Introduction to Optimal Control via LMIs

Matthew M. Peet
Arizona State University

Lecture 01: Optimal Control via LMIs

Summary

Powerful New Tools

- Convex Optimization
  - LMIs
  - Sum-of-Squares

Many old problems have been solved

- H_∞ and H_2 optimal control
- Nonlinear stability analysis
- Analysis and Control of delayed and PDE systems

Many questions are still unresolved

- Control of nonlinear Systems
- Nonlinear Programming (partially resolved)

Question: What is meant by a "solution"?

Outline

Lectures 1-2

1. Linear Systems
2. Convex Optimization and Linear Matrix Inequalities
3. Optimal Control
4. LMI Solutions to the H_∞ and H_2 Optimal Control Problems

Definition 1.

$L_2[0,\infty)$ is the Hilbert space of functions $f : \mathbb{R}^+ \to \mathbb{R}^n$ with inner product

$$\langle u, y \rangle_{L_2} = \frac{1}{2\pi} \int_0^\infty u(t)^T v(t) dt$$

$L_2[0,\infty)$ inherits the norm

$$\|u\|_{L_2}^2 = \int_0^\infty \|u(t)\|^2 dt$$

Definition 2.

The normed space of bounded linear operators from X to Y is denoted $\mathcal{L}(\mathbf{X},\mathbf{Y})$ with norm

$$\|P\|_{\mathcal{L}(X,Y)} := \sup_{\substack{x\in X\\ x\neq 0}} \frac{\|Px\|_Y}{\|x\|_X} = K$$

- Satisfies the properties of a norm
- This type of norm is called an “induced” norm
- Notation: $\mathcal{L}(X) := \mathcal{L}(X,X)$
- If X is a Banach space, then $\mathcal{L}(X,Y)$ is a Banach space

Properties: Suppose $G_1 \in \mathcal{L}(X,Y)$ and $G_2 \in \mathcal{L}(Y,Z)$

- Then $G_2 \odot G_1 \in \mathcal{L}(X,Z)$.
- $\|G_2 \odot G_1\|_{\mathcal{L}(X,Z)} \leq \|G_2\|_{\mathcal{L}(Y,Z)}\|G_1\|_{\mathcal{L}(X,Y)}$.
- Composition forms an *algebra*.

Definition 3.

Given $u \in L_2[0,\infty)$, the Laplace Transform of u is $\hat{u} = \Lambda u$, where

$$\hat{u}(s) = (\Lambda u)(s) = \lim_{T \to \infty} \int_0^T u(t) e^{-st} dt$$

if this limit exists.

Λ is a *bounded linear operator* - $\Lambda \in \mathcal{L}(L_2, H_2)$.

- $\Lambda : L_2 \to H_2$.
- The norm $\|\Lambda\|_{\mathcal{L}(L_2,H_2)}$ is

$$\|\Lambda\| = \sup_{u \in L_2} \frac{\|\Lambda u\|_{H_2}}{\|u\|_{L_2}} = ???$$

H_2 - A Space of Integrable Analytic Functions

Definition 4.

A complex function is **analytic** if it is continuous and bounded.

A function is analytic if the Taylor series converges everywhere in the domain.

Definition 5.

A function $\hat{u} : \bar{\mathbb{C}}^+ \to \mathbb{C}^n$ is in H_2 if

1. $\hat{u}(s)$ is analytic on the Open RHP (denoted $\mathbb{C}^+$)
2. For almost every real ω,

$$\lim_{\sigma \to 0^+} \hat{u}(\sigma + \imath\omega) = \hat{u}(\imath\omega)$$

   - Which means continuous on the imaginary axis

3. $$\int_{-\infty}^{\infty} \sup_{\sigma \geq 0} \|\hat{u}(\sigma + \imath\omega)\|_2^2 < \infty$$

   - Which means integrable on every vertical line.

The Maximum Modulus Principle

Theorem 6 (Maximum Modulus).

An analytic function cannot obtain its extrema in the interior of the domain.

Hence if $\hat{u}$ satisfies 1) and 2), then

$$\int_{-\infty}^{\infty} \sup_{\sigma \geq 0} \|\hat{u}(\sigma + \imath\omega)\|_2^2 = \int_{-\infty}^{\infty} \|\hat{u}(\imath\omega)\|_2^2 d\omega$$

We equip H_2 with a norm and inner product

$$\|\hat{u}\|_{H_2} = \int_{-\infty}^{\infty} \|\hat{u}(\imath\omega)\|_2^2 d\omega, \qquad \langle \hat{u}, \hat{y} \rangle_{H_2} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{u}(\imath\omega)^* \hat{v}(\imath\omega) d\omega$$

Paley-Wiener

Theorem 7.

1. *If $u \in L_2[0,\infty)$, then $\Lambda u \in H_2$.*
2. *If $\hat{u} \in H_2$, then there exists a $u \in L_2[0,\infty)$ such that $\hat{u} = \Lambda u$ (**Onto**).*

- Shows that H_2 is exactly the image of Λ on $L_2[0,\infty)$
- Shows the map is invertible

Definition 8.

The inverse of the Laplace transform, $\Lambda^{-1} : H_2 \to L_2[0,\infty)$ is

$$u(t) = (\Lambda^{-1}\hat{u})(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{\sigma t} \cdot e^{\imath\omega t}\hat{u}(\sigma + \imath\omega)d\omega$$

where σ can be any real number.

Corollary

Lemma 9.

$$\langle \Lambda u, \Lambda y \rangle_{H_2} = \langle u, y \rangle_{L_2}$$

- *Thus* Λ *is unitary.*
- $L_2[0,\infty)$ *and* H_2 *are isomorphic.*

$$\|\Lambda\| = \sup_{u \in L_2} \frac{\|\Lambda u\|_{H_2}}{\|u\|_{L_2}} = ???$$

H_∞ - A Space of Bounded Analytic Functions

Definition 10.

A function $\hat{G} : \bar{\mathbb{C}}^+ \to \mathbb{C}^{n\times m}$ is in H_∞ if

1. $\hat{G}(s)$ is analytic on the CRHP, $\mathbb{C}^+$.
2. $$\lim_{\sigma \to 0^+} \hat{G}(\sigma + \imath\omega) = \hat{G}(\imath\omega)$$
3. $$\sup_{s\in\mathbb{C}^+} \bar{\sigma}(\hat{G}(s)) < \infty$$

- A **Banach Space** with norm

$$\|\hat{G}\|_{H_\infty} = \operatorname{ess}\sup_{\omega\in\mathbb{R}} \bar{\sigma}(\hat{G}(\imath\omega))$$

H_∞ (A Signal Space) and Multiplier Operators

Every element of H_∞ defines a multiplication operator.

Definition 11.

Given $\hat{G} \in H_\infty$, define $M_{\hat{G}} \in \mathcal{L}(H_\infty)$

$$(M_{\hat{G}}\hat{u})(s) = \hat{G}(s)\hat{u}(s)$$

for $\hat{u} \in H_2$.

Functions vs. Operators

- $\hat{G}$ is a *function* of a complex variable.
- $M_{\hat{G}}$ is an *operator* (a function of functions...).

Causal LTI Systems map to H_∞

For any analytic functions, $\hat{u}$ and $\hat{G}$, the function

$$\hat{y}(s) = \hat{G}(s)\hat{u}(s)$$

is analytic.

- Thus $M_{\hat{G}} : H_2 \to H_2$.
- Thus $\Lambda^{-1} M_{\hat{G}} \Lambda$ maps $L_2[0,\infty) \to L_2[0,\infty)$.

Theorem 12.

G is a Causal, Linear, Time-Invariant Operator on L_2 if and only if there exists some $\hat{G} \in H_\infty$ such that $G = \Lambda^{-1} M_{\hat{G}} \Lambda$.

$$(\Lambda G u)(\imath\omega) = \hat{G}(\imath\omega)\hat{u}(\imath\omega)$$

H_∞ is the space of **transfer functions** for linear time-invariant systems.

H_∞ - The space of "Transfer Functions"

From Paley-Wiener, if $G = \Lambda^{-1} M_{\hat{G}} \Lambda$

Theorem 13.

$$\|G\|_{\mathcal{L}(L_2)} = \|M_{\hat{G}}\|_{\mathcal{L}(H_2)} = \|\hat{G}\|_{H_\infty}$$

The **Gain** of the system G can be calculated as $\|\hat{G}\|_{H_\infty}$

- This is the motivation for H_∞ control
- minimize $\sup_u \frac{\|Gu\|_{L_2}}{\|u\|_{L_2}}$.
  - minimize maximum energy of the output.

Conclusion: H_∞ provides a complete parametrization of the space of causal bounded linear time-invariant operators.

Rational Transfer Functions (RH_∞)

The space of bounded analytic functions, H_∞ is infinite-dimensional.

- this makes it hard to design optimal controllers.

We usually restrict ourselves to state-space systems and state-space controllers.

Definition 14.

The space of rational functions is defined as

$$R := \left\{ \frac{p(s)}{q(s)} \; : \; p, q \text{ are polynomials} \right\}$$

We define the following rational subspaces.

$$RH_2 = R \cap H_2$$
$$RH_\infty = R \cap H_\infty$$

Note that RH_2 and RH_∞ are not **complete**(Banach) spaces.

Rational Transfer Functions (RH_∞)

RH_∞ is the set of proper rational functions with no poles in the closed right half-plane (CRHP).

Definition 15.

- A rational function $r(s) = \frac{p(s)}{q(s)}$ is **Proper** if the degree of p is less than or equal to the degree of q.
- A rational function $r(s) = \frac{p(s)}{q(s)}$ is **Strictly Proper** if the degree of p is less than the degree of q.

Proposition 1.

1. *$\hat{G} \in RH_\infty$ if and only if $\hat{G}$ is proper with no poles on the closed right half-plane.*

State-Space Systems

Define a **State-Space System** $G : L_2 \to L_2$ by $y = Gu$ if

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

Theorem 16.

- *For any stable state-space system, G, there exists some $\hat{G} \in RH_\infty$ such that*
$$G = \Lambda^{-1} M_{\hat{G}} \Lambda$$
- *For any $\hat{G} \in RH_\infty$, the operator $G = \Lambda^{-1} M_{\hat{G}} \Lambda$ can be represented in state-space for some A, B, C and D where A is Hurwitz.*

For state-space system, (A, B, C, D),

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

State-Space is NOT Unique. For any invertible T,

- $\hat{G} = C(sI - A)^{-1}B + D = CT^{-1}(sI - TAT^{-1})^{-1}TB + D.$
  - (A, B, C, D) and $(TAT^{-1}, TB, CT^{-1}, D)$ both represent the system G.

Optimal Control Framework

2-input 2-output Framework

We introduce the control framework by separating internal signals from external signals.

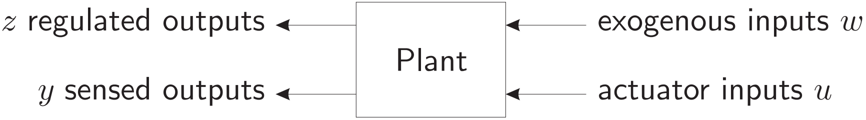
Output Signals:

- **z:** Output to be controlled/minimized
  - Regulated output
- **y:** Output used by the controller
  - Measured in real-time by sensor

The same signal may appear in both outputs.

- e.g. if you can measure what you want to minimize.

2-input 2-output Framework

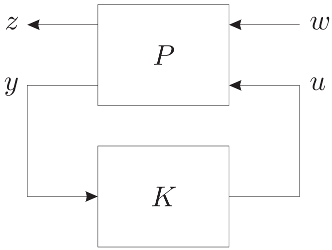


Input Signals:

- **w:** Disturbance, Tracking Signal, etc.
  - exogenous input
- **u:** Output from controller
  - Input to actuator
  - Not related to external input

The Optimal Control Framework

The controller closes the loop from y to u.



For a linear system P, we have 4 subsystems.

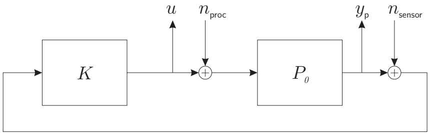
$$\begin{bmatrix} z \\ y \end{bmatrix} = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \begin{bmatrix} w \\ u \end{bmatrix}$$

$$P_{11} : w \mapsto z \qquad\qquad P_{12} : u \mapsto z$$

$$P_{21} : w \mapsto y \qquad\qquad P_{22} : u \mapsto y$$

Note that all P_{ij} can themselves be MIMO.

