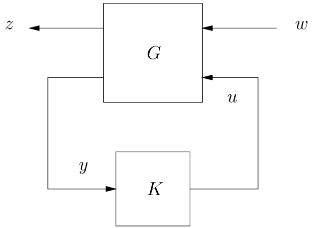
11: $\mathcal{H}_2$ Optimal Control

- defining the closed-loop map (D&P §5.1–5.2)
- the $\mathcal{H}_2$ system norm (D&P §6)
- the $\mathcal{H}_2$ optimal control problem (D&P §6.4)
- the linear quadratic regulator

General feedback arrangement



the plant G is described by the state-space system:

$$\dot{x}(t) = Ax(t) + \begin{bmatrix} B_1 & B_2 \end{bmatrix} \begin{bmatrix} w(t) \\ u(t) \end{bmatrix}$$

$$\begin{bmatrix} z(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} x(t) + \begin{bmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{bmatrix} \begin{bmatrix} w(t) \\ u(t) \end{bmatrix}$$

and controller K by

$$\dot{x}_K(t) = A_K x_K(t) + B_K y(t)$$
$$u(t) = C_K x_K(t) + D_K y(t)$$

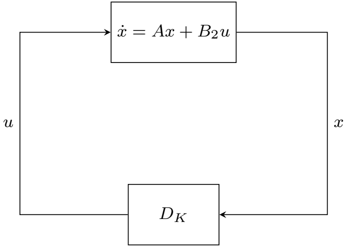
State feedback

- G is described by

$$\dot{x}(t) = Ax(t) + \begin{bmatrix} 0 & B_2 \end{bmatrix} \begin{bmatrix} w(t) \\ u(t) \end{bmatrix}$$
$$\begin{bmatrix} z(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} I \\ I \end{bmatrix} x(t) + \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} w(t) \\ u(t) \end{bmatrix}$$

- K is described by

$$\dot{x}_K(t) = 0x_K(t) + 0y(t)$$
$$u(t) = 0x_K(t) + D_K y(t)$$



Luenberger-based controller

- G has an output map

$$\dot{x}(t) = Ax(t) + \begin{bmatrix} 0 & B_2 \end{bmatrix} \begin{bmatrix} w(t) \\ u(t) \end{bmatrix}$$

$$\begin{bmatrix} z(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} I \\ C_2 \end{bmatrix} x(t) + \begin{bmatrix} 0 & 0 \\ 0 & D_{22} \end{bmatrix} \begin{bmatrix} w(t) \\ u(t) \end{bmatrix}$$

- K is dynamic

$$\left[\begin{array}{c|c} A_K & B_K \\ \hline C_K & D_K \end{array}\right] = \left[\begin{array}{c|c} A + LC + BF + LDF & -L \\ \hline F & 0 \end{array}\right]$$

Assumptions

it will be assumed that:

- (A, B) and (A_K, B_K) are **stabilizable**, *i.e.,*

$$A + BK \quad \text{and} \quad A_K + B_K F \quad \text{can be made stable}$$

- (A, C) and (A_K, C_K) are **detectable**, *i.e.,*

$$A + LC \quad \text{and} \quad A_K + HC_K \quad \text{can be made stable}$$

Definition

The feedback interconnection G, K is said to be **well-posed** if solutions exist for

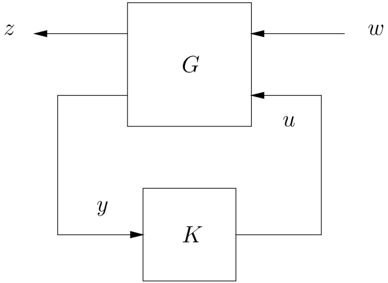
$$x(t), \quad x_K(t), \quad y(t), \quad u(t),$$

for all initial conditions $x(0)$, $x_K(0)$ and inputs $w(t)$.

Note: all "physical" systems are well-posed

Theorem

The feedback interconnection



is well posed if $(I - D_{22}D_K)^{-1}$ *exists.*

Note:

- if $D_{22} = 0$ or $D_K = 0$ then the closed-loop system is always well-posed
- $(I - D_{22}D_K)^{-1}$ exists iff $(I - D_K D_{22})^{-1} =: Q$ exists

Closed-loop stability

Definition

The closed-loop system G, K is **internally stable** if it is well posed, and for all initial conditions $x(0)$ and $x_K(0)$, the limits

$$x(t), x_K(t) \to 0 \text{ as } t \to \infty \text{ hold}$$

when $w = 0$.

internal stability requires:

1. $(I - D_{22}D_K)^{-1}$ to be invertible
2. the eigenvalues of

$$\begin{bmatrix} A & 0 \\ 0 & A_K \end{bmatrix} + \begin{bmatrix} B_2 & 0 \\ 0 & B_K \end{bmatrix} \begin{bmatrix} I & -D_K \\ -D_{22} & I \end{bmatrix}^{-1} \begin{bmatrix} 0 & C_2 \\ C_K & 0 \end{bmatrix}$$

