Convex Optimization Theory and Applications

Topic 3 - Convex Programming Problems

Li Li

Department of Automation
Tsinghua University

Fall, 2009-2021.

3.0. Outline

3.1. Definition 基本定义

3.2. Examples 例子

3.3. Multicriterion Optimization 多目标规划

3.1. Definitions

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \ldots, m \\ & h_i(x) = 0, \quad i = 1, \ldots, p \end{array}$$

- $x \in \mathbf{R}^n$ is the optimization variable
- $f_0 : \mathbf{R}^n \to \mathbf{R}$ is the objective or cost function
- $f_i : \mathbf{R}^n \to \mathbf{R}$, $i = 1, \ldots, m$, are the inequality constraint functions
- $h_i : \mathbf{R}^n \to \mathbf{R}$ are the equality constraint functions

optimal value:

$$p^\star = \inf\{f_0(x) \mid f_i(x) \leq 0,\ i = 1, \ldots, m,\ h_i(x) = 0,\ i = 1, \ldots, p\}$$

- $p^\star = \infty$ if problem is infeasible (no x satisfies the constraints)
- $p^\star = -\infty$ if problem is unbounded below

3.1. Definitions

x is **feasible** if $x \in \mathbf{dom}\, f_0$ and it satisfies the constraints

a feasible x is **optimal** if $f_0(x) = p^\star$; X_{opt} is the set of optimal points

x is **locally optimal** if there is an $R > 0$ such that x is optimal for

$$
\begin{array}{ll}
\text{minimize (over } z) & f_0(z) \\
\text{subject to} & f_i(z) \le 0, \quad i = 1, \ldots, m, \quad h_i(z) = 0, \quad i = 1, \ldots, p \\
& \|z - x\|_2 \le R
\end{array}
$$

examples (with $n = 1$, $m = p = 0$)

- $f_0(x) = 1/x$, $\mathbf{dom}\, f_0 = \mathbf{R}_{++}$: $p^\star = 0$, no optimal point
- $f_0(x) = -\log x$, $\mathbf{dom}\, f_0 = \mathbf{R}_{++}$: $p^\star = -\infty$
- $f_0(x) = x \log x$, $\mathbf{dom}\, f_0 = \mathbf{R}_{++}$: $p^\star = -1/e$, $x = 1/e$ is optimal
- $f_0(x) = x^3 - 3x$, $p^\star = -\infty$, local optimum at $x = 1$

3.1. Definitions

the standard form optimization problem has an **implicit constraint**

$$x \in \mathcal{D} = \bigcap_{i=0}^{m} \mathbf{dom}\, f_i \;\cap\; \bigcap_{i=1}^{p} \mathbf{dom}\, h_i,$$

- we call $\mathcal{D}$ the **domain** of the problem
- the constraints $f_i(x) \leq 0$, $h_i(x) = 0$ are the explicit constraints
- a problem is **unconstrained** if it has no explicit constraints ($m = p = 0$)

example:

$$\text{minimize} \quad f_0(x) = -\textstyle\sum_{i=1}^{k} \log(b_i - a_i^T x)$$

is an unconstrained problem with implicit constraints $a_i^T x < b_i$

3.1. Definitions

非凸规划问题有可能转化为凸规划问题

$$
\begin{array}{ll}
\text{minimize} & f_0(x) = x_1^2 + x_2^2 \\
\text{subject to} & f_1(x) = x_1/(1 + x_2^2) \leq 0 \\
& h_1(x) = (x_1 + x_2)^2 = 0
\end{array}
$$

- f_0 is convex; feasible set $\{(x_1, x_2) \mid x_1 = -x_2 \leq 0\}$ is convex
- not a convex problem (according to our definition): f_1 is not convex, h_1 is not affine
- equivalent (but not identical) to the convex problem

$$
\begin{array}{ll}
\text{minimize} & x_1^2 + x_2^2 \\
\text{subject to} & x_1 \leq 0 \\
& x_1 + x_2 = 0
\end{array}
$$

3.1. Definitions

假设我们考虑求最小值的优化问题。

局部最优点的定义：如果一个点 x 服从优化问题的约束，且对于其邻域内所有服从优化问题约束的点 z，都满足 $f(x) \leq f(z)$，那么点 x 可以被称为局部最优点。

全局最优点的定义：如果一个点 x 服从优化问题的约束，且对于所有服从优化问题约束的点 z 都存在 $f(x) \leq f(z)$，那么点 x 可以被称为全局最优点。

凸优化问题的局部最优解也是全局最优解。

3.1. Definitions

any locally optimal point of a convex problem is (globally) optimal

proof: suppose x is locally optimal and y is optimal with $f_0(y) < f_0(x)$

x locally optimal means there is an $R > 0$ such that

$$z \text{ feasible}, \quad \|z - x\|_2 \leq R \quad \Longrightarrow \quad f_0(z) \geq f_0(x)$$

consider $z = \theta y + (1 - \theta)x$ with $\theta = R/(2\|y - x\|_2)$

- $\|y - x\|_2 > R$, so $0 < \theta < 1/2$
- z is a convex combination of two feasible points, hence also feasible
- $\|z - x\|_2 = R/2$ and

$$f_0(z) \leq \theta f_0(x) + (1 - \theta) f_0(y) < f_0(x)$$

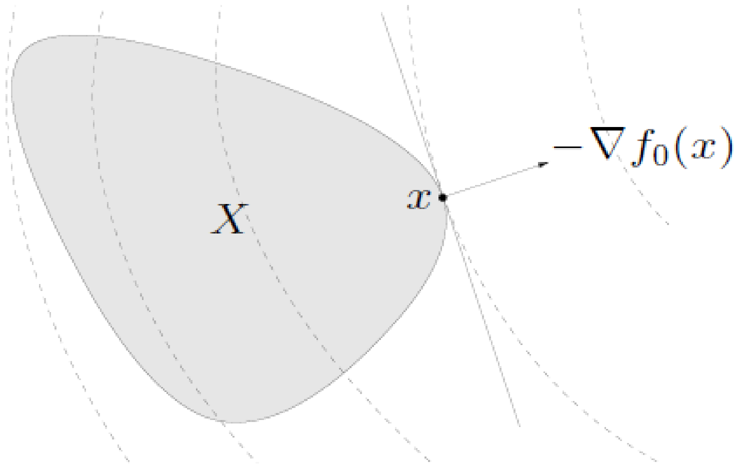
which contradicts our assumption that x is locally optimal