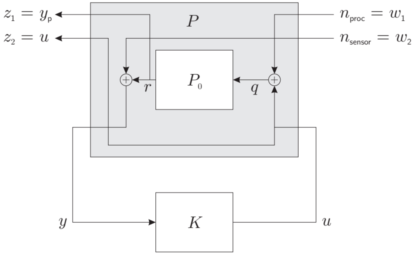
The Regulator



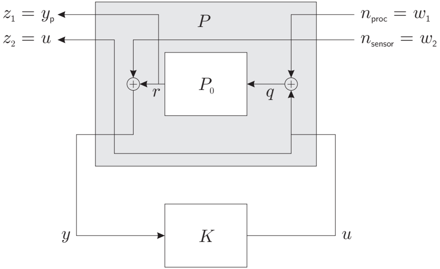
If we define $q = w_1 + u$ and $r = P_0 q$, then

$$z = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} y_p \\ u \end{bmatrix} \qquad\qquad w = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} \eta_{proc} \\ \eta_{sensor} \end{bmatrix}$$

$$y = \begin{bmatrix} y_1 \end{bmatrix} = r + w_2$$



The Regulator



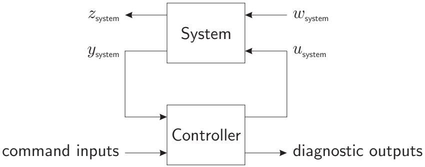
The reconfigured plant P is given by

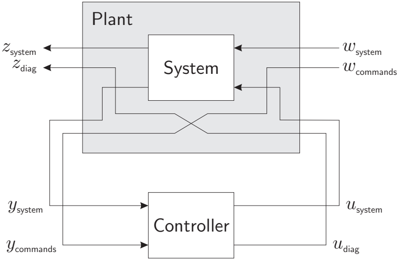
$$\begin{bmatrix} z_1(t) \\ z_2(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} P_0 & 0 & P_0 \\ 0 & 0 & I \\ P_0 & I & P_0 \end{bmatrix} \begin{bmatrix} w_1(t) \\ w_2(t) \\ u(t) \end{bmatrix}$$

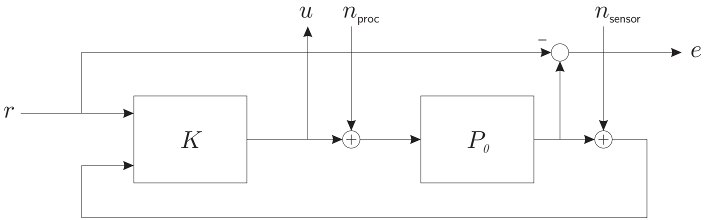
If $P_0 = (A, B, C, D)$, then

$$P = \left[\begin{array}{c|ccc} A & B & 0 & B \\ \hline C & D & 0 & D \\ 0 & 0 & 0 & I \\ C & D & I & D \end{array}\right]$$

Diagnostics







Define $q = n_{proc} + u$, then

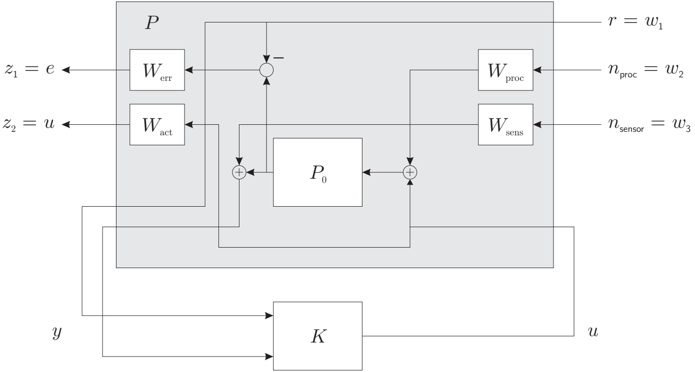
$$z = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} e \\ u \end{bmatrix} \qquad e = r - P_0 q$$

$$y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} r \\ n_{sensor} + P_0 q \end{bmatrix} \qquad w = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} r \\ n_{proc} \\ n_{sensor} \end{bmatrix}$$

$e =$ tracking error $\qquad r =$ tracking input

$n_{proc} =$ process noise $\qquad n_{sensor} =$ sensor noise

Tracking Control

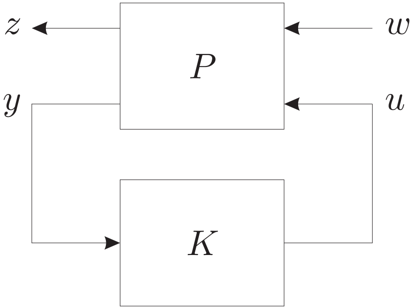


$$P = \begin{bmatrix} I & -P_0 & 0 & -P_0 \\ 0 & 0 & 0 & I \\ I & 0 & 0 & 0 \\ 0 & P_0 & I & P_0 \end{bmatrix}$$

$$z_1 = r - P_0(n_{proc} + u)$$
$$z_2 = u$$
$$y_1 = r$$
$$y_2 = w_3 + P_0(n_{proc} + u)$$

Linear Fractional Transformation

Close the loop



Plant:

$$\begin{bmatrix} z \\ y \end{bmatrix} = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \begin{bmatrix} w \\ u \end{bmatrix} \qquad \text{where} \qquad P = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & D_{22} \end{array}\right]$$

Controller:

$$u = Ky \qquad \text{where} \qquad K = \left[\begin{array}{c|c} A_K & B_K \\ \hline C_K & D_K \end{array}\right]$$

Linear Fractional Transformation

$$z = P_{11}w + P_{12}u$$
$$y = P_{21}w + P_{22}u$$
$$u = Ky$$

Solving for u,

$$u = KP_{21}w + KP_{22}u$$

Thus

$$(I - KP_{22})u = KP_{21}w$$
$$u = (I - KP_{22})^{-1}KP_{21}w$$

Now we solve for z:

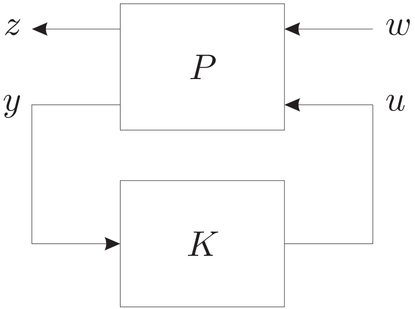
$$z = \left[P_{11} + P_{12}(I - KP_{22})^{-1}KP_{21}\right]w$$

Linear Fractional Transformation

This expression is called the Linear Fractional Transformation of (P, K), denoted

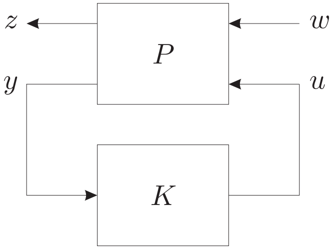
$$\underline{\mathsf{S}}(P, K) := P_{11} + P_{12}(I - KP_{22})^{-1}KP_{21}$$

AKA: Lower Star Product



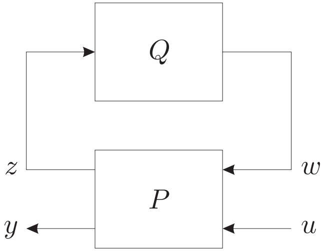
Other Fractional Transformations

Lower LFT:



$$\underline{S}(P,K) := P_{11}+P_{12}(I-KP_{22})^{-1}KP_{21}$$

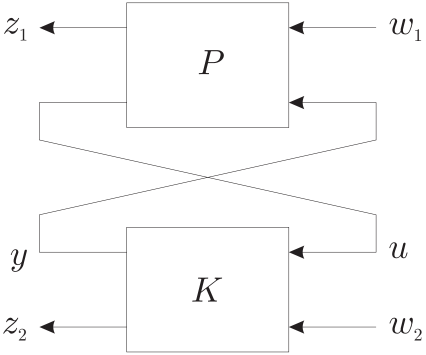
Upper LFT:



$$\bar{S}(P,K) := P_{22}+P_{21}Q(I-P_{11}K)^{-1}P_{12}$$

Other Fractional Transformations

Star Product:



$$S(P,K) := \begin{bmatrix} \underline{S}(P,K_{11}) & P_{12}(I-K_{11}P_{22})^{-1}K_{12} \\ K_{21}(I-P_{22}K_{11})^{-1}P_{21} & \bar{S}(K,P_{22}) \end{bmatrix}$$

Well-Posedness

The interconnection doesn't always make sense. Suppose

$$P = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & D_{22} \end{array}\right] \qquad \text{and} \qquad K = \left[\begin{array}{c|c} A_K & B_K \\ \hline C_K & D_K \end{array}\right].$$

Definition 17.

The interconnection $\underline{S}(P, K)$ is **well-posed** if for any smooth w and any $x(0)$ and $x_K(0)$, there exist functions x, x_K, u, y, z such that

$$\begin{aligned} \dot{x}(t) &= Ax(t) + B_1 w(t) + B_2 u(t) & \dot{x}_K(t) &= A_K x(t) + B_K y(t) \\ z(t) &= C_1 x(t) + D_{11} w(t) + D_{12} u(t) & u(t) &= C_K x(t) + D_K y(t) \\ y(t) &= C_2 x(t) + D_{21} w(t) + D_{22} u(t) & & \end{aligned}$$

Note: The solution does not need to be in L_2.

- Says nothing about stability.

Well-Posedness

In state-space format, the closed-loop system is:

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

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

From

$$u(t) = D_K y(t) + C_K x_K(t)$$
$$y(t) = D_{22} u(t) + C_2 x(t) + D_{21} w(t)$$

We have

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

Because the rest is state-space, the interconnection is well-posed if and only if the matrix $\begin{bmatrix} I & -D_K \\ -D_{22} & I \end{bmatrix}$ is invertible.

Well-Posedness

Question: When is

$$\begin{bmatrix} I & -D_K \\ -D_{22} & I \end{bmatrix}$$

invertible?
Answer: 2x2 matrices have a closed-form inverse

$$\begin{bmatrix} I & -D_K \\ -D_{22} & I \end{bmatrix}^{-1} = \begin{bmatrix} I + D_K Q D_{22} & D_K Q \\ Q D_{22} & Q \end{bmatrix}$$

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

Proposition 2.

The interconnection $\underline{S}(P, K)$ is well-posed if and only if $(I - D_{22} D_K)$ is invertible.

- Equivalently $(I - D_K D_{22})$ is invertible.
- Sufficient conditions: $D_K = 0$ or $D_{22} = 0$.
- To optimize over K, we will need to enforce this constraint somehow.

Definition 18.

The **Optimal H_∞-Control Problem** is

$$\min_{K \in H_\infty} \|\underline{\mathsf{S}}(P,K)\|_{H_\infty}$$

- This is the Optimal H_∞ **Dynamic-Output-Feedback** Control Problem

Another class of optimal control problem:

Definition 19.

The **Optimal H_2-Control Problem** is

$$\min_{K \in H_\infty} \|\underline{\mathsf{S}}(P,K)\|_{H_2} \quad \text{such that}$$

$$\underline{\mathsf{S}}(P,K) \in H_\infty.$$

Optimal Control

Choose K to minimize

$$\|P_{11} + P_{12}(I - KP_{22})^{-1}KP_{21}\|_{H_\infty}$$

Equivalently choose $\left[\begin{array}{c|c} A_K & B_K \\ \hline C_K & D_K \end{array}\right]$ to minimize

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

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

In either case, the problem is Nonlinear.

Optimal Control

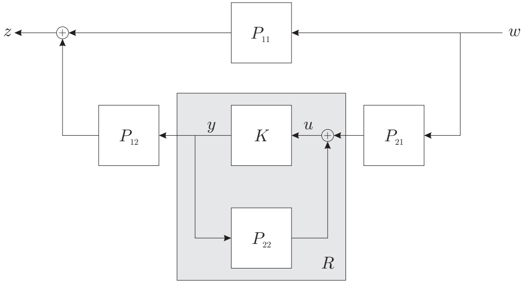
There are several ways to address the problem of nonlinearity.

$$\|P_{11} + P_{12}(I - KP_{22})^{-1}KP_{21}\|_{H_\infty}$$

Variable Substitution: The easiest way to make the problem linear is by declaring a new variable $R := (I - KP_{22})^{-1}K$

The optimization problem becomes: Choose R to minimize

$$\|P_{11} + P_{12}RP_{21}\|_{H_\infty}$$



Optimal Control

We optimize

$$\|P_{11} + P_{12}(I - KP_{22})^{-1}KP_{21}\|_{H_\infty} = \|P_{11} + P_{12}RP_{21}\|_{H_\infty}$$

Once, we have the optimal R, we can recover the optimal K as

$$K = R(I + RP_{22})^{-1}$$

Problems:

- how to optimize $\|\cdot\|_{H_\infty}$.
- Is the controller stable?
  - Does the inverse $(I + RP_{22})^{-1}$ exist? Yes.
  - Is it a bounded linear operator?
  - In which space?
- An important branch of control.
  - Coprime factorization
  - Youla parameterization
- We will sidestep this body of work.

What is Optimization?

Optimization can be posed in functional form:

$$\min_{x \in \mathbb{F}} \; \texttt{objective function}: \qquad \text{subject to}$$

$$\texttt{inequality constraints}$$

which may have the form

$$\min_{x \in \mathbb{F}} f_0(x): \qquad \text{subject to}$$

$$f_i(x) \geq 0 \qquad i = 1, \cdots k$$

Special Cases:

- **Linear Programming**
  - $f_i(x) = Ax - b$ (Affine functions with $f_i : \mathbb{R}^n \to \mathbb{R}^m$)
  - EASY: Simplex/Ellipsoid Algorithm
- **Polynomial Programming**
  - The $f_i : \mathbb{R}^n \to \mathbb{R}^m$ are polynomials. (NP-HARD)
- **Semidefinite Programming**
  - The $f_i : \mathbb{R}^n \to \mathbb{R}^{m \times m}$ are affine. (EASY)

For semidefinite programming, what does $f_i(x) \geq 0$ mean?

How Hard is Optimization?

Why is Linear Programming easy and polynomial programming hard?

$$\min_{x\in\mathbb{F}} f_0(x): \qquad \text{subject to}$$
$$f_i(x) \ge 0 \qquad i = 1, \cdots k$$

The **Geometric Representation** is equivalent:

$$\min_{x\in\mathbb{F}}\; f_0(x): \qquad \text{subject to} \qquad x \in S$$

where $S := \{x \,:\, f_i(x) \ge 0,\; i = 1, \cdots, k\}$.

The **Pure Geometric Representation**:

$$\min_{\gamma, x\in\mathbb{F}}\; \gamma: \qquad \text{subject to}$$
$$(\gamma, x) \in S'$$

where $S' := \{(\gamma,\, x) \,:\, \gamma - f_0(x) \ge 0,\; f_i(x) \ge 0,\; i = 1, \cdots, k\}$.

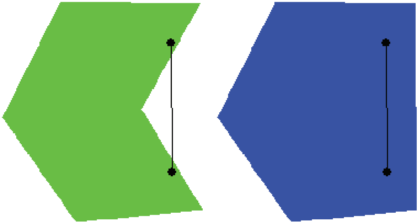
- Two optimization problems are Equivalent if a solution to one can be used to construct a solution to the other.

Definition 20.

A set is **convex** if for any $x, y \in Q$,

$$\{\mu x + (1-\mu)y \,:\, \mu \in [0,1]\} \subset Q.$$

The line connecting any two points lies in the set.