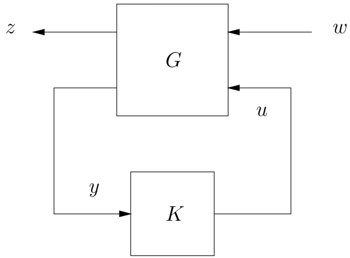
   to have negative real parts

Assumptions

for the remainder of the course, unless otherwise stated, we will assume that

- all state space descriptions are minimal realizations
- all feedback connections are well-posed
- the appropriate matrices are stabilizable and detectable

Closed-loop dynamics



the closed-loop system defines the map from w to z with states $x_{\mathrm{cl}} = \begin{bmatrix} x \\ x_K \end{bmatrix}$

the system has the form

$$\begin{aligned} \dot{x}_{\mathrm{cl}}(t) &= \mathcal{A}x_{\mathrm{cl}}(t) + \mathcal{B}w(t) \\ z(t) &= \mathcal{C}x_{\mathrm{cl}}(t) + \mathcal{D}w(t) \end{aligned}$$

compactly we write the map as

$$z = \mathcal{F}_l(G, K)w$$

Closed-loop dynamics: $\mathcal{A}, \mathcal{B}$

eliminating u and y:

$$\begin{bmatrix} I & -D_K \\ -D_{22} & I \end{bmatrix} \begin{bmatrix} u(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} 0 & C_K \\ C_2 & 0 \end{bmatrix} \begin{bmatrix} x(t) \\ x_K(t) \end{bmatrix} + \begin{bmatrix} 0 \\ D_{21} \end{bmatrix} w(t)$$

Closed-loop dynamics: $\mathcal{A}, \mathcal{B}$

eliminating u and y:

$$\begin{bmatrix} I & -D_K \\ -D_{22} & I \end{bmatrix} \begin{bmatrix} u(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} 0 & C_K \\ C_2 & 0 \end{bmatrix} \begin{bmatrix} x(t) \\ x_K(t) \end{bmatrix} + \begin{bmatrix} 0 \\ D_{21} \end{bmatrix} w(t)$$

substitute the above expression for u and y into

$$\begin{bmatrix} \dot{x} \\ \dot{x}_K \end{bmatrix} = \begin{bmatrix} A & 0 \\ 0 & A_K \end{bmatrix} \begin{bmatrix} x \\ x_K \end{bmatrix} + \begin{bmatrix} B_2 & 0 \\ 0 & B_K \end{bmatrix} \begin{bmatrix} u \\ y \end{bmatrix} + \begin{bmatrix} B_1 \\ 0 \end{bmatrix} w$$

to get

$$\begin{bmatrix} \dot{x} \\ \dot{x}_K \end{bmatrix} = \underbrace{\left(\begin{bmatrix} A & 0 \\ 0 & A_K \end{bmatrix} + \begin{bmatrix} B_2 & 0 \\ 0 & B_K \end{bmatrix} \begin{bmatrix} I & -D_K \\ -D_{22} & I \end{bmatrix}^{-1} \begin{bmatrix} 0 & C_K \\ C_2 & 0 \end{bmatrix} \right)}_{\mathcal{A}} \begin{bmatrix} x \\ x_K \end{bmatrix}$$

$$+ \underbrace{\begin{bmatrix} B_1 + B_2 D_K Q D_{21} \\ B_K Q D_{21} \end{bmatrix}}_{\mathcal{B}} w$$

Closed-loop dynamics: $\mathcal{C}, \mathcal{D}$

now we construct the output equation $z(t) = \mathcal{C}x_{\text{cl}}(t) + \mathcal{D}w(t)$:

$$z = \begin{bmatrix} C_1 & 0 \end{bmatrix} \begin{bmatrix} x \\ x_K \end{bmatrix} + \begin{bmatrix} D_{12} & 0 \end{bmatrix} \begin{bmatrix} u \\ y \end{bmatrix} + D_{11}w.$$

Closed-loop dynamics: $\mathcal{C}, \mathcal{D}$

now we construct the output equation $z(t) = \mathcal{C}x_{\mathrm{cl}}(t) + \mathcal{D}w(t)$:

$$z = \begin{bmatrix} C_1 & 0 \end{bmatrix} \begin{bmatrix} x \\ x_K \end{bmatrix} + \begin{bmatrix} D_{12} & 0 \end{bmatrix} \begin{bmatrix} u \\ y \end{bmatrix} + D_{11}w.$$

Again, substitute

$$\begin{bmatrix} I & -D_K \\ -D_{22} & I \end{bmatrix} \begin{bmatrix} u \\ y \end{bmatrix} = \begin{bmatrix} 0 & C_K \\ C_2 & 0 \end{bmatrix} \begin{bmatrix} x \\ x_K \end{bmatrix} + \begin{bmatrix} 0 \\ D_{21} \end{bmatrix} w,$$