3.1. Definitions

Optimality criterion for differentiable f_0

x is optimal if and only if it is feasible and

$$\nabla f_0(x)^T(y - x) \geq 0 \quad \text{for all feasible } y$$



if nonzero, $\nabla f_0(x)$ defines a supporting hyperplane to feasible set X at x

3.1. Definitions

- **unconstrained problem**: x is optimal if and only if

$$x \in \mathbf{dom}\, f_0, \qquad \nabla f_0(x) = 0$$

- **equality constrained problem**

$$\text{minimize} \quad f_0(x) \quad \text{subject to} \quad Ax = b$$

x is optimal if and only if there exists a ν such that

$$x \in \mathbf{dom}\, f_0, \qquad Ax = b, \qquad \nabla f_0(x) + A^T \nu = 0$$

- **minimization over nonnegative orthant**

$$\text{minimize} \quad f_0(x) \quad \text{subject to} \quad x \succeq 0$$

x is optimal if and only if

$$x \in \mathbf{dom}\, f_0, \qquad x \succeq 0, \qquad \begin{cases} \nabla f_0(x)_i \geq 0 & x_i = 0 \\ \nabla f_0(x)_i = 0 & x_i > 0 \end{cases}$$

3.1. Definitions

Equivalent convex problems

two problems are (informally) **equivalent** if the solution of one is readily obtained from the solution of the other, and vice-versa

some common transformations that preserve convexity:

- **eliminating equality constraints**

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \ldots, m \\ & Ax = b \end{array}$$

  is equivalent to

$$\begin{array}{ll} \text{minimize (over } z\text{)} & f_0(Fz + x_0) \\ \text{subject to} & f_i(Fz + x_0) \leq 0, \quad i = 1, \ldots, m \end{array}$$

  where F and x_0 are such that

$$Ax = b \quad \Longleftrightarrow \quad x = Fz + x_0 \text{ for some } z$$

3.1. Definitions

- **introducing equality constraints**

$$\begin{array}{ll} \text{minimize} & f_0(A_0x + b_0) \\ \text{subject to} & f_i(A_ix + b_i) \leq 0, \quad i = 1, \dots, m \end{array}$$

is equivalent to

$$\begin{array}{ll} \text{minimize (over } x,\ y_i\text{)} & f_0(y_0) \\ \text{subject to} & f_i(y_i) \leq 0, \quad i = 1, \dots, m \\ & y_i = A_ix + b_i, \quad i = 0, 1, \dots, m \end{array}$$

- **introducing slack variables for linear inequalities**

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & a_i^T x \leq b_i, \quad i = 1, \dots, m \end{array}$$

is equivalent to

$$\begin{array}{ll} \text{minimize (over } x,\ s\text{)} & f_0(x) \\ \text{subject to} & a_i^T x + s_i = b_i, \quad i = 1, \dots, m \\ & s_i \geq 0, \quad i = 1, \dots m \end{array}$$

3.1. Definitions

- **epigraph form**: standard form convex problem is equivalent to

$$\begin{array}{ll} \text{minimize (over } x, t) & t \\ \text{subject to} & f_0(x) - t \leq 0 \\ & f_i(x) \leq 0, \quad i = 1, \ldots, m \\ & Ax = b \end{array}$$

- **minimizing over some variables**

$$\begin{array}{ll} \text{minimize} & f_0(x_1, x_2) \\ \text{subject to} & f_i(x_1) \leq 0, \quad i = 1, \ldots, m \end{array}$$

is equivalent to

$$\begin{array}{ll} \text{minimize} & \tilde{f}_0(x_1) \\ \text{subject to} & f_i(x_1) \leq 0, \quad i = 1, \ldots, m \end{array}$$

where $\tilde{f}_0(x_1) = \inf_{x_2} f_0(x_1, x_2)$

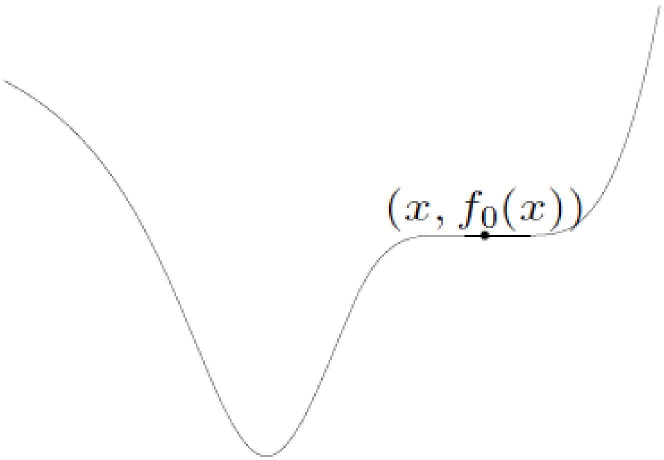
3.1. Definitions

Quasiconvex optimization

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \ldots, m \\ & Ax = b \end{array}$$

with $f_0 : \mathbf{R}^n \to \mathbf{R}$ quasiconvex, $f_1, \ldots, f_m$ convex

can have locally optimal points that are not (globally) optimal



3.1. Definitions

convex representation of sublevel sets of f_0

if f_0 is quasiconvex, there exists a family of functions ϕ_t such that:

- $\phi_t(x)$ is convex in x for fixed t
- t-sublevel set of f_0 is 0-sublevel set of ϕ_t, $i.e.$,

$$f_0(x) \leq t \quad \Longleftrightarrow \quad \phi_t(x) \leq 0$$

example

$$f_0(x) = \frac{p(x)}{q(x)}$$

with p convex, q concave, and $p(x) \geq 0$, $q(x) > 0$ on $\mathbf{dom}\, f_0$

can take $\phi_t(x) = p(x) - tq(x)$:

- for $t \geq 0$, ϕ_t convex in x
- $p(x)/q(x) \leq t$ if and only if $\phi_t(x) \leq 0$

3.1. Definitions

quasiconvex optimization via convex feasibility problems

$$\phi_t(x) \le 0, \qquad f_i(x) \le 0, \quad i = 1, \dots, m, \qquad Ax = b \qquad (1)$$

- for fixed t, a convex feasibility problem in x
- if feasible, we can conclude that $t \ge p^\star$; if infeasible, $t \le p^\star$

Bisection method for quasiconvex optimization

given $l \le p^\star$, $u \ge p^\star$, tolerance $\epsilon > 0$.
repeat
1. $t := (l + u)/2$.
2. Solve the convex feasibility problem (1).
3. **if** (1) is feasible, $u := t$; **else** $l := t$.

until $u - l \le \epsilon$.

---

requires exactly $\lceil \log_2((u - l)/\epsilon) \rceil$ iterations (where u, l are initial values)

3.1. Definitions

In a finite sequence of real numbers the sum of any seven successive terms is negative, and the sum of any eleven successive terms is positive. Please determine the maximum number of terms in the sequence. (IMO 1977)

We arrange it into a matrix of size 11*7:

$$\begin{pmatrix} a_1 & a_2 & \cdots & a_7 \\ a_2 & a_3 & \cdots & a_8 \\ \vdots & \vdots & \vdots & \vdots \\ a_{11} & a_{12} & \cdots & a_{17} \end{pmatrix}$$

Sum all elements of this matrix by rows and we get:

$$S = \sum_{i=1}^{7} a_i + \sum_{i=2}^{8} a_i + \cdots + \sum_{i=11}^{17} a_i < 0$$

Sum all elements of this matrix by columns and we get:

$$S = \sum_{i=1}^{11} a_i + \sum_{i=2}^{12} a_i + \cdots + \sum_{i=7}^{17} a_i > 0$$

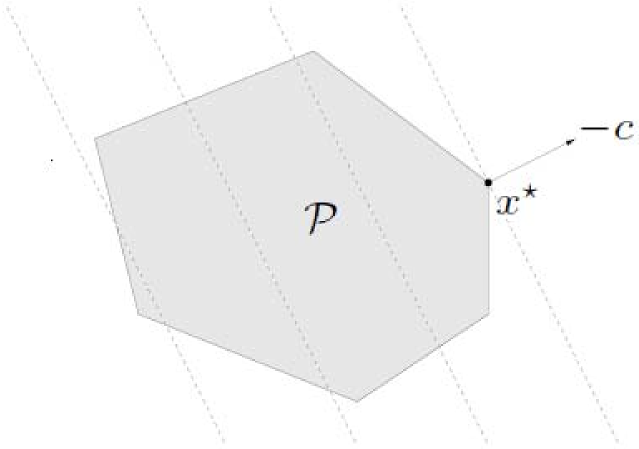
This contradiction implies that the number of terms is no larger than 16. Explore more in https://core.ac.uk/download/pdf/196617294.pdf

3.2. Examples

线性规划问题：

$$\begin{aligned} \min \quad & z = c^T x + d \\ \text{s.t.} \quad & Ax = b \\ & Gx \le h \end{aligned}$$

其中 $x \in R^n$， $c \in R^n$， $b \in R^m$， $A \in R^{m \times n}$。affine objective and constraints (feasible set is a polyhedron)



3.2. Examples

(Generalized) linear-fractional program

$$
\begin{array}{ll}
\text{minimize} & f_0(x) \\
\text{subject to} & Gx \preceq h \\
& Ax = b
\end{array}
$$

linear-fractional program

$$
f_0(x) = \frac{c^T x + d}{e^T x + f}, \qquad \mathbf{dom}\, f_0(x) = \{x \mid e^T x + f > 0\}
$$

- a quasiconvex optimization problem; can be solved by bisection
- also equivalent to the LP (variables y, z)

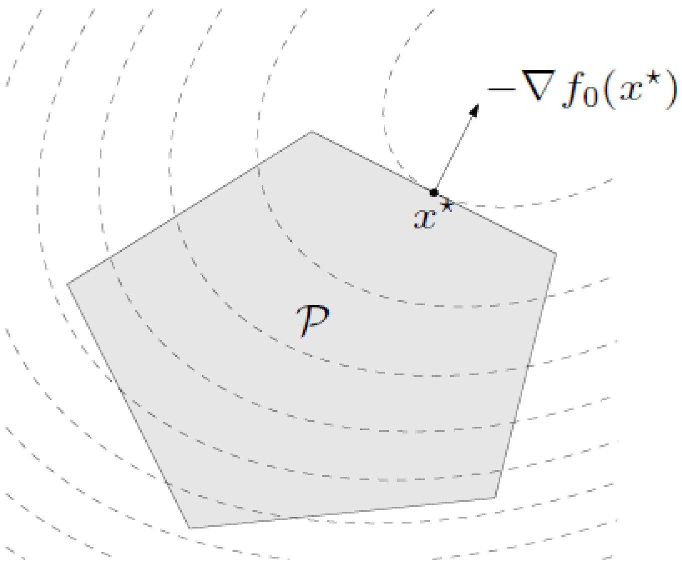
$$
\begin{array}{ll}
\text{minimize} & c^T y + dz \\
\text{subject to} & Gy \preceq hz \\
& Ay = bz \\
& e^T y + fz = 1 \\
& z \geq 0
\end{array}
$$

3.2. Examples

Quadratic program (QP)

$$\begin{array}{ll} \text{minimize} & (1/2)x^T P x + q^T x + r \\ \text{subject to} & Gx \preceq h \\ & Ax = b \end{array}$$

- $P \in \mathbf{S}^n_+$, so objective is convex quadratic
- minimize a convex quadratic function over a polyhedron



