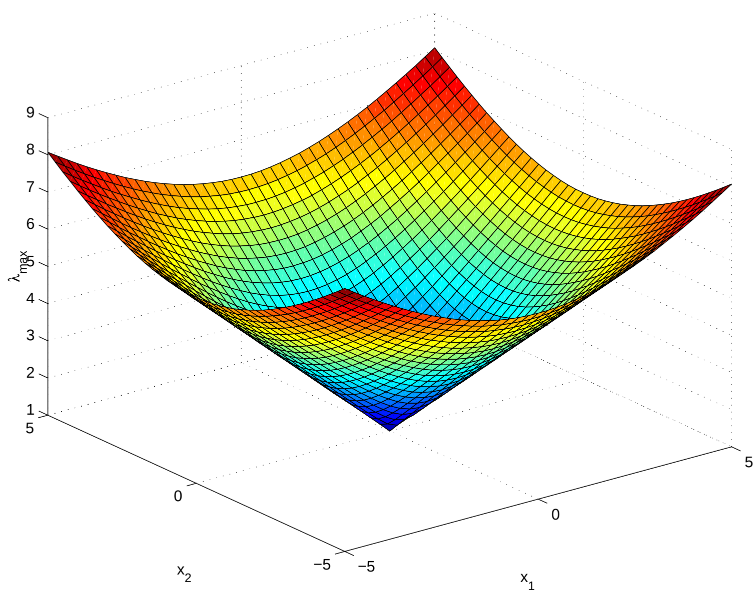
COURSE ON LMI OPTIMIZATION WITH APPLICATIONS IN CONTROL

Didier HENRION
www.laas.fr/∼henrion
henrion@laas.fr

Denis ARZELIER
www.laas.fr/∼arzelier
arzelier@laas.fr



May 2004

Course outline

I LMI optimization

I.1 Introduction: What is an LMI ? What is SDP ?

historical survey - applications - convexity - cones - polytopes

I.2 SDP duality

Lagrangian duality - SDP duality - KKT conditions

I.3 What can be represented as an LMI ?

SDP representability - geometry - algebraic tricks

I.4 LMI relaxations of non-convex problems

hierarchies of relaxations - liftings - reformulations

I.5 Solving LMIs

interior point methods - solvers - interfaces

II LMIs in control

II.1 State-space analysis methods

Lyapunov stability - pole placement in LMI regions - robustness

II.2 State-space design methods

H_2, H_∞, robust state-feedback and output-feedback design

II.3 Polynomial analysis methods

polynomials in control - robust stability of polynomials

II.4 Polynomial design methods

robust fixed-order controller design

Course material

Very good references on convex optimization:

- S. Boyd, L. Vandenberghe. Convex Optimization, Lecture Notes Stanford & UCLA, CA, 2002
- H. Wolkowicz, R. Saigal, L. Vandenberghe. Handbook of semidefinite programming, Kluwer, 2000
- A. Ben-Tal, A. Nemirovskii. Lectures on Modern Convex Optimization, SIAM, 2001

Modern state-space LMI methods in control:

- C. Scherer, S. Weiland. Course on LMIs in Control, Lecture Notes Delft & Eindhoven Univ Tech, NL, 2002
- S. Boyd, L. El Ghaoui, E. Feron, V. Balakrishnan. Linear Matrix Inequalities in System and Control Theory, SIAM, 1994
- M. C. de Oliveira. Linear Systems Control and LMIs, Lecture Notes Univ Campinas, BR, 2002.

Results on LMI and algebraic optimization in control:

- P. A. Parrilo, S. Lall. Mini-Course on SDP Relaxations and Algebraic Optimization in Control. European Control Conference, Cambridge, UK, 2003
- P. A. Parrilo, S. Lall. Semidefinite Programming Relaxations and Algebraic Optimization in Control, Workshop presented at the 42nd IEEE Conference on Decision and Control, Maui HI, USA, 2003

COURSE ON LMI OPTIMIZATION
WITH APPLICATIONS IN CONTROL
PART I.1

WHAT IS AN LMI ?
WHAT IS SDP ?

Didier HENRION
www.laas.fr/∼henrion
henrion@laas.fr

Denis ARZELIER
www.laas.fr/∼arzelier
arzelier@laas.fr

Professeur Jan C Willems

Mai 2004

LMI - Linear Matrix Inequality

$$F(x) = F_0 + \sum_{i=1}^{n} x_i F_i \succeq \mathbf{0}$$

- $F_i \in \mathbb{S}_m$ given symmetric matrices
- $x_i \in \mathbb{R}^n$ decision variables

Fundamental property: feasible set is convex

$$\mathcal{S} = \{x \in \mathbb{R}^n \, : \, F(x) \succeq \mathbf{0}\}$$

$\mathcal{S}$ is the Spectrahedron

Nota : $\succeq 0$ ($\succ 0$) means positive semidefinite (positive definite) e.g. real nonnegative eigenvalues (strictly positive eigenvalues) and defines generalized inequalities on PSD cone

