Introduction to Modern Control Systems
Convex Optimization & Linear Matrix Inequalities

Kostas Margellos

University of Oxford

Convex Optimization

- Optimization programs
- Convex sets
- Convex functions
- Operations that preserve convexity
- Convex optimization programs

Linear Matrix Inequalities (LMIs)

- How do they look like?
- Are they convex?
- Why are they interesting

References

Convex Optimization & Duality Theory:

Boyd & Vandenberghe (2004)
Convex Optimization, *Cambridge University Press*.

Bertsekas (2009)
Convex Optimization Theory, *Athena Scientific*.

Rockafellar (1970)
Convex Analysis, *Princeton, NJ: Princeton University Press*.

Linear Matrix Inequalities (LMIs):

Boyd, El Ghaoui, Feron & Balakrishnan (1994)
Linear Matrix Inequalities in System and Control Theory, *SIAM*.

VanAntwerp & Braatz (2000)
A tutorial on linear and bilinear matrix inequalities, *J. Process Control*.

Optimization program - General description

A more common problem format:

$$\begin{aligned} \min_{x \in \mathcal{X}} \quad & f_0(x) \\ \text{subject to:} \quad & f_i(x) \leq 0 \quad i = 1, \dots, m \\ & h_i(x) = 0 \quad i = 1, \dots, p \end{aligned}$$

- **Objective function** $f_0 : \mathcal{X} \to \mathbb{R}$
- **Domain** $\mathcal{X} \subseteq \mathbb{R}^n$ of the objective function, from which the decision variable $x := (x_1; x_2; \dots; x_n)$ must be chosen.
- **Inequality constraint functions** $f_i : \mathbb{R}^n \to \mathbb{R}$, for $i = 1, \dots, m$
- **Equality constraint functions** $h_i : \mathbb{R}^n \to \mathbb{R}$, for $i = 1, \dots, p$

$\Rightarrow$ *Maximization* fit the framework with a change of sign.

Optimization program – Possible outcomes

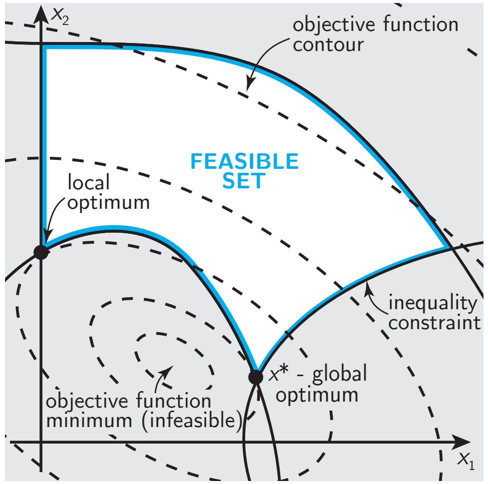
Consider the problem

$$p^* = \min_{x \in \mathcal{X}} f(x)$$

- If $p^* = -\infty$, then the problem is **unbounded below**.
- If the set $\mathcal{X}$ is empty, then the problem is **infeasible** (and we set $p^* = +\infty$).
- If $\mathcal{X} = \mathbb{R}^n$, the problem is **unconstrained**.
- There might be more than one solution. The set of solutions is:

$$\arg\min_{x \in \mathcal{X}} f(x) := \{x \in \mathcal{X} \mid f(x) = p^*\}$$

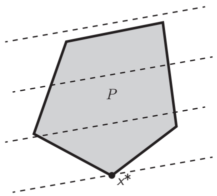
Geometric view



Under convexity it is easier ...

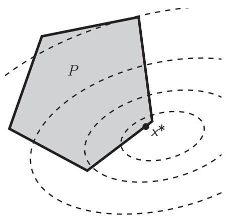
Linear Program (LP):

$$\begin{aligned} \min_x \quad & c^\top x \\ \text{subject to:} \quad & Gx \leq h \\ & Ax = b \end{aligned}$$



Convex Quadratic Program (QP) – $P \succeq 0$:

$$\begin{aligned} \min_x \quad & \tfrac{1}{2}x^\top P x + q^\top x \\ \text{subject to:} \quad & Gx \leq h \\ & Ax = b \end{aligned}$$



$\Rightarrow$ *Convex programs:* Local optimum = Global optimum

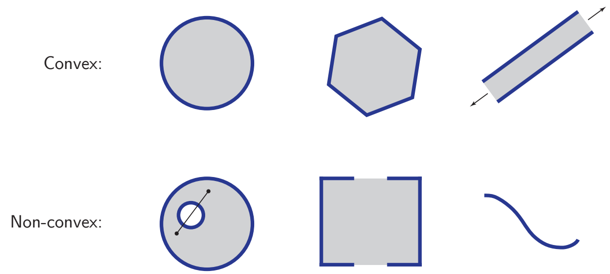
Convex sets

Definition (Convex Set)

A set $\mathcal{X}$ is convex *if and only if* for any pair of points x and y in $\mathcal{X}$, any **convex combination** of x and y lies in $\mathcal{X}$:

$$\mathcal{X} \text{ is convex} \quad \Leftrightarrow \quad \lambda x + (1-\lambda)y \in \mathcal{X},\ \forall \lambda \in [0,1],\ \forall x, y \in \mathcal{X}$$

Interpretation: All line segments starting and ending in $\mathcal{X}$ stay within $\mathcal{X}$.



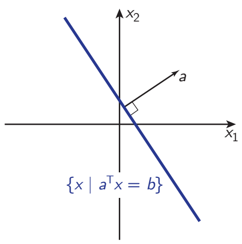
Convex sets

Definitions (Hyperplanes and halfspaces)

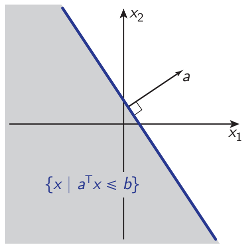
A hyperplane is defined by $\{x \in \mathbb{R}^n \mid a^\top x = b\}$ for $a \neq 0$, where $a \in \mathbb{R}^n$ is the normal vector to the hyperplane.

A halfspace is defined by $\{x \in \mathbb{R}^n \mid a^\top x \leq b\}$ for $a \neq 0$. It can either be **open** (strict inequality) or **closed** (non-strict inequality).

For $n = 2$, hyperplanes define lines. For $n = 3$, hyperplanes define planes.



A hyperplane



A closed halfspace

Convex sets

Definitions (Polyhedra and polytopes)

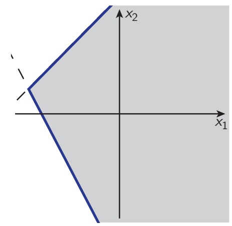
A polyhedron is the intersection of a *finite* number of closed halfspaces:

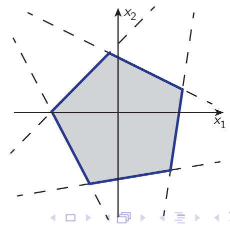
$$\mathcal{X} = \{x \mid a_1^\top x \leq b_1,\ a_2^\top x \leq b_2, \ldots, a_m^\top \leq b_m\} = \{x \mid Ax \leq b\}$$

where $A := [a_1, a_2, \ldots, a_m]^\top$ and $b := [b_1, b_2, \ldots, b_m]^\top$.

A polytope is a *bounded* polyhedron.

Polyhedra and polytopes are always convex.





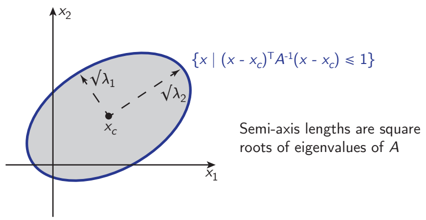
Ellipsoid - Generalized norm ball

Definition (Ellipsoid)

An ellipsoid is a set defined as

$$\mathcal{E} = \{x \mid (x - x_c)^\top A^{-1}(x - x_c) \leq 1\},$$

where x_c is the centre of the ellipsoid, and $A \succ 0$.