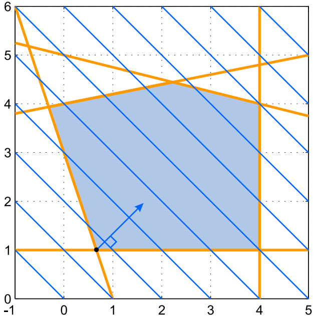
Convex Optimization:

Definition 21.

Consider the optimization problem

$$\min_{\gamma, x \in \mathbb{F}} \; \gamma : \qquad \text{subject to}$$
$$(\gamma, x) \in S'.$$

The problem is **Convex Optimization** if the set S' is convex.



Convex optimization problems have the property that the Gradient projection algorithm (or Newton iteration with barrier functions) will always converge to the global optimal.

The question is, of course, when is the set S' convex?

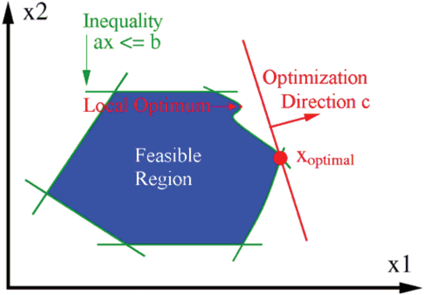
- For polynomial optimization, a sufficient condition is that all functions f_i are convex.
  - The level set of a convex function is a convex set.

Non-Convexity and Local Optima

Newton's Algorithm: Designed to solve $f(x^*) = 0$ (is $\min f(x) \geq 0$?)

$$x_{k+1} = x_k - t\frac{f(x_k)}{f'(x_k)}$$

where t is the step-size. (From $df/dx \cong \frac{f(x)-f(x^*)}{x-x^*}$)
For non-convex optimization, Newton descent may get stuck at local optima.



For constrained optimization, constraints are represented by barrier functions.

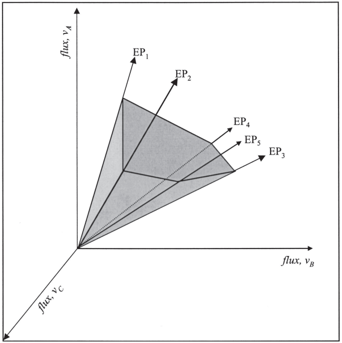
Definition 22.

A set is a **cone** if for any $x \in Q$,

$$\{\mu x \,:\, \mu \geq 0\} \subset Q.$$

A subspace is a cone but not all cones are subspaces.

- If the cone is also convex, it is a convex cone.
- Cones are convex if they are closed under addition.



What is an **Inequality Constraint**?

Question: What does $f(x) \geq 0$ mean.

- What does $y \geq 0$ mean?

If y is a **Scalar** ($y \in \mathbb{R}$), then $y \geq 0$ if $y \in [0, \infty]$.

Question: What if y is a vector ($y \in \mathbb{R}^n$)?

- Then we have several options...

Examples: Let $y \in \mathbb{R}^n$.

- **Positive Orthant:** $y \geq 0$ if $y_i \geq 0$ for $i = 1, \cdots, n$.
- **Half-space:** $y \geq 0$ if $\sum y_i \geq 0$ ($\mathbf{1}^T y \geq 0$).
  - More generally, $y \geq 0$ if $\mathbf{a}^T y + b \geq 0$.
- **Intersection of Half-spaces:** $y \geq 0$ if $a_i^T y + b_i \geq 0$ for $i = 1, \cdots, n$.
  - The positive orthant is the intersection of half-spaces with $b_i = 0$ and $a_i = e_i$ (unit vectors).

Question: What if y is a matrix???

Positivity

What is an inequality? What does ≥ 0 mean?

- An inequality implies a **partial ordering**:
  - $x \geq y$ if $x - y \geq 0$
- Any convex cone, C defines a partial ordering:
  - $x - y \geq 0$ if $x - y \in C$
- The ordering is only partial because $x \not\leq 0$ does not imply $x \geq 0$
  - $-x \notin C$ does not imply $x \in C$.
  - x may be indefinite.

Conclusion:

- Convex Optimization includes positivity induced from any partial ordering.
- In particular, we focus on **Matrix Positivity**.

Matrix Positivity

Definition 23.

A symmetric matrix $P \in \mathbb{S}^n$ is **Positive Semidefinite**, denoted $P \geq 0$ if

$$x^T P x \geq 0 \qquad \text{for all } x \in \mathbb{R}^n$$

Definition 24.

A symmetric matrix $P \in \mathbb{S}^n$ is **Positive Definite**, denoted $P > 0$ if

$$x^T P x > 0 \qquad \text{for all } x \neq 0$$

- P is **Negative Semidefinite** if $-P \geq 0$
- P is **Negative Definite** if $-P > 0$
- A matrix which is neither Positive nor Negative Semidefinite is **Indefinite**

The set of positive or negative matrices is a *convex cone*.

Positive Matrices

Lemma 25.

$P \in \mathbb{S}^n$ is positive definite if and only if all its eigenvalues are positive.

Things which are easy to prove:

- A Positive Definite matrix is invertible.
- The inverse of a positive definite matrix is positive definite.
- If $P > 0$, then $TPT^T \geq 0$ for any T. If T is invertible, then $TPT^T > 0$.

Lemma 26.

For any $P > 0$, there exists a positive square root, $P^{\frac{1}{2}} > 0$ such that $P = P^{\frac{1}{2}} P^{\frac{1}{2}}$.

Semidefinite Programming - Dual Form

$$\begin{array}{lll} \text{minimize} & \operatorname{trace} CX & \\ \text{subject to} & \operatorname{trace} A_i X = b_i & \text{for all } i \\ & X \succeq 0 & \end{array}$$

- The variable X is a symmetric matrix
- $X \succeq 0$ means X is positive semidefinite
- The feasible set is the intersection of an affine set with the *positive semidefinite cone*

$$\left\{ X \in \mathbb{S}^n \mid X \succeq 0 \right\}$$

Recall $\operatorname{trace} CX = \sum_{i,j} C_{i,j} X_{j,i}$.

SDPs with Explicit Variables - Primal Form

We can also explicitly parametrize the affine set to give

$$\begin{aligned} \text{minimize} \quad & c^T x \\ \text{subject to} \quad & F_0 + x_1 F_1 + x_2 F_2 + \cdots + x_n F_n \preceq 0 \end{aligned}$$

where $F_0, F_1, \ldots, F_n$ are symmetric matrices.

The inequality constraint is called a *Linear Matrix Inequality (LMI)*; e.g.,

$$\begin{bmatrix} x_1 - 3 & x_1 + x_2 & -1 \\ x_1 + x_2 & x_2 - 4 & 0 \\ -1 & 0 & x_1 \end{bmatrix} \preceq 0$$

which is equivalent to

$$\begin{bmatrix} -3 & 0 & -1 \\ 0 & -4 & 0 \\ -1 & 0 & 0 \end{bmatrix} + x_1 \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} + x_2 \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \preceq 0$$

Linear Matrix Inequalities

Linear Matrix Inequalities are often a *Simpler* way to solve control problems.
Common Form:

$$\text{Find } X:$$
$$\sum_i A_i X B_i + Q > 0$$

The most important Linear Matrix Inequality is the *Lyapunov Inequality*.

There are several very efficient **LMI/SDP Solvers** for Matlab:

- **SeDuMi**
  - Fast, but somewhat unreliable.
  - See `http://sedumi.ie.lehigh.edu/`
- **LMI Lab** (Part of Matlab's Robust Control Toolbox)
  - Universally disliked
  - See `http://www.mathworks.com/help/robust/lmis.html`
- **YALMIP** (a parser for other solvers)
  - See `http://users.isy.liu.se/johanl/yalmip/`

I recommend YALMIP with solver SeDuMi.

Some Simple examples of LMI conditions in control include:

- Stability

$$A^T X + XA \prec 0$$
$$X \succ 0$$

- Stabilization

$$AX + BZ + XA^T + Z^T B^T \prec 0$$
$$X \succ 0$$

- H_2 Synthesis

$$\min Tr(W)$$
$$\begin{bmatrix} A & B_2 \end{bmatrix} \begin{bmatrix} X \\ Z \end{bmatrix} + \begin{bmatrix} X & Z^T \end{bmatrix} \begin{bmatrix} A^T \\ B_2^T \end{bmatrix} + B_1 B_1^T \prec 0$$
$$\begin{bmatrix} X & (CX + DZ)^T \\ (CX + DZ) & W \end{bmatrix} \succ 0$$

We will go beyond these examples.

LMIs unite time-domain and frequency-domain analysis

$$\dot{x}(t) = f(x(t))$$

Theorem 27 (Lyapunov).

Suppose there exists a continuously differentiable function V for which $V(0) = 0$ and $V(x) > 0$ for $x \neq 0$. Furthermore, suppose $\lim_{\|x\| \to \infty} V(x) = \infty$ and

$$\lim_{h \to 0^+} \frac{V(x(t+h)) - V(x(t))}{h} = \frac{d}{dt} V(x(t)) < 0$$

for any x such that $\dot{x}(t) = f(x(t))$. Then for any $x(0) \in \mathbb{R}$ the system of equations

$$\dot{x}(t) = f(x(t))$$

has a unique solution which is stable in the sense of Lyapunov.

The Lyapunov Inequality (Our First LMI)

Lemma 28.

A is Hurwitz if and only if there exists a $P > 0$ such that

$$A^T P + PA < 0$$

Proof.

Suppose there exists a $P > 0$ such that $A^T P + PA < 0$.

- Define the Lyapunov function $V(x) = x^T P x$.
- Then $V(x) > 0$ for $x \neq 0$ and $V(0) = 0$.
- Furthermore,

$$\begin{aligned} \dot{V}(x(t)) &= \dot{x}(t)^T P x(t) + x(t)^T P \dot{x}(t) \\ &= x(t)^T A^T P x(t) + x(t)^T P A x(t) \\ &= x(t)^T \left(A^T P + PA\right) x(t) \end{aligned}$$

- Hence $\dot{V}(x(t)) < 0$ for all $x \neq 0$. Thus the system is globally stable.
- Global stability implies A is Hurwitz.

The Lyapunov Inequality

Proof.

For the other direction, if A is Hurwitz, let

$$P = \int_0^\infty e^{A^T s} e^{As} ds$$

- Converges because A is Hurwitz.
- Furthermore
$$\begin{aligned} PA &= \int_0^\infty e^{A^T s} e^{As} A ds \\ &= \int_0^\infty e^{A^T s} A e^{As} ds = \int_0^\infty e^{A^T s} \frac{d}{ds}\left(e^{As}\right) ds \\ &= \left[e^{A^T s} e^{As} \right]_0^\infty - \int_0^\infty \frac{d}{ds} e^{A^T s} e^{As} \\ &= -I - \int_0^\infty A^T e^{A^T s} e^{As} = -I - A^T P \end{aligned}$$
- Thus $PA + A^T P = -I < 0$.

Other Versions:

Lemma 29.

(A, B) is controllable if and only if there exists a $X > 0$ such that

$$A^T X + XA + BB^T \leq 0$$

Lemma 30.

(C, A) is observable if and only if there exists a $X > 0$ such that

$$AX + XA^T + C^T C \leq 0$$

The Static State-Feedback Problem

Lets start with the problem of stabilization.

Definition 31.

The **Static State-Feedback Problem** is to find a feedback matrix K such that

$$\dot{x}(t) = Ax(t) + Bu(t)$$
$$u(t) = Kx(t)$$

is stable

- Find K such that $A + BK$ is Hurwitz.

Can also be put in LMI format:

$$\textbf{Find } X > 0,\ K\ :$$
$$X(A + BK) + (A + BK)^T X < 0$$

Problem: Bilinear in K and X.

The Static State-Feedback Problem

- The bilinear problem in K and X is a common paradigm.
- Bilinear optimization is not convex.
- To convexify the problem, we use a change of variables.

Problem 1:

$$\textbf{Find } X > 0, K \; :$$
$$X(A+BK)+(A+BK)^T X < 0$$

Problem 2:

$$\textbf{Find } P > 0, Z \; :$$
$$AP + BZ + PA^T + Z^T B^T < 0$$

Definition 32.

Two optimization problems are equivalent if a solution to one will provide a solution to the other.

Theorem 33.

Problem 1 is equivalent to Problem 2.

The Dual Lyapunov Equation

Problem 1:

Find $X > 0, :$

$XA + A^T X < 0$

Problem 2:

Find $Y > 0, :$

$YA^T + AY < 0$

Lemma 34.

Problem 1 is equivalent to problem 2.

Proof.

First we show 1) solves 2). Suppose $X > 0$ is a solution to Problem 1. Let $Y = X^{-1} > 0$.

- If $XA + A^T X < 0$, then

$$X^{-1}(XA + A^T X)X^{-1} < 0$$

- Hence

$$X^{-1}(XA + A^T X)X^{-1} = AX^{-1} + X^{-1}A^T = AY + YA^T < 0$$

- Therefore, Problem 2 is feasible with solution $Y = X^{-1}$.

The Dual Lyapunov Equation

Problem 1:

$$\textbf{Find } X > 0, :$$
$$XA + A^T X < 0$$

Problem 2:

$$\textbf{Find } Y > 0, :$$
$$YA^T + AY < 0$$

Proof.

Now we show 2) solves 1) in a similar manner. Suppose $Y > 0$ is a solution to Problem 1. Let $X = Y^{-1} > 0$.

- Then

$$\begin{aligned} XA + A^T X &= X(AX^{-1} + X^{-1}A^T)X \\ &= X(AY + YA^T)X < 0 \end{aligned}$$

□

Conclusion: If $V(x) = x^T P x$ proves stability of $\dot{x} = Ax$,

- Then $V(x) = x^T P^{-1} x$ proves stability of $\dot{x} = A^T x$.

The Stabilization Problem

Thus we rephrase Problem 1

Problem 1:

Find $P > 0, K$:

$(A + BK)P + P(A + BK)^T < 0$

Problem 2:

Find $X > 0, Z$:

$AX + BZ + XA^T + Z^T B^T < 0$

Theorem 35.