Terminology coined out by Jan Willems in 1971

$$F(P) = \begin{bmatrix} A'P + PA + Q & PB + C' \\ B'P + C & R \end{bmatrix} \succeq \mathbf{0}$$

"The basic importance of the LMI seems to be largely unappreciated. It would be interesting to see whether or not it can be exploited in computational algorithms"

Lyapunov's LMI

Historically, the first LMIs appeared around 1890 when Lyapunov showed that the autonomous system with LTI model:

$$\frac{d}{dt}x(t) = \dot{x}(t) = Ax(t)$$

is stable (all trajectories converge to zero) iff there exists a solution to the matrix inequalities

$$A'P + PA \prec \mathbf{0} \qquad P = P' \succ \mathbf{0}$$

which are linear in unknown matrix P

Aleksandr Mikhailovich Lyapunov
(1857 Yaroslavl - 1918 Odessa)

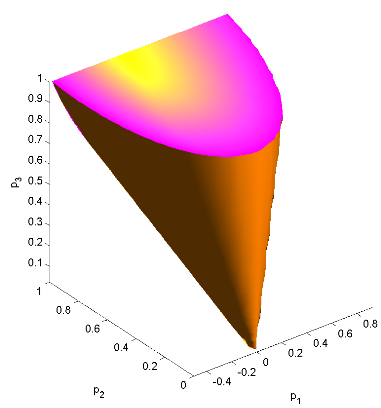
Example of Lyapunov's LMI

$$A = \begin{bmatrix} -1 & 2 \\ 0 & -2 \end{bmatrix} \quad P = \begin{bmatrix} p_1 & p_2 \\ p_2 & p_3 \end{bmatrix}$$

$$A'P + PA \prec \mathbf{0} \qquad P \succ \mathbf{0}$$

$$\begin{bmatrix} -2p_1 & 2p_1 - 3p_2 \\ 2p_1 - 3p_2 & 4p_2 - 4p_3 \end{bmatrix} \prec \mathbf{0}$$

$$\begin{bmatrix} p_1 & p_2 \\ p_2 & p_3 \end{bmatrix} \succ \mathbf{0}$$



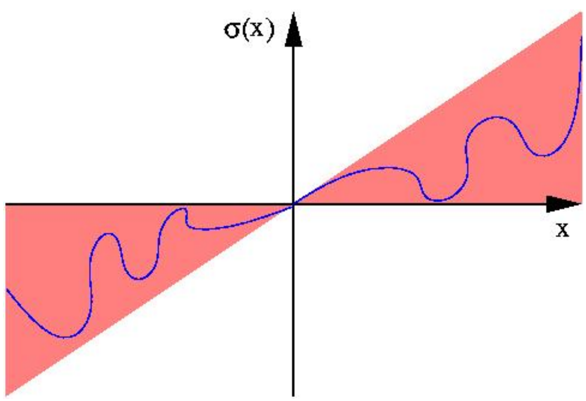
Matrices P satisfying Lyapunov LMI's

$$\begin{bmatrix} 2 & -2 & 0 & 0 \\ -2 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} p_1 + \begin{bmatrix} 0 & 3 & 0 & 0 \\ 3 & -4 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} p_2 + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} p_3 \succ \mathbf{0}$$

Some history

1940s - Absolute stability problem: Lu're, Postnikov et al applied Lyapunov's approach to control problems with nonlinearity in the actuator

$$\dot{x} = Ax + b\sigma(x)$$



Sector-type nonlinearity

- Stability criteria in the form of LMIs solved analytically by hand

- Reduction to Polynomial (frequency dependent) inequalities (small size)

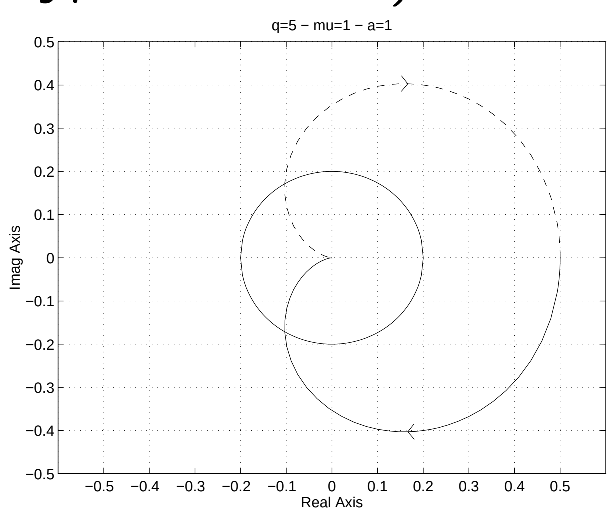
Some history (2)

1960s: Yakubovich, Popov, Kalman, Anderson et al obtained the positive real lemma