Alternatively, $\mathcal{E} = \{x \mid T(x) \leq 0\}$ where

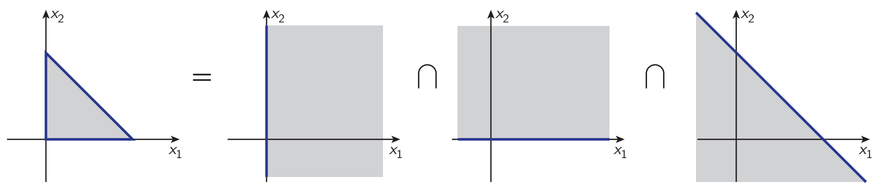
$$T(x) = x^\top Ax + 2x^\top b + c, \text{ with } A = A^\top \succ 0.$$

Intersection of convex sets

Theorem

The intersection of two or more convex sets is itself convex.

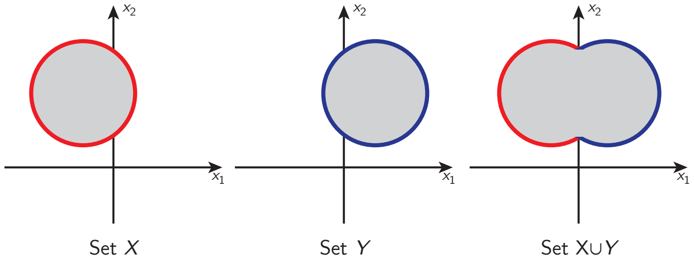
Proof (for two sets): Consider any two points a and b which *both* lie in *both* of two convex sets $\mathcal{X}$ and $\mathcal{Y}$. For any $\lambda \in [0, 1]$, $\lambda a + (1 - \lambda)b$ is in both $\mathcal{X}$ and $\mathcal{Y}$. Therefore $\lambda a + (1 - \lambda)b \in \mathcal{X} \cap \mathcal{Y}$, $\forall \lambda \in [0, 1]$. This satisfies the definition of convexity for set $\mathcal{X} \cap \mathcal{Y}$.



Think of simultaneous constraint satisfaction.

Union of convex sets

Note that the **union** of two sets is **not** convex in general, regardless of whether the original sets were convex!



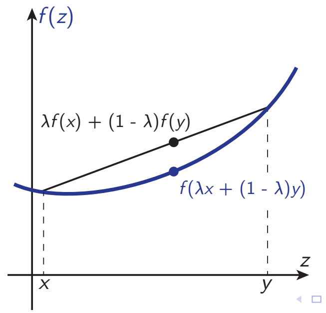
Convex functions

Definitions (Convex function)

A function $f : \mathrm{dom}(f) \to \mathbb{R}$ is convex *if and only if* its domain $\mathrm{dom}(f)$ is convex *and*

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y), \quad \forall \lambda \in (0,1), \quad \forall x, y \in \mathrm{dom}(f)$$

The function f is strictly convex if this inequality is strict.



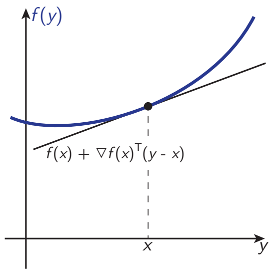
Convex functions – 1st-order condition

A differentiable function $f : \mathrm{dom}(f) \to \mathbb{R}$ with a convex domain is **convex** *if and only if*

$$f(y) \geq f(x) + \nabla f(x)^\top (y - x), \quad \forall x, y \in \mathrm{dom}(f)$$

i.e. a first order approximator of f around any point x is a global underestimator of f.

The gradient is given by $\nabla f(x) = \left[\frac{\partial f(x)}{\partial x_1}, \frac{\partial f(x)}{\partial x_2}, \dots, \frac{\partial f(x)}{\partial x_n}\right]^\top$



Convex functions – 2nd-order condition

A twice-differentiable function $f : \mathrm{dom}(f) \to \mathbb{R}$ is **convex** *if and only if* its domain $\mathrm{dom}(f)$ is convex *and*

$$\nabla^2 f(x) \succeq 0, \quad \forall x \in \mathrm{dom}(f),$$

where the Hessian $\nabla^2 f(x)$ is defined by

$$\nabla^2 f(x)_{ij} = \frac{\partial^2 f(x)}{\partial x_i \partial x_j}$$

If $\mathrm{dom}(f)$ is convex *and* $\nabla^2 f(x) \succ 0$ for all $x \in \mathrm{dom}(f)$, then f is **strictly convex**.

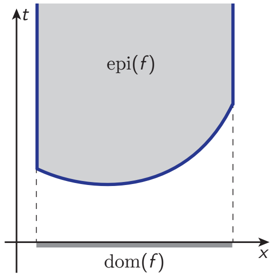
Convex functions – Epigraph

The **epigraph** of a function $f : \mathrm{dom}(f) \to \mathbb{R}$ is the **set**

$$\mathrm{epi}(f) = \left\{ \begin{bmatrix} x \\ t \end{bmatrix} \middle| \; x \in \mathrm{dom}(f),\ f(x) \le t \right\} \subseteq \mathrm{dom}(f) \times \mathbb{R}$$

It has dimension one higher than the domain of f.

A function is convex if and only if its epigraph is a convex set.



Operations that preserve convexity

Theorem (Non-negative weighted sum)

If f is a function convex, then αf is convex for $\alpha \ge 0$. For several convex functions f_i, $\sum_i \alpha_i f_i$ is convex if all $\alpha_i \ge 0$.

Theorem (Composition with affine function)

If f is a convex function, then $f(Ax + b)$ is convex.

Example: $||Ax - b||$ is convex for any norm; Exponential functions.

Theorem (Pointwise maximum)

If $f_1, \dots, f_m$ are convex functions, then $f(x) = \max\{f_1(x), \dots, f_m(x)\}$ is convex.

Example: Piecewise linear functions $\max_{i=1,\dots,m}\{a_i^\top x + b\}$ are convex.

Convex optimization program – standard form

A standard form **convex** optimization problem:

$$\begin{aligned} \min_{x \in \mathcal{X}} \quad & f_0(x) \\ \text{subject to:} \quad & f_i(x) \le 0 \quad i = 1, \dots, m \\ & a_i^\top x = b_i \quad i = 1, \dots, p \end{aligned}$$

This problem is convex if:

- The domain $\mathcal{X}$ is a convex set.
- The objective function f_0 is a convex function.
- The inequality constraint functions f_i are all convex.
- The equality constraint functions $h_i(x) = a_i^\top x$ are all affine.

Convex optimization program – standard form

A standard form **convex** optimization problem:

$$\begin{aligned} \min_{x \in \mathcal{X}} \quad & f_0(x) \\ \text{subject to:} \quad & f_i(x) \le 0 \quad i = 1, \dots, m \\ & Ax = b \quad A \in \mathbb{R}^{p \times m} \end{aligned}$$

This problem is convex if:

- The domain $\mathcal{X}$ is a convex set.
- The objective function f_0 is a convex function.
- The inequality constraint functions f_i are all convex.
- The equality constraint functions $h_i(x) = a_i^\top x$ are all affine.