C

3.2. Examples

Quadratically constrained quadratic program (QCQP)

$$
\begin{array}{ll}
\text{minimize} & (1/2)x^T P_0 x + q_0^T x + r_0 \\
\text{subject to} & (1/2)x^T P_i x + q_i^T x + r_i \leq 0, \quad i = 1, \ldots, m \\
& Ax = b
\end{array}
$$

- $P_i \in \mathbf{S}_+^n$; objective and constraints are convex quadratic
- if $P_1, \ldots, P_m \in \mathbf{S}_{++}^n$, feasible region is intersection of m ellipsoids and an affine set

3.2. Examples

Second-order cone programming

$$\begin{array}{ll} \text{minimize} & f^T x \\ \text{subject to} & \|A_i x + b_i\|_2 \leq c_i^T x + d_i, \quad i = 1, \ldots, m \\ & Fx = g \end{array}$$

$(A_i \in \mathbf{R}^{n_i \times n}$, $F \in \mathbf{R}^{p \times n})$

- inequalities are called second-order cone (SOC) constraints:

$$(A_i x + b_i, c_i^T x + d_i) \in \text{second-order cone in } \mathbf{R}^{n_i+1}$$

- for $n_i = 0$, reduces to an LP; if $c_i = 0$, reduces to a QCQP
- more general than QCQP and LP

3.2. Examples

Geometric programming

monomial function

$$f(x) = cx_1^{a_1} x_2^{a_2} \cdots x_n^{a_n}, \qquad \mathbf{dom}\, f = \mathbf{R}^n_{++}$$

with $c > 0$; exponent α_i can be any real number

posynomial function: sum of monomials

$$f(x) = \sum_{k=1}^{K} c_k x_1^{a_{1k}} x_2^{a_{2k}} \cdots x_n^{a_{nk}}, \qquad \mathbf{dom}\, f = \mathbf{R}^n_{++}$$

geometric program (GP)

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 1, \quad i = 1, \ldots, m \\ & h_i(x) = 1, \quad i = 1, \ldots, p \end{array}$$

with f_i posynomial, h_i monomial

3.2. Examples

change variables to $y_i = \log x_i$, and take logarithm of cost, constraints

- monomial $f(x) = cx_1^{a_1} \cdots x_n^{a_n}$ transforms to

$$\log f(e^{y_1}, \ldots, e^{y_n}) = a^T y + b \qquad (b = \log c)$$

- posynomial $f(x) = \sum_{k=1}^{K} c_k x_1^{a_{1k}} x_2^{a_{2k}} \cdots x_n^{a_{nk}}$ transforms to

$$\log f(e^{y_1}, \ldots, e^{y_n}) = \log \left(\sum_{k=1}^{K} e^{a_k^T y + b_k} \right) \qquad (b_k = \log c_k)$$

- geometric program transforms to convex problem

$$\begin{array}{ll} \text{minimize} & \log \left(\sum_{k=1}^{K} \exp(a_{0k}^T y + b_{0k}) \right) \\ \text{subject to} & \log \left(\sum_{k=1}^{K} \exp(a_{ik}^T y + b_{ik}) \right) \leq 0, \quad i = 1, \ldots, m \\ & Gy + d = 0 \end{array}$$

3.2. Examples

Generalized inequality constraints

convex problem with generalized inequality constraints

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \preceq_{K_i} 0, \quad i = 1, \dots, m \\ & Ax = b \end{array}$$

- $f_0 : \mathbf{R}^n \to \mathbf{R}$ convex; $f_i : \mathbf{R}^n \to \mathbf{R}^{k_i}$ K_i-convex w.r.t. proper cone K_i
- same properties as standard convex problem (convex feasible set, local optimum is global, etc.)

conic form problem: special case with affine objective and constraints

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Fx + g \preceq_K 0 \\ & Ax = b \end{array}$$

extends linear programming ($K = \mathbf{R}^m_+$) to nonpolyhedral cones

3.2. Examples

Semidefinite program (SDP)

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & x_1 F_1 + x_2 F_2 + \cdots + x_n F_n + G \preceq 0 \\ & Ax = b \end{array}$$

with F_i, $G \in \mathbf{S}^k$

- inequality constraint is called linear matrix inequality (LMI)
- includes problems with multiple LMI constraints: for example,

$$x_1 \hat{F}_1 + \cdots + x_n \hat{F}_n + \hat{G} \preceq 0, \qquad x_1 \tilde{F}_1 + \cdots + x_n \tilde{F}_n + \tilde{G} \preceq 0$$

is equivalent to single LMI

$$x_1 \begin{bmatrix} \hat{F}_1 & 0 \\ 0 & \tilde{F}_1 \end{bmatrix} + x_2 \begin{bmatrix} \hat{F}_2 & 0 \\ 0 & \tilde{F}_2 \end{bmatrix} + \cdots + x_n \begin{bmatrix} \hat{F}_n & 0 \\ 0 & \tilde{F}_n \end{bmatrix} + \begin{bmatrix} \hat{G} & 0 \\ 0 & \tilde{G} \end{bmatrix} \preceq 0$$

3.2. Examples

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

Symmetric matrices F_i are given, and decision variables x_i are to be found. It is actually an affine matrix constraint.

The convex feasible set defines a closed LMI set $\{X \in \mathbf{R}^n : F(x) \geq 0\}$, while the strict positive condition $F(x) > 0$ define a open LMI set.

LMI optimization can be regarded as a generalization of linear programming (LP) to cone of positive semidefinite matrices

3.2. Examples

Consider an example $A = \begin{bmatrix} -1 & 2 \\ 0 & -2 \end{bmatrix}$ $P = \begin{bmatrix} p_1 & p_2 \\ p_2 & p_3 \end{bmatrix}$. If require $A^T P + PA < 0$, $P > 0$ indicates that

$$\begin{bmatrix} -2p_1 & 2p_1 - 3p_2 \\ 2p_1 - 3p_2 & 4p_2 - 4p_3 \end{bmatrix} < 0, \quad \begin{bmatrix} p_1 & p_2 \\ p_2 & p_3 \end{bmatrix} > 0$$

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

3.2. Examples

Consider a linear difference equation (i.e. a discrete-time linear system) given by

$$x(k+1) = Ax(k), \quad x(0) = x_0 \tag{3.1}$$

It is well-known (and relatively simple to prove) that $x(k)$ converges to zero for all initial conditions x_0 if $|\lambda_i(A)| < 1$, $i = 1$, …, n..

