Semidefinite programming for polynomial optimization and robust control

Didier HENRION

www.laas.fr/∼henrion

LAAS-CNRS Toulouse, FR
Czech Tech Univ Prague, CZ

CERFACS Sparse Days
Toulouse, October 2007

Main points

Convex optimisation can be used to solve numerically non-convex optimisation problems, specifically polynomial matrix inequalities

Numerical linear algebra is a key ingredient

Applications in many branches of engineering and applied maths, in particular systems control

PMIs and LMIs

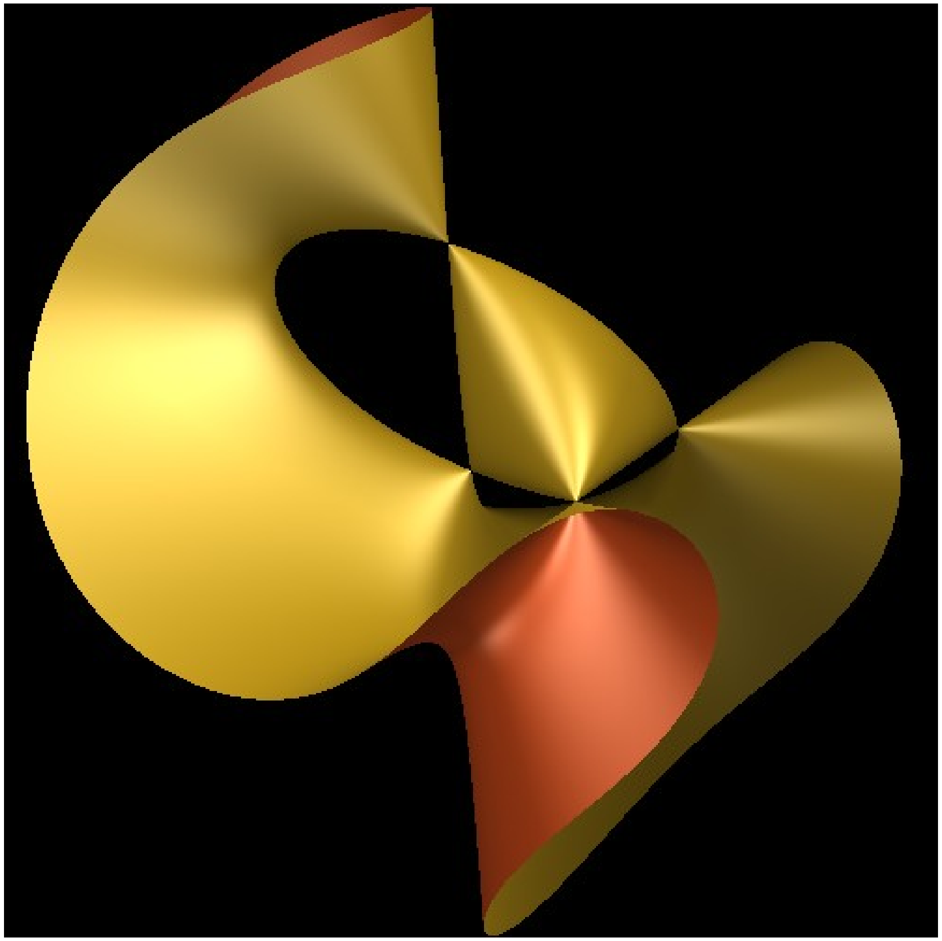
Polynomial programming: objective function and constraints are multivariate polynomials, or equivalently, we minimize a linear function on a (non-convex, non-connected, discrete..) semialgebraic set

Typically, the semialgebraic set is described by polynomial matrix inequalities (PMIs)

Semidefinite programming (SDP): linear programming in the (convex but non-polyhedral) cone of positive semidefinite matrices, also called linear matrix inequality (LMI) optimisation

$$\begin{array}{ll} \min & c^T x \\ \text{s.t.} & F(x) = \sum_\alpha F_\alpha x^\alpha \succeq 0 \end{array}$$

An LMI set (and something more)



