DISC Course on Linear Matrix Inequalities in Control

Siep Weiland
Department of Electrical Engineering
Eindhoven University of Technology

Course 2004 - Class 4

Robustness: an example

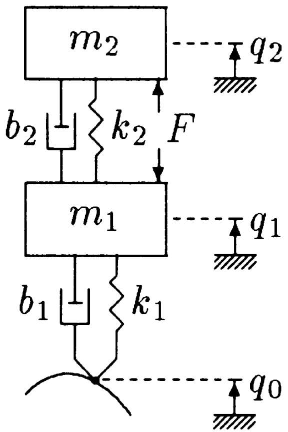
1: leaf-spring
2: shock absorber
3: stabilizer bar
4: rear axle

Aim:

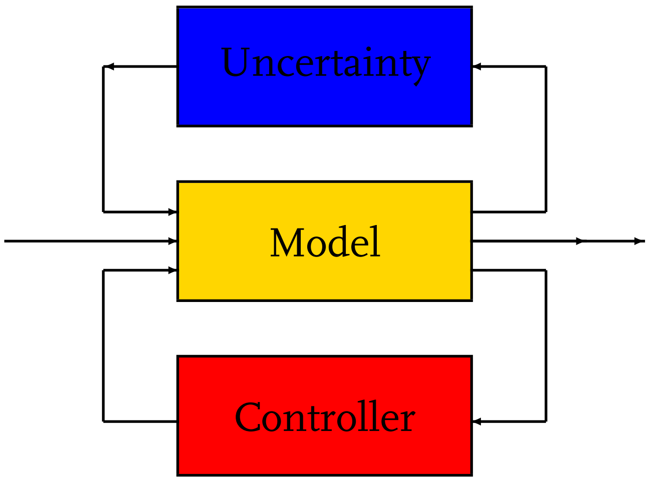
- isolate oscillations and vibrations
- balance comfort vs. road grip
- *passive suspension*: spring-damper specs defined by manufacturer
- *active suspension*: controlled reduction of undesirable deflections

Vehicle suspension

Mass-spring-damper system

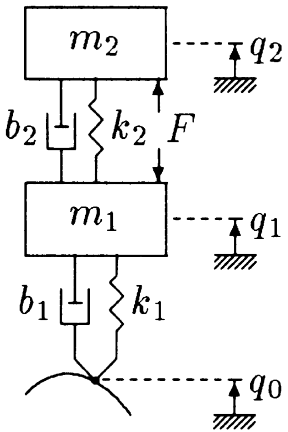


Control configuration



Vehicle suspension

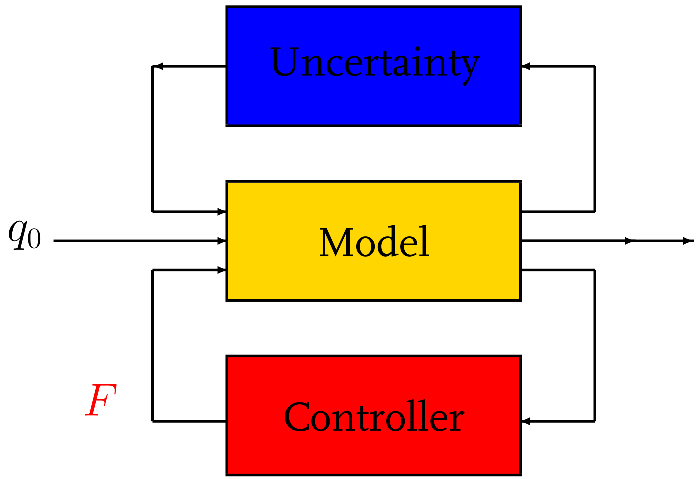
Mass-spring-damper system



Inputs:

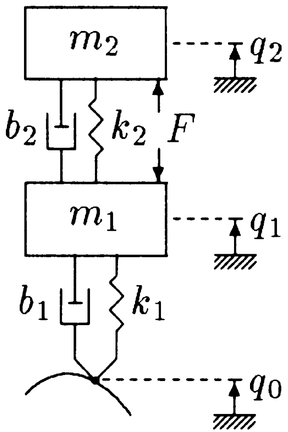
(q_0, F)