Problem 1 is equivalent to Problem 2.

Proof.

We will show that 2) Solves 1). Suppose $X > 0$, Z solves 2). Let $P = X > 0$ and $K = ZP^{-1}$. Then $Z = KP$ and

$$\begin{aligned}(A + BK)P + P(A + BK)^T &= AP + PA^T + BKP + PK^T B^T \\ &= AP + PA^T + BZ + Z^T B^T < 0\end{aligned}$$

Now suppose that $P > 0$ and K solve 1). Let $X = P > 0$ and $Z = KP$. Then

$$AP + PA^T + BZ + Z^T B^T = (A + BK)P + P(A + BK)^T < 0$$

The result can be summarized more succinctly

Theorem 36.

(A, B) *is static-state-feedback stabilizable if and only if there exists some* $P > 0$ *and* Z *such that*

$$AP + PA^T + BZ + Z^T B^T < 0$$

with $u(t) = ZP^{-1}x(t)$.

Standard Format:

$$\begin{bmatrix} A & B \end{bmatrix} \begin{bmatrix} P \\ Z \end{bmatrix} + \begin{bmatrix} P & Z^T \end{bmatrix} \begin{bmatrix} A^T \\ B^T \end{bmatrix} < 0$$

The Schur complement

Before we get to the main result, recall the Schur complement.

Theorem 37 (Schur Complement).

For any $S \in \mathbb{S}^n$, $Q \in \mathbb{S}^m$ and $R \in \mathbb{R}^{n \times m}$, the following are equivalent.

1. $\begin{bmatrix} M & R \\ R^T & Q \end{bmatrix} > 0$
2. $Q > 0$ *and* $M - RQ^{-1}R^T > 0$

A commonly used property of positive matrices.
Also Recall: If $X > 0$,

- then $X - \epsilon I > 0$ for ϵ sufficiently small.

The KYP Lemma (AKA: The Bounded Real Lemma)

The most important theorem in this lecture.

Lemma 38 (KYP Lemma).

Suppose

$$\hat{G}(s) = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right].$$

Then the following are equivalent.

- $\|G\|_{H_\infty} \le \gamma$.
- *There exists a* $X > 0$ *such that*

$$\begin{bmatrix} A^T X + XA & XB \\ B^T X & -\gamma I \end{bmatrix} + \frac{1}{\gamma}\begin{bmatrix} C^T \\ D^T \end{bmatrix}\begin{bmatrix} C & D \end{bmatrix} < 0$$

Can be used to calculate the H_∞-norm of a system

- Originally used to solve LMI's using graphs. (Before Computers)
- Now used directly instead of graphical methods like Bode.

The feasibility constraints are linear

- Can be combined with other methods.

The KYP Lemma

Proof.

We will only show that ii) implies i). The other direction requires the Hamiltonian, which we have not discussed.

- We will show that if $y = Gu$, then $\|y\|_{L_2} \le \gamma \|u\|_{L_2}$.
- From the 1 x 1 block of the LMI, we know that $A^TX + XA < 0$, which means A is Hurwitz.
- Because the inequality is strict, there exists some $\epsilon > 0$ such that

$$\begin{bmatrix} A^TX + XA & XB \\ B^TX & -(\gamma - \epsilon)I \end{bmatrix} + \frac{1}{\gamma}\begin{bmatrix} C^T \\ D^T \end{bmatrix}\begin{bmatrix} C & D \end{bmatrix}$$
$$= \begin{bmatrix} A^TX + XA & XB \\ B^TX & -\gamma I \end{bmatrix} + \frac{1}{\gamma}\begin{bmatrix} C^T \\ D^T \end{bmatrix}\begin{bmatrix} C & D \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & \epsilon I \end{bmatrix} < 0$$

- Let $y = Gu$. Then the state-space representation is

$$y(t) = Cx(t) + Du(t)$$
$$\dot{x}(t) = Ax(t) + Bu(t) \qquad\qquad x(0) = 0$$

The KYP Lemma

Proof.

- Let $V(x) = x^T X x$. Then the LMI implies

$$\begin{bmatrix} x(t) \\ u(t) \end{bmatrix}^T \left[\begin{bmatrix} A^T X + XA & XB \\ B^T X & -(\gamma - \epsilon) I \end{bmatrix} + \frac{1}{\gamma} \begin{bmatrix} C^T \\ D^T \end{bmatrix} \begin{bmatrix} C & D \end{bmatrix} \right] \begin{bmatrix} x(t) \\ u(t) \end{bmatrix}$$

$$= \begin{bmatrix} x \\ u \end{bmatrix}^T \begin{bmatrix} A^T X + XA & XB \\ B^T X & -(\gamma - \epsilon) I \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} + \frac{1}{\gamma} \begin{bmatrix} x \\ u \end{bmatrix}^T \begin{bmatrix} C^T \\ D^T \end{bmatrix} \begin{bmatrix} C & D \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix}$$

$$= \begin{bmatrix} x \\ u \end{bmatrix}^T \begin{bmatrix} A^T X + XA & XB \\ B^T X & -(\gamma - \epsilon) I \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} + \frac{1}{\gamma} y^T y$$

$$= x^T (A^T X + XA) x + x^T X B u + u^T B^T X x - (\gamma - \epsilon) u^T u + \frac{1}{\gamma} y^T y$$

$$= (Ax + Bu)^T X x + x^T X (Ax + Bu) - (\gamma - \epsilon) u^T u + \frac{1}{\gamma} y^T y$$

$$= \dot{x}(t)^T X x(t) + x(t)^T X \dot{x}(t) - (\gamma - \epsilon) \|u(t)\|^2 + \frac{1}{\gamma} \|y(t)\|^2$$

$$= \dot{V}(x(t)) - (\gamma - \epsilon) \|u(t)\|^2 + \frac{1}{\gamma} \|y(t)\|^2 < 0$$

The KYP Lemma

Proof.

- Now we have $\dot{V}(x(t)) - (\gamma - \epsilon)\|u(t)\|^2 + \frac{1}{\gamma}\|y(t)\|^2 < 0$
- Integrating in time, we get

$$\int_0^T \left(\dot{V}(x(t)) - (\gamma - \epsilon)\|u(t)\|^2 + \frac{1}{\gamma}\|y(t)\|^2\right) dt$$
$$= V(x(T)) - V(x(0)) - (\gamma - \epsilon)\int_0^T \|u(t)\|^2 dt + \frac{1}{\gamma}\int_0^T \|y(t)\|^2\Big) dt < 0$$

- Because A is Hurwitz, $\lim_{T\to\infty} x(T) = 0$.
- Hence $\lim_{T\to\infty} V(x(T)) = 0$.
- Likewise, because $x(0) = 0$, we have $V(x(0)) = 0$.

$\square$

The KYP Lemma

Proof.

- Since $V(x(0)) = V(x(\infty)) = 0$,

$$\lim_{T\to\infty}\left[V(x(T)) - V(x(0)) - (\gamma-\epsilon)\int_0^T \|u(t)\|^2 dt + \frac{1}{\gamma}\int_0^T \|y(t)\|^2\Big) dt\right]$$
$$= 0 - 0 - (\gamma-\epsilon)\int_0^\infty \|u(t)\|^2 dt + \frac{1}{\gamma}\int_0^\infty \|y(t)\|^2 dt$$
$$= -(\gamma-\epsilon)\|u\|_{L_2}^2 + \frac{1}{\gamma}\|y\|_{L_2}^2 dt < 0$$

- Thus

$$\|y\|_{L_2}^2 dt < (\gamma^2 - \epsilon\gamma)\|u\|_{L_2}^2$$

- By definition, this means $\|G\|_{H_\infty}^2 \le (\gamma^2 - \epsilon\gamma) < \gamma^2$ or

$$\|G\|_{H_\infty} < \gamma$$

A Variation on the KYP lemma is the positive-real lemma

Lemma 39.

Suppose

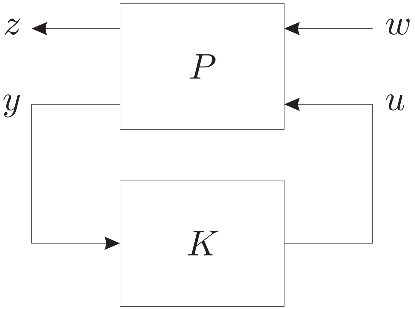
$$\hat{G}(s) = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right].$$

Then the following are equivalent.

- G *is passive. i.e.* $(\langle u, Gu \rangle_{L_2} \geq 0)$*.*
- *There exists a* $P > 0$ *such that*

$$\begin{bmatrix} A^T P + PA & PB - C^T \\ B^T P - C & -D^T - D \end{bmatrix} \leq 0$$

Recall: Linear Fractional Transformation



Plant:

$$\begin{bmatrix} z \\ y \end{bmatrix} = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \begin{bmatrix} w \\ u \end{bmatrix} \qquad \text{where} \qquad P = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & D_{22} \end{array}\right]$$

Controller:

$$u = Ky \qquad \text{where} \qquad K = \left[\begin{array}{c|c} A_K & B_K \\ \hline C_K & D_K \end{array}\right]$$

Choose K to minimize

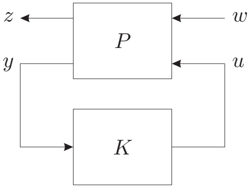
$$\|P_{11} + P_{12}(I - KP_{22})^{-1}KP_{21}\|$$

Equivalently choose $\left[\begin{array}{c|c} A_K & B_K \\ \hline C_K & D_K \end{array}\right]$ to minimize

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

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

Optimal Full-State Feedback Control



For the full-state feedback case, we consider a controller of the form

$$u(t) = Fx(t)$$

Controller:

$$u = Ky \qquad \text{where} \qquad K = \left[\begin{array}{c|c} 0 & 0 \\ \hline 0 & F \end{array}\right]$$

Plant:

$$\begin{bmatrix} z \\ y \end{bmatrix} = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \begin{bmatrix} w \\ u \end{bmatrix} \qquad \text{where} \qquad P = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & D_{11} & D_{12} \\ I & 0 & 0 \end{array}\right]$$

Optimal Full-State Feedback Control

Thus the closed-loop state-space representation is

$$\underline{\mathsf{S}}(\hat{P},\hat{K}) = \left[\begin{array}{c|c} A+B_2F & B_1 \\ \hline C_1+D_{12}F & D_{11} \end{array}\right]$$

By the KYP lemma, $\|\underline{\mathsf{S}}(\hat{P},\hat{K})\|_{H_\infty} < \gamma$ if and only if there exists some $X > 0$ such that

$$\begin{bmatrix} (A+B_2F)^T X + X(A+B_2F) & XB_1 \\ B_1^T X & -\gamma I \end{bmatrix} + \frac{1}{\gamma}\begin{bmatrix} (C_1+D_{12}F)^T \\ D_{11}^T \end{bmatrix}\begin{bmatrix} (C_1+D_{12}F) & D_{11} \end{bmatrix} < 0$$

This is a matrix inequality, but is nonlinear

- Quadratic (Not Bilinear)
- May NOT apply variable substitution trick.

Schur Complement

The KYP condition is

$$\begin{bmatrix} A^T X + XA & XB \\ B^T X & -\gamma I \end{bmatrix} + \frac{1}{\gamma} \begin{bmatrix} C^T \\ D^T \end{bmatrix} \begin{bmatrix} C & D \end{bmatrix} < 0$$

Recall the Schur Complement

Theorem 40 (Schur Complement).

For any $S \in \mathbb{S}^n$, $Q \in \mathbb{S}^m$ and $R \in \mathbb{R}^{n \times m}$, the following are equivalent.

1. $\begin{bmatrix} M & R \\ R^T & Q \end{bmatrix} < 0$
2. $Q < 0$ *and* $M - RQ^{-1}R^T < 0$

In this case, let $Q = -\frac{1}{\gamma} I < 0$,

$$M = \begin{bmatrix} A^T X + XA & XB \\ B^T X & -\gamma I \end{bmatrix} \qquad R = \begin{bmatrix} C & D \end{bmatrix}^T$$

Note we are making the LMI Larger.

Schur Complement

The Schur Complement says that

$$\begin{bmatrix} A^TX + XA & XB \\ B^TX & -\gamma I \end{bmatrix} + \frac{1}{\gamma}\begin{bmatrix} C^T \\ D^T \end{bmatrix}\begin{bmatrix} C & D \end{bmatrix} < 0$$

if and only if

$$\begin{bmatrix} A^TX + XA & XB & C^T \\ B^TX & -\gamma I & D^T \\ C & D & -\gamma I \end{bmatrix} < 0$$

This leads to the
Full-State Feedback Condition

$$\begin{bmatrix} (A + B_2F)^TX + X(A + B_2F) & XB_1 & (C_1 + D_{12}F)^T \\ B_1^TX & -\gamma I & D_{11}^T \\ (C_1 + D_{12}F) & D_{11} & -\gamma I \end{bmatrix} < 0$$

which is now bilinear in X and F.

Dual KYP Lemma

To apply the variable substitution trick, we must also construct the dual form of this LMI.

Lemma 41 (KYP Dual).

Suppose

$$\hat{G}(s) = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right].$$

Then the following are equivalent.

- $\|G\|_{H_\infty} \le \gamma$.
- *There exists a* $Y > 0$ *such that*

$$\begin{bmatrix} YA^T + AY & B & YC^T \\ B^T & -\gamma I & D^T \\ CY & D & -\gamma I \end{bmatrix} < 0$$

Dual KYP Lemma

Proof.