Outline

1. Quasi-Hankel matrices

2. Generalized companion matrices

3. Systems control applications

Outline

1. Quasi-Hankel matrices

2. Generalized companion matrices

3. Systems control applications

Convex vs. non-convex optimisation

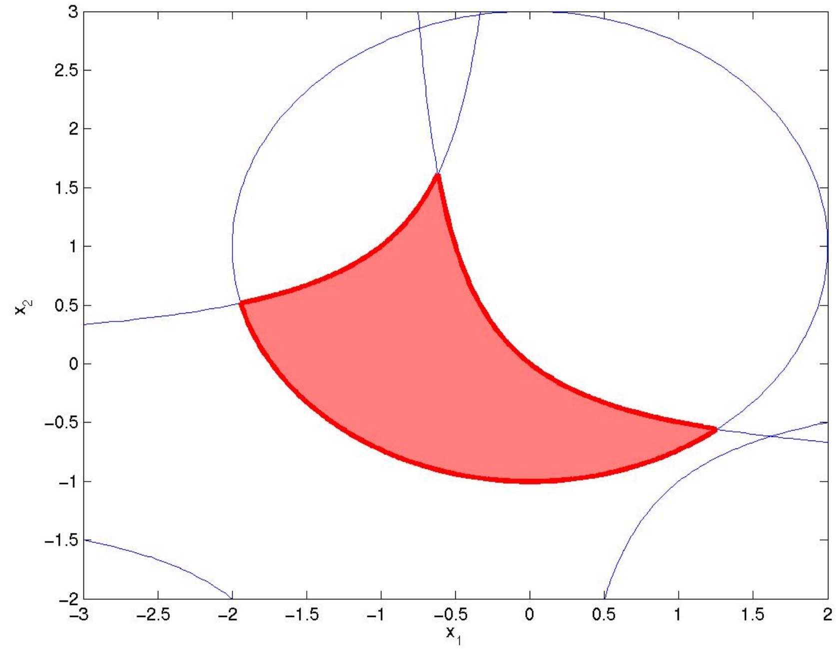
Build and solve a hierarchy of successive convex LMI relaxations of increasing size, whose optima are guaranteed to converge asymptotically to the global optimum

Relaxations are built thanks to results on the primal theory of moments and the dual theory of sum-of-squares (SOS) decompositions of multivariate polynomials

LMI relaxations: an example

Quadratic problem

$$\begin{array}{ll} \text{max} & x_2 \\ \text{s.t.} & 3 - 2x_2 - x_2^1 - x_2^2 \geq 0 \\ & -x_1 - x_2 - x_1x_2 \geq 0 \\ & 1 + x_1x_2 \geq 0 \end{array}$$



Non-convex feasible set delimited by circular and hyperbolic arcs

Lifting

LMI relaxation built by replacing each monomial $x_1^i x_2^j$ with a lifting variable y_{ij}

For example, quadratic expression $3 - 2x_2 - x_2^1 - x_2^2 \geq 0$ replaced with linear expression $3 - 2y_{01} - y_{20} - y_{02} \geq 0$

Lifting variables y_{ij} satisfy non-convex relations such as $y_{10}y_{01} = y_{11}$ or $y_{20} = y_{10}^2$

Relax these non-convex relations by enforcing LMI constraint

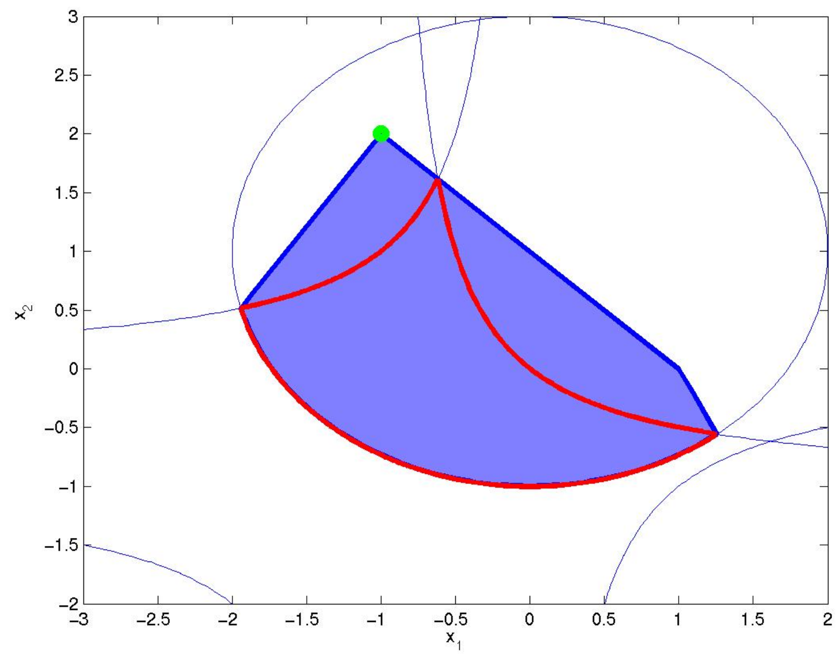
$$M_1^1(y) = \left[\begin{array}{c|cc} 1 & y_{10} & y_{01} \\ \hline y_{10} & y_{20} & y_{11} \\ y_{01} & y_{11} & y_{02} \end{array}\right] \succeq 0$$