Control configuration

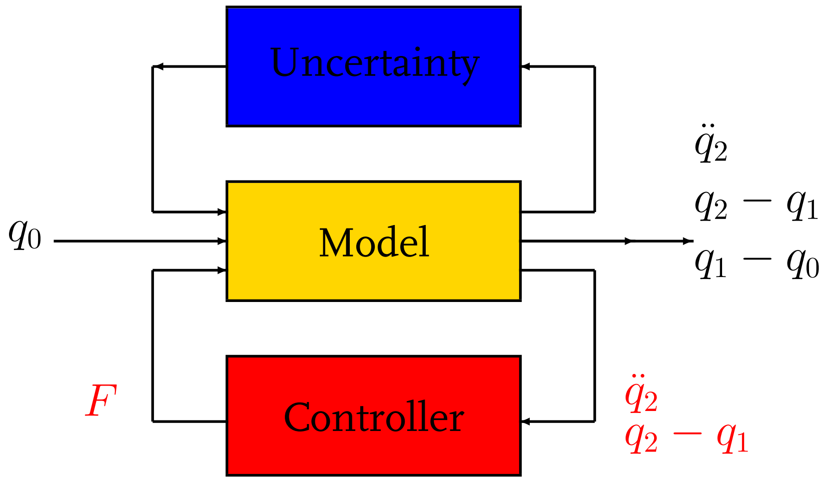


Vehicle suspension

Mass-spring-damper system



Control configuration

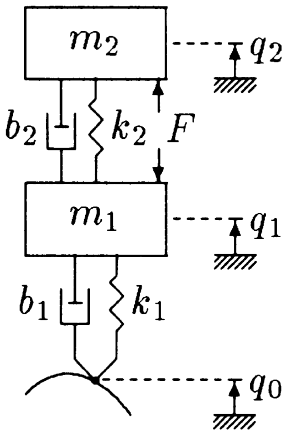


Inputs: **Outputs:**

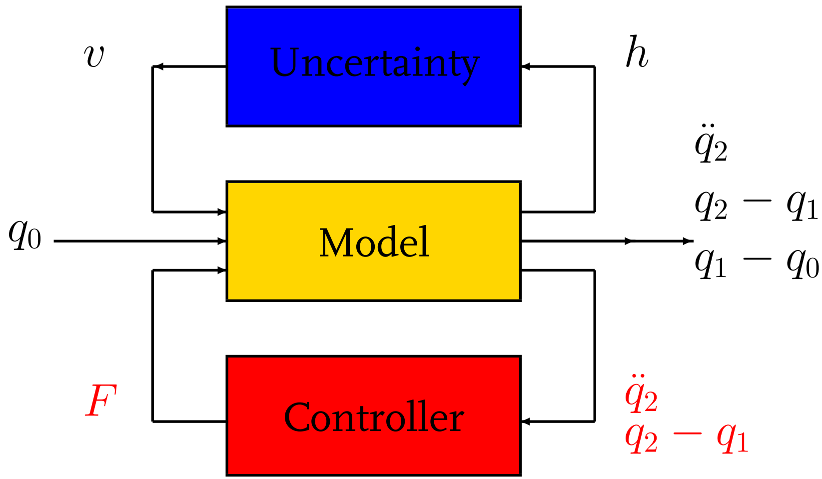
$$(q_0, F) \quad \mapsto \quad (\ddot{q}_2, q_2 - q_1, q_1 - q_0, \ddot{q}_2, q_2 - q_1)$$

Vehicle suspension

Mass-spring-damper system



Control configuration



Inputs: **Outputs:**

$$(q_0, F) \quad \mapsto \quad (\ddot{q}_2, q_2 - q_1, q_1 - q_0, \ddot{q}_2, q_2 - q_1)$$

Uncertainty: in load m_2:

$v = \delta h$ with δ in some uncertainty set Δ.

The model

$$\begin{pmatrix} q_0 \\ F \end{pmatrix} \longrightarrow \boxed{\text{Plant}} \longrightarrow \begin{pmatrix} \ddot{q}_2 \\ q_2 - q_1 \\ q_1 - q_0 \end{pmatrix}$$

Described by **differential equations**

$$\begin{aligned} 0 &= m_2\ddot{q}_2 + b_2(\dot{q}_2 - \dot{q}_1) + k_2(q_2 - q_1) - F \\ 0 &= m_1\ddot{q}_1 + b_2(\dot{q}_1 - \dot{q}_2) + k_2(q_1 - q_2) + k_1(q_1 - q_0) + b_1(\dot{q}_1 - \dot{q}_0) + F. \end{aligned}$$

Described by **state space equations**

$$\dot{x} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -(k_1+k_2)/m_1 & k_2/m_1 & -(b_1+b_2)/m_1 & b_2/m_1 \\ k_2/m_2 & -k_2/m_2 & b_2/m_2 & -b_2/m_2 \end{pmatrix} x + \begin{pmatrix} b_1/m_1 & 0 \\ 0 & 0 \\ (-b_1^2-b_1b_2)/m_1^2+k_1/m_1 & -1/m_1 \\ b_1b_2/m_1m_2 & 1/m_2 \end{pmatrix} w$$

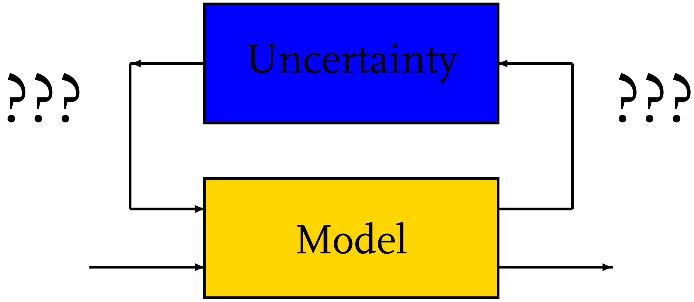
$$z = \begin{pmatrix} k_2/m_2 & -k_2/m_2 & b_2/m_2 & -b_2/m_2 \\ -1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} x + \begin{pmatrix} b_1b_2/m_1m_2 & 1/m_2 \\ 0 & 0 \\ -1 & 0 \end{pmatrix} w$$

Physical specifications

	m_1	m_2	k_1	k_2	b_1	b_2
unloaded	1.5×10^3	1.5×10^3	5.0×10^6	5.0×10^5	1.5×10^{-3}	50×10^3
loaded	1.5×10^3	1.0×10^4	5.0×10^6	5.0×10^5	1.5×10^{-3}	50×10^3

Variations in mass m_2

Can this be written in the form:



Uncertainty: $\mathrm{col}(v_1, v_2) = \delta\, \mathrm{col}(x, u)$ with $\delta = \frac{1}{m_2} \in [\frac{1}{m_2^{\max}}, \frac{1}{m_2^{\min}}]$.

