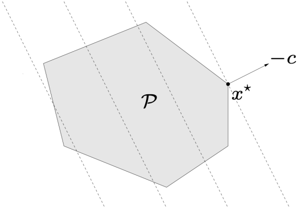
6: Convex Optimization: Linear Matrix Inequalities in Control

- LP recap
- semidefinite programming
  - autonomous system stability
  - nonlinear/robust stability analysis

Linear programming

$$
\begin{aligned}
\text{minimize} \quad & c^T x \\
\text{s.t.} \quad & Gx \preceq h \\
& Ax = b
\end{aligned}
\tag{1}
$$



- feasible set is a **polytope**
- the optimal point, if it exists, is, w.l.o.g at a vertex

[Figure from Boyd & Vandenberghe]

Alternative form of an LP

LPs are sometimes expressed in an equivalent form:

$$\begin{aligned} \text{minimize} \quad & \hat{c}^T \hat{x} \\ \text{s.t.} \quad & \hat{A}\hat{x} = \hat{b} \\ & \hat{x} \succeq 0 \end{aligned} \tag{2}$$

- the set $x \succeq 0$ is called the positive orthant, denoted $\mathbb{R}^k_+$
- the positive orthant is a convex cone
- reformulating the LP in this form introduces additional variables and and constraints
- equivalent means we can recover the solution of one problem from the other

expressing LP (1) in the form of (2) can be done in two stages

1. **inequality constraint elimination:** introduce slack variables to get rid of $Gx \preceq h$

$$Gx \preceq h \quad \Longleftrightarrow \quad Gx + s = h, \quad s \succeq 0$$

expressing LP (1) in the form of (2) can be done in two stages

1. **inequality constraint elimination:** introduce slack variables to get rid of $Gx \preceq h$

$$Gx \preceq h \quad \Longleftrightarrow \quad Gx + s = h, \quad s \succeq 0$$

2. **express decision vector as difference of two +ve variables**

$$x := x^+ - x^- \quad \text{where} \quad x^+, x^- \succeq 0$$

expressing LP (1) in the form of (2) can be done in two stages

1. **inequality constraint elimination:** introduce slack variables to get rid of $Gx \preceq h$

$$Gx \preceq h \quad \Longleftrightarrow \quad Gx + s = h, \quad s \succeq 0$$

2. **express decision vector as difference of two +ve variables**

$$x := x^+ - x^- \quad \text{where} \quad x^+, x^- \succeq 0$$

the resulting problem is an LP in the form of (2) with decision variables in red

$$\begin{array}{ll} \text{minimize} & c^T x^+ - c^T x^- \\ \text{s.t.} & Gx^+ - Gx^- + s = h \\ & Ax^+ - Ax^- = b \\ & x^+ \succeq 0, \quad x^- \succeq 0, \quad s \succeq 0 \end{array}$$

Semidefinite programming

most general form of a convex optimization problem is a *semidefinite program* (SDP)

an SDP consists of

- cost function: linear
- equality constraints: affine equality constraints
- inequality constraints: linear matrix inequalities (LMIs)

Primal form SDP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & F_0 + x_1 F_1 + x_2 F_2 + \ldots x_n F_n \succeq 0 \\ & Ax = b \end{array} \tag{3}$$

Problem data: $F_i, \in \mathbb{S}^k$, $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, and $c \in \mathbb{R}^n$

Linear matrix inequalities

matrix valued constraints of the form

$$\underbrace{F_0 + x_1 F_1 + x_2 F_2 + \ldots x_n F_n}_{F(x)} \succeq 0$$

where

- decision vector x takes values from in $\mathbb{R}^n$
- the operator $F : \mathbb{R}^n \to \mathbb{S}^k$ is affine
- $F_0, F_1, \ldots, F_n \in \mathbb{S}^k$

Note

- looks abstract and not "user friendly"
- very general formulation, the most expressive convex constraint that exists
- most control problems we formulate in this class will be defined by LMIs

Symmetric matrices

the set of **symmetric** $n \times n$ matrices

$$\mathbb{S}^n = \{X \in \mathbb{R}^{n \times n} \mid X = X^T\}$$

is a subspace of $\mathbb{R}^{n \times n}$

the matrix $W \in \mathbb{S}^n$ is said to be **positive semidefinite** if, for all $x \in \mathbb{R}^n$,

$$x^T W x \geq 0$$

- $W \succeq 0$ denotes that W is psd
- $W \in \mathbb{S}^n$ is **positive definite** if, for all $x \neq 0$, $x^T W x > 0$
- interpret $X \succeq Y$ as $X - Y \succeq 0$

Theorem

The set of $n \times n$ **positive semidefinite matrices** *is a convex cone.*

Eigenvalues of symmetric matrices

Theorem

The eigenvalues of $A \in \mathbb{S}^n$ are real.