Convex programs: Local optimum = Global optimum

Theorem

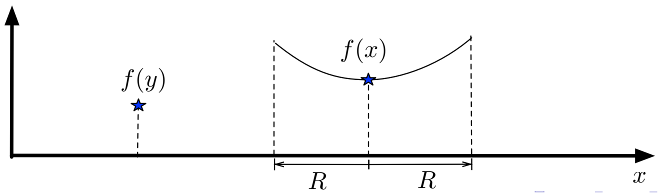
For a convex optimization problem, **every** *locally optimal solution is globally optimal.*

Proof:

- Assume that x is locally optimal, but not globally optimal.
- Therefore there is some other point y such that $f(y) < f(x)$.
- x locally optimal implies that there is some $R > 0$ such that

$$\|z - x\|_2 \le R \Rightarrow f(x) \le f(z)$$

- The problem can't be convex.

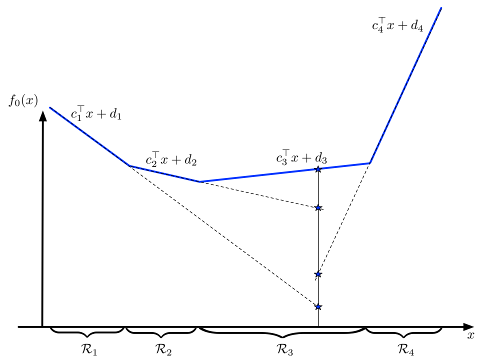


Example: Piecewise affine minimization

Piecewise affine minimization:

$$\min_x \left[\max_{i=1,\ldots,m} \{c_i^\top x + d_i\} \right]$$
$$\text{subject to: } \; Gx \le h$$

- The function is affine on each region $\mathcal{R}_i$.
- Any convex and piecewise affine function can be written this way (e.g. 1st norm).
- Can be reformulated as an LP.



Example : Piecewise affine minimization (con'd)

Piecewise affine minimization:

$$\min_x \left[\max_{i=1,\ldots,m} \{c_i^\top x + d_i\} \right]$$
$$\text{subject to: } \; Gx \le h$$

is **equivalent** to an LP:

$$\begin{aligned} \min_{x,t} \quad & t \\ \text{subject to:} \quad & c_i^\top x + d_i \le t \quad \forall i = 1, \ldots, m \\ & Gx \le h \end{aligned}$$

Add variables and write the problem in epigraph form $\Rightarrow$ **epigraphic reformulation**.

What are LMIs?

A **Linear Matrix Inequality** (LMI) is a constraint of the form:

$$x_1 A_1 + x_2 A_2 + \cdots x_n A_n \preceq B$$

where the matrices $A_1, \ldots, A_n, B \in \mathbb{R}^{m \times m}$ are all symmetric.

- This is a constraint that imposes matrix

$$B - \sum_i^n x_i A_i$$

  to be positive semidefinite (positive definite if $\preceq$ replaced by $\prec$).
- It is equivalent to imposing m polynomial inequalities
  - *Not* element-wise constraints.
  - All leading principle minors are positive (for positive definite matrices).

What are LMIs?

A **Linear Matrix Inequality** (LMI) is a constraint of the form:

$$x_1 A_1 + x_2 A_2 + \cdots x_n A_n \preceq B$$

where the matrices $(A_1, \ldots, A_n, B)$ are all symmetric.

- Consider the constraint

$$Q = \begin{bmatrix} x_1 & x_2 \\ x_2 & x_3 \end{bmatrix} \succeq 0$$

- This is equivalent to 2 inequalities:

$$\begin{aligned} &x_1 \geq 0 \\ &\det(Q) \geq 0 \Leftrightarrow x_1 x_3 - x_2^2 \geq 0 \end{aligned}$$

General form LMIs

Example 1: $y - x^2 > 0,\ y > 0 \iff \begin{bmatrix} y & x \\ x & 1 \end{bmatrix} \succ 0$

- Check leading principle minors
- That is an LMI; rewrite as

$$\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} + y \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + x \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \succ 0$$

Example 2: $x_1^2 + x_2^2 < 1 \iff \begin{bmatrix} 1 & 0 & x_1 \\ 0 & 1 & x_2 \\ x_1 & x_2 & 1 \end{bmatrix} \succ 0$

- Leading principle minors are: $1 > 0$, $\det \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} > 0$, and

$$1 \cdot \det \begin{bmatrix} 1 & x_2 \\ x_2 & 1 \end{bmatrix} - 0 \cdot \det \begin{bmatrix} 0 & x_2 \\ x_1 & 1 \end{bmatrix} + x_1 \cdot \det \begin{bmatrix} 0 & 1 \\ x_1 & x_2 \end{bmatrix} > 0$$

LMIs are convex constraints

Theorem

The following LMI constraint is convex.

$$F(x) = B - \sum_i^n x_i A_i \succeq 0$$

Proof: Let $x,\ y$ such that $F(x), F(y) \succeq 0$, and $\lambda \in (0, 1)$.

$$\begin{aligned} F(\lambda x + (1-\lambda) y) &= B - \sum_i (\lambda x_i + (1-\lambda) y_i) A_i \\ &= \lambda B + (1-\lambda) B - \lambda \sum_i x_i A_i - (1-\lambda) \sum_i y_i A_i \\ &= \lambda F(x) + (1-\lambda) F(y) \\ &\succeq 0 \end{aligned}$$

LMIs are convex constraints

Theorem

The following LMI constraint is convex.

$$F(x) = B - \sum_i^n x_i A_i \succeq 0$$

Alternative proof: We want to show that the set $\{x : \ F(x) \succeq 0\}$ is convex. We have that ...

$$\begin{aligned} \{x : \ F(x) \succeq 0\} &= \{x : \ z^\top F(x) z \geq 0, \text{ for all } z\} \\ &= \bigcap_z \{x : \ z^\top F(x) z \geq 0\} \end{aligned}$$

... but this is an infinite intersection of sets affine in x ... so it is convex!

- LMI much harder than linear constraints – an infinite number of them!
- Result can be piecewise affine – LMIs nonlinear!

Why are LMIs interesting?

Linear Matrix Inequalities:

- Appear in many common control design problems (more later on)
- Most of the problems presented so far can be written using LMI constraints

Linear constraints

$$Ax \leq b \quad \iff \quad \mathrm{diag}(Ax) \preceq \mathrm{diag}(\mathrm{b})$$

Quadratic constraints (It will be clear later on)

$$x^\top Qx + b^\top x + c \leq 0, \quad Q \succ 0 \quad \iff \quad \begin{bmatrix} c + b^\top x & x^\top \\ x & -Q^{-1} \end{bmatrix} \preceq 0$$

Summary

1. Introduction to convex optimization
   - Under convexity: local = global optima
   - Recognizing convexity makes life easier
   - Interplay between convex functions and sets (epigraphic reformulation)
2. Linear Matrix Inequalities (LMIs)
   - Nonlinear constraints
   - LMI constraints are convex!
   - Generalize many of the well known constraints (e.g. linear, quadratic)

Duality Theory

- The Lagrangian function
- The dual problem
- Weak and strong duality
- Optimality conditions
- Game theoretic view

LMIs in optimization

- Semidefinite programming (SDP)
- The dual of an SDP

LMIs in optimization

Consider the following optimization program

$$(\mathrm{SDP}): \quad \begin{array}{c} \min \quad c^\top x \\ \text{subject to:} \;\; x_1A_1 + x_2A_2 + \cdots x_nA_n \preceq B \end{array}$$

where the matrices $(A_1, \ldots, A_n, B)$ are all symmetric.

- We could also have equality constraints
- Optimization over LMI constraints

Why is this class of optimization programs interesting?

- Semidefinite programming (SDP)
- Many control analysis and synthesis problems can be written as SDPs
- Most of the problems presented so far can be written as SDPs

Semidefinite optimization programs (SDPs)

Consider the following optimization program

$$(\text{SDP}): \quad \begin{array}{ll} \min & c^\top x \\ \text{subject to:} & x_1 A_1 + x_2 A_2 + \cdots x_n A_n \preceq B \end{array}$$

where the matrices $(A_1, \ldots, A_n, B)$ are all symmetric.