Closed-loop dynamics: $\mathcal{C}, \mathcal{D}$

now we construct the output equation $z(t) = \mathcal{C}x_{\mathrm{cl}}(t) + \mathcal{D}w(t)$:

$$z = \begin{bmatrix} C_1 & 0 \end{bmatrix} \begin{bmatrix} x \\ x_K \end{bmatrix} + \begin{bmatrix} D_{12} & 0 \end{bmatrix} \begin{bmatrix} u \\ y \end{bmatrix} + D_{11}w.$$

Again, substitute

$$\begin{bmatrix} I & -D_K \\ -D_{22} & I \end{bmatrix} \begin{bmatrix} u \\ y \end{bmatrix} = \begin{bmatrix} 0 & C_K \\ C_2 & 0 \end{bmatrix} \begin{bmatrix} x \\ x_K \end{bmatrix} + \begin{bmatrix} 0 \\ D_{21} \end{bmatrix} w,$$

to obtain

$$z = \overbrace{\left(\begin{bmatrix} C_1 & 0 \end{bmatrix} + \begin{bmatrix} D_{12} & 0 \end{bmatrix} \begin{bmatrix} I & -D_K \\ -D_{22} & I \end{bmatrix}^{-1} \begin{bmatrix} 0 & C_K \\ C_2 & 0 \end{bmatrix} \right)}^{\mathcal{C}} \begin{bmatrix} x \\ x_K \end{bmatrix}$$
$$+ \underbrace{(D_{11} + D_{12}D_K Q D_{21})}_{\mathcal{D}} w.$$

Optimal control

Choose $\left[\begin{array}{c|c} A_K & B_K \\ \hline C_K & D_K \end{array}\right]$ in order to minimize $\|\mathcal{F}_l(G, K)\|$:

$$\left\| \left[\begin{array}{c|c} \begin{bmatrix} A & 0 \\ 0 & A_K \end{bmatrix} + \begin{bmatrix} B_2 & 0 \\ 0 & B_K \end{bmatrix} \begin{bmatrix} I & -D_K \\ -D_{22} & I \end{bmatrix}^{-1} \begin{bmatrix} 0 & C_K \\ C_2 & 0 \end{bmatrix} & \begin{matrix} B_1 + B_2 D_K Q D_{21} \\ B_K Q D_{21} \end{matrix} \\ \hline \begin{bmatrix} C_1 & 0 \end{bmatrix} + \begin{bmatrix} D_{12} & 0 \end{bmatrix} \begin{bmatrix} I & -D_K \\ -D_{22} & I \end{bmatrix}^{-1} \begin{bmatrix} 0 & C_K \\ C_2 & 0 \end{bmatrix} & D_{11} + D_{12} D_K Q D_{21} \end{array}\right] \right\|$$

where $Q = (I - D_K D_{22})^{-1}$

- clearly not convex in (A_K, B_K, C_K, D_K)
- which norm should we use?

System norms

- in the context of BIBO stability, we looked at $L_\infty \to L_\infty$ norm
  - clearly (!) this is called the L_1-induced norm

- minimizing the closed-loop in this norm is called L_1-optimal control

- L_1 optimal control in general is difficult to solve

instead, we will look at two other systems norms:

- the $\mathcal{H}_2$-norm: average energy
- the $\mathcal{H}_\infty$-norm: quantifies the peak energy of a stable system

The $\mathcal{H}_2$-norm: Impulse response interpretation

consider the system

$$\dot{x}(t) = Ax(t) + Bu(t)$$
$$y(t) = Cx(t)$$

which has a transfer matrix

$$\hat{G}(s) = C(sI - A)^{-1}B$$

and impulse response

$$H(t) = Ce^{At}B$$

Definition

Assume $G = (A, B, C)$ is controllable and observable, the $\mathcal{H}_2$-norm of G is

$$\|G\|_{\mathcal{H}_2} = \left(\int_0^\infty \|H(t)\|_F^2 \, \mathrm{d}t\right)^{\frac{1}{2}} = \left(\frac{1}{2\pi}\int_{-\infty}^{\infty} \|G(j\omega)\|_F^2 \, \mathrm{d}\omega\right)^{\frac{1}{2}}.$$

The $\mathcal{H}_2$-norm: White noise response

consider the system

$$\dot{x}(t) = Ax(t) + Bu(t)$$
$$y(t) = Cx(t)$$

$\mathcal{H}_2$ norm interpreted as the average energy of y when w is white noise:

$$\|G\|_{\mathcal{H}_2} = \mathbb{E} \int_0^\infty \|y(t)\|_2^2 = \mathbb{E}\ \|y\|_{L_2[0,\infty)]}$$

- $w(t)$ satisfies $\mathbb{E}w(t) = 0$, $\mathbb{E}[w(t)w(t)^T] = Q\delta(t-\tau)$
- in discrete time $w_k \sim N(0, Q)$

The $\mathcal{H}_2$-norm: White noise response

consider the system

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) \end{aligned}$$

$\mathcal{H}_2$ norm interpreted as the average energy of y when w is white noise:

$$\|G\|_{\mathcal{H}_2} = \mathbb{E} \int_0^\infty \|y(t)\|_2^2 = \mathbb{E}\ \|y\|_{L_2[0,\infty)]}$$