The linear system $\dot{x} = Ax + Bu, \quad y = Cx + Du$ is passive $H(s) + H(s)^* \geq 0 \ \forall \ s + s^* > 0$ iff

$$P \succ \mathbf{0} \quad \begin{bmatrix} A'P + PA & PB - C' \\ B'P - C & -D - D' \end{bmatrix} \preceq \mathbf{0}$$

- Solution via a simple graphical criterion (Popov, circle and Tsypkin criteria)



Mathieu equation: $\ddot{y} + 2\mu\dot{y} + (\mu^2 + a^2 - q\cos\omega_0 t)y = 0$
$q < 2\mu a$

Some history (3)

1971: Willems focused on solving algebraic Riccati equations (AREs)

$$A'P + PA - (PB + C')R^{-1}(B'P + C) + Q = \mathbf{0}$$

Numerical algebra

$$H = \begin{bmatrix} A - BR^{-1}C & BR^{-1}B' \\ -C'R^{-1}C & -A' + C'R^{-1}B' \end{bmatrix} \quad V = \begin{bmatrix} V_1 \\ V_2 \end{bmatrix}$$

$$P_{are} = V_2 V_1^{-1}$$

By 1971, methods for solving LMIs:

- Direct for small systems
- Graphical methods
- Solving Lyapunov or Riccati equations

Some history (4)

1963: Bellman-Fan: infeasibility criteria for multiple Lyapunov inequalities (duality theory)

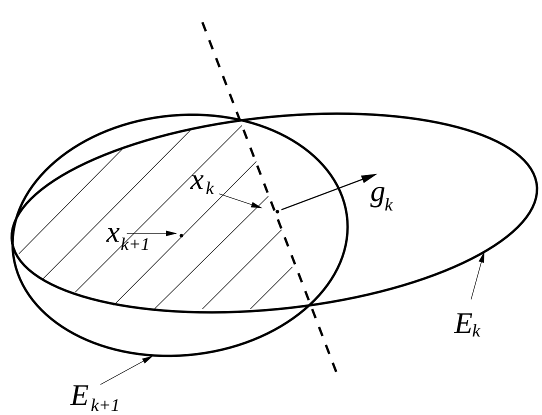
On Systems of Linear Inequalities in hermitian Matrix Variables

1975: Cullum-Donath-Wolfe: properties of criterion and algorithm for minimization of maximum eigenvalues

The minimization of certain nondifferentiable sums of eigenvalues of symmetric matrices

1979: Khachiyan: polynomial bound on worst case iteration count for LP ellipsoid algorithm

A polynomial algorithm in linear programming



Some history (5)

1981: Craven-Mond: Duality theory

Linear Programming with Matrix variables

1984: Karmarkar introduces interior-point (IP) methods for LP: improved complexity bound and efficiency

1985: Fletcher: Optimality conditions for non-differentiable optimization

Semidefinite matrix constraints in optimization

1988: Overton: Nondifferentiable optimization

On minimizing the maximum eigenvalue of a symmetric matrix

1988: Nesterov, Nemirovski, Alizadeh extend IP methods for convex programming

Interior-Point Polynomial Algorithms in Convex Programming

1990s: most papers on SDP are written (control theory, combinatorial optimization, approximation theory...)

LMI and SDP formalisms

In mathematical programming terminology LMI optimization = semidefinite programming (SDP)

LMI (SDP dual)

$$\begin{array}{ll} \text{min} & c'x \\ \text{under} & F_0 + \sum_{i=1}^{n} x_i F_i \prec \mathbf{0} \end{array}$$

SDP (primal)

$$\begin{array}{ll} \text{min} & -\text{Tr}(F_0 Z) \\ \text{under} & -\text{Tr}(F_i Z) = c_i \\ & Z \succeq \mathbf{0} \end{array}$$

$x \in \mathbb{R}^n$, $Z \in \mathbb{S}_m$, $F_i \in \mathbb{S}_m$, $c \in \mathbb{R}^n$, $\quad i = 1, \cdots, n$

Nota:

In a typical control LMI

$$A'P + PA = F_0 + \sum_{i=1}^{n} x_i F_i \prec \mathbf{0}$$

individual matrix entries are decision variables

LMI and SDP formalisms (2)

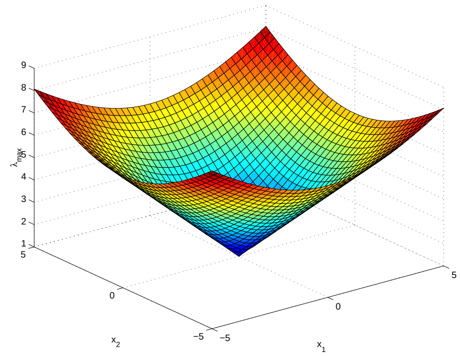
$$\exists\ x \in \mathbb{R}^n \mid \underbrace{F_0 + \sum_{i=1}^{n} x_i F_i}_{F(x)} \prec \mathbf{0} \quad \Leftrightarrow \quad \min_{x \in \mathbb{R}^n} \ \lambda_{max}(F(x))$$

The LMI feasibility problem is a convex and non differentiable optimization problem.