There is another presentation of this spectral radius condition in terms of a quadratic Lyapunov function $V(x(k)) = x(k)^T Px(k)$.

3.2. Examples

$$\left|\lambda_i(A)\right| < 1 \;\; \forall i \Leftrightarrow \exists P > 0 \;\; A^T P A - P < 0 \tag{3.2}$$

3.2. Examples

$$\left|\lambda_i(A)\right| < 1 \ \forall i \Leftrightarrow \exists P > 0 \ \ A^T P A - P < 0 \tag{3.3}$$

Proof: $(\Leftarrow)$ Let $Av = \lambda v$, $0 > v^T (A^T P A - P) v = (\left|\lambda\right|^2 - 1)\underbrace{v^T P v}_{>0}$.

And therefore $\left|\lambda\right| < 1$

$(\Rightarrow)$ Let $P = \sum_{i=0}^{\infty} (A^i)^T Q A^i$ with $Q > 0$, the sum converges by the eigenvalue assumption.

$$A^T P A - P = \sum_{i=1}^{\infty} (A^i)^T Q A^i - \sum_{i=0}^{\infty} (A^i)^T Q A^i = -Q < 0 \tag{3.4}$$

3.2. Examples

By using the Schur complement, we have

$$\begin{bmatrix} P^{-1} & A \\ A^T & P \end{bmatrix} > 0 \qquad (3.5)$$

3.2. Examples

By using the Schur complement, we have

$$\begin{bmatrix} P^{-1} & A \\ A^T & P \end{bmatrix} > 0 \qquad (3.6)$$

Or more interestingly (why?)

$$\begin{bmatrix} P & PA \\ A^T P & P \end{bmatrix} > 0 \qquad (3.7)$$

3.2. Examples

Issai Schur

January 10, 1875 Mogilyov - January 10, 1941 Tel Aviv

3.2. Examples

Consider now the case where A is not stable, but we can use linear state feedback, i.e., $A(K) = A + BK$, where K is a fixed matrix. We want to find a matrix K such that $A + BK$ is stable, i.e., all its eigenvalues have absolute value smaller than one.

Use Schur complements to rewrite the condition:

$$(A + BK)^T P(A + BK) - P < 0, \quad P > 0 \tag{3.8}$$

Or in LMI formulation

3.2. Examples

$$\begin{bmatrix} P & (A+BK)^T P \\ P(A+BK) & P \end{bmatrix} > 0 \tag{3.9}$$

Condition is nonlinear in (P,K). However, we can apply a congruence transformation with $Q = P^{-1}$, and obtain

$$\begin{bmatrix} Q & Q(A+BK)^T \\ (A+BK)Q & Q \end{bmatrix} > 0 \tag{3.10}$$

3.2. Examples

Now, defining a new variable $Y = KQ$ we have

$$\begin{bmatrix} Q & QA^T + Y^T B^T \\ AQ + BY & Q \end{bmatrix} > 0 \tag{3.11}$$

This problem is now linear in (Q, Y). In fact, it is an SDP problem. After solving it, we can recover controller K via $K = Q^{-1}Y$.

Why we do not define $Y = PBK$?

3.2. Examples

Please discuss the continuous case

$$\frac{dx(t)}{dt} = Ax(t) \tag{3.12}$$

3.2. Examples

Please discuss the continuous case

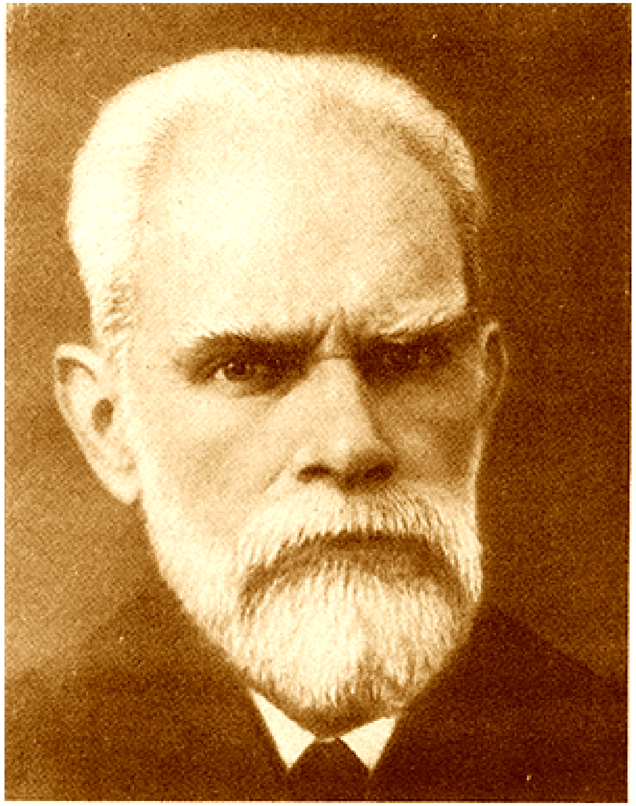
$$\frac{dx(t)}{dt} = Ax(t) \tag{3.13}$$

Clearly we have

$$PA + A^T P < 0\,, \quad P > 0 \tag{3.14}$$

Please prove the equivalence to the condition $\mathrm{Re}(A) < 0$, and consider the controller and observer for this contimuous system.