See simulation

Time-invariant parametric uncertainty

Consider linear time-invariant system

$$\dot{x}(t) = A(\delta)x(t)$$

where $A(\cdot)$ is a **continuous** function of the parameter vector

$$\delta = \mathrm{col}(\delta_1, \cdots, \delta_p)$$

which is known to be contained in the **uncertainty set**

$$\Delta \subset \mathbb{R}^p$$

Robust stability analysis
Is system asymptotically stable for all possible $\delta \in \Delta$?

Example

Consider system with rational parameter dependence

$$\dot{x}(t) = \begin{pmatrix} -1 & 2\delta_1 & 2 \\ \delta_2 & -2 & 1 \\ 3 & -1 & \frac{\delta_3 - 10}{\delta_1 + 1} \end{pmatrix} x(t)$$

where parameters δ_1, δ_2, δ_3 are bounded as

$$\delta_1 \in [-0.5, 1], \qquad \delta_2 \in [-2, 1], \qquad \delta_3 \in [-0.5, 2].$$

Hence δ belongs to a **polytopic** uncertainty set

$$\Delta = [-0.5, 1] \times [-2, 1] \times [-0.5, 2] =$$
$$= \mathrm{co}\left\{ \begin{pmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \end{pmatrix} : \delta_1 \in \{-0.5, 1\}, \quad \delta_2 \in \{-2, 1\}, \quad \delta_3 \in \{-0.5, 2\} \right\}$$

Relation to optimization

The **spectral abscissa** of square matrix A is $\rho_a(A) = \max_{\lambda \in \lambda(A)} \frac{1}{2}(\lambda + \bar{\lambda})$.

$A(\delta)$ is Hurwitz for all $\delta \in \Delta$ **if and only if**

$$\rho_a(A(\delta)) < 0 \quad \text{for all } \delta \in \Delta$$

Two main sources for trouble:

- spectral abscissa $\rho_a(A(\delta))$ is not convex/concave in δ.
- inequality has to hold at infinitely many points.

Consequences:
Computational approaches therefore fail.

$\longrightarrow$ Cannot find global maximum of $\rho_a(A(\delta))$ over Δ.
$\longrightarrow$ Even if δ is a polytope, not sufficient to check its generators.
$\longrightarrow$ Even more trouble if Δ is not a polytope.

Quadratic stability

The uncertain system $\dot{x} = A(\delta)x$ with $\delta \in \Delta$ is defined to be **quadratically stable** if there exists $X \succ 0$ with

$$A(\delta)^\top X + XA(\delta) \prec 0 \quad \text{for all} \quad \delta \in \Delta$$

Why is this relevant?
Function $V(x) := x^\top X x$ serves as quadratic Lyapunov function.

Worthwhile to understand details:
With above X there exists $\varepsilon > 0$ such that

$$A(\delta)^\top X + XA(\delta) + \varepsilon X \preceq 0.$$

Then, abbreviating $V(x(t))$ as $V(t)$ we get for all $t \in \mathbb{R}$ and all $\delta \in \Delta$

$$\begin{aligned} \dot{V}(t) + \varepsilon V(t) &= x(t)^\top \left[A(\delta)^\top X + XA(\delta) \right] x(t) + \varepsilon x^\top(t) X x(t) \\ &\leq 0 \end{aligned}$$

Integrating over $[t_0, t_1]$, this yields that V has exponential decay

$$V(t_1) \quad \leq \quad V(t_0)e^{-\varepsilon(t_1-t_0)} \qquad \text{for all } \delta \in \Delta$$

Now use that

$$\lambda_{\min}(X)\|x\|^2 \leq x^\top X x \leq \lambda_{\max}(X)\|x\|^2$$

and infer that $\|x(t)\|$ has exponential decay

$$\boxed{\|x(t)\|^2 \quad \leq \quad \|x_0\|^2 \frac{\lambda_{\max}(X)}{\lambda_{\min}(X)} e^{-\varepsilon t} \text{ for all } \delta \in \Delta}$$

Conclude that $\lim_{t\to\infty} x(t) = 0$

- with exponential decay rate ε.
- irrespective of initial condition x_0.
- for all $\delta \in \Delta$

How to verify quadratic stability?

Theorem: If $A(\delta)$ is affine in δ and the uncertainty set $\Delta = \mathrm{co}(\delta^1, \cdots, \delta^N)$ is a polytope then $\dot{x} = A(\delta)x$ is quadratically stable **if and only if** there exists $X \succ 0$ such that

$$A(\delta^k)^\top X + XA(\delta^k) \prec 0 \quad \text{for } k = 1, \ldots, N$$

Comments:

- Converts verification of quadratic stability to feasibility problem in **finite set of LMI's**
- Routine `quadstab` in LMI toolbox
- implies that $A(\delta)$ is Hurwitz for all $\delta \in \Delta$.

Proof:
Nice application of convexity of the function
$f_x(\delta) := x^\top[A(\delta)^\top X + XA(\delta)]x$.

