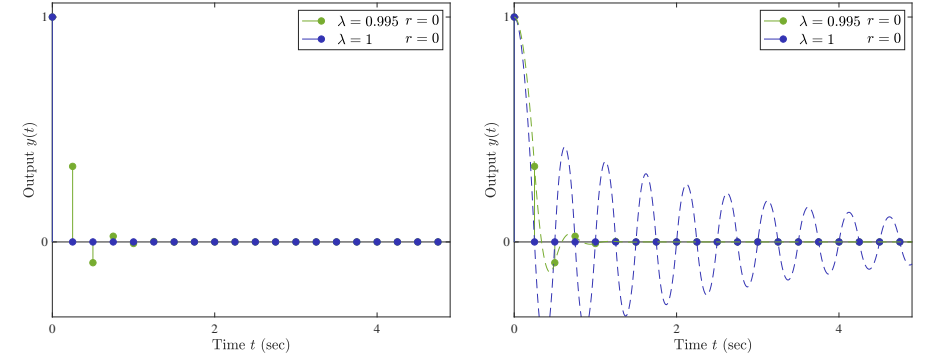


- under $\lambda = 1$ the response is faster, but overshoots (complex poles)
- as λ decreases, the response slows down, but smooths
- control effort decreases as λ decreases
- intersample behavior is as expected

67/84

Example: simulations of digital controller (contd)

With $r = 0$:



- if $\dot{y}$ is not penalized, we have the dead-beat discrete response
- the intersample behavior is poor
oscillations are invisible in $\mathcal{T}$, so that it pays off to place pole at the sampling zero
- even a tiny penalty of $\dot{y}$ changes that (the dead-beat $\dot{y}(ih)$ is huge)

68/84

Example: finite-horizon formulation

The same discretized DC motor, under the same sampling period,

$$\begin{cases} x[t+1] = \begin{bmatrix} 1 & 1 - e^{-h} \\ 0 & e^{-h} \end{bmatrix} x[t] + \begin{bmatrix} h - 1 + e^{-h} \\ 1 - e^{-h} \end{bmatrix} u[t] \\ y[t] = \begin{bmatrix} 1 & 0 \end{bmatrix} x[t] \end{cases}$$

and the same cost weights,

$$\begin{bmatrix} Q & S \\ S' & R \end{bmatrix} = \begin{bmatrix} \lambda & 0 & 0 \\ 0 & 1 - \lambda & 0 \\ 0 & 0 & r \end{bmatrix}.$$

But now

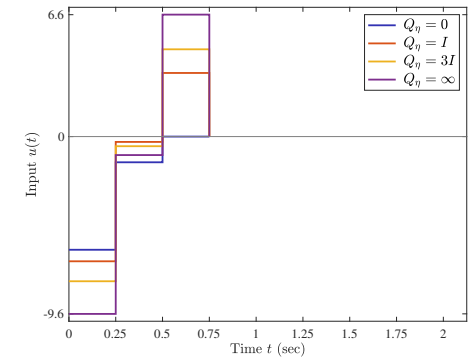
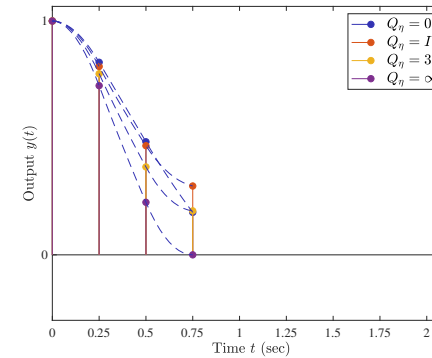
$$\mathcal{J} = x'[\eta] Q_{\eta} x[\eta] + \sum_{t=0}^{\eta-1} (\lambda y^2(ih) + (1 - \lambda) \dot{y}^2(ih) + r u^2[i])$$

(remember, $x_1[i] = y(ih)$ and $x_2[i] = \dot{y}(ih)$).

69/84

Example: finite-horizon simulations

With $\lambda = 1$, $r = 0.01$, and $\eta = 3$:

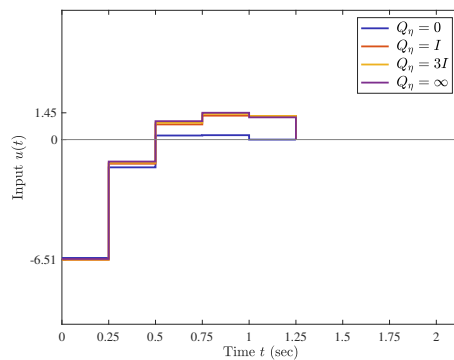
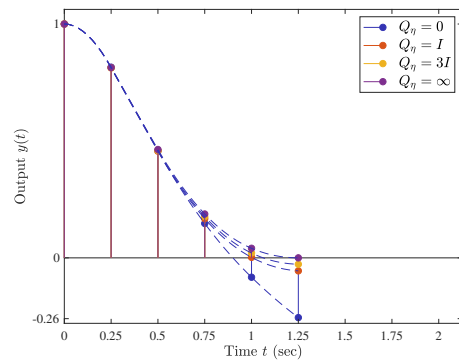