Proof: By definition

$$Ax = \lambda x \quad \text{for } x \neq 0.$$

Left multiply by x^* to get

$$\lambda x^* x = x^* A x \quad \Rightarrow \quad \lambda = \frac{x^* A x}{x^* x}.$$

Take conjugate of λ:

$$\lambda^* = \frac{x^* A^* x}{x^* x} = \frac{x^* A x}{x^* x},$$

and we conclude $\lambda = \lambda^*$. $\square$

Eigenvectors and diagonalization

Theorem

If $A = A^T \in \mathbb{R}^{n \times n}$, then there exist n mutually orthogonal eigenvectors $\{u_1, \ldots, u_n\}$ that satisfy

$$Au_i = \lambda_i u_i \quad u_i^T u_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases},$$

In matrix form, there exists a U such that $U^T U = I$ and

$$U^{-1} A U = U^T A U = \Lambda.$$

Testing for positive semidefiniteness

recall, $W \succeq 0 \iff x^T W x \geq 0$ for all $x \neq 0$

- all symmetric matrices admit an **orthogonal** decomposition

$$W = U \Lambda U^T, \quad \text{where} \quad \Lambda = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

- U orthogonal means $U^T U = U U^T = I$
- $\langle u_i, u_j \rangle = 0$ when $i \neq j$, $\langle u_i, u_i \rangle = 1$
- define the new coordinates $z = U^T x$ (and so $x = Uz$)
- it follows that

$$x^T W x = \sum_{i=1}^{n} \lambda_i z_i^2,$$

and so testing if $W \succeq 0$ reduces to checking the sign of **all** the eigenvalues

Partial ordering

for scalars $x, y \in \mathbb{R}$, the query $x \overset{?}{\geq} y \mapsto \{\text{true}, \text{false}\}$.

- the inequality defines a **complete ordering**

for $X, Y \in \mathbb{S}^n$, evaluating X "*greater than or equal to*" Y is more involved

$$X \succeq Y \quad \Longleftrightarrow \quad X - Y \succeq 0$$

- $\succeq$ defines a *partial ordering*
- $X \not\succeq 0$ does not imply $X \succeq 0$

more generally, a convex cone $\mathcal{C}$ generates a **partial ordering**

$$X - Y \preceq 0 \quad \text{if} \quad X - Y \in \mathcal{C}$$

SDP convexity

Theorem

The set $\{x \in \mathbb{R}^n \mid F(x) \succeq 0\}$ *is a convex set.*

- multiple LMIs can be combined into a single constraint

$$\begin{array}{c} F_0 + x_1 F_1 + x_2 F_2 + \dots x_n F_n \succeq 0 \\ \hat{F}_0 + x_1 \hat{F}_1 + x_2 \hat{F}_2 + \dots x_n \hat{F}_n \succeq \hat{0} \end{array}$$

  is equivalent to

$$\begin{bmatrix} F_0 & \\ & \hat{F}_0 \end{bmatrix} + x_1 \begin{bmatrix} F_1 & \\ & \hat{F}_1 \end{bmatrix} + x_2 \begin{bmatrix} F_2 & \\ & \hat{F}_2 \end{bmatrix} + \cdots + x_n \begin{bmatrix} F_n & \\ & \hat{F}_n \end{bmatrix} \succeq 0$$

LMI example

- recall the definition on an LMI: $F_0 + x_1F_1 + x_2F_2 + \ldots x_nF_n \succeq 0$

  the LMI constraint

$$\begin{bmatrix} 4x_2 & 2x_1 + x_2 & 5 \\ 2x_1 + x_2 & x_2 & 2x_2 + 8 \\ 5 & 2x_2 + 8 & -3x_1 - 6 \end{bmatrix} \succeq 0$$