Proof

Let $x \in \mathbb{R}^n$ be arbitrary and define $f_x : \Delta \to \mathbb{R}$ by

$$f_x(\delta) := x^\top [A(\delta)^\top X + XA(\delta)]x$$

Observation: A affine implies f_x convex.

But then

$$A(\delta)^\top X + XA(\delta) \prec 0 \quad \text{for all } \delta \in \Delta$$

if and only if (by definition) for all $x \in \mathbb{R}^n$

$$f_x(\delta) < 0 \quad \text{for all } \delta \in \Delta$$

if and only if (by slide 24 class 1) for all $x \in \mathbb{R}^n$

$$f_x(\delta) < 0 \quad \text{for all } \delta \in \Delta_0$$

if and only if

$$A(\delta)^\top X + XA(\delta) \prec 0 \quad \text{for all } \delta \in \Delta$$

Example

If $A(\delta)$ is not affine in δ, a parameter transformation often helps!

In example, let $\delta_4 = \frac{\delta_3 - 10}{\delta_1 + 1} + 12$. Then

$$A(\delta) = \begin{pmatrix} -1 & 2\delta_1 & 2 \\ \delta_2 & -2 & 1 \\ 3 & -1 & \delta_4 - 12 \end{pmatrix}, \quad \begin{pmatrix} \delta_1 \\ \delta_2 \\ \delta_4 \end{pmatrix} \in \Delta = [-0.5, 1] \times [-2, 1] \times [-9, 8]$$

LMI toolbox `quadstab` yields system quadratically stable for

$$(\delta_1, \delta_2, \delta_4) \in r\Delta \quad \text{with largest possible scaling factor } r \approx 0.45$$

Call number r^* the **quadratic stability margin** if system is quadratically stable for uncertainty set $r\Delta$ with $r < r^*$, and not quadratically stable for $r\Delta$ with $r > r^*$.

An application in state-feedback robust control

Consider uncertain control system

$$\dot{x} = A(\delta)x + B(\delta)u, \qquad \delta \in \Delta$$

Problem: Find state feedback $u = Fx$ such that the controlled system

$$\dot{x}(t) = (A(\delta) + B(\delta)F)x(t)$$

is quadratically stable for all $\delta \in \Delta$.

An application in state-feedback robust control

Consider uncertain control system

$$\dot{x} = A(\delta)x + B(\delta)u, \qquad \delta \in \Delta$$

Problem: Find state feedback $u = Fx$ such that the controlled system

$$\dot{x}(t) = (A(\delta) + B(\delta)F)x(t)$$

is quadratically stable for all $\delta \in \Delta$.

Equivalently: find F and $X \succ 0$ such that

$$(A(\delta) + B(\delta)F)^\top X + X(A(\delta) + B(\delta)F) \prec 0 \quad \text{for all} \quad \delta \in \Delta$$

An application in state-feedback robust control

Consider uncertain control system

$$\dot{x} = A(\delta)x + B(\delta)u, \qquad \delta \in \Delta$$

Problem: Find state feedback $u = Fx$ such that the controlled system

$$\dot{x}(t) = (A(\delta) + B(\delta)F)x(t)$$

is quadratically stable for all $\delta \in \Delta$.

Equivalently: find K and $Y \succ 0$ such that

$$A(\delta)Y + YA(\delta)^\top + B(\delta)K + (B(\delta)K)^\top \prec 0 \quad \text{for all} \quad \delta \in \Delta.$$

Transformation: $Y = X^{-1}$ and $K = FX^{-1}$.

An application in state-feedback robust control

Consider uncertain control system

$$\dot{x} = A(\delta)x + B(\delta)u, \qquad \delta \in \Delta$$

Problem: Find state feedback $u = Fx$ such that the controlled system

$$\dot{x}(t) = (A(\delta) + B(\delta)F)x(t)$$

is quadratically stable for all $\delta \in \Delta$.

Equivalently: find K and $Y \succ 0$ such that

$$A(\delta)Y + YA(\delta)^\top + B(\delta)K + (B(\delta)K)^\top \prec 0 \quad \text{for all} \quad \delta \in \Delta.$$

Then set $F = KY^{-1}$.

- $A(\delta)$ and $B(\delta)$ affine, Δ finitely generated: **LMI feasibility problem!**
- Won't work for output feedback!!

Time-varying parametric uncertainties

Now assume that the parameters $\delta(t)$ vary in time, and that they are known to satisfy $\delta(t) \in \Delta$ for all t. Check stability of the system

$$\dot{x}(t) = A(\delta(t))x(t), \qquad \delta(t) \in \Delta$$

The uncertain system with time-varying parametric uncertainties is **quadratically stable** if there exists $X \succ 0$ with

$$A(\delta)^\top X + XA(\delta) \prec 0 \quad \text{for all } \delta \in \Delta.$$

Proof will be given for more general result in full detail.

Quadratic stability therefore implies roust stability for **arbitrary fast** time-varying parametric uncertainties in Δ. If bounds on velocities $\dot{\delta}$ are known, this test is **conservative**.

Rate-bounded parametric uncertainties

Let us assume that the parameter curves $\delta(\cdot)$ are continuously differentiable and are only known to satisfy

$$\delta(t) \in \Delta \quad \text{and} \quad \dot{\delta}(t) \in \Lambda \quad \text{for all time.}$$

Here, $\Delta \in \mathbb{R}^p$ and $\Lambda \in \mathbb{R}^p$ are given **compact** sets (e.g., polytopes).