- $w(t)$ satisfies $\mathbb{E}w(t) = 0$, $\mathbb{E}[w(t)w(t)^T] = Q\delta(t-\tau)$
- in discrete time $w_k \sim N(0, Q)$

Note

in both interpretations, the initial condition does not affect the $\mathcal{H}_2$ norm

Computing the $\mathcal{H}_2$-norm

Theorem

Assume A is Hurwitz and (A, B, C) controllable and observable. Then

$$\|G\|_{\mathcal{H}_2} = \Big(\mathrm{trace}(B^T P B)\Big)^{\frac{1}{2}} \,,$$

where P solves the Lyapunov equation $A^T P + PA + C^T C = 0$.

Equivalently,

$$\|G\|_{\mathcal{H}_2} = \Big(\mathrm{trace}(C Q C^T)\Big)^{\frac{1}{2}} \,,$$

where where Q solves the Lyapunov equation $AQ + QA^T + BB^T = 0$.

Proof.

Apply the definition and make use of the linearity of the trace function. □

Equations and gramians

Controllability gramian: $AX_\mathrm{c} + X_\mathrm{c}A^T + BB^T = 0$

Observability gramian: $A^T Y_\mathrm{o} + Y_\mathrm{o}A + C^T C = 0$

Generalized gramians:

$$AX + XA^T + BB^T \preceq 0 \qquad A^T Y + YA + C^T C \preceq 0$$

Equations and gramians

Controllability gramian: $AX_c + X_c A^T + BB^T = 0$

Observability gramian: $A^T Y_o + Y_o A + C^T C = 0$

Generalized gramians:

$$AX + XA^T + BB^T \preceq 0 \qquad A^T Y + YA + C^T C \preceq 0$$

Lemma

Suppose A is Hurwitz and X_c satisfies

$$AX_c + X_c A^T + Q = 0$$

where Q is an arbitrary symmetric matrix. If X satisfies

$$AX + XA^T + Q \preceq 0,$$

then $X \succeq X_c$.

Useful matrix inequality facts

Monotonicity of the trace

given two symmetric matrices X and Y, then

$$X \preceq Y \quad \Longrightarrow \quad \mathbf{Tr}\, X \leq \mathbf{Tr}\, Y$$

n.b. the converse is not true

Useful matrix inequality facts

Monotonicity of the trace

given two symmetric matrices X and Y, then

$$X \preceq Y \quad \Longrightarrow \quad \mathbf{Tr}\, X \leq \mathbf{Tr}\, Y$$

n.b. the converse is not true

Schur complement

given matrices Q, M and R with M and Q being symmetric, then the following statements are equivalent:

1. the following two matrix inequalities hold:

$$Q \succ 0 \quad \text{and} \quad M - RQ^{-1}R^T \succ 0.$$

2. the linear matrix inequality

$$\begin{bmatrix} M & R \\ R^T & Q \end{bmatrix} \succ 0$$

is satisfied.

Computing the $\mathcal{H}_2$-norm with an LMI

Theorem

Let $G = \left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right]$*, then* A *is Hurwitz and* $\|G\|_{\mathcal{H}_2} < 1$ *if and only if there exists a symmetric matrix* $X \succ 0$ *such that*

$$\mathrm{trace}(CXC^T) < 1 \quad \textit{and} \quad AX + XA^T + BB^T \prec 0.$$

Proof.

("only if")

1. by hypothesis A is Hurwitz and $\mathrm{trace}(CXC^T) < 1$

2. we know $X \succeq X_c$, this implies $\mathrm{trace}(CX_cC^T) < 1$

3. define
$$X = \int_0^\infty e^{tA}(BB^T + \epsilon I)e^{tA^T}\ \mathrm{d}t \quad \text{for } \epsilon > 0,$$
R.H.S. is a continuous function of ϵ and $X = X_c$ when $\epsilon = 0$

4. thus, X satisfies
$$AX + XA^T + BB^T + \epsilon I = 0.$$

("if")

1. same idea, start with $A^TX + XA + BB^T \prec 0$ and work backwards.

□

The $\mathcal{H}_2$ optimal control problem

- we only cover the **full information** control problem

- given a system of the form

$$G = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & 0 & D_{12} \\ I & 0 & 0 \end{array}\right], \quad \text{where } A \in \mathbb{R}^{n\times n}, B_1 \in \mathbb{R}^{n\times p}, B_2 \in \mathbb{R}^{n\times m}, C_1 \in \mathbb{R}^{q\times n}$$

  find a controller K with structure

$$K = \left[\begin{array}{c|c} 0 & 0 \\ \hline 0 & F \end{array}\right]$$

  that minimizes $\|\mathcal{F}_l(G, K)\|_{\mathcal{H}_2}$

- no loss in achievable performance by using a static controller (skip proof of this)