LMI example

- recall the definition on an LMI: $F_0 + x_1 F_1 + x_2 F_2 + \ldots x_n F_n \succeq 0$

  the LMI constraint

$$\begin{bmatrix} 4x_2 & 2x_1 + x_2 & 5 \\ 2x_1 + x_2 & x_2 & 2x_2 + 8 \\ 5 & 2x_2 + 8 & -3x_1 - 6 \end{bmatrix} \succeq 0$$

  in standard form $F(x)$, is

$$\begin{bmatrix} 0 & 0 & 5 \\ 0 & 0 & 8 \\ 5 & 8 & -6 \end{bmatrix} + x_1 \begin{bmatrix} 0 & 2 & 0 \\ 2 & 0 & 0 \\ 0 & 0 & -3 \end{bmatrix} + x_2 \begin{bmatrix} 4 & 1 & 0 \\ 1 & 1 & 2 \\ 0 & 2 & 0 \end{bmatrix}$$

Alternative form form SDP

in the same way that there were different, but equivalent ways of expressing an LP: (1) and (2), the same is true for SDPs

Dual form SDP

$$\begin{array}{rl} \text{minimize} & \textbf{trace}\ CX \\ \text{s.t.} & \textbf{trace}\ A_i X = b_i, \quad i = 1, \ldots, m \\ & X \succeq 0 \end{array} \tag{4}$$

where the problem data is $C, A_1, \ldots A_m, \in \mathbb{S}^n$ and $X \in \mathbb{S}^n$ is the decision variable

- **trace** XY defines an inner product over the space of symmetric matrices
- the feasible set is the intersection of a convex cone and and an affine set
- generalizes LP from positive orthant to the psd cone
- rich theory that links problems (3) and (4), we mostly not use it

Autonomous system stability LMI

- the autonomous linear system $\dot{x} = Ax$ is stable iff

$$A^T P + PA + Q = 0$$

  returns $P \succ 0$ when $Q \succ 0$

- $V(z) = z^T Pz$ is a Lyapunov function if

$$z^T Pz > 0 \quad \text{and} \quad z^T (A^T P + PA)z < 0 \quad \text{for all } z \tag{5}$$

Autonomous system stability LMI

- the autonomous linear system $\dot{x} = Ax$ is stable iff

$$A^T P + PA + Q = 0$$

  returns $P \succ 0$ when $Q \succ 0$

- $V(z) = z^T P z$ is a Lyapunov function if

$$z^T P z > 0 \quad \text{and} \quad z^T(A^T P + PA)z < 0 \quad \text{for all } z \tag{5}$$

- the inequalities (5) are equivalent to the existence of a matrix P where

$$P \succ 0, \quad A^T P + PA \prec 0 \tag{6}$$

  an LMI!

- the LMIs in (6) can be expressed in the standard form $\sum x_i F_i \preceq G$

Theorem

The autonomous linear system $\dot{x} = Ax$ *is stable iff*

$$\begin{bmatrix} -P & 0 \\ 0 & A^T P + PA \end{bmatrix} \prec 0$$

has a feasible solution.

Notes

- we cannot directly enforce $X \succ 0$, if this is needed (as above), we implement

  $$X - \epsilon I \succeq 0$$

  where ϵ is a small positive constant

- if $F(x)$ is linear in x, then $F(x) \succ 0 \iff F(x) \succeq I$

Another Lyapunov LMI

if the system $\dot{x} = Ax$ is stable, then there exists an $\alpha > 0$, such that

$$V(x) = x^T P x > 0, \ (P \succ 0) \quad \text{and} \quad \dot{V} \leq -\alpha V \tag{7}$$

- Specifically, if P and $Q \succ 0$ solve the Lyapunov equation, then

$$\alpha = \frac{\lambda_{\min}(Q)}{\lambda_{\max}(P)}$$

- follows from the fact that for any $W \succeq 0$

$$\lambda_{\min}(W)\|z\|^2 \leq z^T W z \leq \lambda_{\max}(W)\|z\|^2$$

- condition (7) is only an LMI if α (or a bound) is known
- expressing (7) as an LMI

$$P \succ 0 \quad A^T P + PA + \alpha P \prec 0$$

S-procedure

it is often the case that enforcing a matrix inequality on all x is too restrictive, i.e., we only require