Let $X = Y^{-1}$. Then

$$\begin{bmatrix} YA^T + AY & XB & YC^T \\ B^TX & -\gamma I & D^T \\ CY & D & -\gamma I \end{bmatrix} < 0 \qquad \text{and} \qquad Y > 0$$

if and only if $X > 0$ and

$$\begin{bmatrix} Y^{-1} & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix} \begin{bmatrix} YA^T + AY & B & YC^T \\ B^T & -\gamma I & D^T \\ CY & D & -\gamma I \end{bmatrix} \begin{bmatrix} Y^{-1} & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix}$$

$$= \begin{bmatrix} A^TX + XA & XB & C^T \\ B^TX & -\gamma I & D^T \\ C & D & -\gamma I \end{bmatrix} < 0.$$

By the Schur complement this is equivalent to

$$\begin{bmatrix} A^TX + XA & XB \\ B^TX & -\gamma I \end{bmatrix} + \frac{1}{\gamma} \begin{bmatrix} C^T \\ D^T \end{bmatrix} \begin{bmatrix} C & D \end{bmatrix} < 0$$

By the KYP lemma, this is equivalent to $\|G\|_{H_\infty} \le \gamma$. $\square$

Full-State Feedback Optimal Control

We can now apply this result to the state-feedback problem.

Theorem 42.

The following are equivalent:

- *There exists an F such that $\|\underline{S}(P, K(0,0,0,F))\|_{H_\infty} \le \gamma$.*
- *There exist $Y > 0$ and Z such that*

$$\begin{bmatrix} YA^T + AY + Z^T B_2^T + B_2 Z & B_1 & YC_1^T + Z^T D_{12}^T \\ B_1^T & -\gamma I & D_{11}^T \\ C_1 Y + D_{12} Z & D_{11} & -\gamma I \end{bmatrix} < 0$$

One may use $F = ZY^{-1}$.

Full-State Feedback Optimal Control

Proof.

Suppose there exists an F such that $\|\underline{S}(P, K(0,0,0,F))\|_{H_\infty} \le \gamma$. By the Dual KYP lemma, this implies there exists a $Y > 0$ such that

$$\begin{bmatrix} Y(A+B_2F)^T + (A+B_2F)Y & B_1 & Y(C_1+D_{12}F)^T \\ B_1^T & -\gamma I & D_{11}^T \\ (C_1+D_{12}F)Y & D_{11} & -\gamma I \end{bmatrix} < 0$$

Let $Z = FY$. Then

$$\begin{bmatrix} YA^T + Z^TB_2^T + AY + B_2Z & B_1 & YC_1^T + Z^TD_{12}^T)^T \\ B_1^T & -\gamma I & D_{11}^T \\ C_1Y + D_{12}Z & D_{11} & -\gamma I \end{bmatrix}$$

$$= \begin{bmatrix} YA^T + YF^TB_2^T + AY + B_2FY & B_1 & YC_1^T + YF^TD_{12}^T)^T \\ B_1^T & -\gamma I & D_{11}^T \\ C_1Y + D_{12}FY & D_{11} & -\gamma I \end{bmatrix}$$

$$= \begin{bmatrix} Y(A+B_2F)^T + (A+B_2F)Y & B_1 & Y(C_1+D_{12}F)^T \\ B_1^T & -\gamma I & D_{11}^T \\ (C_1+D_{12}F)Y & D_{11} & -\gamma I \end{bmatrix} < 0.$$

Full-State Feedback Optimal Control

Proof.

Now suppose there exists a $Y > 0$ and Z such that

$$\begin{bmatrix} YA^T + Z^T B_2^T + AY + B_2 Z & B_1 & YC_1^T + Z^T D_{12}^T \\ B_1^T & -\gamma I & D_{11}^T \\ C_1 Y + D_{12} Z & D_{11} & -\gamma I \end{bmatrix} < 0$$

Let $F = ZY^{-1}$. Then

$$\begin{aligned} &\begin{bmatrix} Y(A + B_2 F)^T + (A + B_2 F)Y & B_1 & Y(C_1 + D_{12} F)^T \\ B_1^T & -\gamma I & D_{11}^T \\ (C_1 + D_{12} F)Y & D_{11} & -\gamma I \end{bmatrix} \\ &= \begin{bmatrix} YA^T + YF^T B_2^T + AY + B_2 FY & B_1 & YC_1^T + YF^T D_{12}^T \\ B_1^T & -\gamma I & D_{11}^T \\ C_1 Y + D_{12} FY & D_{11} & -\gamma I \end{bmatrix} \\ &= \begin{bmatrix} YA^T + Z^T B_2^T + AY + B_2 Z & B_1 & YC_1^T + Z^T D_{12}^T \\ B_1^T & -\gamma I & D_{11}^T \\ C_1 Y + D_{12} Z & D_{11} & -\gamma I \end{bmatrix} < 0 \end{aligned}$$

Full-State Feedback Optimal Control

Therefore the following optimization problems are equivalent
Form A

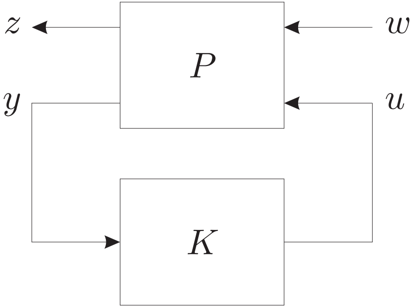
$$\min_F \|\underline{\mathsf{S}}(P, K(0,0,0,F))\|_{H_\infty}$$

Form B

$$\min_{\gamma, Y, Z} \gamma :$$

$$\begin{bmatrix} -Y & 0 & 0 & 0 \\ 0 & YA^T + AY + Z^T B_2^T + B_2 Z & B_1 & YC_1^T + Z^T D_{12}^T \\ 0 & B_1^T & -\gamma I & D_{11}^T \\ 0 & C_1 Y + D_{12} Z & D_{11} & -\gamma I \end{bmatrix} < 0$$

The optimal controller is given by $F = ZY^{-1}$.
Next: Optimal Output Feedback



Plant:

$$\begin{bmatrix} z \\ y \end{bmatrix} = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \begin{bmatrix} w \\ u \end{bmatrix} \qquad \text{where} \qquad P = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & D_{22} \end{array}\right]$$

Controller:

$$u = Ky \qquad \text{where} \qquad K = \left[\begin{array}{c|c} A_K & B_K \\ \hline C_K & D_K \end{array}\right]$$

Choose K to minimize

$$\|P_{11} + P_{12}(I - KP_{22})^{-1}KP_{21}\|$$

Equivalently choose $\left[\begin{array}{c|c} A_K & B_K \\ \hline C_K & D_K \end{array}\right]$ to minimize

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

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

Optimal Control

Recall that

$$\begin{bmatrix} I & -D_K \\ -D_{22} & I \end{bmatrix}^{-1} = \begin{bmatrix} I + D_K Q D_{22} & D_K Q \\ Q D_{22} & Q \end{bmatrix}$$

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

$$\begin{aligned} A_{cl} &:= \begin{bmatrix} A & 0 \\ 0 & A_K \end{bmatrix} + \begin{bmatrix} B_2 & 0 \\ 0 & B_K \end{bmatrix} \begin{bmatrix} I & -D_K \\ -D_{22} & I \end{bmatrix}^{-1} \begin{bmatrix} 0 & C_K \\ C_2 & 0 \end{bmatrix} \\ &= \begin{bmatrix} A & 0 \\ 0 & A_K \end{bmatrix} + \begin{bmatrix} B_2 & 0 \\ 0 & B_K \end{bmatrix} \begin{bmatrix} I + D_K Q D_{22} & D_K Q \\ Q D_{22} & Q \end{bmatrix} \begin{bmatrix} 0 & C_K \\ C_2 & 0 \end{bmatrix} \\ &= \begin{bmatrix} A + B_2 D_K Q C_2 & B_2(I + D_K Q D_{22}) C_K \\ B_K Q C_2 & A_K + B_K Q D_{22} C_K \end{bmatrix} \end{aligned}$$

Likewise

$$\begin{aligned} C_{cl} &:= \begin{bmatrix} C_1 & 0 \end{bmatrix} + \begin{bmatrix} D_{12} & 0 \end{bmatrix} \begin{bmatrix} I + D_K Q D_{22} & D_K Q \\ Q D_{22} & Q \end{bmatrix} \begin{bmatrix} 0 & C_K \\ C_2 & 0 \end{bmatrix} \\ &= \begin{bmatrix} C_1 + D_{12} D_K Q C_2 & D_{12}(I + D_K Q D_{22}) C_K \end{bmatrix} \end{aligned}$$

Thus we have

$$\left[\begin{array}{cc|c} A + B_2 D_K Q C_2 & B_2(I + D_K Q D_{22}) C_K & B_1 + B_2 D_K Q D_{21} \\ B_K Q C_2 & A_K + B_K Q D_{22} C_K & B_K Q D_{21} \\ \hline \left[C_1 + D_{12} D_K Q C_2 \right.& \left. D_{12}(I + D_K Q D_{22}) C_K\right] & D_{11} + D_{12} D_K Q D_{21} \end{array}\right]$$

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

- This is nonlinear in (A_K, B_K, C_K, D_K).
- Hence we make a change of variables (First of several).

$$\begin{aligned} A_{K2} &= A_K + B_K Q D_{22} C_K \\ B_{K2} &= B_K Q \\ C_{K2} &= (I + D_K Q D_{22}) C_K \\ D_{K2} &= D_K Q \end{aligned}$$

Optimal Output Feedback Control

This yields the system

$$\left[\begin{array}{c|c} \begin{bmatrix} A + B_2 D_{K2} C_2 & B_2 C_{K2} \\ B_{K2} C_2 & A_{K2} \end{bmatrix} & \begin{matrix} B_1 + B_2 D_{K2} D_{21} \\ B_{K2} D_{21} \end{matrix} \\ \hline \begin{bmatrix} C_1 + D_{12} D_{K2} C_2 & D_{12} C_{K2} \end{bmatrix} & D_{11} + D_{12} D_{K2} D_{21} \end{array}\right]$$

Which is affine in $\left[\begin{array}{c|c} A_{K2} & B_{K2} \\ \hline C_{K2} & D_{K2} \end{array}\right]$.

Optimal Output Feedback Control

Hence we can optimize over our new variables.

- However, the change of variables must be invertible.

If we recall that

$$(I - QM)^{-1} = I + Q(I - MQ)^{-1}M$$

then we get

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

Examine the variable C_{K2}

$$\begin{aligned} C_{K2} &= (I + D_K(I - D_{22}D_K)^{-1}D_{22})C_K \\ &= (I - D_K D_{22})^{-1}C_K \end{aligned}$$

Hence, given C_{K2}, we can recover C_K as

$$C_K = (I - D_K D_{22})C_{K2}$$

Optimal Output Feedback Control

Now suppose we have D_{K2}. Then

$$D_{K2} = D_K Q = D_K (I - D_{22} D_K)^{-1}$$

implies that

$$D_K = D_{K2}(I - D_{22} D_K) = D_{K2} - D_{K2} D_{22} D_K$$

or

$$(I + D_{K2} D_{22}) D_K = D_{K2}$$

which can be inverted to get

$$D_K = (I + D_{K2} D_{22})^{-1} D_{K2}$$

Optimal Output Feedback Control

Once we have C_K and D_K, the other variables are easily recovered as

$$B_K = B_{K2}Q^{-1} = B_{K2}(I - D_{22}D_K)$$
$$A_K = A_{K2} - B_K(I - D_{22}D_K)^{-1}D_{22}C_K$$

To summarize, the original variables can be recovered as

$$D_K = (I + D_{K2}D_{22})^{-1}D_{K2}$$
$$B_K = B_{K2}(I - D_{22}D_K)$$
$$C_K = (I - D_K D_{22})C_{K2}$$
$$A_K = A_{K2} - B_K(I - D_{22}D_K)^{-1}D_{22}C_K$$

Optimal Output Feedback Control

$$\left[\begin{array}{c|c} A_{cl} & B_{cl} \\ \hline C_{cl} & D_{cl} \end{array}\right] := \left[\begin{array}{c|c} \begin{bmatrix} A + B_2 D_{K2} C_2 & B_2 C_{K2} \\ B_{K2} C_2 & A_{K2} \end{bmatrix} & \begin{matrix} B_1 + B_2 D_{K2} D_{21} \\ B_{K2} D_{21} \end{matrix} \\ \hline \begin{bmatrix} C_1 + D_{12} D_{K2} C_2 & D_{12} C_{K2} \end{bmatrix} & D_{11} + D_{12} D_{K2} D_{21} \end{array}\right]$$

$$\begin{bmatrix} A_{cl} & B_{cl} \\ C_{cl} & D_{cl} \end{bmatrix} = \begin{bmatrix} A & 0 & B_1 \\ 0 & 0 & 0 \\ C_1 & 0 & D_{11} \end{bmatrix} + \begin{bmatrix} 0 & B_2 \\ I & 0 \\ 0 & D_{12} \end{bmatrix} \begin{bmatrix} A_{K2} & B_{K2} \\ C_{K2} & D_{K2} \end{bmatrix} \begin{bmatrix} 0 & I & 0 \\ C_2 & 0 & D_{21} \end{bmatrix}$$

Or

$$A_{cl} = \begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & B_2 \\ I & 0 \end{bmatrix} \begin{bmatrix} A_{K2} & B_{K2} \\ C_{K2} & D_{K2} \end{bmatrix} \begin{bmatrix} 0 & I \\ C_2 & 0 \end{bmatrix}$$

$$B_{cl} = \begin{bmatrix} B_1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 & B_2 \\ I & 0 \end{bmatrix} \begin{bmatrix} A_{K2} & B_{K2} \\ C_{K2} & D_{K2} \end{bmatrix} \begin{bmatrix} 0 \\ D_{21} \end{bmatrix}$$

$$C_{cl} = \begin{bmatrix} C_1 & 0 \end{bmatrix} + \begin{bmatrix} 0 & D_{12} \end{bmatrix} \begin{bmatrix} A_{K2} & B_{K2} \\ C_{K2} & D_{K2} \end{bmatrix} \begin{bmatrix} 0 & I \\ C_2 & 0 \end{bmatrix}$$

$$D_{cl} = \begin{bmatrix} D_{11} \end{bmatrix} + \begin{bmatrix} 0 & D_{12} \end{bmatrix} \begin{bmatrix} A_{K2} & B_{K2} \\ C_{K2} & D_{K2} \end{bmatrix} \begin{bmatrix} 0 \\ D_{21} \end{bmatrix}$$

Lemma 43 (Transformation Lemma).