LMI solution to the full-information $\mathcal{H}_2$ control problem

Theorem

There exists a feedback law $u = Kx$ that internally stabilizes G and satisfies

$$\|\mathcal{F}_l(G, K)\| < 1$$

if and only if there exist symmetric matrices $X \in \mathbb{R}^{n \times n}$, $W \in \mathbb{R}^{q \times q}$, and a rectangular matrix $Z \in \mathbb{R}^{m \times n}$ such that

$$F = ZX^{-1}$$

and the following inequalities are satisfied:

$$\begin{aligned}
\begin{bmatrix} A & B_2 \end{bmatrix} \begin{bmatrix} X \\ Z \end{bmatrix} + \begin{bmatrix} X & Z^T \end{bmatrix} \begin{bmatrix} A^T \\ B_2^T \end{bmatrix} + B_1 B_1^T \prec 0, \\
\begin{bmatrix} W & C_1 X + D_{12} Z \\ (C_1 X + D_{12} Z)^T & X \end{bmatrix} \succ 0, \\
\operatorname{trace}(W) < 1.
\end{aligned}$$

Proof.

$\mathcal{H}_2$ lemma, monotonicity of trace, Schur complement. □

LQR as a special case of $\mathcal{H}_2$-control

The Linear Quadratic Regulator:

the LQR problem requires finding a control input $u(t)$ such that

$$\int_0^\infty \Big(x(t)^T Q x(t) + u(t) R u(t)\Big)\,\mathrm{d}t$$

is minimized, subject to $\dot{x}(t) = Ax(t) + Bu(t)$, $Q \succeq 0$ and $R \succ 0$ are given

- $x(t)^T Q x(t)$ penalizes the deviation of x from 0 at time t
- $u(t)^T R u(t)$ penalizes the magnitude of the input u at time t
- objectives are in competition with each other

LQR as a special case of $\mathcal{H}_2$-control

The Linear Quadratic Regulator:

the LQR problem requires finding a control input $u(t)$ such that

$$\int_0^\infty \Big(x(t)^T Qx(t) + u(t)Ru(t)\Big)\,\mathrm{d}t$$

is minimized, subject to $\dot{x}(t) = Ax(t) + Bu(t)$, $Q \succeq 0$ and $R \succ 0$ are given

- $x(t)^T Qx(t)$ penalizes the deviation of x from 0 at time t
- $u(t)^T Ru(t)$ penalizes the magnitude of the input u at time t
- objectives are in competition with each other

using the notation for the general problem, the state equation is

$$\dot{x}(t) = A(t) + B_1 w(t) + B_2 u(t)$$

and we assume $w_i(t) = \delta(t) \cdot e_i$

Defining z

construct the **fictitious** output

$$z(t) = \begin{bmatrix} Q^{\frac{1}{2}} \\ 0 \end{bmatrix} x(t) + \begin{bmatrix} 0 \\ R^{\frac{1}{2}} \end{bmatrix} u(t)$$

where $Q = Q^{\frac{1}{2}} Q^{\frac{1}{2}}, R = R^{\frac{1}{2}} R^{\frac{1}{2}}$

$\mathcal{H}_2$-control is then simply finding K to minimize $\|z\|_{L_2[0,\infty)}$

using the control law $u(t) = Kx(t)$, the closed-loop system is

$$\begin{aligned} x(t) &= (A + BK)x(t) + w(t) \\ z(t) &= \begin{bmatrix} Q^{\frac{1}{2}} \\ R^{\frac{1}{2}} K \end{bmatrix} x(t) \\ y(t) &= x(t) \end{aligned}$$

Finite-horizon, discrete-time LQR

consider the discrete time system

$$x_{t+1} = Ax_t + Bu_t, \quad t = 0, \dots, N$$

with initial condition $x_0 = x^{\text{init}}$ over **time horizon** N

control objective: pick inputs $u_0, u_1, \dots, u_{N-1}$ in order to make

- $x_0, x_1, \dots$ small (i.e., good regulation **regulation** or **control**)
- $u_0, u_1, \dots$ small (i.e., input efficiency)

cost function

define the **quadratic** cost function

$$J(u) := \underset{u_0, u_1, \ldots u_{N-1}}{\text{minimize}} \quad \sum_{i=1}^{N-1} (x_t^T Q x_t + u_t^T R u_t) + x_N^T Q_f x_N$$

where $Q \succeq 0$, $Q_f \succeq 0$, and $R \succ 0$ are given

positive (semi)definiteness ensure the minimum possible cost is non-negative

- Q determines the state cost and Q_f the terminal state cost
- R determines the input cost, $R \succ 0$ means any (non-zero) input adds to J

LQR solution via dynamic programming

- **dynamic programming** provides a **recursive** method for solving the LQR problem
- broadly applicable to many problems sequential decision making
- **idea:** break problem into a sequence of problems, solve backwards in time