Example :

$$F(x) = \begin{bmatrix} -x_1 - 1 & -x_2 \\ -x_2 & -1 + x_1 \end{bmatrix}$$

$$\lambda_{max}(F(x)) = 1 + \sqrt{(x_1^2 + x_2^2)}$$



LMI and SDP formalisms (3)

$$\begin{array}{ll} \min & c'x \\ \text{s.t.} & \\ & b - A'x \in \mathcal{K} \end{array} \qquad \begin{array}{ll} \min & b'y \\ \text{s.t.} & \\ & Ay = c \\ & y \in \mathcal{K} \end{array}$$

Conic programming in cone $\mathcal{K}$

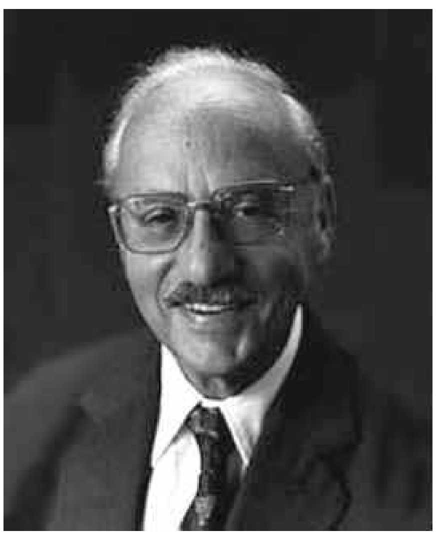
- positive orthant (LP)
- Lorentz (second-order) cone (SOCP)
- positive semidefinite cone (SDP)

Hierarchy: LP cone $\subset$ SOCP cone $\subset$ SDP cone

LMI and SDP formalisms (3)

LMI optimization = generalization of linear programming (LP) to cone of positive semidefinite matrices = semidefinite programming (SDP)

Linear programming pioneered by
• Dantzig and its simplex algorithm (1947, ranked in the top 10 algorithms by SIAM Review in 2000)
• Kantorovich (co-winner of the 1975 Nobel prize in economics)



George Dantzig
(1914 Portland, Oregon)

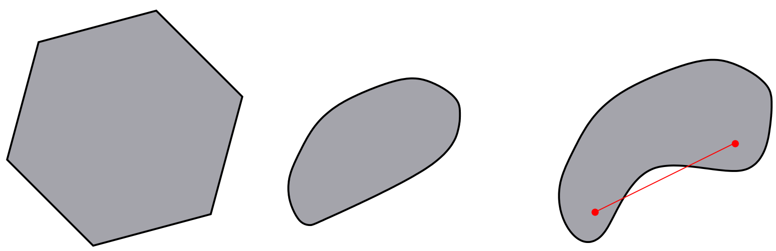
Leonid V Kantorovich
(1921 St Petersburg - 1986)

Unfortunately, SDP has not reached maturity of LP or SOCP so far..

Mathematical preliminaries

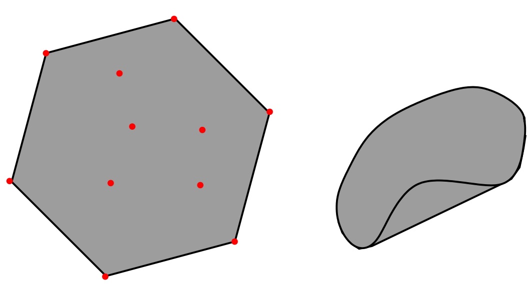
A set $\mathcal{C}$ is convex if the line segment between any two points in $\mathcal{C}$ lies in $\mathcal{C}$

$$\forall\ x_1, x_2 \in \mathcal{C} \quad \lambda x_1 + (1-\lambda)x_2 \in \mathcal{C} \quad \forall\lambda \quad 0 \leq \lambda \leq 1$$



The convex hull of a set $\mathcal{C}$ is the set of all convex combinations of points in $\mathcal{C}$

$$\mathrm{co}\,\mathcal{C} = \{\sum_i \lambda_i x_i\ :\ x_i \in \mathcal{C} \quad \lambda_i \geq 0 \quad \sum_i \lambda_i = 1\}$$