Suppose that

$$\begin{bmatrix} Y_1 & I \\ I & X_1 \end{bmatrix} > 0$$

Then there exist X_2, X_3, Y_2, Y_3 *such that*

$$X = \begin{bmatrix} X_1 & X_2 \\ X_2^T & X_3 \end{bmatrix} = \begin{bmatrix} Y_1 & Y_2 \\ Y_2^T & Y_3 \end{bmatrix}^{-1} = Y^{-1} > 0$$

where $Y_{cl} = \begin{bmatrix} Y_1 & I \\ Y_2^T & 0 \end{bmatrix}$ *has full rank.*

Transformation Lemma

Proof.

- Since
$$\begin{bmatrix} Y_1 & I \\ I & X_1 \end{bmatrix} > 0,$$
by the Schur complement $X_1 > 0$ and $X_1^{-1} - Y_1 > 0$. Since $I - X_1 Y_1 = X_1(X_1^{-1} - Y_1)$, we conclude that $I - X_1 Y_1$ is invertible.
- Choose any two square invertible matrices X_2 and Y_2 such that
$$X_2 Y_2^T = I - X_1 Y_1$$
- Because X_2 and Y_2 are non-singular,
$$Y_{cl}^T = \begin{bmatrix} Y_1 & Y_2 \\ I & 0 \end{bmatrix} \text{ and } X_{cl} = \begin{bmatrix} I & 0 \\ X_1 & X_2 \end{bmatrix}$$
are also non-singular.

□

Transformation Lemma

Proof.

- Now define X and Y as

$$X = Y_{cl}^{-T} X_{cl} \qquad \text{and} \qquad Y = X_{cl}^{-1} Y_{cl}^{T}.$$

Then

$$XY = Y_{cl}^{-1} X_{cl} X_{cl}^{-1} Y_{cl} = I$$

Likewise, $YX = I$. Hence, $Y = X^{-1}$. $\square$

Lemma 44 (Converse Transformation Lemma).

Given $X = \begin{bmatrix} X_1 & X_2 \\ X_2^T & X_3 \end{bmatrix} > 0$ *where* X_2 *has full column rank. Let*

$$X^{-1} = Y = \begin{bmatrix} Y_1 & Y_2 \\ Y_2^T & Y_3 \end{bmatrix}$$

then

$$\begin{bmatrix} Y_1 & I \\ I & X_1 \end{bmatrix} > 0$$

and $Y_{cl} = \begin{bmatrix} Y_1 & I \\ Y_2^T & 0 \end{bmatrix}$ *has full column rank.*

Converse Transformation Lemma

Proof.

Since X_2 is full rank, $X_{cl} = \begin{bmatrix} I & 0 \\ X_1 & X_2 \end{bmatrix}$ also has full column rank. Note that $XY = I$ implies

$$Y_{cl}^T X = \begin{bmatrix} Y_1 & Y_2 \\ I & 0 \end{bmatrix} \begin{bmatrix} X_1 & X_2 \\ X_2^T & X_3 \end{bmatrix} = \begin{bmatrix} I & 0 \\ X_1 & X_2 \end{bmatrix} = X_{cl}.$$

Hence

$$Y_{cl}^T = \begin{bmatrix} Y_1 & Y_2 \\ I & 0 \end{bmatrix} = \begin{bmatrix} I & 0 \\ X_1 & X_2 \end{bmatrix} Y = X_{cl} Y$$

has full column rank. Now, since $XY = I$ implies $X_1 Y_1 + X_2 Y_2^T = I$, we have

$$X_{cl} Y_{cl} = \begin{bmatrix} I & 0 \\ X_1 & X_2 \end{bmatrix} \begin{bmatrix} Y_1 & I \\ Y_2^T & 0 \end{bmatrix} = \begin{bmatrix} Y_1 & I \\ X_1 Y_1 + X_2 Y_2^T & X_1 \end{bmatrix} = \begin{bmatrix} Y_1 & I \\ I & X_1 \end{bmatrix}$$

Furthermore, because Y_{cl} has full rank,

$$\begin{bmatrix} Y_1 & I \\ I & X_1 \end{bmatrix} = X_{cl} Y_{cl} = X_{cl} Y X_{cl}^T = Y_{cl}^T X Y_{cl} > 0$$

Theorem 45.

The following are equivalent.

- *There exists a* $\hat{K} = \left[\begin{array}{c|c} A_K & B_K \\ \hline C_K & D_K \end{array}\right]$ *such that* $\|S(P,K)\|_{H_\infty} < \gamma$.
- *There exist* $X_1, Y_1, A_n, B_n, C_n, D_n$ *such that* $\begin{bmatrix} X_1 & I \\ I & Y_1 \end{bmatrix} > 0$

$$\begin{bmatrix} AY_1+Y_1A^T+B_2C_n+C_n^TB_2^T & *^T & *^T & *^T \\ A^T+A_n+[B_2D_nC_2]^T & X_1A+A^TX_1+B_nC_2+C_2^TB_n^T & *^T & *^T \\ [B_1+B_2D_nD_{21}]^T & [XB_1+B_nD_{21}]^T & -\gamma I & \\ C_1Y_1+D_{12}C_n & C_1+D_{12}D_nC_2 & D_{11}+D_{12}D_nD_{21} & -\gamma I \end{bmatrix} < 0$$

Moreover,

$$\left[\begin{array}{c|c} A_{K2} & B_{K2} \\ \hline C_{K2} & D_{K2} \end{array}\right] = \begin{bmatrix} X_2 & X_1B_2 \\ 0 & I \end{bmatrix}^{-1} \left[\begin{bmatrix} A_n & B_n \\ C_n & D_n \end{bmatrix} - \begin{bmatrix} X_1AY_1 & 0 \\ 0 & 0 \end{bmatrix}\right] \begin{bmatrix} Y_2^T & 0 \\ C_2Y_1 & I \end{bmatrix}^{-1}$$

for any full-rank X_2 *and* Y_2 *such that*

$$\begin{bmatrix} X_1 & X_2 \\ X_2^T & X_3 \end{bmatrix} = \begin{bmatrix} Y_1 & Y_2 \\ Y_2^T & Y_3 \end{bmatrix}^{-1}$$

Optimal Output Feedback Control

Proof: If.

Suppose there exist $X_1, Y_1, A_n, B_n, C_n, D_n$ such that the LMI is feasible. Since

$$\begin{bmatrix} X_1 & I \\ I & Y_1 \end{bmatrix} > 0,$$

by the transformation lemma, there exist X_2, X_3, Y_2, Y_3 such that

$$X := \begin{bmatrix} X & X_2 \\ X_2^T & X_3 \end{bmatrix} = \begin{bmatrix} Y & Y_2 \\ Y_2^T & Y_3 \end{bmatrix}^{-1} > 0$$

where $Y_{cl} = \begin{bmatrix} Y & I \\ Y_2^T & 0 \end{bmatrix}$ has full row rank. Let $K = \left[\begin{array}{c|c} A_K & B_K \\ \hline C_K & D_K \end{array}\right]$ where

$$\begin{aligned}
D_K &= (I + D_{K2}D_{22})^{-1}D_{K2} \\
B_K &= B_{K2}(I - D_{22}D_K) \\
C_K &= (I - D_K D_{22})C_{K2} \\
A_K &= A_{K2} - B_K(I - D_{22}D_K)^{-1}D_{22}C_K.
\end{aligned}$$

Optimal Output Feedback Control

Proof: If.

and where

$$\left[\begin{array}{c|c} A_{K2} & B_{K2} \\ \hline C_{K2} & D_{K2} \end{array}\right] = \begin{bmatrix} X_2 & X_1 B_2 \\ 0 & I \end{bmatrix}^{-1} \left[\begin{bmatrix} A_n & B_n \\ C_n & D_n \end{bmatrix} - \begin{bmatrix} X_1 A Y_1 & 0 \\ 0 & 0 \end{bmatrix} \right] \begin{bmatrix} Y_2^T & 0 \\ C_2 Y_1 & I \end{bmatrix}^{-1}.$$

□

Proof: If.

As discussed previously, this means the closed-loop system is

$$\begin{bmatrix} A_{cl} & B_{cl} \\ C_{cl} & D_{cl} \end{bmatrix} = \begin{bmatrix} A & 0 & B_1 \\ 0 & 0 & 0 \\ C_1 & 0 & D_{11} \end{bmatrix} + \begin{bmatrix} 0 & B_2 \\ I & 0 \\ 0 & D_{12} \end{bmatrix} \begin{bmatrix} A_{K2} & B_{K2} \\ C_{K2} & D_{K2} \end{bmatrix} \begin{bmatrix} 0 & I & 0 \\ C_2 & 0 & D_{21} \end{bmatrix}$$

$$= \begin{bmatrix} A & 0 & B_1 \\ 0 & 0 & 0 \\ C_1 & 0 & D_{11} \end{bmatrix} + \begin{bmatrix} 0 & B_2 \\ I & 0 \\ 0 & D_{12} \end{bmatrix}$$

$$\begin{bmatrix} X_2 & X_1 B_2 \\ 0 & I \end{bmatrix}^{-1} \left[\begin{bmatrix} A_n & B_n \\ C_n & D_n \end{bmatrix} - \begin{bmatrix} X_1 A Y_1 & 0 \\ 0 & 0 \end{bmatrix} \right] \begin{bmatrix} Y_2^T & 0 \\ C_2 Y_1 & I \end{bmatrix}^{-1} \begin{bmatrix} 0 & I & 0 \\ C_2 & 0 & D_{21} \end{bmatrix}$$

Now look at the LMI from the KYP lemma.

□

Proof: If.

Expanding out, we obtain

$$\begin{bmatrix} Y_{cl}^T & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix} \begin{bmatrix} A_{cl}^T X + X A_{cl} & X B_{cl} & C_{cl}^T \\ B_{cl}^T X & -\gamma I & D_{cl}^T \\ C_{cl} & D_{cl} & -\gamma I \end{bmatrix} \begin{bmatrix} Y_{cl} & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix} =$$

$$\begin{bmatrix} AY_1 + Y_1 A^T + B_2 C_n + C_n^T B_2^T & *^T & *^T & *^T \\ A^T + A_n + [B_2 D_n C_2]^T & X_1 A + A^T X_1 + B_n C_2 + C_2^T B_n^T & *^T & *^T \\ [B_1 + B_2 D_n D_{21}]^T & [X B_1 + B_n D_{21}]^T & -\gamma I & \\ C_1 Y_1 + D_{12} C_n & C_1 + D_{12} D_n C_2 & D_{11} + D_{12} D_n D_{21} & -\gamma I \end{bmatrix} < 0$$

Hence, by the KYP lemma, $\underline{S}(P,K) = \left[\begin{array}{c|c} A_{cl} & B_{cl} \\ \hline C_{cl} & D_{cl} \end{array}\right]$ satisfies $\|\underline{S}(P,K)\|_{H_\infty} < \gamma$. □

Proof: Only If.

Now suppose that $\|\underline{\mathsf{S}}(P,K)\|_{H_\infty} < \gamma$ for some $K = \left[\begin{array}{c|c} A_K & B_K \\ \hline C_K & D_K \end{array}\right]$. Since $\|\underline{\mathsf{S}}(P,K)\|_{H_\infty} < \gamma$, by the KYP lemma, there exists a

$$X = \begin{bmatrix} X_1 & X_2 \\ X_2^T & X_3 \end{bmatrix} > 0$$

such that

$$\begin{bmatrix} A_{cl}^T X + X A_{cl} & X B_{cl} & C_{cl}^T \\ B_{cl}^T X & -\gamma I & D_{cl}^T \\ C_{cl} & D_{cl} & -\gamma I \end{bmatrix} < 0$$

Because the inequalities are strict, we can assume that X_2 has full row rank. Define

$$Y = \begin{bmatrix} Y_1 & Y_2 \\ Y_2^T & Y_3 \end{bmatrix} = X^{-1} \qquad \text{and} \qquad Y_{cl} = \begin{bmatrix} Y_1 & I \\ Y_2^T & 0 \end{bmatrix}$$

Then, according to the converse transformation lemma, Y_{cl} has full row rank and

$$\begin{bmatrix} X_1 & I \\ I & Y_1 \end{bmatrix} > 0.$$

Optimal Output Feedback Control

Proof: Only If.

Now, using the given A_K, B_K, C_K, D_K, define the variables

$$\begin{bmatrix} A_n & B_n \\ C_n & D_n \end{bmatrix} = \begin{bmatrix} X_2 & X_1 B_2 \\ 0 & I \end{bmatrix} \begin{bmatrix} A_{K2} & B_{K2} \\ C_{K2} & D_{K2} \end{bmatrix} \begin{bmatrix} Y_2^T & 0 \\ C_2 Y_1 & I \end{bmatrix} + \begin{bmatrix} X_1 A Y_1 & 0 \\ 0 & 0 \end{bmatrix}.$$

where

$$A_{K2} = A_K + B_K (I - D_{22} D_K)^{-1} D_{22} C_K \qquad B_{K2} = B_K (I - D_{22} D_K)^{-1}$$
$$C_{K2} = (I + D_K (I - D_{22} D_K)^{-1} D_{22}) C_K \qquad D_{K2} = D_K (I - D_{22} D_K)^{-1}$$

Then as before

$$\begin{bmatrix} A_{cl} & B_{cl} \\ C_{cl} & D_{cl} \end{bmatrix} = \begin{bmatrix} A & 0 & B_1 \\ 0 & 0 & 0 \\ C_1 & 0 & D_{11} \end{bmatrix} + \begin{bmatrix} 0 & B_2 \\ I & 0 \\ 0 & D_{12} \end{bmatrix}$$

$$\begin{bmatrix} X_2 & X_1 B_2 \\ 0 & I \end{bmatrix}^{-1} \left[\begin{bmatrix} A_n & B_n \\ C_n & D_n \end{bmatrix} - \begin{bmatrix} X_1 A Y_1 & 0 \\ 0 & 0 \end{bmatrix} \right] \begin{bmatrix} Y_2^T & 0 \\ C_2 Y_1 & I \end{bmatrix}^{-1} \begin{bmatrix} 0 & I & 0 \\ C_2 & 0 & D_{21} \end{bmatrix}$$

□

Proof: Only If.