Value function

for $t = 0, \ldots, N$ define the value function is defined as

$$V_t(z) = \underset{u_t, \ldots, u_{N-1}}{\text{minimize}} \quad \sum_{\tau=t}^{N-1} (x_\tau^T Q x_\tau + {u_\tau}^T R u_\tau) + x_N^T Q_f x_N$$

$$\text{s.t.} \quad x_t = z, \quad x_{\tau+1} = A x_\tau + B u_\tau, \quad \tau = t, \ldots, N-1$$

Value function

for $t = 0, \ldots, N$ define the value function is defined as

$$V_t(z) = \underset{u_t,\ldots,u_{N-1}}{\text{minimize}} \quad \sum_{\tau=t}^{N-1} (x_\tau^T Q x_\tau + {u_\tau}^T R u_\tau) + x_N^T Q_f x_N$$

$$\text{s.t.} \quad x_t = z, \quad x_{\tau+1} = A x_\tau + B u_\tau, \quad \tau = t, \ldots, N-1$$

- V_t aka *value function*, *Bellman equation*, *Hamilton-Jacobi Equation*...
- $V_t(z)$ gives the minimum LQR cost-to-go, starting from state z at time t
- $V_0(x_0)$ recovers the LQR problem exactly
- the larger t is, the fewer decisions there are to be made
- at the $t = N$, there is no decision to make

$$V_N(z) = z^T Q_f z$$

- **dynamic programming**: solve for V_k in terms of V_{k+1} until we hit V_0

Dynamic programming

begin by taking the value function and pull the first stage cost out:

$$V_t(z) = \min_{u_t} \min_{u_{t+1},\dots u_{N-1}} \left(x_t^T Q x_t + u_t^T R u_t + \sum_{k=t+1}^{N-1} (x_k^T Q x_k + u_k^T R u_k) + x_N^T Q_f x_N \right)$$

$$= \min_{u_t} x_t^T Q x_t + u_t^T R u_t + \min_{u_{t+1},\dots u_{N-1}} \left(\sum_{k=t+1}^{N-1} (x_k^T Q x_k + u_k^T R u_k) + x_N^T Q_f x_N \right)$$

Dynamic programming

begin by taking the value function and pull the first stage cost out:

$$V_t(z) = \min_{u_t} \min_{u_{t+1},\ldots u_{N-1}} \left(x_t^T Q x_t + u_t^T R u_t + \sum_{k=t+1}^{N-1} (x_k^T Q x_k + u_k^T R u_k) + x_N^T Q_f x_N \right)$$

$$= \min_{u_t} x_t^T Q x_t + u_t^T R u_t + \min_{u_{t+1},\ldots u_{N-1}} \left(\sum_{k=t+1}^{N-1} (x_k^T Q x_k + u_k^T R u_k) + x_N^T Q_f x_N \right)$$

to simplify, relabel u_t as w to get

$$V_t(z) = \min_{w} z^T Q z + w^T R w + \underbrace{\min_{u_{t+1},\ldots u_{N-1}} \left(\sum_{k=t+1}^{N-1} (x_k^T Q x_k + u_k^T R u_k) + x_N^T Q_f x_N \right)}_{V_{t+1}(Az+Bw)}$$

where $x_{t+1} = Az + Bw$

the value function can now be written recursively as

$$V_t(z) = \min_{w} \left(z^T Q z + w^T R w + V_{t+1}(Az + Bw) \right)$$

the value function can now be written recursively as

$$V_t(z) = \min_w \left(z^T Q z + w^T R w + V_{t+1}(Az + Bw) \right)$$

it will now be shown that V_t is a **quadratic function** (for all t, not just $t = N$):

assume that $V_t(z) = z^T P_t z$ for some P_t, then

$$\begin{aligned} V_t(z) &= \min_w \left(z^T Q z + w^T R w + V_{t+1}(Az + Bw) \right) \\ &= \min_w \left(z^T Q z + w^T R w + (Az + Bw)^T P_{t+1} (Az + Bw) \right) \end{aligned}$$

the value function can now be written recursively as

$$V_t(z) = \min_w \left(z^T Q z + w^T R w + V_{t+1}(Az + Bw) \right)$$

it will now be shown that V_t is a **quadratic function** (for all t, not just $t = N$):

assume that $V_t(z) = z^T P_t z$ for some P_t, then

$$\begin{aligned} V_t(z) &= \min_w \left(z^T Q z + w^T R w + V_{t+1}(Az + Bw) \right) \\ &= \min_w \left(z^T Q z + w^T R w + (Az + Bw)^T P_{t+1} (Az + Bw) \right) \end{aligned}$$

solving for w by expanding, and setting the derivative to zero gives

$$2w^T R + 2(Az + Bw)^T P_{t+1} B = 0$$

from which

$$w^\star = -(R + B^T P_{t+1} B)^{-1} B^T P_{t+1} A z$$

finally, to show V_t is quadratic, substitute $w^\star$ into value function

$$
\begin{aligned}
V_t(z) &= z^T Q z + w^{\star T} R w^\star + (Az + Bw^\star)^T P_{t+1} (Az + Bw^\star) \\
&= z^T \underbrace{\left(A^T P_{t+1} A + Q - A^T P_{t+1} B (B^T P_{t+1} B + R)^{-1} B^T P_{t+1} A \right)}_{P_t} z
\end{aligned}
$$