- Assume we are interested in the optimal value p^* of (SDP)
- Can we construct a lower bound for p^*, i.e. $d^* \leq p^*$, by solving another problem?
- This problem, called *dual*, might sometimes be easier to solve

To do this we first need some machinery – Duality Theory

The Lagrangian function

Recall our standard form (primal) optimization problem:

$$(\mathcal{P}): \quad \begin{array}{lll} \min\limits_{x \in \mathcal{X}} & f_0(x) & \\ \text{subject to:} & f_i(x) \leq 0 & i = 1 \ldots m \\ & h_i(x) = 0 & i = 1 \ldots p \end{array}$$

with (primal) decision variable x, domain $\mathcal{X}$ and optimal value p^*.

Lagrangian Function: $L : \mathcal{X} \times \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$

$$L(x, \lambda, \nu) = f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) + \sum_{i=1}^p \nu_i h_i(x)$$

- λ_i : inequality Lagrange multiplier for $f_i(x) \leq 0$.
- ν_i : equality Lagrange multiplier for $h_i(x) = 0$.
- Lagrangian: weighted sum of the objective and constraint functions.

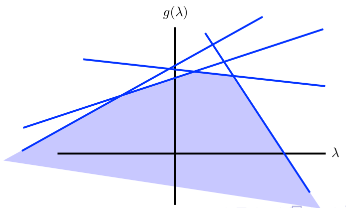
Lagrange dual function

The **dual function** $g : \mathbb{R}^m \times \mathbb{R}^p$ is

$$\begin{aligned} g(\lambda, \nu) &= \inf_{x \in \mathcal{X}} L(x, \lambda, \nu) \\ &= \inf_{x \in \mathcal{X}} \left[f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) + \sum_{i=1}^p \nu_i h_i(x) \right] \end{aligned}$$

The dual function $g(\lambda, \nu)$ is always a **concave** function.

- $g(\lambda, \nu)$ is the pointwise infimum of affine functions
  Do you recall pointwise maximum?



Lagrange dual function

The **dual function** $g : \mathbb{R}^m \times \mathbb{R}^p$ is

$$\begin{aligned} g(\lambda, \nu) &= \inf_{x \in \mathcal{X}} L(x, \lambda, \nu) \\ &= \inf_{x \in \mathcal{X}} \left[f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) + \sum_{i=1}^p \nu_i h_i(x) \right] \end{aligned}$$

The dual function generates lower bounds for the primal optimal value, i.e. $g(\lambda, \nu) \leq p^*$ for $\lambda \geq 0$:
Proof:
For any primal feasible solution $\bar{x}$: $\sum_{i=1}^m \lambda_i f_i(\bar{x}) + \sum_{i=1}^p \nu_i h_i(\bar{x}) \leq 0$

$$\begin{aligned} g(\lambda, \nu) &= \inf_{x \in \mathcal{X}} L(x, \lambda, \nu) \leq L(\bar{x}, \lambda, \nu) \leq f_0(\bar{x}) \text{ for all } \bar{x} \\ g(\lambda, \nu) &\leq \inf_{x \in \mathcal{X}} f_0(x) \leq p^* \end{aligned}$$

- $g(\lambda, \nu)$ might be $-\infty$; Non-trivial if $\operatorname{dom} g := \{\lambda, \nu \mid g(\lambda, \nu) > -\infty\}$

The dual problem

Every $\nu \in \mathbb{R}^p$, $\lambda \geq 0$ produces a lower bound for p^* using the dual function.
Which is the best?

$$(\mathcal{D}): \quad \begin{aligned} \max_{\lambda,\nu} \quad & g(\lambda,\nu) \\ \text{subject to:} \quad & \lambda \geq 0 \end{aligned}$$

- Problem $(\mathcal{D})$ is **convex**, even if $(\mathcal{P})$ is not.
- Problem $(\mathcal{D})$ has optimal value $d^* \leq p^*$.
- The point (λ,ν) is **dual feasible** if $\lambda \geq 0$ and $(\lambda,\nu) \in \operatorname{dom} g$.
- Often impose the constraint $(\lambda,\nu) \in \operatorname{dom} g$ explicitly in $(\mathcal{D})$.

Example : Dual of LPs

$$(\mathcal{P}): \quad \begin{aligned} \min_{x\in\mathbb{R}^n} \quad & c^\top x \\ \text{subject to:} \quad & Ax = b \\ & Cx \leq d \end{aligned}$$

The **dual function** is

$$\begin{aligned} g(\lambda,\nu) &= \min_{x\in\mathbb{R}^n} \left[c^\top x + \nu^\top(Ax-b) + \lambda^\top(Cx-d)\right] \\ &= \min_{x\in\mathbb{R}^n} \left[(A^\top\nu + C^\top\lambda + c)^\top x - b^\top\nu - d^\top\lambda\right] \\ &= \begin{cases} -b^\top\nu - d^\top\lambda & \text{if } A^\top\nu + C^\top\lambda + c = 0 \\ -\infty & \text{otherwise} \end{cases} \end{aligned}$$

Example : Dual of LPs – (cont'd)

$$(\mathcal{P}): \quad \begin{aligned} \min_{x\in\mathbb{R}^n} \quad & c^\top x \\ \text{subject to:} \quad & Ax = b \\ & Cx \leq d \end{aligned}$$

The **dual problem** is

$$(\mathcal{D}): \quad \begin{aligned} \max_{\lambda,\nu} \quad & -b^\top\nu - d^\top\lambda \\ \text{subject to:} \quad & A^\top\nu + C^\top\lambda + c = 0 \\ & \lambda \geq 0 \end{aligned}$$

- Lower bound property:
  $-b^\top\nu - d^\top\lambda \leq p^*$ whenever $\lambda \geq 0$.
- The dual of a linear program is also a linear program.

Example : Dual of a mixed-integer linear program (MILP)

$$(\mathcal{P}): \quad \begin{aligned} \min_{x\in\mathcal{X}} \quad & c^\top x \\ \text{subject to:} \quad & Ax \leq b \\ & \mathcal{X} = \{-1, 1\}^n \end{aligned}$$

The **dual function** is

$$\begin{aligned} g(\lambda) &= \min_{x_i\in\{-1,1\}} \left[c^\top x + \lambda^\top(Ax-b)\right] \\ &= -\|A^\top\lambda + c\|_1 - b^\top\lambda \end{aligned}$$

The **dual problem** is

$$(\mathcal{D}): \quad \begin{aligned} \max_{\lambda} \quad & -\|A^\top\lambda + c\|_1 - b^\top\lambda \\ \text{subject to:} \quad & \lambda \geq 0 \end{aligned}$$

The dual of a mixed-integer linear program is a linear program!

Weak and strong duality

Weak Duality

- It is **always** true that $d^* \leq p^*$.
- Sometimes the dual is much easier to solve than the primal (or vice-versa).
- Example: The dual of an MILP (difficult to solve) is a standard LP (easy to solve).

Strong Duality

- It is **sometimes** true that $d^* = p^*$.
- Strong duality usually holds for convex problems.
- Strong duality usually does not hold for non-convex problems.
- Can impose conditions on convex problems to guarantee that $d^* = p^*$.

Strong duality for convex problems

An optimization problem with f_0 and all f_i convex:

$$(\mathcal{P}): \quad \begin{array}{rll} \min & f_0(x) & \\ \text{subject to:} & f_i(x) \leq 0 & i = 1 \dots m \\ & Ax = b & A \in \mathbb{R}^{p \times n} \end{array}$$

Slater Condition

If there is at least one **strictly feasible point**, i.e.

$$\left\{ x \;\middle|\; Ax = b,\ f_i(x) < 0,\ \forall i \in \{1, \dots, m\} \right\} \neq \emptyset$$

Then $p^* = d^*$.