Expanding out the LMI, we find

$$\begin{bmatrix} AY_1+Y_1A^T+B_2C_n+C_n^TB_2^T & *^T & *^T & *^T \\ A^T+A_n+[B_2D_nC_2]^T & X_1A+A^TX_1+B_nC_2+C_2^TB_n^T & *^T & *^T \\ [B_1+B_2D_nD_{21}]^T & [XB_1+B_nD_{21}]^T & -\gamma I & \\ C_1Y_1+D_{12}C_n & C_1+D_{12}D_nC_2 & D_{11}+D_{12}D_nD_{21} & -\gamma I \end{bmatrix}$$
$$= \begin{bmatrix} Y_{cl}^T & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix} \begin{bmatrix} A_{cl}^TX+XA_{cl} & XB_{cl} & C_{cl}^T \\ B_{cl}^TX_{cl} & -\gamma I & D_{cl}^T \\ C_{cl} & D_{cl} & -\gamma I \end{bmatrix} \begin{bmatrix} Y_{cl} & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix} < 0$$

□

Conclusion

To solve the H_∞-optimal state-feedback problem, we solve

$$\min_{\gamma, X_1, Y_1, A_n, B_n, C_n, D_n} \gamma \qquad \text{such that}$$

$$\begin{bmatrix} X_1 & I \\ I & Y_1 \end{bmatrix} > 0$$

$$\begin{bmatrix} AY_1 + Y_1 A^T + B_2 C_n + C_n^T B_2^T & *^T & *^T & *^T \\ A^T + A_n + [B_2 D_n C_2]^T & X_1 A + A^T X_1 + B_n C_2 + C_2^T B_n^T & *^T & *^T \\ [B_1 + B_2 D_n D_{21}]^T & [X B_1 + B_n D_{21}]^T & -\gamma I & \\ C_1 Y_1 + D_{12} C_n & C_1 + D_{12} D_n C_2 & D_{11} + D_{12} D_n D_{21} & -\gamma I \end{bmatrix} < 0$$

Conclusion

Then, we construct our controller using

$$\begin{aligned} D_K &= (I + D_{K2}D_{22})^{-1}D_{K2} \\ B_K &= B_{K2}(I - D_{22}D_K) \\ C_K &= (I - D_K D_{22})C_{K2} \\ A_K &= A_{K2} - B_K(I - D_{22}D_K)^{-1}D_{22}C_K. \end{aligned}$$

where

$$\left[\begin{array}{c|c} A_{K2} & B_{K2} \\ \hline C_{K2} & D_{K2} \end{array}\right] = \begin{bmatrix} X_2 & X_1 B_2 \\ 0 & I \end{bmatrix}^{-1} \left[\begin{bmatrix} A_n & B_n \\ C_n & D_n \end{bmatrix} - \begin{bmatrix} X_1 A Y_1 & 0 \\ 0 & 0 \end{bmatrix}\right] \begin{bmatrix} Y_2^T & 0 \\ C_2 Y_1 & I \end{bmatrix}^{-1}.$$

and where X_2 and Y_2 are any matrices which satisfy $X_2 Y_2 = I - X_1 Y_1$.

- e.g. Let $Y_2 = I$ and $X_2 = I - X_1 Y_1$.
- The optimal controller is NOT uniquely defined.
- Don't forget to check invertibility of $I - D_{22}D_K$

Conclusion

The H_∞-optimal controller is a dynamic system.

- Transfer Function $\hat{K}(s) = \left[\begin{array}{c|c} A_K & B_K \\ \hline C_K & D_K \end{array}\right]$

Minimizes the effect of external input (w) on external output (z).

$$\|z\|_{L_2} \le \|\underline{\mathsf{S}}(P,K)\|_{H_\infty}\|w\|_{L_2}$$

- Minimum Energy Gain

H_2-optimal control

Motivation

H_2-optimal control minimizes the H_2-norm of the transfer function.

- The H_2-norm has no direct interpretation.

$$\|G\|_{H_2}^2 = \frac{1}{2\pi}\int_{-\infty}^{\infty} \text{Trace}(\hat{G}(\imath\omega)^*\hat{G}(\imath\omega))d\omega$$

Motivation: Assume external input is Gaussian noise with signal variance S_w

$$E[w(t)^2] = \frac{1}{2\pi}\int_{-\infty}^{\infty} \text{Trace}(\hat{S}_w(\imath\omega))d\omega$$

Theorem 46.

For an LTI system P, if w is noise with spectral density $\hat{S}_w(\imath\omega)$ and $z = Pw$, then z is noise with density

$$\hat{S}_z(\imath\omega) = \hat{P}(\imath\omega)\hat{S}(\imath\omega)\hat{P}(\imath\omega)^*$$

Then the output $z = Pw$ has signal variance (Power)

$$E[z(t)^2] = \frac{1}{2\pi}\int_{-\infty}^{\infty} \text{Trace}(\hat{G}(\imath\omega)^* S(\imath\omega)\hat{G}(\imath\omega))d\omega$$
$$\leq \|S\|_{H_\infty}\|G\|_{H_2}^2$$

If the input signal is white noise, then $\hat{S}(\imath\omega) = I$ and

$$E[z(t)^2] = \|G\|_{H_2}^2$$

Now suppose the noise is colored with variance $\hat{S}_w(\imath\omega)$. Now define $\hat{H}$ as $\hat{H}(\imath\omega)\hat{H}(\imath\omega)^* = \hat{S}_w(\imath\omega)$ and the filtered system.

$$\hat{P}_s(s) = \begin{bmatrix} \hat{P}_{11}(s)\hat{H}(s) & \hat{P}_{12}(s) \\ \hat{P}_{21}(s)\hat{H}(s) & \hat{P}_{22}(s) \end{bmatrix}$$

Now, applying feedback to the filtered plant, we get

$$\underline{\mathsf{S}}(P_s, K)(s) = P_{11}H + P_{12}(I - KP_{22})^{-1}KP_{21}H = \underline{\mathsf{S}}(P, K)H$$

Now the spectral density, $\hat{S}_z$ of the output of the true plant using colored noise equals the output of the artificial plant under white noise. i.e.

$$\begin{aligned} S_z(s) &= \underline{\mathsf{S}}(P, K)(s)\hat{S}_w(s)\underline{\mathsf{S}}(P, K)(s)^* \\ &= \underline{\mathsf{S}}(P, K)(s)\hat{H}(s)\hat{H}(s)^*\underline{\mathsf{S}}(P, K)(s)^* = \underline{\mathsf{S}}(P_s, K)(s)\hat{S}(P_s, K)(s)^* \end{aligned}$$

Thus if K minimizes the H_2-norm of the filtered plant ($\|\hat{S}(P_s, K)\|_{H_2}^2$), it will minimize the variance of the true plant under the influence of colored noise with density $\hat{S}_w$.

Theorem 47.

Suppose $\hat{P}(s) = C(sI - A)^{-1}B$. Then the following are equivalent.

1. *A is Hurwitz and $\|\hat{P}\|_{H_2} < \gamma$.*
2. *There exists some $X > 0$ such that*

$$\operatorname{trace} CXC^T < \gamma^2$$
$$AX + XA^T + BB^T < 0$$

H_2-optimal control

Proof.

Suppose A is Hurwitz and $\|\hat{P}\|_{H_2} < \gamma$. Then the Controllability Grammian is defined as

$$X_c = \int_0^\infty e^{At} B B^T e^{A^T} dt$$

Now recall the Laplace transform

$$\begin{aligned}
\left(\Lambda e^{At}\right)(s) &= \int_0^\infty e^{At} e^{-ts} dt \\
&= \int_0^\infty e^{-(sI-A)t} dt \\
&= -(sI-A)^{-1} e^{-(sI-A)t} dt \Big|_{t=0}^{t=-\infty} \\
&= (sI-A)^{-1}
\end{aligned}$$

Hence $\left(\Lambda C e^{At} B\right)(s) = C(sI-A)^{-1}B$. □

Proof.

$\left(\Lambda C e^{At} B\right)(s) = C(sI - A)^{-1}B$ implies

$$\begin{aligned}
\|\hat{P}\|_{H_2}^2 &= \|C(sI-A)^{-1}B\|_{H_2}^2 \\
&= \frac{1}{2\pi}\int_0^\infty \text{Trace}((C(\imath\omega I - A)^{-1}B)^*(C(\imath\omega I - A)^{-1}B))d\omega \\
&= \frac{1}{2\pi}\int_0^\infty \text{Trace}((C(\imath\omega I - A)^{-1}B)(C(\imath\omega I - A)^{-1}B)^*)d\omega \\
&= \text{Trace}\int_{-\infty}^\infty C e^{At} B B^* e^{A^* t} C^* dt \\
&= \text{Trace} C X_c C^T
\end{aligned}$$

Thus $X_c \geq 0$ and Trace$C X_c C^T = \|\hat{P}\|_{H_2}^2 < \gamma^2$. □

Proof.

Likewise Trace$B^T X_o B = \|\hat{P}\|_{H_2}^2$ where X_o is the observability Grammian. To show that we can take strict the inequality $X > 0$, we simply let

$$X = \int_0^\infty e^{At}\left(BB^T + \epsilon I\right) e^{A^T} dt$$

for sufficiently small $\epsilon > 0$. Furthermore, we already know the controllability grammian X_c and thus X_ϵ satisfies the Lyapunov inequality.

$$A^T X_\epsilon + X_\epsilon A + BB^T < 0$$

These steps can be reversed to obtain necessity. □

Lets consider the full-state feedback problem

$$\hat{G}(s) = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & 0 & D_{12} \\ I & 0 & 0 \end{array}\right]$$

- D_{12} is the weight on control effort.
- $D_{11} = 0$ is neglected as the feed-through term.
- $C_2 = I$ as this is state-feedback.

$$\hat{K}(s) = \left[\begin{array}{c|c} 0 & 0 \\ \hline 0 & K \end{array}\right]$$

Theorem 48.

The following are equivalent.

1. $\|S(K,P)\|_{H_2} < \gamma$.
2. $K = ZX^{-1}$ *for some* Z *and* $X > 0$ *where*

$$\begin{bmatrix} A & B_2 \end{bmatrix} \begin{bmatrix} X \\ Z \end{bmatrix} + \begin{bmatrix} X & Z^T \end{bmatrix} \begin{bmatrix} A^T \\ B_2^T \end{bmatrix} + B_1 B_1^T < 0$$

$$\textit{Trace}\left[C_1 X + D_{12} Z\right] X^{-1} \left[C_1 X + D_{12} Z\right] < \gamma^2$$

However, this is nonlinear, so we need to reformulate using the Schur Complement.

Theorem 49.

The following are equivalent.

1. $\|S(K,P)\|_{H_2} < \gamma$.
2. $K = ZX^{-1}$ *for some* Z *and* $X > 0$ *where*

$$\begin{bmatrix} A & B_2 \end{bmatrix} \begin{bmatrix} X \\ Z \end{bmatrix} + \begin{bmatrix} X & Z^T \end{bmatrix} \begin{bmatrix} A^T \\ B_2^T \end{bmatrix} + B_1 B_1^T < 0$$

$$\begin{bmatrix} X & (C_1 X + D_{12} Z)^T \\ C_1 X + D_{12} Z & W \end{bmatrix} > 0$$

$$\mathit{Trace} W < \gamma^2$$

Thus we can solve the H_2-optimal static full-state feedback problem.

Applying the Schur Complement gives the alternative formulation convenient for control.

Theorem 50.

Suppose $\hat{P}(s) = C(sI - A)^{-1}B$. Then the following are equivalent.

1. A *is Hurwitz and* $\|\hat{P}\|_{H_2} < \gamma$.
2. *There exists some* $X, Z > 0$ *such that*

$$\begin{bmatrix} A^TX + XA & XB \\ B^TX & -\gamma I \end{bmatrix} < 0, \qquad \begin{bmatrix} X & C^T \\ C & Z \end{bmatrix} > 0, \qquad \textit{Trace} Z < \gamma^2$$

H_2-optimal control

Relationship to LQR

The LQR Problem:

- Full-State Feedback
- Choose K to minimize the cost function

$$\int_0^\infty x(t)^T Q x(t) + u(t)^T R u(t) dt$$

subject to dynamic constraints

$$\dot{x}(t) = Ax(t) + Bu(t)$$
$$u(t) = Kx(t), \qquad x(0) = x_0$$

To solve the LQR problem using H_2 optimal state-feedback control, let

- $C_1 = \begin{bmatrix} Q^{\frac{1}{2}} \\ 0 \end{bmatrix}$
- $D_{12} = \begin{bmatrix} 0 \\ R^{\frac{1}{2}} \end{bmatrix}$
- $B_2 = B$ and $B_1 = I$.

So that

$$\underline{\mathsf{S}}(\hat{P}, \hat{K}) = \left[\begin{array}{c|c} A + B_2 K & B_1 \\ \hline C_1 + D_{12} K & D_{11} \end{array}\right] = \left[\begin{array}{c|c} A + BK & I \\ \hline \begin{matrix} Q^{\frac{1}{2}} \\ R^{\frac{1}{2}} K \end{matrix} & 0 \end{array}\right]$$

And solve the H_2 full-state feedback problem. Then if

$$\begin{aligned} \dot{x}(t) &= A_{CL} x(t) = (A + BK) x(t) = Ax(t) + Bu(t) \\ u(t) &= Kx(t), \qquad x(0) = x_0 \end{aligned}$$

Then $x(t) = e^{A_{CL} t} x_0$

If

$$\dot{x}(t) = A_{CL}x(t) = (A + BK)x(t) = Ax(t) + Bu(t)$$
$$u(t) = Kx(t), \qquad x(0) = x_0$$

then $x(t) = e^{A_{CL}t}x_0$ and

$$\int_0^\infty x(t)^T Q x(t) + u(t)^T R u(t) dt = \int_0^\infty x_0^T e^{A_{CL}^T t}(Q + K^T R K) e^{A_{CL}t} x_0 dt$$
$$= \text{Trace} \int_0^\infty x_0^T e^{A_{CL}^T t} \begin{bmatrix} Q^{\frac{1}{2}} \\ R^{\frac{1}{2}} K \end{bmatrix}^T \begin{bmatrix} Q^{\frac{1}{2}} \\ R^{\frac{1}{2}} K \end{bmatrix} e^{A_{CL}t} x_0 dt$$
$$= \|x_0\|^2 \text{Trace} \int_0^\infty B_1 e^{A_{CL}^T t}(C_1 + D_{12}K)^T (C_1 + D_{12}K) e^{A_{CL}t} B_1^T dt$$
$$= \|x_0\|^2 \|S(K,P)\|_{H_2}^2$$

Thus LQR reduces to a special case of H_2 static state-feedback.