Moment matrix of first order

Projecting

First LMI relaxation

$$\begin{array}{ll} \max & y_{01} \\ \text{s.t.} & \begin{bmatrix} 1 & y_{10} & y_{01} \\ y_{10} & y_{20} & y_{11} \\ y_{01} & y_{11} & y_{02} \end{bmatrix} \succeq 0 \\ & 3 - 2y_{01} - y_{20} - y_{02} \geq 0 \\ & -y_{10} - y_{01} - y_{11} \geq 0 \\ & 1 + y_{11} \geq 0 \end{array}$$



Projection onto the plane y_{10}, y_{01} of first-order moments

LMI optimum = 2 = upper-bound on global optimum

Higher order relaxation

To build second LMI relaxation, the moment matrix must capture expressions of degrees up to 4

$$M_2^2(y) = \left[\begin{array}{c|cc|ccc} 1 & y_{10} & y_{01} & y_{20} & y_{11} & y_{02} \\ \hline y_{10} & y_{20} & y_{11} & y_{30} & y_{21} & y_{12} \\ y_{01} & y_{11} & y_{02} & y_{21} & y_{12} & y_{03} \\ \hline y_{20} & y_{30} & y_{21} & y_{40} & y_{31} & y_{22} \\ y_{11} & y_{21} & y_{12} & y_{31} & y_{22} & y_{13} \\ y_{02} & y_{12} & y_{03} & y_{22} & y_{13} & y_{04} \end{array}\right]$$