- Stronger version: Only the nonlinear functions $f_i(x)$ must be strictly satisfiable (non-empty interior).
- Other **constraint qualification** conditions exist.

Duality – A geometric view

Assume one inequality constraint only:

$$\mathcal{G} := \{(u, t) \mid t = f_0(x),\ u = f_1(x),\ x \in \mathcal{X}\}$$

Primal problem:

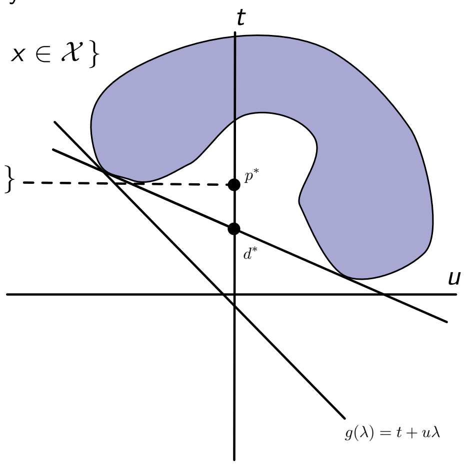
$$p^* = \min \{t \mid (u, t) \in \mathcal{G},\ u \leq 0\}$$

Dual function:

$$g(\lambda) = \min_{(u,t) \in \mathcal{G}} (t + \lambda u)$$

Dual problem:

$$d^* = \max_{\lambda \geq 0} g(\lambda)$$



The quantity $p^* - d^*$ is the **duality gap**.

Primal and dual solution properties

Assume that strong duality holds, with optimal solution x^* and (λ^*, ν^*).

- From strong duality, $d^* = p^* \;\Rightarrow\; g(\lambda^*, \nu^*) = f_0(x^*)$.
- From the definition of the dual function:

$$f_0(x^*) = g(\lambda^*, \nu^*) = \min_x \left\{ f_0(x) + \sum_{i=1}^m \lambda_i^* f_i(x) + \sum_{i=1}^p \nu_i^* h_i(x) \right\}$$

$$\leq f_0(x^*) + \sum_{i=1}^m \lambda_i^* f_i(x^*) + \sum_{i=1}^p \nu_i^* h_i(x^*) \leq f_0(x^*)$$

[weak duality]

$$\implies f_0(x^*) = g(\lambda^*, \nu^*) = f_0(x^*) + \sum_{i=1}^m \lambda_i^* f_i(x^*) + \sum_{i=1}^p \nu_i^* h_i(x^*)$$

$$\implies \left. \begin{array}{l} \lambda_i^* = 0 \text{ for every } f_i(x^*) < 0. \\ f_i(x^*) = 0 \text{ for every } \lambda_i^* > 0. \end{array} \right\} \textbf{Complementary slackness}$$

Karush-Kuhn-Tucker (KKT) optimality conditions

Assume that all f_i and h_i are differentiable. **Necessary** conditions for optimality:

1. Primal Feasibility:

$$\begin{aligned} f_i(x^*) &\leq 0 \quad i = 1, \ldots, m \\ h_i(x^*) &= 0 \quad i = 1 \ldots, p \end{aligned}$$

2. Dual Feasibility:

$$\lambda^* \geq 0$$

3. Complementary Slackness:

$$\lambda_i^* f_i(x^*) = 0 \quad i = 1, \ldots, m$$

4. Stationarity:

$$\nabla_x L(x^*, \lambda^*, \nu^*) = \nabla f_0(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla f_i(x^*) + \sum_{i=1}^{p} \nu_i^* \nabla h_i(x^*) = 0$$

KKT optimality conditions

Assume that all f_i and h_i are differentiable and problem is convex:

1. If (x^*, λ^*, ν^*) satisfy the KKT conditions, then
   - they are primal and dual optimal
   - they result in zero duality gap, i.e. $p^* = d^*$
2. If in addition Slater's condition holds, then
   - duality gap is zero and the dual optimum is attained (existence of (λ^*, ν^*) is guaranteed)
   - x^* is optimal **if and only if** there exists (λ^*, ν^*) that, together with x^*, satisfy the KKT conditions

Game theoretic view

Assume inequality constraints only.

We have that for all x

$$\begin{aligned} \max_{\lambda \geq 0} L(x, \lambda) &= \max_{\lambda \geq 0} \left(f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) \right) \\ &= \begin{cases} f_0(x) & \text{if } f_i(x) \leq 0 \textit{ for all } \text{i}; \\ \infty & \text{otherwise.} \end{cases} \end{aligned}$$

Since this holds for all x, we then have that

$$\begin{aligned} p^* &= \min_{x \in \mathcal{X}} \max_{\lambda \geq 0} L(x, \lambda) \\ d^* &= \max_{\lambda \geq 0} \min_{x \in \mathcal{X}} L(x, \lambda) \end{aligned}$$

Game theoretic view

- Game between primal (Peter) and dual (Debbie) variables:

$$\begin{aligned} p^* &= \min_{x} \max_{\lambda} L(x, \lambda) \\ d^* &= \max_{\lambda} \min_{x} L(x, \lambda) \end{aligned}$$

- Consider the d^* game – Debbie plays first, Peter plays second

$$\begin{aligned} d^* &= \max_{\lambda} \quad \min_{x} \quad L(x, \lambda) \leq \text{ any value} \\ &= \forall \lambda \quad \exists x \quad L(x, \lambda) \leq \text{ any value} \\ &= \exists x(\lambda) \quad \forall \lambda \quad L(x, \lambda) \leq \text{ any value} \quad [x(\cdot) \text{ is parametric in } \lambda] \\ &\leq \exists x \quad \forall \lambda \quad L(x, \lambda) \leq \text{ any value} \\ &= \min_{x} \quad \max_{\lambda} \quad L(x, \lambda) \\ &= p^* \end{aligned}$$

Game theoretic view

- Game between primal (Peter) and dual (Debbie) variables:

$$p^* = \min_x \max_\lambda L(x,\lambda)$$
$$d^* = \max_\lambda \min_x L(x,\lambda)$$

- If Peter plays second $\Rightarrow$

$$d^* \leq p^* \quad \text{[weak duality]}$$

- Duality gap corresponds to the advantage of Peter
- Strong duality = Zero duality gap
  $\Rightarrow$ No advantage for any of the players

Semidefinite programming

Primal SDP problem (all matrices are symmetric):

$$\min \quad c^\top x$$
$$\text{subject to:} \quad x_1A_1 + x_2A_2 + \cdots x_nA_n \preceq B$$

Lagrangian:

$$\mathcal{L}(x,\Lambda) = c^\top x + \sum_i \langle \Lambda, A_i\rangle x_i - \langle \Lambda, B\rangle,$$

where $\langle X, Y\rangle = \mathrm{trace}(X^\top Y) = \sum_{i,j} X_{ij}Y_{ij}$.

This fact relies on "dual cone" arguments, and the fact that trace is the inner product for matrices. Alternatively, recall that

$$F(x) = \sum_i x_iA_i - B \preceq 0 \Leftrightarrow z^\top F(x)z \leq 0, \ \forall z \neq 0$$
$$\Leftrightarrow \max_{z\neq 0} z^\top F(x)z \leq 0$$

Semidefinite programming – Interpretation of the dual

We can lift this constraint in the objective with Lagrange multiplier $\lambda \geq 0$:

$$\begin{aligned}\mathcal{L}(x,\lambda) &= c^\top x + \lambda \max_{z\neq 0} z^\top F(x)z \\ &= c^\top x + \lambda \bar{z}_x^\top F(x)\bar{z}_x\end{aligned}$$

where $\bar{z}_x$ is the one that achieves the maximum (assuming it is attained); merging it with λ we have

$$\begin{aligned}\mathcal{L}(x,\Lambda) &= c^\top x + z^\top F(x)z \\ &= c^\top x + \langle zz^\top, F(x)\rangle = c^\top x + \langle \Lambda, F(x)\rangle\end{aligned}$$

where $\Lambda = zz^\top \succeq 0$.