- larger Q_{η} reduces $x[\eta]$
- with $Q_{\eta} = \infty$ end up with deadbeat behavior

70/84

Example: finite-horizon simulations (contd)

With $\lambda = 1$, $r = 0.01$, and $\eta = 5$:

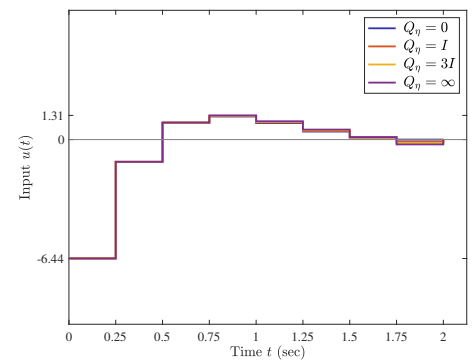
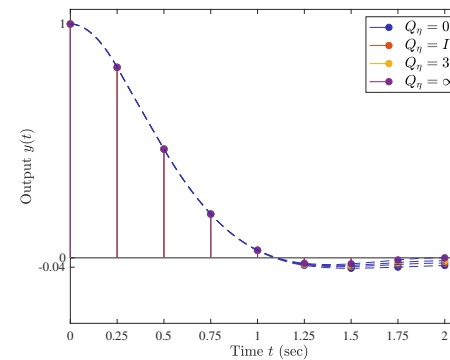


- larger Q_{η} reduces $x[\eta]$
- with $Q_{\eta} = \infty$ end up with deadbeat behavior

71/84

Example: finite-horizon simulations (contd)

With $\lambda = 1$, $r = 0.01$, and $\eta = 8$:



- larger Q_{η} reduces $x[\eta]$
- with $Q_{\eta} = \infty$ end up with deadbeat behavior
- as η increases, approach infinite-horizon behavior

72/84

Outline

Sampled-data control systems and discretization

Structural properties of discrete state-space realizations

Discrete-time state-feedback

Finite-horizon LQR

Discrete state-space realizations *ab initio* (for curious)

73/84

Stacked form

The convolution $y[t] = \sum_{s=-\infty}^{\infty} g[t-s]u[s]$ can be equivalently written as

$$\begin{bmatrix} \vdots \\ y[t-2] \\ y[t-1] \\ y[t] \\ y[t+1] \\ \vdots \end{bmatrix} = \begin{bmatrix} \vdots & \vdots & \vdots & \vdots & \vdots \\ \cdots & g[0] & g[-1] & g[-2] & g[-3] & \cdots \\ \cdots & g[1] & g[0] & g[-1] & g[-2] & \cdots \\ \cdots & g[2] & g[1] & g[0] & g[-1] & \cdots \\ \cdots & g[3] & g[2] & g[1] & g[0] & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} \vdots \\ u[t-2] \\ u[t-1] \\ u[t] \\ u[t+1] \\ \vdots \end{bmatrix}$$

(for causal systems $g[t] = 0$ whenever $t < 0$). Of a particular interest is the relation between past inputs and present and future outputs:

$$\begin{bmatrix} y[t] \\ y[t+1] \\ y[t+2] \\ \vdots \end{bmatrix} = \underbrace{\begin{bmatrix} g[1] & g[2] & g[3] & \cdots \\ g[2] & g[3] & g[4] & \cdots \\ g[3] & g[4] & g[5] & \cdots \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}}_{\text{past}} \begin{bmatrix} u[t-1] \\ u[t-2] \\ u[t-3] \\ \vdots \end{bmatrix} + \underbrace{f(u[t], u[t+1], \dots)}_{\text{present + future}}$$

74/84

Hankel operator

The semi-infinite matrix

$$\begin{bmatrix} g[1] & g[2] & g[3] & \cdots \\ g[2] & g[3] & g[4] & \cdots \\ g[3] & g[4] & g[5] & \cdots \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

is the matrix representation of the

- **Hankel operator** $H_G : u_{\text{past}} \mapsto y_{\text{present+future}}$

associated with G . The Hankel operator represents the **memory** of G , which may be thought of as a key factor in understanding its complexity.

Systems having

- **finite-rank Hankel** operators are said to be **finite dimensional**

and $n = \text{rank } H_G$ is known as the dimension of G .