Problem: robust stability analysis
Verify whether the linear time-varying system

$$\dot{x}(t) = A(\delta(t))x(t)$$

is exponentially stable for all parameter curves $\delta(\cdot)$ that satisfy the above bounds on value and velocity.

Amounts searching for Lyapunov function.

Main stability result

Theorem: Suppose $X(\delta)$ is continuously differentiable on Δ and satisfies

$$X(\delta) \succ 0, \qquad \sum_{k=1}^{p} \partial_k X(\delta) v_k + A(\delta)^\top X(\delta) + X(\delta) A(\delta) \prec 0$$

for all $\delta \in \Delta$ and $v \in \Lambda$.

Then there exist constants $M > 0$, $a > 0$ such that all trajectories of the uncertain time-varying system satisfy

$$\|x(t)\| \leq M e^{-a(t-t_0)} \|x(t_0)\| \quad \text{for all } t \geq t_0.$$

- Covers **many** tests in the literature.
- Condition for robust stability is **sufficient only**!
- Is also necessary in case $\Lambda = \{0\}$: time-invariant uncertainty.

Proof

Same idea as before!! Exists $\alpha, \beta, \gamma > 0$ such that for all $\delta \in \Delta$, $v \in \Lambda$:

$$\alpha I \preceq X(\delta) \preceq \beta I, \quad \sum_{k=1}^{p} \partial_k X(\delta) v_k + A(\delta)^\top X(\delta) + X(\delta) A(\delta) \prec -\gamma I$$

Suppose that $\delta(t)$ is admissible parameter curve and let $x(t)$ be corresponding state trajectory. **Here's the crux**:

$$\frac{d}{dt} x(t)^\top X(\delta(t)) = x(t)^\top \left[\sum_{k=1}^{p} \partial_k X(\delta(t)) \dot{\delta}_k(t) \right] x(t) + \\ + x(t)^\top \left[A(\delta(t))^\top X(\delta(t)) + X(\delta(t)) A(\delta(t)) \right] x(t)$$

Since $\delta(t) \in \Delta$ and $\dot{\delta}(t) \in \Lambda$ we conclude that

$$\alpha \|x(t)\|^2 \leq x(t)^\top X(\delta(t)) x(t) \leq \beta \|x(t)\|^2, \quad \frac{d}{dt} x(t)^\top X(\delta(t)) x(t) \leq -\gamma \|x(t$$

Proof (ctd.)

This is similar to what we have seen before.

Define $V(t) := x(t)^\top X(\delta(t))x(t)$ to infer that

$$\|x(t)\|^2 \leq \frac{1}{\alpha}V(t), \quad V(t) \leq \beta\|x(t)\|^2, \quad \dot{V}(t) \leq -\frac{\gamma}{\beta}V(t)$$

So that $V(t)$ has exponential decay

$$V(t) \leq V(t_0)e^{-\frac{\gamma}{\beta}(t-t_0)} \quad \text{for all } \ t \geq t_0.$$

This yields that $\|x(t)\|$ has exponential decay

$$\|x(t)\|^2 \leq \frac{\beta}{\alpha}e^{-\frac{\gamma}{\beta}(t-t_0)}\|x(t_0)\|^2 \quad \text{for all } \ t \geq t_0.$$

so that we can choose $M = \sqrt{\frac{\beta}{\alpha}}$ and $a = \frac{\gamma}{2\beta}$.

Extreme cases

- Parameters are **time-invariant**: $\Lambda = \{0\}$.

  Have to find $X(\delta)$ satisfying

  $$X(\delta) \succ 0, \quad A(\delta)^\top X(\delta) + X(\delta)A(\delta) \prec 0$$

- Parameters vary **arbitrary fast**:

  Have to find parameter-independent X satisfying

  $$X \succ 0, \quad A(\delta)^\top X + XA(\delta) \prec 0 \quad \text{for all } \delta \in \Delta$$

  is identical to **quadratic stability** test!

  Idea of proof: If inequality holds for $\Lambda = [-r, r]^p$ and all $r > 0$ then $\partial_k X(\delta)$ must vanish for all $k = 1, \ldots, p$ and all $\delta \in \Delta$

Remarks

- We have derived general results based on Lyapunov functions which still depend **quadratically** on the state (restrictive!) but which allow for non-linear (smooth) dependence on the uncertain parameters.

- Tests are purely algebraic and do not involve system- or parameter trajectories

- Not easy to apply:
  - ◇ Have to find **function** satisfying partial differential LMI
  - ◇ Have to make sure that inequality holds for all $\delta \in \Delta$, $\dot{\delta} \in \Lambda$.

- Allows to easily derive specialization which are or can be implemented with LMI solvers.