Notice that we used the fact that

$$z^\top F(x)z = \sum_{i,j} z_iz_jF(x)_{ij} = \sum_{i,j}(zz^\top)_{ij}F(x)_{ij} = \langle zz^\top, F(x)\rangle$$

Semidefinite programming

Primal SDP problem:

$$\min \quad c^\top x$$
$$\text{subject to:} \quad x_1A_1 + x_2A_2 + \cdots x_nA_n \preceq B$$

where the matrices $(A_1, \ldots, A_n, B)$ are all symmetric.

Lagrangian:

$$\begin{aligned}\mathcal{L}(x,\Lambda) &= c^\top x + \sum_i \langle \Lambda, A_i\rangle x_i - \langle \Lambda, B\rangle \\ &= \sum_i \left(c_i + \langle \Lambda, A_i\rangle\right)x_i - \langle \Lambda, B\rangle\end{aligned}$$

Dual function:

$$g(\lambda) = \begin{cases} -\langle \Lambda, B\rangle & \text{if } c_i + \langle \Lambda, A_i\rangle = 0 \text{ for } i = 1\ldots n \\ -\infty & \text{otherwise}\end{cases}$$

Semidefinite programming

Primal SDP problem:

$$\min \quad c^\top x$$
$$\text{subject to:} \quad x_1 A_1 + x_2 A_2 + \cdots x_n A_n \preceq B$$

where the matrices $(A_1, \ldots, A_n, B)$ are all symmetric.

Dual function:

$$g(\lambda) = \begin{cases} -\langle \Lambda, B \rangle & \text{if } c_i + \langle \Lambda, A_i \rangle = 0 \text{ for } i = 1 \ldots n \\ -\infty & \text{otherwise} \end{cases}$$

The dual problem:

$$\max \quad -\langle B, \Lambda \rangle$$
$$\text{subject to:} \quad \langle A_i, \Lambda \rangle = -c_i, \text{ for all } i$$
$$\Lambda \succeq 0$$

Semidefinite programming

Primal SDP problem:

$$\min \quad c^\top x$$
$$\text{subject to:} \quad x_1 A_1 + x_2 A_2 + \cdots x_n A_n \preceq B$$

The dual problem:

$$\max \quad -\langle B, \Lambda \rangle$$
$$\text{subject to:} \quad \langle A_i, \Lambda \rangle = -c_i, \text{ for all } i$$
$$\Lambda \succeq 0$$

Weak duality:

$$\begin{aligned} p^* - d^* &= c^\top x + \langle B, \Lambda \rangle && \text{[primal feasibility]} \\ &\geq c^\top x + \sum_i \langle A_i, \Lambda \rangle x_i && \text{[dual feasibility]} \\ &= \sum_i c_i x_i - \sum_i c_i x_i = 0 \end{aligned}$$

Semidefinite programming

Primal SDP problem:

$$\min \quad c^\top x$$
$$\text{subject to:} \quad x_1 A_1 + x_2 A_2 + \cdots x_n A_n \preceq B$$

The dual problem:

$$\max \quad -\langle B, \Lambda \rangle$$
$$\text{subject to:} \quad \langle A_i, \Lambda \rangle = -c_i, \text{ for all } i$$
$$\Lambda \succeq 0$$

Weak duality: $p^* - d^* \geq 0$

Strong duality:

Under Slater's condition, i.e. constraints in the primal need to be satisfied with $\prec$ instead of $\preceq$. For SDPs the *dual of the dual* is the primal.

Summary

1. Duality Theory
   - Construct $d^* \leq p^*$ in three steps
     1. Construct the Lagrangian (lift and weight constraints in the objective)
     2. Construct dual function and "eliminate" primal variables
     3. Formulate dual problem (don't forget constraints on dual variables)
   - Optimality conditions
   - Geometric and gaming interpretation of duality
2. LMIs in optimization
   - Semidefinite programming (SDP)
   - Construct the dual of an SDP (similar procedure with linear programs)
   - Weak duality, strong duality under Slater's condition

Reformulation in LMIs

- The Schur complement
  - Non-obvious LMIs
  - From nonlinear constraints to LMIs
- The S-procedure
  - From quadratic implications to LMIs
  - Turning set containment arguments in LMIs

LMIs for stability & controller synthesis

- Recap of stability theorems
- Lyapunov matrix inequality
- Controller synthesis by means of an example

Non-obvious LMIs

Some cases (like the QP) are harder to write as LMIs.

The Schur complement provides the means to do so

Schur complement: Turns a nonlinear constraint into an LMI

Theorem (Schur complement)

Assume that $Q(x) = Q(x)^\top$, $R(x) = R(x)^\top$: affine functions of x. We then have that

$$R(x) \succ 0 \text{ and } Q(x) - S(x)R(x)^{-1}S(x)^\top \succ 0$$
$$\Leftrightarrow \begin{bmatrix} Q(x) & S(x) \\ S(x)^\top & R(x) \end{bmatrix} \succ 0$$

Schur complement

Schur complement: The non-strict case

Assume that $Q(x) = Q(x)^\top, R(x) = R(x)^\top \succ 0$: affine functions of x

We then have that

$$Q(x) - S(x)R(x)^{-1}S(x)^\top \succeq 0 \Leftrightarrow \begin{bmatrix} Q(x) & S(x) \\ S(x)^\top & R(x) \end{bmatrix} \succeq 0$$

Example 1:

$$\|A\|_2 \le t \quad \Leftrightarrow \quad A^\top A \preceq t^2 I,\ t \ge 0 \quad \Leftrightarrow \quad \begin{bmatrix} tI & A^\top \\ A & tI \end{bmatrix} \succeq 0$$

Example 2: The QP (we have seen this before)

$$x^\top Qx + b^\top x + c \le 0, \quad Q \succ 0 \quad \Leftrightarrow \quad \begin{bmatrix} c + b^\top x & x^\top \\ x & -Q^{-1} \end{bmatrix} \preceq 0$$

Schur complement – Proof for the strict case

Proof of ($\Leftarrow$):

Assume $\begin{bmatrix} Q(x) & S(x) \\ S(x)^\top & R(x) \end{bmatrix} \succ 0$. For all $[u\ v] \ne 0$ we have

$$F(u,v) = \begin{bmatrix} u \\ v \end{bmatrix}^\top \begin{bmatrix} Q(x) & S(x) \\ S(x)^\top & R(x) \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} > 0$$

Considering $u = 0$ we have

$$F(0,v) = v^\top R(x)v > 0, \text{ for all } v \ne 0 \Rightarrow R(x) \succ 0$$

Consider now $v = -R(x)^{-1}S(x)^\top u$, with $u \ne 0$

$$F(u,v) = u^\top \left(Q(x) - S(x)R(x)^{-1}S(x)^\top\right)u > 0, \text{ for all } u \ne 0$$
$$\Rightarrow Q(x) - S(x)R(x)^{-1}S(x)^\top \succ 0$$

Schur complement – Proof for the strict case

Proof of ($\Rightarrow$):

Now assume $R(x) \succ 0$ and $Q(x) - S(x)R(x)^{-1}S(x)^\top \succ 0$, and as before

$$F(u,v) = \begin{bmatrix} u \\ v \end{bmatrix}^\top \begin{bmatrix} Q(x) & S(x) \\ S(x)^\top & R(x) \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} > 0$$

Fix u and minimize over v: $\nabla_v F(u,v) = 2R(x)v + 2S(x)^\top u = 0$. Since $R(x) \succ 0$, we have that $v^* = -R(x)^{-1}S(x)^\top u$. Substitute it in the expression of $F(u,v)$ to obtain

$$F(u) = u^\top \left(Q(x) - S(x)R(x)^{-1}S(x)^\top\right) u$$

Since $Q(x) - S(x)R(x)^{-1}S(x)^\top \succ 0$, $u^* = 0$ minimizes $F(u)$. As a result, $(u^*, v^*) = (0,0)$ and $F(u^*, v^*) = 0$.