H_2-optimal output feedback control

Theorem 51 (Lall).

The following are equivalent.

- *There exists a* $\hat{K} = \left[\begin{array}{c|c} A_K & B_K \\ \hline C_K & D_K \end{array}\right]$ *such that* $\|S(K,P)\|_{H_2} < \gamma$.
- *There exist* $X_1, Y_1, Z, A_n, B_n, C_n, D_n$ *such that*

$$\begin{bmatrix} AY_1 + Y_1A^T + B_2C_n + C_n^TB_2^T & *^T & *^T \\ A^T + A_n + [B_2D_nC_2]^T & X_1A + A^TX_1 + B_nC_2 + C_2^TB_n^T & *^T \\ [B_1 + B_2D_nD_{21}]^T & [X_1B_1 + B_nD_{21}]^T & -I \end{bmatrix} < 0,$$

$$\begin{bmatrix} Y_1 & I & *^T \\ I & X_1 & *^T \\ C_1Y_1 + D_{12}C_n & C_1 + D_{12}D_nC_2 & Z \end{bmatrix} > 0,$$

$$D_{11} + D_{12}D_nD_{21} = 0, \qquad \mathrm{trace}(Z) < \gamma^2$$

H_2-optimal output feedback control

As before, the controller can be recovered as

$$\left[\begin{array}{c|c} A_{K2} & B_{K2} \\ \hline C_{K2} & D_{K2} \end{array}\right] = \begin{bmatrix} X_2 & X_1 B_2 \\ 0 & I \end{bmatrix}^{-1} \left[\begin{bmatrix} A_n & B_n \\ C_n & D_n \end{bmatrix} - \begin{bmatrix} X_1 A Y_1 & 0 \\ 0 & 0 \end{bmatrix}\right] \begin{bmatrix} Y_2^T & 0 \\ C_2 Y_1 & I \end{bmatrix}^{-1}$$

for any full-rank X_2 and Y_2 such that

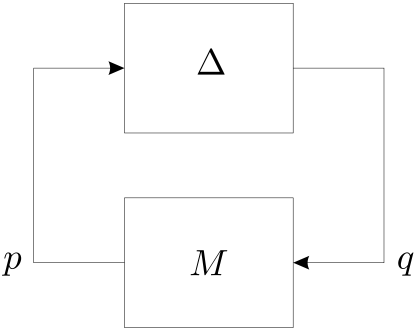
$$\begin{bmatrix} X_1 & X_2 \\ X_2^T & X_3 \end{bmatrix} = \begin{bmatrix} Y_1 & Y_2 \\ Y_2^T & Y_3 \end{bmatrix}^{-1}$$

To find the actual controller, we use the identities:

$$\begin{aligned} D_K &= (I + D_{K2} D_{22})^{-1} D_{K2} \\ B_K &= B_{K2}(I - D_{22} D_K) \\ C_K &= (I - D_K D_{22}) C_{K2} \\ A_K &= A_{K2} - B_K (I - D_{22} D_K)^{-1} D_{22} C_K \end{aligned}$$

Robust Control

Before we finish, let us briefly touch on the use of LMIs in Robust Control.



Questions:

- Is $\underline{\mathsf{S}}(\Delta, M)$ stable for all $\Delta \in \boldsymbol{\Delta}$?
- Determine

$$\sup_{\Delta \in \boldsymbol{\Delta}} \|\underline{\mathsf{S}}(\Delta, M)\|_{H_\infty}.$$

Suppose we have the system M

$$M = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix}$$

Definition 52.

We say the pair $(M, \mathbf{\Delta})$ is **Robustly Stable** if $(I - M_{22}\Delta)$ is invertible for all $\Delta \in \mathbf{\Delta}$.

$$S_l(M, \Delta) = M_{11} + M_{12}\Delta(I - M_{22}\Delta)^{-1}M_{21}$$

Robust Control

The structure of $\mathbf{\Delta}$ makes a lot of difference. e.g.

- **Unstructured, Dynamic, norm-bounded:**

$$\mathbf{\Delta} := \{\Delta \in \mathcal{L}(L_2) \,:\, \|\Delta\|_{H_\infty} < 1\}$$

- **Structured, Dynamic, norm-bounded:**

$$\mathbf{\Delta} := \{\Delta_1, \Delta_2, \cdots \in \mathcal{L}(L_2) \,:\, \|\Delta_i\|_{H_\infty} < 1\}$$

- **Unstructured, Parametric, norm-bounded:**

$$\mathbf{\Delta} := \{\Delta \in \mathbb{R}^{n\times n} \,:\, \|\Delta\| \le 1\}$$

- **Unstructured, Parametric, polytopic:**

$$\mathbf{\Delta} := \{\Delta \in \mathbb{R}^{n\times n} \,:\, \Delta = \sum_i \alpha_i H_i,\, \alpha_i \ge 0,\, \sum_i \alpha_i \le 1\}$$

Let's consider a simple question: *Additive Uncertainty*.

$$M_{11} = 0, \quad M_{12} = M_{21} = I$$

Question: Is $\dot{x} = A(t)x(t)$ stable if $A(t) \in \mathbf{\Delta}$ for all $t \geq 0$.

Definition 53 (Quadratic Stability).

$\dot{x} = A(t)x(t)$ is **Quadratically Stable** for $A(t) \in \mathbf{\Delta}$ if there exists some $P > 0$ such that

$$A^T P + PA < 0 \qquad \text{for all } A \in \mathbf{\Delta}$$

Theorem 54.

If $\dot{x} = A(t)x(t)$ is Quadratically Stable, then it is stable for $A \in \mathbf{\Delta}$.

We examine this problem for:

- **Parametric, Polytopic Uncertainty:**

$$\boldsymbol{\Delta} := \{\Delta \in \mathbb{R}^{n\times n} \,:\, \Delta = \sum_i \alpha_i A_i,\, \alpha_i \geq 0,\, \sum_i \alpha_i = 1\}$$

Parametric, Polytopic Uncertainty

For the polytopic case, we have the following result

Theorem 55 (Quadratic Stability).

Let

$$\mathbf{\Delta} := \{\Delta \in \mathbb{R}^{n\times n} \,:\, \Delta = \sum_i \alpha_i H_i,\, \alpha_i \ge 0,\, \sum_i \alpha_i = 1\}$$

Then $\dot{x}(t) = A(t)x(t)$ is quadratically stable for all $A \in \mathbf{\Delta}$ if and only if there exists some $P > 0$ such that

$$A_i^T P + P A_i < 0 \qquad \textit{for } i = 1, \cdots ,$$

Thus quadratic stability of systems with polytopic uncertainty is equivalent to an LMI.

Parametric, Norm-Bounded Uncertainty

A more complex uncertainty set is:

$$\begin{aligned} \dot{x}(t) &= A_0 x(t) + Mp(t), \qquad & p(t) &= \Delta(t) q(t), \\ q(t) &= Nx(t) + Qp(t), \qquad & \Delta &\in \mathbf{\Delta} \end{aligned}$$

- **Parametric, Norm-Bounded Uncertainty:**

$$\mathbf{\Delta} := \{\Delta \in \mathbb{R}^{n\times n} \,:\, \|\Delta\| \le 1\}$$

Parametric, Norm-Bounded Uncertainty

Quadratic Stability: There exists a $P > 0$ such that

$$P(A_0x(t)+Mp)+(A_0x(t)+Mp)^TP < 0 \text{ for all } p \in \{p : p = \Delta q, q = Nx+Qp\}$$

Theorem 56.

The system

$$\dot{x}(t) = A_0x(t) + Mp(t), \qquad p(t) = \Delta(t)q(t),$$
$$q(t) = Nx(t) + Qp(t), \qquad \Delta \in \boldsymbol{\Delta} := \{\Delta \in \mathbb{R}^{n\times n} : \|\Delta\| \le 1\}$$

is quadratically stable if and only if there exists some $P > 0$ such that

$$\begin{bmatrix} x \\ y \end{bmatrix} \begin{bmatrix} A^TP + PA & PM \\ M^TP & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} < 0$$

$$\text{for all } \begin{bmatrix} x \\ y \end{bmatrix} \in \left\{ \begin{bmatrix} x \\ y \end{bmatrix} : \begin{bmatrix} x \\ y \end{bmatrix} \begin{bmatrix} -N^TN & -N^TQ \\ -Q^TN & I - Q^TQ \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \le 0 \right\}$$

Parametric, Norm-Bounded Uncertainty

If.

If

$$\begin{bmatrix} x \\ y \end{bmatrix} \begin{bmatrix} A^T P + PA & PM \\ M^T P & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} < 0$$

$$\text{for all } \begin{bmatrix} x \\ y \end{bmatrix} \in \left\{ \begin{bmatrix} x \\ y \end{bmatrix} : \begin{bmatrix} x \\ y \end{bmatrix} \begin{bmatrix} -N^T N & -N^T Q \\ -Q^T N & I - Q^T Q \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \leq 0 \right\}$$

then

$$x^T P(Ax + My) + (Ax + My)^T Px < 0$$

for all x, y such that

$$\|Nx + Qy\|^2 \leq \|y\|^2$$

Therefore, since $p = \Delta q$ implies $\|p\| \leq \|q\|$, we have quadratic stability.
The *only if* direction is similar.

□

S-procedure to the rescue!
The S-procedure asks the question:

- Is $z^T F z \geq 0$ for all $z \in \{x : x^T G x \geq 0\}$?

Corollary 57 (S-Procedure).

$z^T F z \geq 0$ for all $z \in \{x : x^T G x \geq 0\}$ if there exists a $\tau \geq 0$ such that $F - \tau G \succeq 0$.

The S-procedure is **Necessary** if $\{x : x^T G x > 0\} \neq \emptyset$.

Parametric, Norm-Bounded Uncertainty

Theorem 58.

The system

$$\dot{x}(t) = Ax(t) + Mp(t), \qquad p(t) = \Delta(t)q(t),$$
$$q(t) = Nx(t) + Qp(t), \qquad \Delta \in \mathbf{\Delta} := \{\Delta \in \mathbb{R}^{n \times n} \,:\, \|\Delta\| \le 1\}$$

is quadratically stable if and only if there exists some $\mu \ge 0$ and $P > 0$ such that

$$\begin{bmatrix} AP + PA^T & PN^T \\ NP & 0 \end{bmatrix} + \mu \begin{bmatrix} MM^T & MQ^T \\ QM^T & QQ^T - I \end{bmatrix} < 0\}$$

These approaches can be readily extended to controller synthesis.

Consider Quadratic Stability in Discrete-Time: $x_{k+1} = S_l(M, \Delta)x_k$.

Definition 59.

$(S_l, \mathbf{\Delta})$ is QS if

$$S_l(M, \Delta)^T P S_l(M, \Delta) - P < 0 \qquad \text{for all } \Delta \in \mathbf{\Delta}$$

Theorem 60 (Packard and Doyle).

Let $M \in \mathbb{R}^{(n+m)\times(n+m)}$ be given with $\rho(M_{11}) \le 1$ and $\sigma(M_{22}) < 1$. Then the following are equivalent.

1. *The pair $(M, \Delta = \mathbb{R}^{m\times m})$ is quadratically stable.*
2. *The pair $(M, \Delta = \mathbb{C}^{m\times m})$ is quadratically stable.*
3. *The pair $(M, \Delta = \mathbb{C}^{m\times m})$ is robustly stable.*

The Structured Singular Value

For the case of structured parametric uncertainty, we define the structured singular value.

$$\mathbf{\Delta} = \{\Delta = \operatorname{diag}(\delta_1 I_{n1}, \cdots, \delta_s I_{ns}, \Delta_{s+1}, \cdots, \Delta_{s+f}) \,:\, \delta_i \in \mathbb{F},\, \Delta \in \mathbb{F}^{n_k \times n_k}\}$$

- δ and Δ represent unknown parameters.
- s is the number of scalar parameters.
- f is the number of matrix parameters.

Definition 61.

Given system $M \in \mathcal{L}(L_2)$ and set $\mathbf{\Delta}$ as above, we define the **Structured Singular Value** of $(M, \mathbf{\Delta})$ as

$$\mu(M, \mathbf{\Delta}) = \frac{1}{\inf_{\substack{\Delta \in \mathbf{\Delta} \\ I - M_{22}\Delta \text{ is singular}}} \|\Delta\|}$$

The Structured Singular Value

Theorem 62.

Let

$$\boldsymbol{\Delta}_n = \{\Delta \in \boldsymbol{\Delta}, \, \|\Delta\| \le \mu(M, \boldsymbol{\Delta})\}.$$

Then the pair $(M, \boldsymbol{\Delta}_n)$ *is robustly stable.*

Conclusion

LMIs are a versatile tool for

- Optimal H_∞ Control
- Optimal H_2 Control (LQR/LQG)
- Robust Control

Next Lecture, we expand the use of LMIs exponentially

1. Nonlinear Systems Theory
2. Sum-of-Squares Nonlinear Stability Analysis

Time permitting, we will explore other applications

1. Stability and Control of Time-Delay Systems
2. Stability and Control of PDE systems.