Example: affine system-affine Lyapunov function

Suppose $A(\delta)$ depend affinely on parameters:

$$A(\delta) = A_0 + \delta_1 A_1 + \cdots + \delta_p A_p$$

Parameter and rate constraints are boxes:

$$\Delta = \{\delta \in \mathbb{R}^p \mid \delta_k \in [\underline{\delta}_k, \bar{\delta}_k]\}, \quad \Lambda = \{v \in \mathbb{R}^p \mid v_k \in [\underline{v}_k, \bar{v}_k]\},$$

These are the convex hulls of

$$\Delta_0 = \{\delta \in \mathbb{R}^p \mid \delta_k \in \{\underline{\delta}_k, \bar{\delta}_k\}\}, \quad \Lambda_0 = \{v \in \mathbb{R}^p \mid v_k \in \{\underline{v}_k, \bar{v}_k\}\},$$

Search for affine parameter dependent $X(\delta)$:

$$X(\delta) = X_0 + \delta_1 X_1 + \cdots + \delta_p X_p$$

Hence,

$$\partial_k X(\delta) = X_k.$$

Example: affine system-affine Lyapunov function

With $\delta_0 = 1$ we find

$$\sum_{k=1}^{p} \partial_k X(\delta) v_k + A(\delta)^\top X(\delta) + X(\delta) A(\delta) =$$

$$= \sum_{k=1}^{p} X_k v_k + \sum_{\nu=0}^{p} \sum_{\mu=0}^{p} \delta_\nu \delta_\mu (A_\nu X_\mu + X_\mu A_\nu).$$

- **affine** in $X_1, \ldots, X_p$
- **affine** in $v_1, \ldots, v_p$
- **quadratic** in $\delta_1, \ldots, \delta_p$ (mixture of constant, linear and quadratic terms).

Consequently: The function

$$f_x(\delta, v) := x^\top \left[\sum_{k=1}^{p} X_k v_k + \sum_{\nu=0}^{p} \sum_{\mu=0}^{p} \delta_\nu \delta_\mu (A_\nu X_\mu + X_\mu A_\nu) \right] x$$

may **not be convex** in $\delta \in \Delta$, $v \in \Lambda$.

Example: affine system-affine Lyapunov function

Main issue: Can we provide conditions such that

$$f_x(\delta, v) < 0 \quad \text{on} \quad \Delta \times \Lambda$$

is implied by

$$f_x(\delta, v) < 0 \quad \text{on} \quad \Delta_0 \times \Lambda_0$$

That is, can we cook up a **generator test**?

- Implies that it suffices to guarantee required inequality at generators.

Relaxation: Include additional constraint $A_\nu^\top X_\nu + X_\nu A_\nu \succeq 0$.

Why? Sufficient condition for generator test is that

$$f_x(\delta, v) = f_x(\delta_1, \ldots, \delta_j, \ldots, \delta_p, v_1, \ldots, v_j, \ldots, v_p)$$

is **partially convex**, i.e., convex in each **individual argument** δ_j and v_j. That is if $\frac{\partial^2 f}{\partial \delta_j^2} \geq 0$ and $\frac{\partial^2 f}{\partial v_j^2} \geq 0$.

Example: affine system-affine Lyapunov function

Robust exponential stability is therefore guaranteed if

There exist $X_0, \ldots, X_p$ with $A_\nu^\top X_\nu + X_\nu A_\nu \succeq 0, nu = 1, \ldots, p$, and

$$\sum_{k=0}^{p} X_k \delta_k \succ 0, \quad \sum_{k=1}^{p} X_k v_k + \sum_{\nu=0}^{p} \sum_{\mu=0}^{p} \delta_\nu \delta_\mu (A_\nu X_\mu + X_\mu A_\nu) \prec 0$$

for all $\delta \in \Delta_0$ and $v \in \Lambda_0$ and with $\delta_0 = 1$.

- A finite test!!
- This test is implemented in the LMI toolbox in `pdstab`.

  For rate-bounded uncertainties often much less conservative than quadratic stability test.
- Need to understand the arguments in the proof to derive your own variants.

General recipe to reduce to finite dimensions

Restrict the search to a chosen finite dimensional subspace.

For example choose scalar continuously differentiable basis functions $b_1(\delta), \ldots, b_N(\delta)$ and search for the coefficient matrices $X_1, \ldots, X_N$ in the expansion

$$X(\delta) = \sum_{\nu=1}^{N} X_\nu b_\nu(\delta) \quad \text{with} \quad \partial_k X(\delta) = \sum_{\nu=1}^{N} X_\nu \partial_k b_\nu(\delta)$$

Have to guarantee that for all $\delta \in \Delta$ and $v \in \Lambda$:

$$\sum_{\nu=1}^{N} X_\nu b_\nu(\delta) \succ 0, \sum_{\nu=1}^{N} \left(\sum_{k=1}^{p} X_\nu \partial_k b_\nu(\delta) v_k + [A(\delta)^\top X_\nu + X_\nu A(\delta)] b_\nu(\delta) \right) \prec 0$$

Is finite dimensional but still semi-infinite LMI problem

Remarks

- If systematically extending the set of basis functions one can improve the sufficient stability conditions. Example: **Polynomial basis**

$$b_{k_1,\ldots,k_p}(\delta) = \delta_1^{\,k_1} \cdots \delta_p^{\,k_p}, \quad k_\nu = 0, 1, 2, \ldots, v = 1, \ldots, p.$$

  If δ is star-shaped one can prove:
  if partial diff. LMI has a continuously differentiable solution, then it also has a polynomial solution (of possibly higher degree).

  Polynomial basis is a generic choice with guaranteed success.

- One can also grid Δ and Λ to arrive at **finite** system of LMI's.