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}^T \underbrace{\begin{bmatrix} A & B \\ C & D \end{bmatrix}}_{F} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \geq 0 \quad \text{for } x \in \mathcal{D} \tag{8}$$

- in general, this is NP-hard, e.g., $\mathcal{D} = \mathbb{R}^n_+$
- however, if $\mathcal{D}$ can be represented by a quadratic inequality, we can derive necessary and sufficient conditions for (8)
- a specific example of $\mathcal{D}$ is the set of x_1, x_2 such that

  $$x_1^T x_1 \geq x_2^T H^T H x_2$$

  which can be expressed as

  $$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}^T \underbrace{\begin{bmatrix} I & 0 \\ 0 & -H^T H \end{bmatrix}}_{G} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \geq 0$$

S-procedure

it is often the case that enforcing a matrix inequality on all x is too restrictive, i.e., we only require

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}^T \underbrace{\begin{bmatrix} A & B \\ C & D \end{bmatrix}}_{F} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \geq 0 \quad \text{for } x \in \mathcal{D}$$

an equivalent way of viewing this is the implication:

$$\text{when does } x \in \mathcal{D} \implies x^T F x \geq 0?$$

S-procedure

the S-Procedure formalizes this for quadratics described by symmetric matrices

$$\text{is} \quad \underbrace{z^T F z \geq 0}_{(\dagger)} \quad \text{for all} \quad z \in \underbrace{\{x \mid x^T G x \geq 0\}}_{\mathcal{D}}$$

S-procedure

the S-Procedure formalizes this for quadratics described by symmetric matrices

$$\text{is} \quad \underbrace{z^T F z \geq 0}_{(\dagger)} \quad \text{for all} \quad z \in \underbrace{\{x \mid x^T G x \geq 0\}}_{\mathcal{D}}$$

Theorem

1. *If there exists a $\tau \in \mathbb{R}_+$ such that $F \succeq \tau G$ then*

$$z^T F z \geq 0 \quad \textit{for all} \quad z \in \{x \mid x^T G x \geq 0\}.$$

2. *Equivalently, for all z such that $z^T F z \geq 0 \implies z^T G z \geq 0$ if there exists a $\tau \geq 0$ such that $F \succeq \tau G$.*

The converse of (1) and (2) holds if there exists exists a vector w such that $w^T G w > 0$.

S-procedure

the S-Procedure formalizes this for quadratics described by symmetric matrices

$$\text{is} \quad \underbrace{z^T F z \geq 0}_{(\dagger)} \quad \text{for all} \quad z \in \underbrace{\{x \mid x^T G x \geq 0\}}_{\mathcal{D}}$$

Theorem

1. *If there exists a $\tau \in \mathbb{R}_+$ such that $F \succeq \tau G$ then*

$$z^T F z \geq 0 \quad \textit{for all} \quad z \in \{x \mid x^T G x \geq 0\}.$$

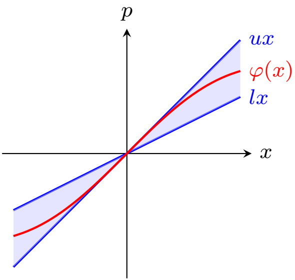
2. *Equivalently, for all z such that $z^T F z \geq 0 \implies z^T G z \geq 0$ if there exists a $\tau \geq 0$ such that $F \succeq \tau G$.*

The converse of (1) and (2) holds if there exists exists a vector w such that $w^T G w > 0$.