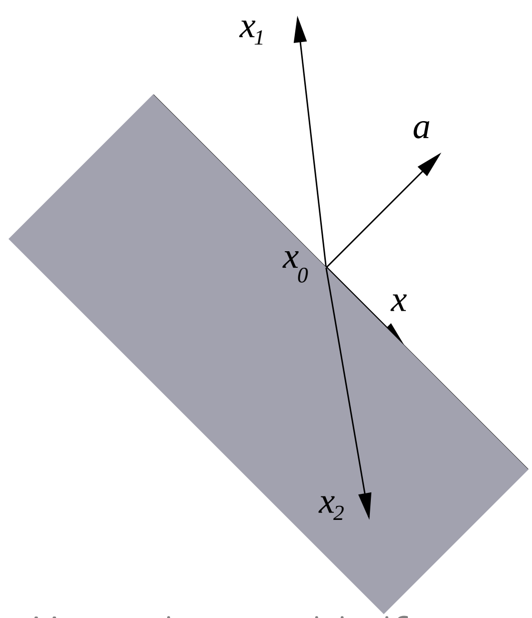
Mathematical preliminaries (2)

A hyperplane is a set of the form:

$$\mathcal{H} = \left\{ x \in \mathbb{R}^n \mid a'(x - x_0) = 0 \right\} \quad a \neq 0 \in \mathbb{R}^n$$

A hyperplane divides $\mathbb{R}^n$ into two halfspaces:

$$\mathcal{H}_- = \left\{ x \in \mathbb{R}^n \mid a'(x - x_0) \leq 0 \right\} \quad a \neq 0 \in \mathbb{R}^n$$



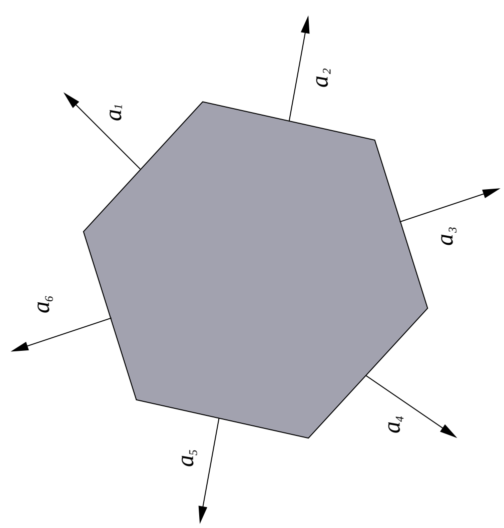
Hyperplane and halfspace
$x \in \mathcal{H}$, $x_1 \notin \mathcal{H}_-$, $x_2 \in \mathcal{H}_-$

Mathematical preliminaries (3)

A polyhedron is defined by a finite number of linear equalities and inequalities

$$\begin{aligned}\mathcal{P} &= \left\{x \in \mathbb{R}^n : \ a_j' x \le b_j, j = 1, \cdots, m, c_i' x = d_i, i = 1, \cdots, p\right\} \\ &= \{x \in \mathbb{R}^n \ : \ Ax \preceq b, \ Cx = d\}\end{aligned}$$

A bounded polyhedron is a polytope



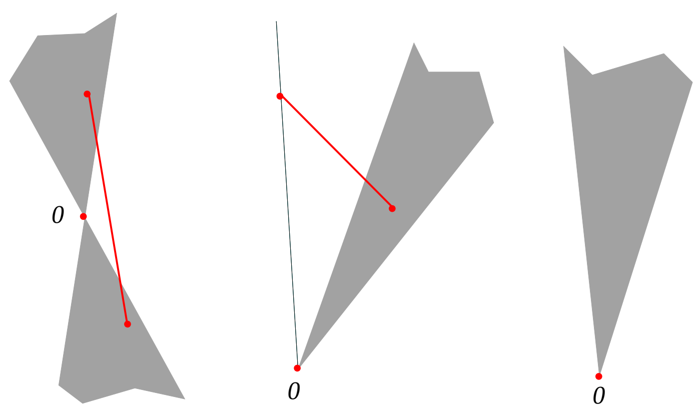
Polytope as an intersection of halfspaces

- positive orthant is a polyhedral cone
- k-dimensional simplexes in $\mathbb{R}^n$

$$\mathcal{X} = \mathrm{co}\{v_0, \cdots, v_k\} = \left\{ \sum_{i=0}^{k} \lambda_i v_i \ \lambda_i \ge 0 \ \sum_{i=0}^{k} \lambda_i = 1 \right\}$$

Mathematical preliminaries (4)