75/84

Hankel operator: examples

Integrator with the step $g[t] = \mathbb{1}[t-1]$ has the Hankel operator

$$\begin{bmatrix} 1 & 1 & 1 & \cdots \\ 1 & 1 & 1 & \cdots \\ 1 & 1 & 1 & \cdots \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ \vdots \end{bmatrix} [1 \ 1 \ 1 \ \cdots] \Rightarrow \text{rank } H_G = 1$$

Double integrator with the ramp $g[t] = t\mathbb{1}[t-1]$ has the Hankel operator

$$\begin{bmatrix} 1 & 2 & 3 & \cdots \\ 2 & 3 & 4 & \cdots \\ 3 & 4 & 5 & \cdots \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ \vdots & \vdots \end{bmatrix} \begin{bmatrix} 0 & 1 & 2 & \cdots \\ 1 & 1 & 1 & \cdots \end{bmatrix} \Rightarrow \text{rank } H_G = 2$$

76/84

Hankel operator: examples (contd)

3-step delay has $g[t] = \delta[t - 3]$, so its Hankel operator

$$\begin{bmatrix} 0 & 0 & 1 & 0 & \cdots \\ 0 & 1 & 0 & 0 & \cdots \\ 1 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 & 0 & \cdots \\ 0 & 1 & 0 & 0 & \cdots \\ 1 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \Rightarrow \text{rank } H_G = 3$$

77/84

State vector

If the rank of (infinite-dimensional) H_G is finite, say n , factorize

$$\begin{bmatrix} g[1] & g[2] & g[3] & \cdots \\ g[2] & g[3] & g[4] & \cdots \\ g[3] & g[4] & g[5] & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} = \begin{bmatrix} O_1 \\ O_2 \\ O_3 \\ \vdots \end{bmatrix} \begin{bmatrix} R_1 & R_2 & R_3 & \cdots \end{bmatrix} =: \mathbf{O}_G \mathbf{R}_G$$

($O_i \in \mathbb{R}^{p \times n}$ and $R_i \in \mathbb{R}^{n \times m}$ for each $i \in \mathbb{N}$), where $\mathbf{O}_G$ and $\mathbf{R}_G$ are rank- n matrices having n columns and n rows, respectively. The n -dimensional

$$\mathbf{x}[t] := \underbrace{\begin{bmatrix} R_1 & R_2 & R_3 & \cdots \end{bmatrix}}_{\mathbf{R}_G} \begin{bmatrix} u[t-1] \\ u[t-2] \\ u[t-3] \\ \vdots \end{bmatrix} = \sum_{s=-\infty}^{t-1} R_{t-s} u[s]$$

is called the **state vector** of G at t .

78/84

State vector (contd)

The state vector is the **history accumulator**. Indeed,

$$\begin{bmatrix} y[t] \\ y[t+1] \\ y[t+2] \\ \vdots \end{bmatrix} = \begin{bmatrix} O_1 \\ O_2 \\ O_3 \\ \vdots \end{bmatrix} \mathbf{x}[t] + \begin{bmatrix} g[0] & 0 & 0 & \cdots \\ g[1] & g[0] & 0 & \cdots \\ g[2] & g[1] & g[0] & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} u[t] \\ u[t+1] \\ u[t+2] \\ \vdots \end{bmatrix},$$

meaning that no knowledge of $u[s]$ for $s < t$ is required to calculate $y[s]$ for $s \geq t$ if $\mathbf{x}[t]$ is known. This also suggests that

- causal systems may be viewed as mappings $(\mathbf{x}[0], u) \mapsto y$

for u and y defined on $\mathbb{Z}_+ := \{t \in \mathbb{Z} \mid t \geq 0\}$, rather than $u \mapsto y$ for u and y on $\mathbb{Z}$, with the

- initial condition $\mathbf{x}[0]$ explaining the effect of prehistoric inputs.

Remark. Having a starting point may be intuitive in control applications, this is the time moment where the control law is applied. Causal systems with zero initial conditions may be viewed as operators $u \mapsto y$ for u and y having support in $\mathbb{Z}_+ := \{t \in \mathbb{Z} \mid t \geq 0\}$.

79/84

Structure of $\mathbf{O}_G$ and $\mathbf{R}_G$

Denote $B := R_1 \in \mathbb{R}^{n \times m}$ and $C := O_1 \in \mathbb{R}^{p \times n}$. Because

$$\text{Im} \begin{bmatrix} O_2 \\ O_3 \\ O_4 \\ \vdots \end{bmatrix} \subset \text{Im} \begin{bmatrix} O_1 \\ O_2 \\ O_3 \\ \vdots \end{bmatrix}, \quad \begin{bmatrix} O_i \\ O_{i+1} \\ O_{i+2} \\ \vdots \end{bmatrix} R_{i+1} = \begin{bmatrix} O_{i+1} \\ O_{i+2} \\ O_{i+3} \\ \vdots \end{bmatrix} R_i, \quad \forall i \in \mathbb{N}$$

and $\mathbf{O}_G$ has full column rank, by construction, there is $A \in \mathbb{R}^{n \times n}$ such that

$$\mathbf{O}_G = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \end{bmatrix} \quad \text{and} \quad \mathbf{R}_G = \begin{bmatrix} B & AB & A^2B & \cdots \end{bmatrix}.$$