  Trouble: Huge LMI system, no guarantees at points outside grid.

Generalized stability regions

So far, quadratic stability meant **exponential decay** of the state trajectory in the sense that

$$\|x(t)\| \leq M\|x(t_0)\|e^{\alpha(t-t_0)}$$

for suitable gain $M > 0$ and decay rate $\alpha < 0$.

Can we influence the exponential decay rate?

Alternative stability regions:

Damping:	$\mathrm{Re}(s) < \alpha$	$s + \bar{s} - 2\alpha < 0$
Circle:	$\|s\| < r$	$\begin{pmatrix} -r & -s \\ -\bar{s} & -r \end{pmatrix} \prec 0$
Strip:	$\alpha_1 < \mathrm{Re}(s) < \alpha_2$	$s + \bar{s} - 2\alpha_2 < 0$ $0 < s + \mathrm{Re}(s) - 2\alpha_1$
Conic:	$\tan(\theta)\,\mathrm{Re}(s) < -\|\mathrm{Im}(s)\|$	$\begin{pmatrix} (s+\bar{s})\sin\theta & (s-\bar{s})\cos\theta \\ (\bar{s}-s)\cos\theta & (s+\bar{s})\sin\theta \end{pmatrix} \prec 0$

Alle these are examples of **LMI regions**

For a real symmetric $2m \times 2m$ matrix P the set of complex numbers

$$L_P := \left\{ s \in \mathbb{C} \mid \begin{pmatrix} I \\ sI \end{pmatrix}^* P \begin{pmatrix} I \\ sI \end{pmatrix} \prec 0 \right\}$$

is called an **LMI region**.

Notation
Define the **Kronecker product**

$$A \otimes B = \begin{pmatrix} A_{11}B & \dots & A_{1n}B \\ \vdots & & \vdots \\ A_{m1}B & \dots & A_{mn}B \end{pmatrix}$$

- $1 \otimes A = A = A \otimes 1$
- $(A + B) \otimes C = (A \otimes C) + (B \otimes C)$

Generalization standard stability criterion

Theorem: All eigenvalues of $A \in \mathbb{R}^{n \times n}$ are contained in the LMI region

$$\left\{ s \in \mathbb{C} \mid \begin{pmatrix} I \\ sI \end{pmatrix}^* \begin{pmatrix} Q & S \\ S^\top & R \end{pmatrix} \begin{pmatrix} I \\ sI \end{pmatrix} \prec 0 \right\}$$

if and only if

there exists $K \succ 0$ such that

$$\begin{pmatrix} I \\ A \otimes I \end{pmatrix}^* \begin{pmatrix} K \otimes Q & K \otimes S \\ K \otimes S^\top & K \otimes R \end{pmatrix} \begin{pmatrix} I \\ A \otimes I \end{pmatrix} \prec 0$$

- An LMI characterization!
- Very nice generalization of usual stability test.
- Also applicable as test for intersections of various stability domains.

Duality in convex programming

Let $\mathcal{S}$ be a subset of a linear vector space $\mathcal{X}$ and let mappings $f : \mathcal{S} \to \mathbb{R}$, $g_i : \mathcal{S} \to \mathbb{R}$, $h_i \mathcal{X} \to \mathbb{R}$ be given.

Many optimization problems involve equality and inequality constraints.

$$\begin{aligned} g_i(x) &\leq 0, \qquad && i = 1, \ldots, k \\ h_i(x) &= 0, \qquad && i = 1, \ldots, \ell. \end{aligned}$$

Let $g = \operatorname{col}(g_1, \ldots, g_k)$ and $h = \operatorname{col}(h_1, \ldots, h_\ell)$.

Problem: Consider primal optimization problem with optimal value P_{opt}:

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & x \in \mathcal{S}, \quad g(x) \leq 0, \quad h(x) = 0 \end{array}$$

Is **convex** if $\mathcal{S}$, f, g_i are convex and h_i affine.

Convex programs

- Examples: saturation constraints, safety margins, physically meaningful variables, constitutive and balance equations assume the form $\mathcal{S}$.
- **linear program:**

$$f(x) = c^\top x, \quad g(x) = g_0 + Gx, \quad h(x) = h_0 + Hx$$

- **quadratic program:**

$$f(x) = x^\top Qx, \quad g(x) = g_0 + Gx, \quad h(x) = h_0 + Hx$$

- **quadratically constraint quadratic program:**

$$f(x) = x^\top Qx + 2s^\top x + r, \quad g_j(x) = x^\top Q_j x + 2s_j^\top x + r_j, \quad h(x) = h_0$$

Upper bound on optimal value

If $x_0 \in \mathcal{S}$ satisfies $g(x_0) \leq 0$ and $h(x_0) = 0$ then $P_{\text{opt}} \leq f(x_0)$ defines an **upper bound** on P_{opt}.