Hence, $F(u,v) > 0$ for all $u, v \neq 0 \;\Rightarrow \begin{bmatrix} Q(x) & S(x) \\ S(x)^\top & R(x) \end{bmatrix} \succ 0$.

Schur complement – Ellipsoidal inequality

Assume that $Q(x) = Q(x)^\top$, $R(x) = R(x)^\top \succ 0$: affine functions of x.
We then have that

$$Q(x) - S(x)R(x)^{-1}S(x)^\top \succeq 0 \Leftrightarrow \begin{bmatrix} Q(x) & S(x) \\ S(x)^\top & R(x) \end{bmatrix} \succeq 0$$

Consider the ellipsoid

$$(x - x_c)^\top A^{-1}(x - x_c) \leq 1, \quad A = A^\top \succ 0$$

(... and recall that it is convex).
Setting $Q(x) = 1$, $R(x) = A$ and $S(x) = (x - x_c)^\top$:

$$\begin{bmatrix} 1 & (x - x_c)^\top \\ (x - x_c) & A \end{bmatrix} \succeq 0$$

Schur complement – Maximum singular value

Assume that $Q(x) = Q(x)^\top$, $R(x) = R(x)^\top \succ 0$: affine functions of x.
We then have that

$$Q(x) - S(x)R(x)^{-1}S(x)^\top \succeq 0 \Leftrightarrow \begin{bmatrix} Q(x) & S(x) \\ S(x)^\top & R(x) \end{bmatrix} \succeq 0$$

Let $A(x)$: affine in x and real valued.
Let also $\bar{\sigma}[A(x)]$ be the maximum singular value of $A(x)$, i.e. the square root of the largest eigenvalue of $A(x)A(x)^\top$, i.e. $\bar{\lambda}[A(x)A(x)^\top]^{\frac{1}{2}}$.

$$\begin{aligned} \bar{\sigma}(A(x)) \leq 1 &\Leftrightarrow \bar{\lambda}[A(x)A(x)^\top] \leq 1 \\ &\Leftrightarrow A(x)A(x)^\top \preceq I \\ &\Leftrightarrow I - A(x)I^{-1}A(x)^\top \succeq 0 \\ &\Leftrightarrow \begin{bmatrix} I & A(x) \\ A(x)^\top & I \end{bmatrix} \succeq 0 \end{aligned}$$

S-procedure

S-**procedure**: Turns quadratic implications to LMIs

Consider two quadratic functions

$$\begin{aligned} f_0(x) &= x^\top A_0 x + 2x^\top b_0 + c_0 \\ f(x) &= x^\top A x + 2x^\top b + c, \end{aligned}$$

where all matrices/vectors are given, and $A_0 = A_0^\top$, $A = A^\top$.

Problem: When is it true that one quadratic inequality implies another?
In other words, when does

$$f(x) \geq 0, \; x \neq 0 \;\Rightarrow\; f_0(x) \geq 0$$

S-procedure (cont'd)

Theorem

The following implication holds

$$f(x) \geq 0,\ x \neq 0 \ \Rightarrow\ f_0(x) \geq 0$$

if *there exists*

$$\tau \geq 0 \text{ such that } f_0(x) - \tau f(x) \geq 0$$

Still not an LMI ... but $f_0(x), f(x)$, are quadratic in x.

S-procedure (cont'd)

Theorem

The following implication holds

$$f(x) \geq 0,\ x \neq 0 \ \Rightarrow\ f_0(x) \geq 0$$

if *there exists*

$$\tau \geq 0 \text{ such that } f_0(x) - \tau f(x) \geq 0$$

For a quadratic function $f(x) = x^\top Ax + 2x^\top b + c$

$$\begin{bmatrix} x \\ 1 \end{bmatrix}^\top \begin{bmatrix} A & b \\ b^\top & c \end{bmatrix} \begin{bmatrix} x \\ 1 \end{bmatrix} \geq 0,\ \forall x \Leftrightarrow \begin{bmatrix} \xi x \\ \xi \end{bmatrix}^\top \begin{bmatrix} A & b \\ b^\top & c \end{bmatrix} \begin{bmatrix} \xi x \\ \xi \end{bmatrix} \geq 0,\ \forall x, \xi$$

$$\Leftrightarrow \begin{bmatrix} A & b \\ b^\top & c \end{bmatrix} \succeq 0$$

S-procedure (cont'd)

Theorem

The following implication holds

$$f(x) \geq 0,\ x \neq 0 \ \Rightarrow\ f_0(x) \geq 0$$

if *there exists*

$$\tau \geq 0 \text{ such that } f_0(x) - \tau f(x) \geq 0$$

Since $f_0(x), f(x)$, are quadratic in x, the condition above is equivalent to an LMI in τ

$$\begin{bmatrix} A_0 & b_0 \\ b_0^\top & c_0 \end{bmatrix} - \tau \begin{bmatrix} A & b \\ b^\top & c \end{bmatrix} \succeq 0$$

S-procedure (cont'd)

Theorem

The following implication holds

$$f(x) \geq 0,\ x \neq 0 \ \Rightarrow\ f_0(x) \geq 0$$

if *there exists*

$$\tau \geq 0 \text{ such that } f_0(x) - \tau f(x) \geq 0$$

Since $f_0(x), f(x)$, are quadratic in x, this is equivalent to an LMI in τ

$$\begin{bmatrix} A_0 & b_0 \\ b_0^\top & c_0 \end{bmatrix} - \tau \begin{bmatrix} A & b \\ b^\top & c \end{bmatrix} \succeq 0$$

The **only if** part also holds true (though non-obvious) if $\exists \bar{x}$ such that $f(\bar{x}) > 0$, i.e. the "ellipsoids" have non-empty interior condition. In that case we get equivalence!

A containment problem

Problem: Determine an ellipsoid $\mathcal{E}$ centered at the origin

$$\mathcal{E} = \{x \mid x^\top A^{-1} x \le 1\},$$

that contains a polytope $\mathcal{P}$ with vertices $v_1, \ldots, v_p$.
In other words, we are looking for $\mathcal{P} \subseteq \mathcal{E}$.

Restate the problem: If $x \in \mathcal{P}$ then $x \in \mathcal{E}$. But $x \in \mathcal{P}$ is equivalent to $v_i \in \mathcal{P}$, for all $i = 1, \ldots, p$. Hence,

$$\begin{aligned} & v_i^\top A^{-1} v_i \le 1, \text{ for all } i = 1, \ldots, p. \\ \Leftrightarrow \ & 1 - v_i^\top A^{-1} v_i \ge 0, \text{ for all } i = 1, \ldots, p. \end{aligned}$$

Using the Schur complement lemma we can turn it into an LMI

$$\begin{bmatrix} 1 & v_i^\top \\ v_i & A \end{bmatrix} \succeq 0, \text{ for all } i = 1, \ldots, p.$$

Stability analysis – Linear systems

Consider the linear, time-invariant (LTI) dynamical system

$$\dot{x}(t) = Ax(t), \quad x(0) = x_0$$

where $x(t) \in \mathbb{R}^n$ is the system state and $A \in \mathbb{R}^{n \times n}$.
It is called *autonomous* since there are no inputs.