Constraints are also lifted and relaxed with the help of localizing matrices

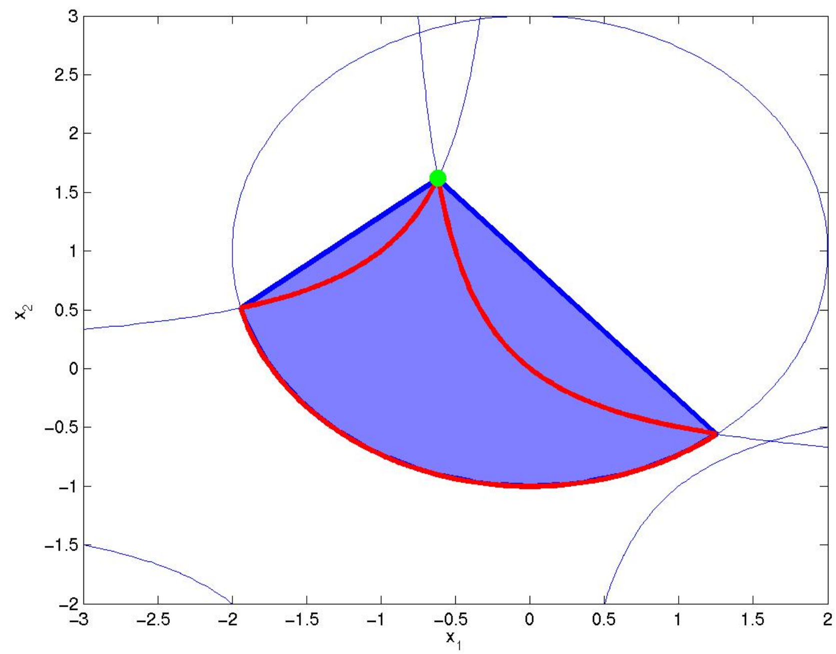
Second LMI provides tighter relaxation

$$
\begin{array}{ll}
\max & y_{01} \\
\text{s.t.} & \begin{bmatrix} 1 & y_{10} & y_{01} & y_{20} & y_{11} & y_{02} \\ y_{10} & y_{20} & y_{11} & y_{30} & y_{21} & y_{12} \\ y_{01} & y_{11} & y_{02} & y_{21} & y_{12} & y_{03} \\ y_{20} & y_{30} & y_{21} & y_{40} & y_{31} & y_{22} \\ y_{11} & y_{21} & y_{12} & y_{31} & y_{22} & y_{13} \\ y_{02} & y_{12} & y_{03} & y_{22} & y_{13} & y_{04} \end{bmatrix} \succeq 0 \\
& \begin{bmatrix} 3 - 2y_{01} - y_{20} - y_{02} & 3y_{10} - 2y_{11} - y_{30} - y_{12} & 3y_{01} - 2y_{02} - y_{21} - y_{03} \\ 3y_{10} - 2y_{11} - y_{30} - y_{12} & 3y_{20} - 2y_{21} - y_{40} - y_{22} & 3y_{11} - 2y_{12} - y_{31} - y_{13} \\ 3y_{01} - 2y_{02} - y_{21} - y_{03} & 3y_{11} - 2y_{12} - y_{31} - y_{13} & 3y_{02} - 2y_{03} - y_{22} - y_{04} \end{bmatrix} \succeq 0 \\
& \begin{bmatrix} -y_{10} - y_{01} - y_{11} & -y_{20} - y_{11} - y_{21} & -y_{11} - y_{02} - y_{12} \\ -y_{20} - y_{11} - y_{21} & -y_{30} - y_{21} - y_{31} & -y_{21} - y_{12} - y_{22} \\ -y_{11} - y_{02} - y_{12} & -y_{21} - y_{12} - y_{22} & -y_{12} - y_{03} - y_{13} \end{bmatrix} \succeq 0 \\
& \begin{bmatrix} 1 + y_{11} & y_{10} + y_{21} & y_{01} + y_{12} \\ y_{10} + y_{21} & y_{20} + y_{31} & y_{11} + y_{22} \\ y_{01} + y_{12} & y_{11} + y_{22} & y_{02} + y_{13} \end{bmatrix} \succeq 0
\end{array}
$$

Global optimality

Optimal value of 2nd LMI relaxation = 1.6180
= global optimum within numerical accuracy
Numerical certificate = moment matrix has rank one

First order moments
$(y^*_{10}, y^*_{01}) = (-0.6180, 1.6180)$
provide optimal solution
of original problem



Quasi-Hankel structure

Structure of moment matrices

univariate = Hankel

multivariate = sparse quasi-Hankel

$$\left[\begin{array}{c|cc|ccc}
1 & x_{10} & x_{01} & x_{20} & x_{11} & x_{02} \\
\hline
x_{10} & x_{20} & x_{11} & x_{30} & x_{21} & x_{12} \\
x_{01} & x_{11} & x_{02} & x_{21} & x_{12} & x_{03} \\
\hline
x_{20} & x_{30} & x_{21} & x_{40} & x_{31} & x_{22} \\
x_{11} & x_{21} & x_{12} & x_{31} & x_{22} & x_{13} \\
x_{02} & x_{12} & x_{03} & x_{22} & x_{13} & x_{04}
\end{array}\right]$$

Known sparsity pattern: LMI matrix coefficients in pre-factored form (Cholesky)

Quasi-Hankel structure

Efficient low-rank algebra to compute gradient and Hessian in interior-point codes ?

Theoretical results by B. Mourrain and V. Pan

So far our implemention in GloptiPoly for Matlab

`www.laas.fr/~henrion/software/gloptipoly`

uses a general purpose SDP solver (SeDuMi by default)

Outline

1. Quasi-Hankel matrices
2. Generalized companion matrices
3. Systems control applications

Detecting global optimality

Global optimisation problem

$$\begin{array}{rcll} p^\star & = & \min_x & g_0(x) \\ & & \text{s.t.} & g_i(x) \geq 0,\ i = 1,2.. \end{array}$$

Let $\deg g_i(x) = 2d_i - 1$ or $2d_i$ and $d = \max_i d_i$

LMI relaxation of order k

$$\begin{array}{rcll} p_k^* & = & \min_y & \sum_\alpha (g_0)_\alpha y_\alpha \\ & & \text{s.t.} & M_k(y) \succeq 0 \\ & & & M_{k-d_i}(g_i y) \succeq 0,\ i = 1,2.. \end{array}$$

with solution y_k^*

Hierarchy with convergence guarantee

$$p_d^* \leq p_{d+1}^* \leq \cdots \leq p_{k^*}^* = p^*.$$

for (generally) small k^*

Sufficient condition for global optimality
rank of moment matrices

$$\text{rank}\, M_k(y_k^*) = \text{rank}\, M_{k-d}(y_k^*)$$

Need for extraction procedures

Because rank condition is only sufficient
any other global optimality certificate is welcome