Lower bound on optimal value

Let $x \in \mathcal{S}$ satisfy $g(x) \leq 0$ and $h(x) = 0$.
Then for arbitrary $y \geq 0$ and z we have

$$\langle y, g(x) \rangle \;\leq\; 0, \qquad \langle z, h(x) \rangle = 0$$

and, in particular,

$$\inf_{x \in \mathcal{S}} f(x) + \langle y, g(x) \rangle + \langle z, h(x) \rangle \;\leq\; \inf_{x \in \mathcal{S},\, g(x) \leq 0,\, h(x) = 0} f(x) + \langle y, g(x) \rangle + \langle z, h(x) \rangle \;\leq\; \inf_{x \in \mathcal{S},\, g(x) \leq 0,\, h(x) = 0} f(x).$$

The best **lower bound** is obtained by maximization over $y \geq 0$ and z:

$$\inf_{x \in \mathcal{S}, g(x) \leq 0, h(x) = 0} f(x) \geq \sup_{y \geq 0, z} \left[\inf_{x \in \mathcal{S}} f(x) + \langle y, g(x) \rangle + \langle z, h(x) \rangle \right].$$

Terminology:

- **Lagrange function:** $L(x,y,z) = f(x) + \langle y, g(x)\rangle + \langle z, h(x)\rangle$.
- **Lagrange dual cost:** $\ell(y,z) = \inf_{x\in\mathcal{S}} L(x,y,z) \in [-\infty, \infty]$.
- **Lagrange dual optimization problem:**

$$\boxed{D_{\text{opt}} := \sup_{y\geq 0,\, z} \ell(y,z)}$$

Remarks:

- $\ell(y,z)$ computed by solving an **unconstrained** optimization problem.
- $\ell(y,z)$ is a concave function.
- Dual problem is concave **maximization** problem.
  Constraints are usual simpler than in primal problem
- **Weak duality:** $D_{\text{opt}} \leq P_{\text{opt}}$. Main question:

$$\boxed{\text{when is } D_{\text{opt}} = P_{\text{opt}}\text{?}}$$

Lagrange duality theorem

Theorem: Weak duality always true: $P_{\text{opt}} \geq \ell(y, z)$ for all $y \geq 0$, z.

If primal optimization problem is convex and satisfies the **constraint qualification**,

i.e. there exist x_0 in the interior of $\mathcal{S}$ with $g(x_0) \leq 0$, $h(x_0) = 0$ such that $g_j(x_0) < 0$ for all component functions g_j that are not affine.

Then **strong duality** holds: exists $y_{\text{opt}} \geq 0$, z_{opt} such that

$$P_{\text{opt}} = \ell(y_{\text{opt}}, z_{\text{opt}}) = D_{\text{opt}}$$

In other words, strong duality means

$$\inf_{x \in \mathcal{S}, g(x) \leq 0, h(x) = 0} f(x) = \max_{y \geq 0, z} \left[\inf_{x \in \mathcal{S}} f(x) + \langle y, g(x) \rangle + \langle z, h(x) \rangle \right].$$

For example, constraint qualification holds if g is affine.

Karush-Kuhn-Tucker

Sufficiency: Suppose there exists $y_{\text{opt}} \geq 0$, z_{opt} such that $x_0 \in \mathcal{S}$ satisfies $g(x_0) \leq 0$ and $h(x_0) = 0$ and is an optimal solution to the **unconstrained problem**:

$$\boxed{\inf_{x \in \mathcal{S}} f(x) + \langle y_{\text{opt}}, g(x) \rangle + \langle z_{\text{opt}}, h(x) \rangle}$$

with **complementary slackness** condition

$$\langle y_{\text{opt}}, g(x_0) \rangle = 0 \tag{3}$$

Then x_0 is an **optimal solution of the primal problem.**

Neccessity: Suppose x_0 is an optimal solution of the primal problem.

If the primal is convex and satisfies the constraint qualification, then there exist $y_{\text{opt}} \geq 0$ z_{opt} such that x_0 is a solution of the **unconstrained problem** and satisfies (3)

Karush-Kuhn-Tucker and duality

Main result in convex optimization

Theorem: (Karush-Kuhn-Tucker) If $P_{\text{opt}} > -\infty$ and the primal problem satisfies the constraint qualification, then

$$\boxed{D_{\text{opt}} = P_{\text{opt}}}$$

and there exist $y_{\text{opt}} \geq 0$ and z_{opt}, such that

$$D_{\text{opt}} = \ell(y_{\text{opt}}, z_{\text{opt}}).$$

Moreover, x_{opt} is an optimal solution of the primal optimization problem and $(y_{\text{opt}}, z_{\text{opt}})$ is an optimal solution of the dual optimization problem, **if and only if**

1. $g(x_{\text{opt}}) \leq 0$, $h(x_{\text{opt}}) = 0$,
2. $y_{\text{opt}} \geq 0$ and x_{opt} minimizes $L(x, y_{\text{opt}}, z_{\text{opt}})$ over all $x \in \mathcal{X}$ and
3. $\langle y_{\text{opt}}, g(x_{\text{opt}}) \rangle = 0$.

Some comments on KKT theorem

- Very general result, strong tool in convex optimization
- Dual problem simpler to solve, $(y_{\text{opt}}, z_{\text{opt}})$ called **Kuhn Tucker point**.
- The triple $(x_{\text{opt}}, y_{\text{opt}}, z_{\text{opt}})$ exists if and only if it defines a **saddle point** of the LAgrangian L in the sense that

$$L(x_{\text{opt}}, y, z) \leq \underbrace{L(x_{\text{opt}}, y_{\text{opt}}, z_{\text{opt}})}_{=D_{\text{opt}}=P_{\text{opt}}} \leq L(x, y_{\text{opt}}, z_{\text{opt}})$$

for all x, $y \geq 0$ and z.

Next class