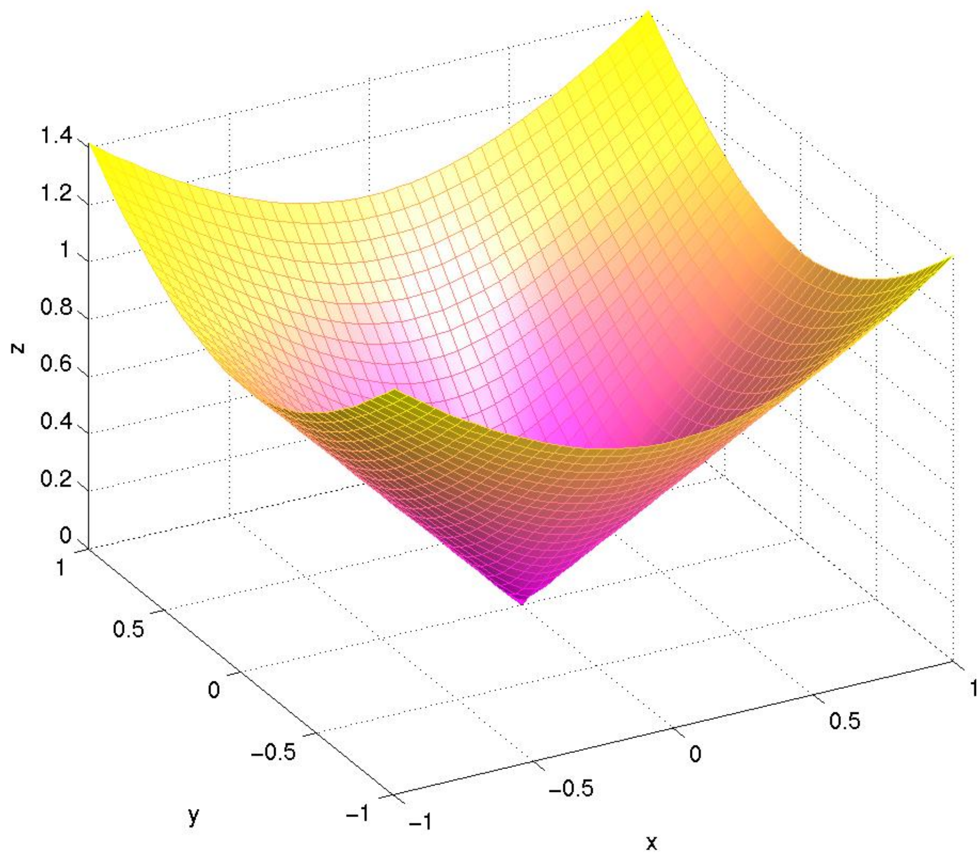
A set $\mathcal{K}$ is a cone if for every $x \in \mathcal{K}$ and $\lambda \geq 0$ we have $\lambda x \in \mathcal{K}$. A set $\mathcal{K}$ is a convex cone if it is convex and a cone



$\mathcal{K} \subseteq \mathbb{R}^n$ is called a proper cone if it is a closed solid pointed convex cone

$$a \in \mathcal{K} \quad \text{and} \quad -a \in \mathcal{K} \quad \Rightarrow \quad a = 0$$

Lorentz cone $\mathbb{L}^n$

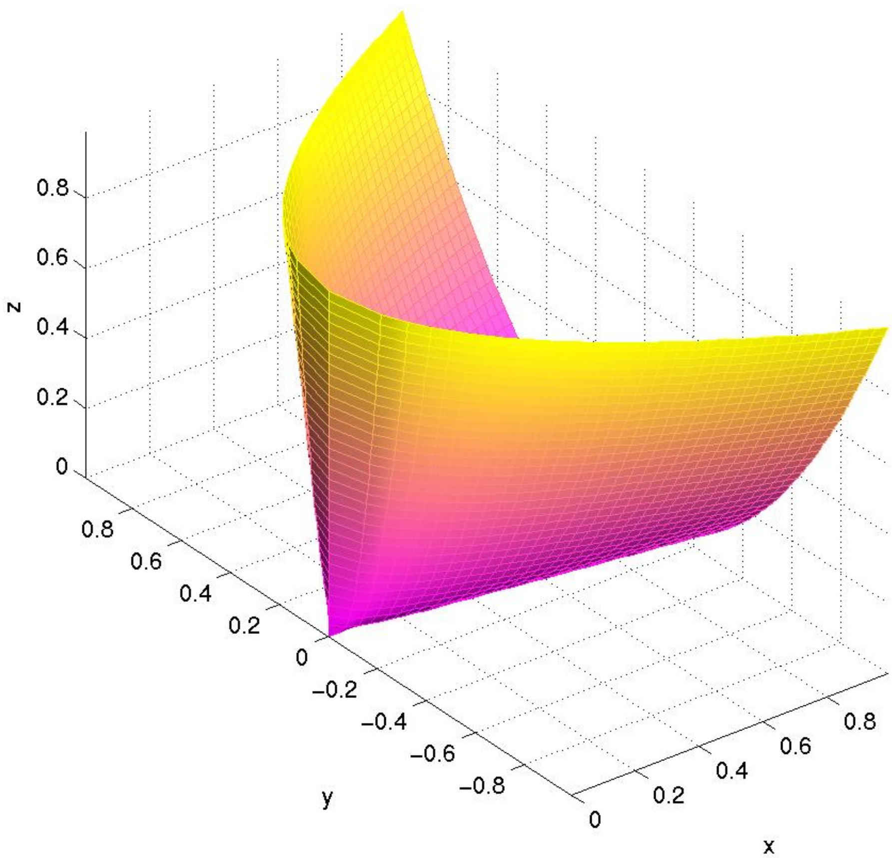


3D Lorentz cone or ice-cream cone

$$x^2 + y^2 \leq z^2 \quad z \geq 0$$

arises in quadratic programming

PSD cone $\mathbb{S}^n_+$



2D positive semidefinite cone

$$\begin{bmatrix} x & y \\ y & z \end{bmatrix} \succeq 0 \iff x \geq 0 \quad z \geq 0 \quad xz \geq y^2$$

arises in semidefinite programming

Mathematical preliminaries (5)