moreover, $P_t \succeq 0$

LQR algorithm

1. $P_N := Q_f$

2. for $t = N$ down to $t = 1$:

$$P_{t-1} := Q + A^T P_t A - A^T P_t B(R + B^T P_t B)^{-1} B^T P_t A$$

3. for $t = 0$ up to $t = N - 1$:

$$K_t := -(R + B^T P_{t+1} B)^{-1} B^T P_{t+1} A$$

4. for $t = 0$ up to $t = N - 1$:

$$u_t^{\mathrm{lqr}} := K_t x_t$$

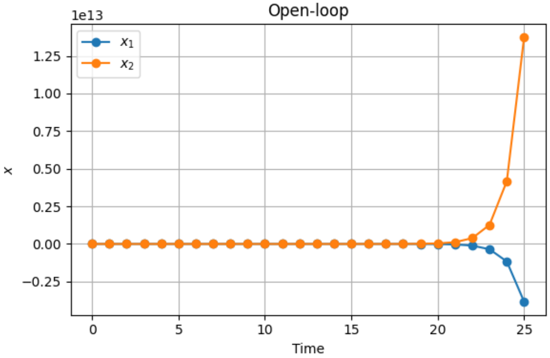
Notes

- optimal control input is **linear** in the state x_t
- P_t and K_t can be computed ahead of time (u_t computed in real time)
- cost to go is cost-to-go is $V_t(z) = z^T P_t z$.

LQR Example

$$x_{t+1} = \begin{bmatrix} 1.5 & -0.5 \\ 0.75 & 3.5 \end{bmatrix} x_t + \begin{bmatrix} -0.5 \\ 0.25 \end{bmatrix} u_t$$

the autonomous part of the system is not internally stable



LQR Example

$$x_{t+1} = \begin{bmatrix} 1.5 & -0.5 \\ 0.75 & 3.5 \end{bmatrix} x_t + \begin{bmatrix} -0.5 \\ 0.25 \end{bmatrix} u_t$$

the autonomous part of the system is not internally stable

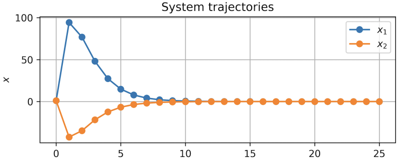
- cost function

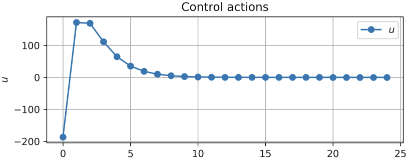
$$\sum_{i=1}^{24} (x_t^T Q x_t + u_t^T R u_t) + x_{25}^T Q x_{25}$$

  with $Q = \rho I, \; R = \mu$

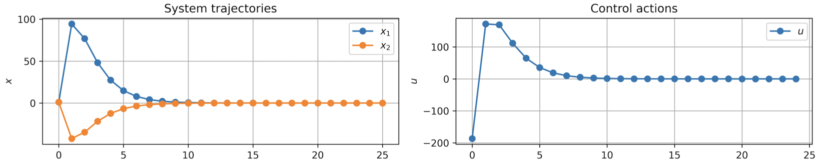
- $x_0 = [1, \; 1]^T$
- will examine the effect of different cost function choices

❶ equal weighting: $\rho = \mu = 1$

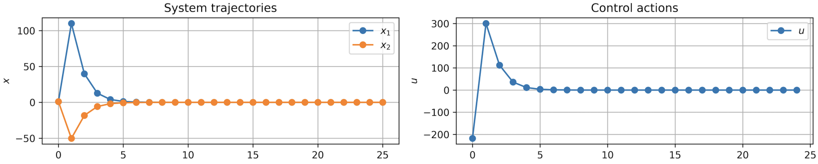




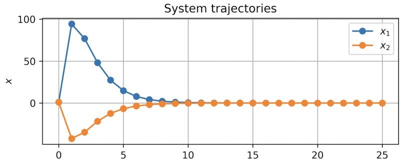
1. equal weighting: $\rho = \mu = 1$

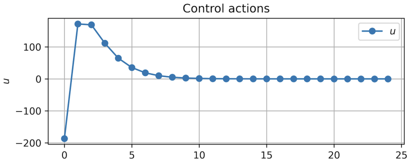


2. up-weighting state penalty: $\rho = 100, \mu = 1$

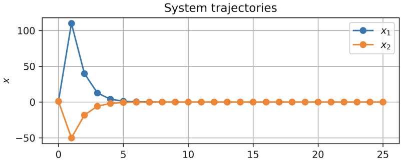


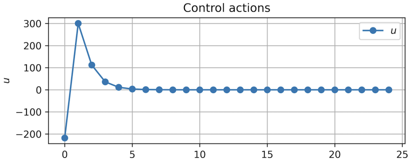
❶ equal weighting: $\rho = \mu = 1$



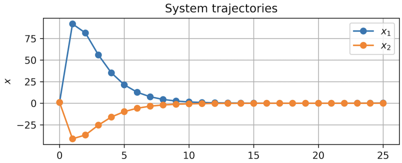


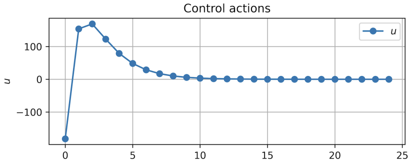
❷ up-weighting state penalty: $\rho = 100, \mu = 1$