Definition: The autonomous LTI system is *asymptotically stable* if, for all $x(0) \in \mathbb{R}^n$,

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

In the scalar case ($n = 1$ and $A = a \in \mathbb{R}$), we can solve the ODE:

$$x(t) = e^{at} x_0$$

If $a < 0$, then the system is asymptotically stable.

Stability analysis recap – Linear systems

Consider the linear, time-invariant (LTI) dynamical system

$$\dot{x}(t) = Ax(t), \quad x(0) = x_0$$

where $x(t) \in \mathbb{R}^n$ is the system state and $A \in \mathbb{R}^{n \times n}$.
It is called *autonomous* since there are no inputs.

Definition: The autonomous LTI system is *asymptotically stable* if, for all $x(0) \in \mathbb{R}^n$,

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

What if $n > 1$? Can we work the same way? The ODE solution is then

$$x(t) = e^{At} x_0,$$

where e^{At} is the matrix exponential, i.e.

$$e^{At} = I + At + \frac{1}{2}A^2 t^2 + \frac{1}{3}A^3 t^3 + \ldots$$

Stability analysis recap – Linear systems

Consider the linear, time-invariant (LTI) dynamical system

$$\dot{x}(t) = Ax(t), \quad x(0) = x_0$$

where $x(t) \in \mathbb{R}^n$ is the system state and $A \in \mathbb{R}^{n \times n}$.
It is called *autonomous* since there are no inputs.

Definition: The autonomous LTI system is *asymptotically stable* if, for all $x(0) \in \mathbb{R}^n$,

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

What if $n > 1$? Can we work the same way? The ODE solution is then

$$x(t) = e^{At} x_0,$$

where e^{At} is the matrix exponential. Can we do without computing e^{At}?

Stability analysis recap – Linear systems

Theorem

An autonomous LTI system is asymptotically stable, i.e. $\lim_{t\to\infty} x(t) = 0$, **if and only if** *$A$ is Hurwitz, i.e. all its eigenvalues have negative real part.*

Moved from matrix exponential to eigenvalue computation – there must be some connection with LMIs.

Theorem

Given some matrix $Q = Q^\top \succ 0$, a matrix A is Hurwitz **if and only if** *there exists $X = X^\top \succ 0$ that satisfies the Lyapunov Matrix Equation*

$$A^\top X + XA = -Q$$

Equivalently, since $Q \succ 0$ and it is arbitrary ...

Stability analysis recap – Linear systems

For asymptotic stability A has to be Hurwitz, i.e.

Theorem

Given some matrix $Q = Q^\top \succ 0$, a matrix A is Hurwitz **if and only if** *there exists $X = X^\top \succ 0$ that satisfies the Lyapunov Matrix Equation*

$$A^\top X + XA = -Q$$

Equivalently, since $Q \succ 0$ and it is arbitrary ...

Theorem

A matrix A is Hurwitz **if and only if** *there exists $X = X^\top \succ 0$ that satisfies the Lyapunov Matrix Inequality*

$$A^\top X + XA \prec 0$$

This is an LMI in X!

Stability analysis recap – Nonlinear systems

Asymptotic stability for nonlinear systems; Lyapunov theory again

Theorem

Let $x = 0$ be an equilibrium of $\dot{x}(t) = f(x(t))$, and let $\mathcal{D} \subset \mathbb{R}^n$ be a domain containing $x = 0$. **If** *there exists a continuous, differentiable function $V : \mathcal{D} \to \mathbb{R}$ such that*

$$\begin{aligned} V(0) = 0,\ V(x) > 0, &\quad \textit{for all } x \in \mathcal{D}\setminus\{0\} \\ \dot{V}(x) < 0, &\quad \textit{for all } x \in \mathcal{D}\setminus\{0\} \end{aligned}$$

then *$x = 0$ is asymptotically stable.*

Linear systems stability comes then as a special case.

Stability analysis recap – Nonlinear systems

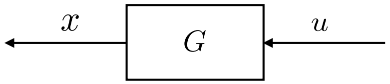
Linear systems stability comes then as a special case. Consider $\dot{x}(t) = Ax(t)$ and let $V(x) = x^\top X x$ be a Lyapunov function. The Lyapunov stability theorem requires

$$\begin{aligned} V(0) = 0 : &\quad \text{satisfied} \\ V(x) > 0, \quad \text{for all } x \in \mathcal{D}\setminus\{0\} : &\quad \Leftrightarrow \quad X \succ 0 \\ \dot{V}(x) < 0, \quad \text{for all } x \in \mathcal{D}\setminus\{0\} : &\quad \Leftrightarrow \quad x^\top (A^\top X + XA) x < 0 \\ &\quad \Leftrightarrow \quad A^\top X + XA \prec 0 \end{aligned}$$

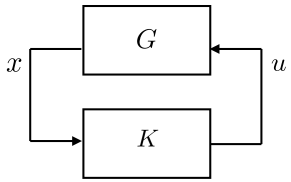
Using a quadratic Lyapunov function we can "prove" Lyapunov Matrix Equation from the nonlinear Lyapunov's stability theorem.

State feedback control design

Consider a system G: $\dot{x} = Ax + Bu$



Determine a feedback gain matrix K such that $u = Kx$ renders the closed loop system stable.



Closed loop system: $\dot{x} = (A + BK)x$.

- Goal: Determine K such that $A + BK$ is Hurwitz.

State feedback control design (cont'd)

Closed loop system: $\dot{x} = (A + BK)x$.

- Goal: Determine K such that $A + BK$ is Hurwitz.

Lyapunov stability (recall from Lecture 3): A matrix A is Hurwitz **if and only if** there exists $P = P^\top \succ 0$ such that

$$A^\top P + PA \prec 0$$

Equivalent representation: Multiply by P^{-1} from the left and right:

$$P^{-1}A^\top PP^{-1} + P^{-1}PAP^{-1} \prec 0$$

and set $X = P^{-1}$. We then have

$$XA^\top + AX \prec 0$$

State feedback control design (cont'd)

Closed loop system: $\dot{x} = (A + BK)x$.

- Goal: Determine K such that $A + BK$ is Hurwitz.

Lyapunov stability: A matrix A is stable **if and only if** there exists $X = X^\top \succ 0$ such that

$$XA^\top + AX \prec 0$$

Enforce this condition with $A + BK$ in place of A and determine K and X:

$$X(A + BK)^\top + (A + BK)X \prec 0$$

which leads to

$$XA^\top + (XK^\top)B^\top + AX + B(KX) \prec 0$$

State feedback control design (cont'd)

Closed loop system: $\dot{x} = (A + BK)x$.

- Goal: Determine K such that $A + BK$ is Hurwitz.

Lyapunov stability: A matrix A is stable **if and only if** there exists $X = X^\top \succ 0$ such that

$$XA^\top + AX \prec 0$$

We are left with this condition which is not nice!

$$XA^\top + (XK^\top)B^\top + AX + B(KX) \prec 0$$

Setting $Z = KX$ we have

$$XA^\top + Z^\top B^\top + AX + BZ \prec 0$$

Solve this LMI to determine X and Z and then compute $K = ZX^{-1}$

Summary

1. Reformulation in LMI constraints
   - Schur complement
     - Commonly used "trick"
     - Appears in quadratic problems, and many others
   - The S-procedure
     - Turns quadratic implications in LMI constraints
     - Useful in set containment problems
2. LMIs for stability & controller synthesis
   - Recap of stability theorems for linear and nonlinear systems
   - Lyapunov stability for linear systems by means of LMIs
   - Example for controller synthesis

Thank you! Questions?

Contact at:
kostas.margellos@eng.ox.ac.uk