Every proper cone $\mathcal{K}$ in $\mathbb{R}^n$ induces a partial ordering $\geq_{\mathcal{K}}$ defining generalized inequalities on $\mathbb{R}^n$

$$a \geq_{\mathcal{K}} b \quad \Leftrightarrow \quad a - b \in \mathcal{K}$$

The positive orthant, the Lorentz cone and the PSD cone are all proper cones

- positive orthant $\mathbb{R}^n_+$: standard coordinatewise ordering (LP)

$$x \geq_{\mathbb{R}^n_+} y \;\Leftrightarrow\; x_i \geq y_i$$

- Lorentz cone $\mathbb{L}^n$

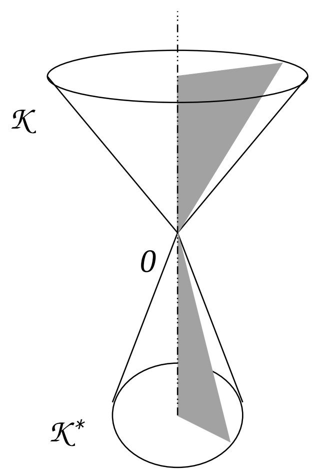
$$x_n \geq \sqrt{\sum_{i=1}^{n-1} x_i^2}$$

- PSD cone $\mathbb{S}^n_+$: Löwner partial order

Mathematical preliminaries (6)

The set $\mathcal{K}^* = \{y \in \mathbb{R}^n \mid x'y \le 0 \;\; \forall\, x \in \mathcal{K}\}$ is called the dual cone of the cone $\mathcal{K}$

- Revolution cone $\mathcal{K}_s(\theta) = \{x \in \mathbb{R}^n : \;\; s'x \le ||x|| \cos\theta\}$



$\mathcal{K}_s(\theta)^* = \mathcal{K}_{-s}(\frac{\pi}{2} - \theta)$

- $(\mathbb{R}^n_+)^* = \mathbb{R}^n_-$

$\mathcal{K}^*$ is closed and convex, $\mathcal{K}_1 \subseteq \mathcal{K}_2 \;\Rightarrow\; \mathcal{K}_2^* \subseteq \mathcal{K}_1^*$

$\preceq_{\mathcal{K}^*}$ is a dual generalized inequality

$$x \preceq_{\mathcal{K}} y \quad \Leftrightarrow \quad \lambda' x \le \lambda' y \;\; \forall\, \lambda \succeq_{\mathcal{K}^*} 0$$

Mathematical preliminaries (6)

$f\ :\ \mathbb{R}^n \rightarrow \mathbb{R}$ is convex if dom$f$ is a convex set and $\forall\ x,\ y \in$ dom$f$ and $0 \leq \lambda \leq 1$

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y)$$

If f is differentiable: dom$f$ is a convex set and $\forall\ x,\ y \in$ dom$f$

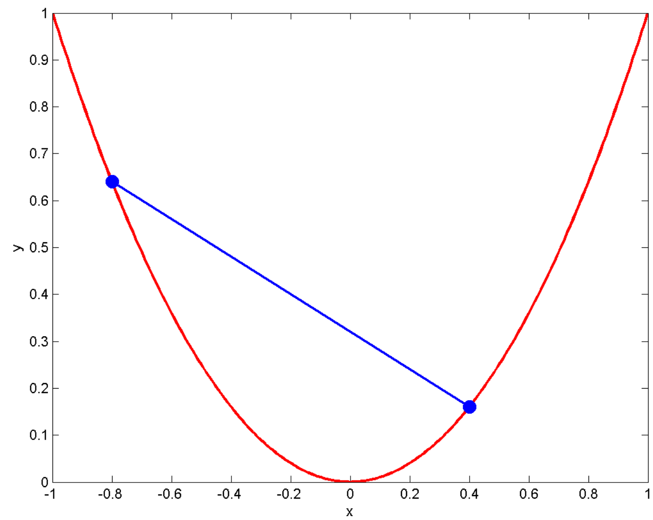
$$f(y) \geq f(x) + \nabla f(x)'(y-x)$$

If f is twice differentiable: dom$f$ is a convex set and $\forall\ x,\ y \in$ dom$f$