3.2. Examples



Aleksandr Mikhailovich Lyapunov
June 6, 1857 Yaroslavl - November 3, 1918 Odessa

3.2. Examples

The algebraic Riccati equation is either of the following matrix equations:

1) the continuous time algebraic Riccati equation (CARE):

$$A^T X + XA - XBB^T X + Q = 0 \tag{3.14}$$

2) the discrete time algebraic Riccati equation (DARE):

$$X = A^T XA - (A^T XB)(R + B^T XB)^{-1}(B^T XA) + Q \tag{3.15}$$

3.2. Examples

X is a unknown n by n symmetric matrix and A, B, Q, R and are known real coefficient matrices.

The name Riccati is given to the CARE equation by analogy to the Riccati differential equation: the unknown appears linearly and in a quadratic term (but no higher-order term). The DARE arises in place of the CARE when studying discrete time systems; it is not obviously related to the differential equation studied by Riccati.

If $X \geq 0$, $Q \leq 0$ and X and Q unknown, we can have the LMI type formulations as

3.2. Examples

Continuous time ARE

$$\begin{bmatrix} I & B^T X \\ XB & A^T X + XA \end{bmatrix} \geq 0 \tag{3.16}$$

Discrete time ARE

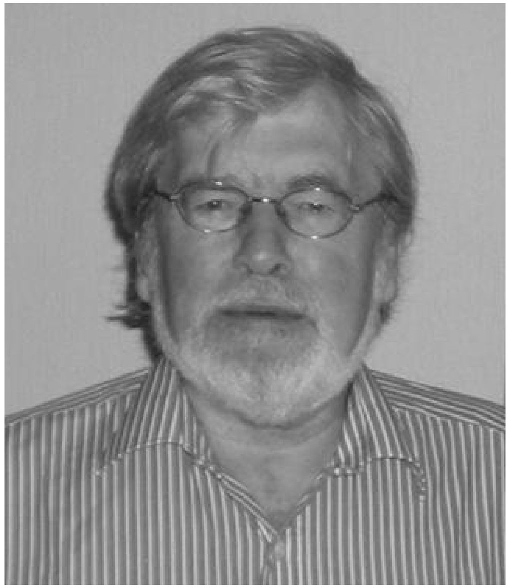
$$\begin{bmatrix} R + B^T XB & B^T XA \\ AXB & A^T XA - X \end{bmatrix} \geq 0 \tag{3.17}$$

3.2. Examples

Jacopo Francesco Riccati
May 28, 1676 Venice - April 15, 1754 Treviso

3.2. Examples

The terminology of LMI was introduced by Jan Willems in 1971. He once stated that "*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*".



Jan Willems
September 18, 1939, Bruges-August 31, 2013

3.2. Examples

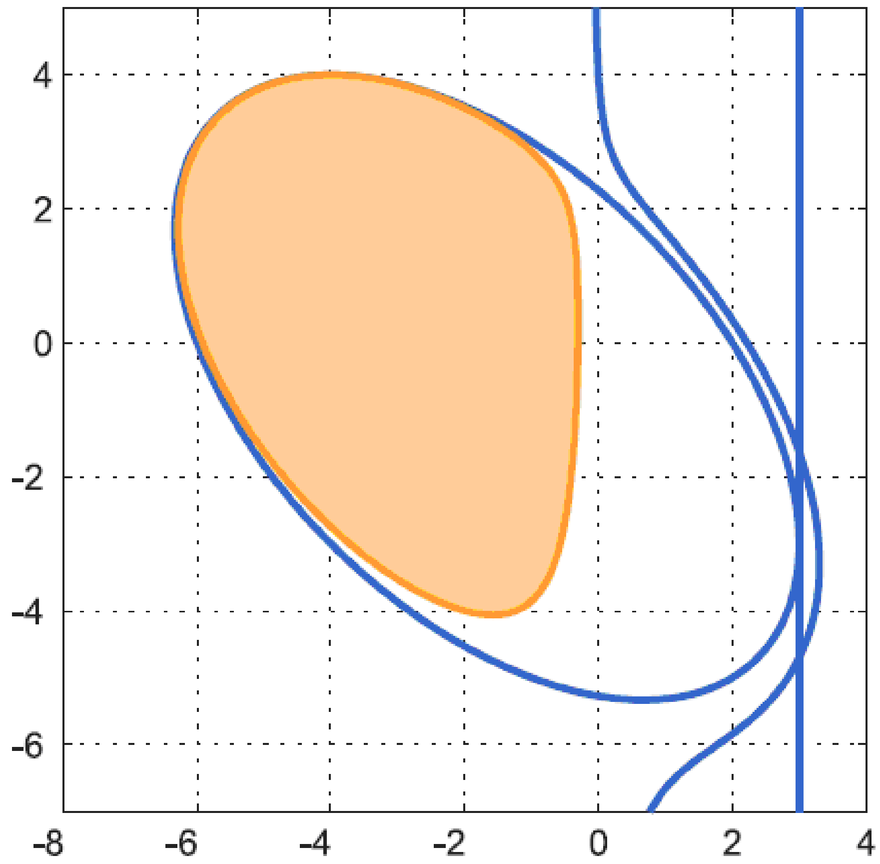
For example

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

is equivalent to the polynomial inequalities

$$0 < 3 - x_1$$
$$0 < (3-x_1)(4-x_2) - (x_1+x_2)^2$$
$$0 < -x_1((3-x_1)(4-x_2) - (x_1+x_2)^2) - (4-x_2)$$

3.2. Examples



Non-negative polynomials!

3.2. Examples

Hilbert's 17th problem is one of the 23 Hilbert problems set out in a celebrated list compiled in 1900 by David Hilbert. Original Hilbert's question was:

Given a multivariate polynomial that takes only non-negative values over the reals, can it be represented as a sum of squares of rational functions?

This was proved by Emil Artin in 1927. His result guaranteed the existence of such a finite representation. But the rational functions, in general, cannot be replaced by polynomials.

3.2. Examples

Any a uni-variate non-negtive polynomial function must be a sum of squares of polynomial functions. But even for a bi-/ tri-variate non-negtive polynomial function, we have the following cases

$$f(x,y) = (x^2 + y^2 - 3)x^2y^2 + 1$$

$$f(x,y,z) = z^6 + x^4y^2 + x^2y^4 - 3x^2y^2z^2$$

are non-negative over reals and yet which cannot be represented as a sum of squares of other polynomials.