- we will not prove this result
- inequality $(\dagger)$ can be made strict if $F \succ \tau G$
- this is often referred to as the **lossless S-procedure**

Robustness analysis

$\phi : \mathbb{R} \to \mathbb{R}$ is said to be $[l, u]$-sector bounded if for all $x \in \mathbb{R}$, $p = \phi(x) \in [lx, ux]$



can express as a quadratic inequality

$$(p - ux)(p - lx) \leq 0 \quad \text{for all } x, p = \phi(x)$$

Special cases:

- $[0, \infty] \iff \mathbf{sign}(p) = \mathbf{sign}(x)$
- $[-1, 1] \iff |p| < |x|$

Lur'e system

consider the **nonlinear** system

$$\dot{x} = Ax + Bp, \quad q = Cx, \quad p = \phi(t, q) \tag{9}$$

Problem Specification

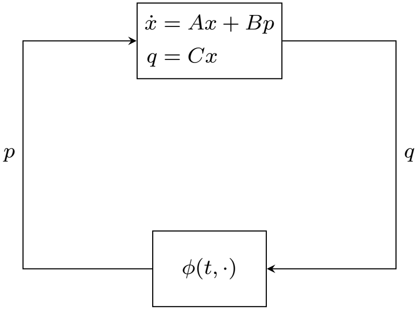
- $\phi(t, \cdot)$ is $[l, u]$-sector bounded for all t
- ϕ is unknown, but $[l, u]$ is given
- all other problem data is known
- for simplicity, we consider scalar outputs, i.e., $q(t) \in \mathbb{R}$

Is this system stable for all $[l, u]$-sector bounded functions ϕ?

Closed-loop representation

system (9) is a linear dynamical system with a static nonlinear feedback term

$$\dot{x} = Ax + B\phi(t, Cx)$$



- can generalize to vector-valued outputs

Global asymptotic stability analysis

our goal is to determine if the Lur'e system (9) is globally asymptotically stable via the Lyapunov function $V(x) = x^T P x$

1. we need to show that $V(x) \geq 0$ and $\dot{V}(x) \leq -\alpha V(x)$ for some given $\alpha > 0$

   $\dot{V}(x) = \dot{x}^T P x + x^T P \dot{x} + \alpha x^T P x \leq 0$ for all x and $[l, u]$-sector bounded functions

2. write out the derivative

$$\dot{V}(x) = (Ax + B\phi(t, Cx))^T P x + x^T P (Ax + B\phi(t, Cx)) + \alpha x^T P x \leq 0$$

3. using $p = \phi(x)$, derivative condition becomes

$$(Ax + Bp)^T P + P(Ax + Bp) + \alpha x^T P x \leq 0$$

   for all x and p that satisfy $(p - uq)(p - lq) \leq 0$ where $q = Cx$

5. the inequality $(p-uq)(p-lq) \leq 0$ can be written as the matrix inequality

$$\left[\begin{array}{c} x \\ p \end{array}\right]^T \left[\begin{array}{cc} \sigma C^T C & -\nu C^T \\ -\nu C & 1 \end{array}\right] \left[\begin{array}{c} x \\ p \end{array}\right] \leq 0$$

where $\sigma = lu$ and $\nu = \frac{l+u}{2}$

6. in matrix inequality form, we need

$$\left[\begin{array}{c} x \\ p \end{array}\right]^T \left[\begin{array}{cc} A^T P + PA + \alpha P & PB \\ B^T P & 0 \end{array}\right] \left[\begin{array}{c} x \\ p \end{array}\right] \leq 0$$

whenever

$$\left[\begin{array}{c} x \\ p \end{array}\right]^T \left[\begin{array}{cc} \sigma C^T C & -\nu C^T \\ -\nu C & 1 \end{array}\right] \left[\begin{array}{c} x \\ p \end{array}\right] \leq 0$$

7. apply the S-procedure

$$\left[\begin{array}{cc} A^T P + PA + \alpha P & PB \\ B^T P & 0 \end{array}\right] \preceq \tau \left[\begin{array}{cc} \sigma C^T C & -\nu C^T \\ -\nu C & 1 \end{array}\right]$$

Lur'e stability summary

we conclude that Lur'e system

$$\dot{x} = Ax + Bp, \quad q = Cx, \quad p = \phi(t, q)$$

with ϕ an $[l, u]$-sector bounded function is stable, if

$$\begin{bmatrix} A^T P + PA + \alpha P - \tau\sigma C^T C & PB + \tau\nu C^T \\ B^T P + \tau\nu C & -\tau \end{bmatrix} \preceq 0, \quad P \succeq I, \; \tau \geq 0$$

- this is a robust result; it guarantees stability for *all* functions in the sector bound
- only requires the nonlinearity to be covered by choosing u and l
- the larger the sector, the more functions are covered, but feasibility becomes less likely