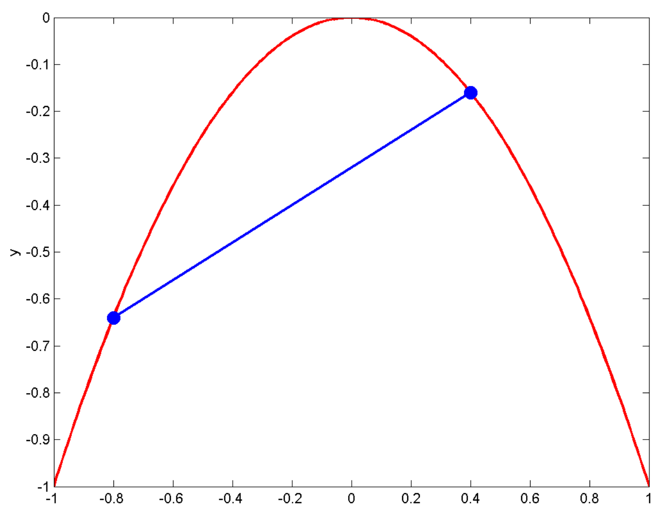
$$\nabla^2 f(x) \succeq \mathbf{0}$$

Quadratic functions:
$f(x) = (1/2)x'Px + q'x + r$ is convex if and only if $P \succeq \mathbf{0}$

Convex function $y = x^2$



Nonconvex function $y = -x^2$



Mind the sign !

Applications of SDP

- control systems (part II of the course)
- robust optimization
- signal processing
- synthesis of antennae arrays
- design of chips
- structural design (trusses)
- geometry (ellipsoids)
- graph theory and combinatorics (MAXCUT, Shannon capacity)

and many others...

See Helmberg's page on SDP

`www-user.tu-chemnitz.de/~helmberg/semidef.html`

Robust optimization

In real-life optimization problems:

- exact values of input data (constraints) are seldom known
- variables may be implemented with errors

Observation: problems are uncertain

Case study by Ben Tal and Nemirovski:
90 LP problems from NETLIB + uncertainty
Conclusion: small errors in data can have strong impact on optimality and even feasibility of optimal solutions

Remedy: robust optimization, with robustly feasible solutions guaranteed to remain feasible

Robust optimization (2)

Optimization problem in conic form:

$$\begin{array}{ll} \text{max} & b'y \\ \text{s.t.} & c - A'y \in \mathcal{K} \end{array}$$

Assume $A \in \mathcal{U}$, uncertainty set

Robust conic problem:

$$\begin{array}{lll} \text{max} & b'y & \\ \text{s.t.} & \forall A \in \mathcal{U}, & c - A'y \in \mathcal{K} \end{array}$$

Still convex, but depending on the structure of $\mathcal{U}$, can be much harder that original conic problem

Robust optimization (3)

Problem	Uncertainty	ellipsoid	LMI
LP		SOCP	SDP
SOCP		SDP	hard
SDP		hard	hard

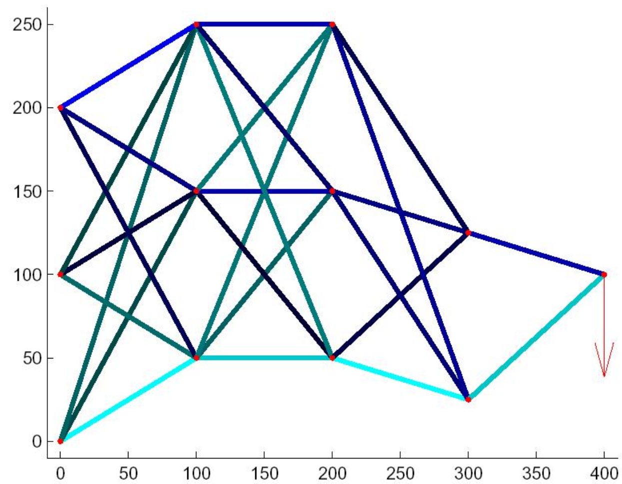
Examples (Laurent El Ghaoui):

LP+ellipse: robust portfolio design in finance

SOCP+ellipse: robust least-squares in identification

Robust SDP can be approximated by SDP

Truss topology design



Connect N nodes with bars of length l_i (fixed) and cross-sections s_i (to be designed)

Construction reacts to external force f (fixed) on each node with displacement vector d satisfying $A(s)d = f$

Linearity assumption: stiffness matrix $A(s)$ affine in s and positive definite

Goal: maximize stiffness = minimize elastic stored energy $f^T d$ s.t. bounds $a \leq s \leq b$ on cross-section and $l^T s = \sum_i l_i s_i \leq v$ on total volume (weight)

Truss topology design (2)

Can be formulated as an LMI

$$\min \gamma \quad \text{s.t.} \quad \begin{bmatrix} \gamma & f^T \\ f & A(\gamma) \end{bmatrix} \succ 0 \quad l^T s \leq v \quad a \leq s \leq b$$

Optimal truss (Scherer)