If there is a finite number of global optimisers,
they generate a zero-dimensional variety

Moment matrices are multiplication matrices in the quotient ideal, a finite-dimension vector space

Key ideas:

- Cholesky decomposition of moment matrix
- column reduced echelon form
- computation of common eigenvalues

Cholesky factorization

Extract Cholesky factor V of moment matrix

$$M_k(y_k^\star) = VV'$$

Matrix V has r columns, corresponding to r globally optimal solutions $x_j^\star$, $j = 1, 2, \ldots, r$ (provided global optimum was reached)

Choose e.g. $v = \left[1 \; x_1 \; x_2 \; \ldots \; x_n \; x_1^2 \; x_1x_2 \; \ldots \; x_1x_n \; x_2^2 \; x_2x_3 \; \ldots \; x_n^2 \; \ldots \; x_n^k\right]'$
as a basis for polynomials of degree at most k

By definition of the moment matrix:

$$M_k(y_k^\star) = V^\star(V^\star)'$$

where

$$V^\star = \left[\begin{array}{cccc} v_1^* & v_2^* & \cdots & v_r^* \end{array}\right]$$

and $v_j^\star$ is polynomial basis v evaluated at solution $x_j^\star$

Extracting solutions amounts to finding linear transformation between Cholesky factors V and $V^\star$

Reduction to column echelon form

Next step is reduction of V into column echelon form

$$U = \begin{bmatrix} 1 & & & & \\ \star & & & & \\ 0 & 1 & & & \\ 0 & 0 & 1 & & \\ \star & \star & \star & & \\ & \vdots & & \ddots & \\ 0 & 0 & 0 & \cdots & 1 \\ \star & \star & \star & \cdots & \star \\ & \vdots & & & \vdots \\ \star & \star & \star & \cdots & \star \end{bmatrix}$$

by Gaussian elimination with column pivoting
Each row in U = monomial x_α in basis v
Pivot entry in U = monomial x_{β_j} in generating basis of the set of solutions

In other words, denoting $w = \begin{bmatrix} x_{\beta_1} & x_{\beta_2} & \ldots & x_{\beta_r} \end{bmatrix}'$ it holds

$$v = Uw$$

for all solutions x_j^*, $j = 1, 2, \ldots, r$

Multiplication matrices

For each first degree monomial x_i extract from U the r-by-r multiplication matrix N_i containing coefficients of product monomials $x_i x_{\beta_j}$ in generating basis w, i.e. such that

$$N_i w = x_i w \quad i = 1, 2, \ldots, n$$

Given matrices N_i finding scalars x_i is an eigenvalue problem

Extracting solutions amounts to solving an eigenvalue problem

Eigenvectors w are shared by commuting matrices N_i so it is a particular common eigenvalue problem

Common eigenvalue problem

Build combination of multiplication matrices

$$N = \sum_{i=1}^{n} \lambda_i N_i$$

where λ_i are random positive numbers (summing up to one)

Compute ordered Schur decomposition

$$N = QTQ'$$

where $Q = \left[\begin{array}{cccc} q_1 & q_2 & \cdots & q_r \end{array} \right]$ is orthogonal and T upper triangular

Finally, due to orthogonality of vectors q_i,
ith entry in solution vector $x_j^\star$ is given by

$$(x_j^*)_i = q_j' N_i q_j,\ i = 1, 2, \ldots, n,\ j = 1, 2, \ldots, r$$

Number of solutions

No easy way to control number of extracted solutions in case of multiple global optima

Number of solutions = rank of moment matrix, but enforcing rank in an LMI is a difficult non-convex problem

By default GloptiPoly minimizes the trace (sum of eigenvalues) of the moment matrix, which may indirectly minimize the rank (number of non-zero eigenvalues)

Practical experiments reveal that low rank moment matrices ensure faster convergence of LMI relaxations to global optimum

Example of extraction

Quadratic problem

$$\begin{array}{rcl} p^* & = & \max_x \;\; (x_1-1)^2+(x_1-x_2)^2+(x_2-3)^2 \\ & & \text{s.t.} \;\; (x_1-1)^2 \leq 1 \\ & & \quad\quad (x_1-x_2)^2 \leq 1 \\ & & \quad\quad (x_2-3)^2 \leq 1 \end{array}$$

First LMI relaxation yields $p_1^\star = 3$ and rank $M_1(y^*) = 3$, extraction algorithm fails due to incomplete monomial basis