($\mathbf{O}_G$ by the first property, $\mathbf{R}_G$ by the others). Hence, $g[t] = CA^{t-1}B$ for all $t \in \mathbb{N}$ and

$$\mathbf{x}[t] = \sum_{s=-\infty}^{t-1} A^{t-1-s} B u[s] = A^t \mathbf{x}[0] + \sum_{s=0}^{t-1} A^{t-1-s} B u[s].$$

80/84

Structure of O_G and R_G : examples

Integrator:

$$\begin{bmatrix} 1 \\ 1 \\ 1 \\ \vdots \end{bmatrix} [1 \ 1 \ 1 \ \dots] \Rightarrow \begin{bmatrix} A & B \\ C \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ \vdots & \vdots \end{bmatrix}$$

Double integrator:

$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ \vdots & \vdots \end{bmatrix} \begin{bmatrix} 0 & 1 & 2 & \dots \\ 1 & 1 & 1 & \dots \end{bmatrix} \Rightarrow \begin{bmatrix} A & B \\ C \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & \vdots \end{bmatrix}$$

3-step delay:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 & 0 & \dots \\ 0 & 1 & 0 & 0 & \dots \\ 1 & 0 & 0 & 0 & \dots \end{bmatrix} \Rightarrow \begin{bmatrix} A & B \\ C \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & \vdots \end{bmatrix}$$

81/84

Consequences of structure of O_G and R_G

1. The state vector can be propagated **recursively** (convenient):

$$\begin{aligned} x[t+1] &= \sum_{s=-\infty}^t A^{t-s} B u[s] = \sum_{s=-\infty}^{t-1} \overbrace{A^{t-s}}^{\geq 1} B u[s] + B u[t] \\ &= A \sum_{s=-\infty}^{t-1} A^{t-1-s} B u[s] + B u[t] \\ &= A x[t] + B u[t]. \end{aligned}$$

2. The output at each t is a linear function of current state and input:

$$\begin{aligned} y[t] &= \sum_{s=-\infty}^t g[t-s] u[s] = C \sum_{s=-\infty}^{t-1} A^{t-1-s} B u[s] + g[0] u[t] \\ &= C x[t] + D u[t], \end{aligned}$$

where $D := g[0]$ is the **feedthrough** term.

82/84

Realizations

Thus, we end up with the following description (with $D := g[0]$):

$$G : \begin{cases} x[t+1] = A x[t] + B u[t], & x[0] = x_0 \\ y[t] = C x[t] + D u[t] \end{cases}$$

called the **state representation** of G and the quadruple (A, B, C, D) is called the **state-space realization** of G .

This is a compact form of

$$\begin{bmatrix} y[0] \\ y[1] \\ y[2] \\ \vdots \end{bmatrix} = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \end{bmatrix} x_0 + \begin{bmatrix} D & 0 & 0 & \dots \\ CB & D & 0 & \dots \\ CAB & CB & D & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} u[0] \\ u[1] \\ u[2] \\ \vdots \end{bmatrix}$$

or

$$y[t] = CA^t x_0 + C \sum_{s=0}^{t-1} A^{t-1-s} B u[s] + D u[t].$$

83/84

Similar realizations

The rank decomposition of the Hankel operator $O_G R_G$ is not unique. All its rank- n factors are $O_G T^{-1}$ and $T R_G$ for a nonsingular $T \in \mathbb{R}^{n \times n}$, viz.

$$O_G T^{-1} = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \end{bmatrix} T^{-1} = \begin{bmatrix} CT^{-1} \\ CT^{-1}(TAT^{-1}) \\ CT^{-1}(TAT^{-1})^2 \\ \vdots \end{bmatrix}$$

and

$$T R_G = T [B \ AB \ A^2 B \ \dots] = [TB \ (TAT^{-1})TB \ (TAT^{-1})^2 TB \ \dots]$$

In other words, given an n -dimensional system G ,

- all its n -dimensional realizations are $(TAT^{-1}, TB, CT^{-1}, D)$

for $\det T \neq 0$, where (A, B, C, D) is any particular n -dimensional realization. The realization $(TAT^{-1}, TB, CT^{-1}, D)$ is said to be **similar** to (A, B, C, D) and its state vector is $Tx[t]$ for all t . The impulse response is not affected by T , as $CA^t B = CT^{-1}(TAT^{-1})^t TB$ for all $t \in \mathbb{N}$.

84/84