3.2. Examples

Several sufficient conditions for a polynomial function to be a sum of squares of other polynomials were found. However, the necessary condition is yet unknown.

We are also interested in: *however every real nonnegative polynomial function can be approximated as closely as desired (in the* l_1 *-norm of its coefficient vector) by a sequence of polynomials that are sums of squares of polynomials?* No final conclusion had been given.

For an arbitrary polynomial function, determining whether it's convex is what's called NP-hard.

3.2. Examples

David Hilbert

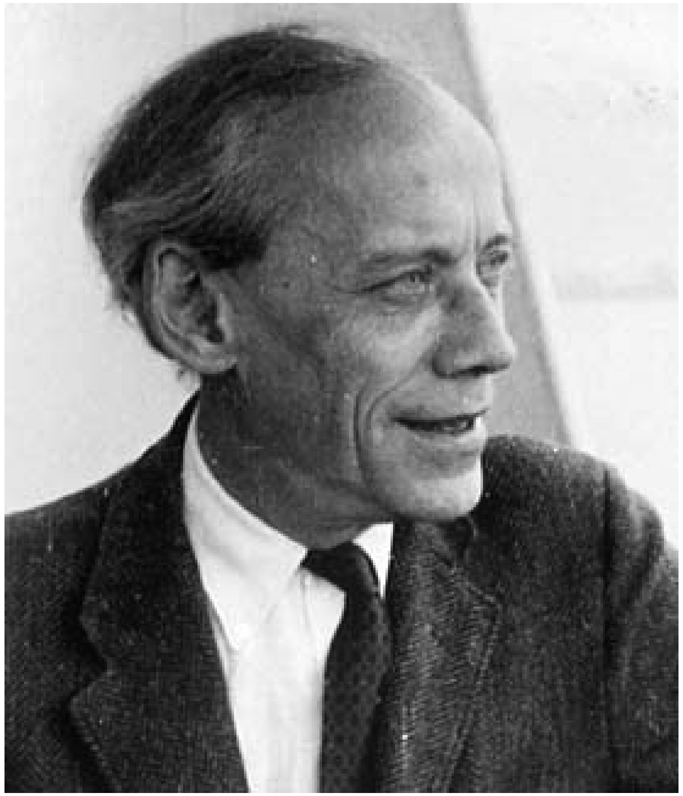
January 23, 1862 Königsberg - February 14, 1943 Göttingen

3.2. Examples

"*Wenn uns die Beantwortung eines mathematischen Problems nicht gelingen will, so liegt häufig der Grund darin, daß wir noch nicht den allgemeineren Gesichtspunkt erkannt haben, von dem aus das vorgelegte Problem nur als einzelnes Glied einer Kette verwandter Probleme erscheint. Nach Auffindung dieses Gesichtspunktes wird häufig nicht nur das vorgelegte Problem unserer Erforschung zugänglicher, sondern wir gelangen so zugleich in den Besitz einer Methode, die auf die verwandten Probleme anwendbar ist. …. Dieser Weg zur Auffindung allgemeiner Methoden ist gewiß der gangbarste und sicherste; denn wer, ohne ein bestimmtes Problem vor Auge zu haben, nach Methoden sucht, dessen Suchen ist meist vergeblich.*"

-David Hilbert

3.2. Examples



Emil Artin

March 3, 1898 Vienna - December 20, 1962 Hamburg

3.2. Examples

Prove or disprove that, for $x, y, z > 0$, $xyz = 1$, we have $\left(x^5 + y^5 + z^5\right)^2 \geq 3\left(x^7 + y^7 + z^7\right)$. (from the American Mathematical Monthly)

3.2. Examples

Prove or disprove that, for $x, y, z > 0$, $xyz = 1$, we have $\left(x^5 + y^5 + z^5\right)^2 \geq 3\left(x^7 + y^7 + z^7\right)$.

Without loss of generality, assume the values are named such that $x \geq y \geq z > 0$. Indeed, we have

$$xy(x-y)^4(x+y)^2\left(x^2+y^2\right)+yz(y-z)^4(y+z)^2\left(y^2+z^2\right)$$

$$+xz(x-z)^4(x+z)^2\left(x^2+z^2\right)+(x-y)^2\left[\left(2z^4-x^4\right)^2+\left(2y^4-x^4\right)^2\right]/2$$

$$+(y-z)^2\left[\left(2x^4-z^4\right)^2+\left(2y^4-z^4\right)^2\right]/2$$

$$+(x-y)(y-z)\left[\left(2z^4-x^4\right)2 \ +3\left(y^8-z^8\right)\right]\geq 0$$

3.2. Examples

LP and equivalent SDP

LP: minimize $c^T x$
subject to $Ax \preceq b$

SDP: minimize $c^T x$
subject to $\mathbf{diag}(Ax - b) \preceq 0$

(note different interpretation of generalized inequality $\preceq$)

SOCP and equivalent SDP

SOCP: minimize $f^T x$
subject to $\|A_i x + b_i\|_2 \leq c_i^T x + d_i, \quad i = 1, \ldots, m$

SDP: minimize $f^T x$
subject to

$$\begin{bmatrix} (c_i^T x + d_i)I & A_i x + b_i \\ (A_i x + b_i)^T & c_i^T x + d_i \end{bmatrix} \succeq 0, \quad i = 1, \ldots, m$$