❸ up-weighting control penalty: $\rho = 1, \mu = 100$





Infinite-horizon, discrete time LQR

- linear system dynamics: $x_{t+1} = Ax_t + Bu_t$ with $x_0 = x^{\mathrm{init}}$ and quadratic cost:

$$J(u) = \sum_{\tau=0}^{\infty} (x_\tau^T Q x_\tau + u_\tau^T R u_\tau)$$

- assume (A, B) controllable
- controllability $\implies$ existence of n such that $u_\tau = 0$ for all $\tau > n$ $\implies$ $J(u) < \infty$

Infinite-horizon, discrete time LQR

- linear system dynamics: $x_{t+1} = Ax_t + Bu_t$ with $x_0 = x^{\text{init}}$ and quadratic cost:

$$J(u) = \sum_{\tau=0}^{\infty} (x_\tau^T Q x_\tau + u_\tau^T R u_\tau)$$

- assume (A, B) controllable
- controllability $\implies$ existence of n such that $u_\tau = 0$ for all $\tau > n$ $\implies$ $J(u) < \infty$

Value function

$$V(z) = \min_{u_0, u_1, \ldots} \sum_{\tau=0}^{\infty} (x_\tau^T Q x_\tau + u_\tau^T R u_\tau)$$
$$\text{s.t.} \quad x_t = z, \quad x_{\tau+1} = Ax_\tau + Bu_\tau$$

- V is not a function of t (c.f. finite-horizon case)

Solution derivation

- the value function is quadratic, i.e., $V(z) = z^T P z$ with $P \succeq 0$
- V again expressed recursively:

$$V(z) = \min_w \left(z^T Q z + w^T R w + V(Az + Bw)\right)$$

- using the quadratic "assumption"

$$z^T P z = \min_w \overbrace{\left(z^T Q z + w^T R w + (Az + Bw)^T P (Az + Bw)\right)}^{(\#)}$$

- $w^{\rm opt} = -(R + B^T P B)^{-1} B^T P A z$
- to obtain P, substitute $w^{\rm opt}$ into $(\#)$ to get

$$z^T (Q + A^T P A - A^T P A (R + B^T P B)^{-1} B^T P A) z$$

Solution

the optimal controller is $u_t = Kx_t$, where

$$K = -(R + B^T PB)^{-1} B^T PA$$

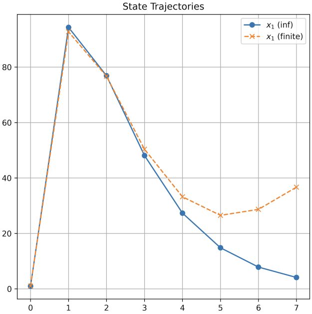
- contrast this to finite-horizon case
- P obtained from solving

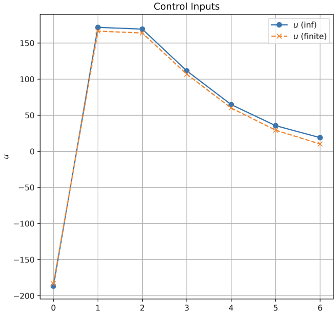
$$P = Q + A^T PA - A^T PA(R + B^T PB)^{-1} B^T PA \qquad (\#)$$

- $(\#)$ is the Discrete Algebraic Riccati Equation (DARE)
- `>>DARE` in MATLAB
- `scipy.linalg.solve_discrete_are` in Python
- does not require a recursive solution for P, solve DARE once

Infinite vs finite horizon

- $T = 7$ shown below
- $T > 12$ almost impossible to distinguish
- in general, P_t converges quickly as t gets further below N





Continuous-time LQR

quadratic cost function and linear dynamics, (A, B) controllable

$$\min_u \int_0^\infty \left(x(t)^T Q x(t) + u(t) R u(t) \right) \mathrm{d}t$$

$$\text{s.t.} \qquad \dot{x}(t) = Ax(t) + Bu(t)$$

many ways to obtain the solution (we omit details)

- optimal controller is static, $u(t) = Kx(t)$ where

$$K = -R^{-1} B^T P$$

  and P is obtained by solving the Algebraic Riccati Equation (ARE)

$$Q + A^T P + PA - PBR^{-1} B^T P = 0$$

- `>>CARE` in MATLAB
- $A + BK$ is stable if (A, B) controllable and (A, Q) detectable