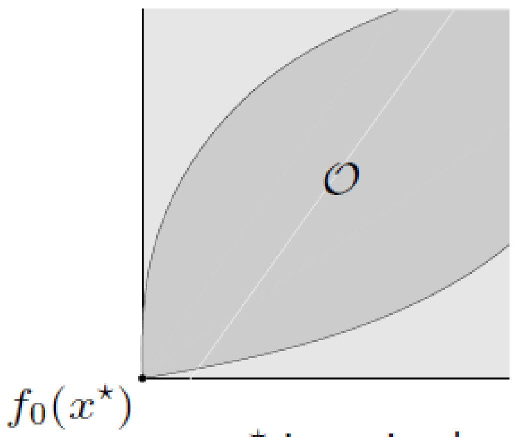
3.3. Multicriterion Optimization

Optimal and Pareto optimal points

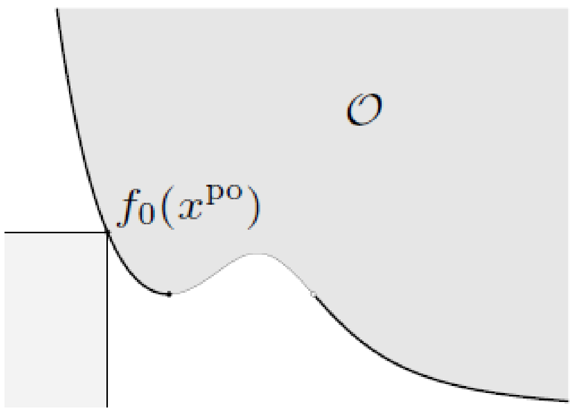
set of achievable objective values

$$\mathcal{O} = \{f_0(x) \mid x \text{ feasible}\}$$

- feasible x is **optimal** if $f_0(x)$ is a minimum value of $\mathcal{O}$
- feasible x is **Pareto optimal** if $f_0(x)$ is a minimal value of $\mathcal{O}$



$x^\star$ is optimal



x^{po} is Pareto optimal

3.3. Multicriterion Optimization

general vector optimization problem

$$
\begin{array}{ll}
\text{minimize (w.r.t. } K) & f_0(x) \\
\text{subject to} & f_i(x) \leq 0, \quad i = 1, \ldots, m \\
& h_i(x) \leq 0, \quad i = 1, \ldots, p
\end{array}
$$

vector objective $f_0 : \mathbf{R}^n \to \mathbf{R}^q$, minimized w.r.t. proper cone $K \in \mathbf{R}^q$

convex vector optimization problem

$$
\begin{array}{ll}
\text{minimize (w.r.t. } K) & f_0(x) \\
\text{subject to} & f_i(x) \leq 0, \quad i = 1, \ldots, m \\
& Ax = b
\end{array}
$$

with f_0 K-convex, $f_1, \ldots, f_m$ convex

3.3. Multicriterion Optimization

vector optimization problem with $K = \mathbf{R}^q_+$

$$f_0(x) = (F_1(x), \ldots, F_q(x))$$

- q different objectives F_i; roughly speaking we want all F_i's to be small
- feasible $x^\star$ is optimal if

$$y \text{ feasible} \quad \Longrightarrow \quad f_0(x^\star) \preceq f_0(y)$$

  if there exists an optimal point, the objectives are noncompeting

- feasible x^{po} is Pareto optimal if

$$y \text{ feasible}, \quad f_0(y) \preceq f_0(x^{\mathrm{po}}) \quad \Longrightarrow \quad f_0(x^{\mathrm{po}}) = f_0(y)$$

  if there are multiple Pareto optimal values, there is a trade-off between the objectives

3.2. References

[1] S. Boyd, L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004. http://www.stanford.edu/~boyd/cvxbook/ http://www.ee.ucla.edu/~vandenbe/cvxbook

[2] S. Boyd, L. El Ghaoui, E. Feron, V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*, Society for Industrial and Applied Mathematics, Philadelphia, MA, 1994. http://stanford.edu/~boyd/lmibook/lmibook.pdf

[3] G. Herrmann, M. C. Turner, I. Postlethwaite, "Linear Matrix Inequalities in control," *Lecture Notes in Control and Information Sciences*, pp. 123-142, 2007.

[4] G. Chesi, "LMI techniques for optimization over polynomials in control: A survey," *IEEE Transactions on Automatic Control*, vol. 55, no. 11, pp. 2500-2510, 2010.

[5] E. Artin, "Über die Zerlegung definiter Funktionen in Quadrate," *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg*, vol. 5, no. 1, pp. 100-115, 1927.

[6] C. N. Delzell, "A continuous, constructive solution to Hilbert's 17th problem," *Inventiones Mathematicae*, vol. 76, no. 3, pp. 365-384, 1984.

[7] C. J. Hillar, J. Nie, "An elementary and constructive solution to Hilbert's 17th Problem for matrices," http://arxiv.org/abs/math/0610388 http://www.msri.org/people/members/chillar/files/matrixhilbert17thfinal.pdf

[8] T. S. Motzkin, "The arithmetic-geometric inequality." *Inequalities*, Proceedings of Symposium at Wright-Patterson Air Force Base, Ohio, 1965, pp. 205-224, 1967.

[9] M.-F. Roy, "The role of Hilbert's problems in real algebraic geometry," *Proceedings of the ninth EWM Meeting*, Loccum, Germany, 1999.

[10] A. Pfister, "Zur Darstellung definiter Funktionen als Summe von Quadraten," *Inventiones Mathematicae*, vol. 4, no. 4, pp. 229-237, 1967.

[11] A. Prestel, C. N. Delzell, *Positive Polynomials: From Hilbert's 17th Problem to Real Algebra*, Springer, 2001.

[12] D. Henrion, A. Garulli, eds., *Positive Polynomials in Control*, Springer, 2005.

[13] A. Ali Ahmadi, A. Olshevsky, P. A. Parrilo, J. N. Tsitsiklis, "NP-hardness of deciding convexity of quartic polynomials and related problems," *Mathematical Programming, Series A*, in press.

[14] http://en.wikipedia.org/wiki/Positive_polynomial

[15] D. Hilbert, Mathematische Probleme, Vortrag, gehalten auf dem internationalen Mathematiker-Kongreß zu Paris 1900 http://www.mathematik.uni-bielefeld.de/~kersten/hilbert/rede.html

[16] Prof. D. Henrion, *LMI, Optimization and Polynomial Methods*, ECCI Graduate School on Control, Supélec, 2009.

[17] Associativity from 1977 IMO. http://www.cut-the-knot.org/arithmetic/algebra/1977IMO-2.shtml