Second LMI relaxation yields $p_2^\star = 2$ and

$$\text{rank}\, M_1(y^*) = \text{rank}\, M_2(y^*) = 3$$

so rank condition ensures global optimality

Cholesky factor

Moment matrix of order $k = 2$ reads

$$M_2(y^*) = \begin{bmatrix} 1.0000 & 1.5868 & 2.2477 & 2.7603 & 3.6690 & 5.2387 \\ 1.5868 & 2.7603 & 3.6690 & 5.1073 & 6.5115 & 8.8245 \\ 2.2477 & 3.6690 & 5.2387 & 6.5115 & 8.8245 & 12.7072 \\ 2.7603 & 5.1073 & 6.5115 & 9.8013 & 12.1965 & 15.9960 \\ 3.6690 & 6.5115 & 8.8245 & 12.1965 & 15.9960 & 22.1084 \\ 5.2387 & 8.8245 & 12.7072 & 15.9960 & 22.1084 & 32.1036 \end{bmatrix}$$

Positive semidefinite with rank 3

Cholesky factor

$$V = \begin{bmatrix} -0.9384 & -0.0247 & 0.3447 \\ -1.6188 & 0.3036 & 0.2182 \\ -2.2486 & -0.1822 & 0.3864 \\ -2.9796 & 0.9603 & -0.0348 \\ -3.9813 & 0.3417 & -0.1697 \\ -5.6128 & -0.7627 & -0.1365 \end{bmatrix}$$

satisfies

$$M_k(y_k^\star) = VV'$$

Column echelon form

Gaussian elimination on V yields

$$U = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ -2 & 3 & 0 \\ -4 & 2 & 2 \\ -6 & 0 & 5 \end{bmatrix} \begin{matrix} 1 \\ x_1 \\ x_2 \\ x_1^2 \\ x_1 x_2 \\ x_2^2 \end{matrix}$$

$$\begin{matrix} 1 & x_1 & x_2 \end{matrix}$$

which means that solutions to be extracted satisfy the following system of polynomial equations

$$\begin{aligned} x_1^2 &= -2 + 3x_1 \\ x_1 x_2 &= -4 + 2x_1 + 2x_2 \\ x_2^2 &= -6 + 5x_2 \end{aligned}$$

Extraction

Multiplication matrices of monomials x_1 and x_2 in polynomial basis 1, x_1, x_2 are extracted from U:

$$N_1 = \begin{bmatrix} 0 & 1 & 0 \\ -2 & 3 & 0 \\ -4 & 2 & 2 \end{bmatrix}, \quad N_2 = \begin{bmatrix} 0 & 0 & 1 \\ -4 & 2 & 2 \\ -6 & 0 & 5 \end{bmatrix}$$

Random linear combination

$$N = 0.6909 N_1 + 0.3091 N_2$$

Schur decomposition of $N = QTQ'$ yields

$$Q = \begin{bmatrix} 0.4082 & 0.1826 & -0.8944 \\ 0.4082 & -0.9129 & -0.0000 \\ 0.8165 & 0.3651 & 0.4472 \end{bmatrix}$$

Projections of orthogonal columns of Q onto N yield the 3 expected globally optimal solutions

$$x_1^* = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad x_2^* = \begin{bmatrix} 2 \\ 2 \end{bmatrix} \quad x_3^* = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

Outline

1. Quasi-Hankel matrices

2. Generalized companion matrices

3. Systems control applications

Robust stability

Robust stability of linear system

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

where uncertain parameter q belongs to a compact set Q

Characteristic polynomial

$$p(s) = \det(sI - A(q)) = \sum_k s^k p_k(q)$$

has roots in the open left half-plane iff

$$H(p) = \sum_i \sum_j p_i(q) p_j(q) H_{ij} \succ 0$$

with Hermite matrix $H(p)$ (Bézoutian)

Robust stability

Assuming nominal system stable ($q = 0$), uncertain system remains stable if and only if the multivariate polynomial

$$\det H(p) = \sum_{\alpha} q^{\alpha} h_{\alpha}$$

remains positive for all $q \in Q$

General framework of μ-analysis (Doyle, Safonov)

Assessing robust stability of linear systems
amounts to polynomial minimisation
over compact sets (typically the unit box)

Gain scheduling

Linear system

$$z(s) = \frac{b(s,q)}{a(s,q)}u(s)$$

whose transfer function depends on
an exogeneous parameter $q \in Q$

Find a linear controller

$$u(s) = \frac{y(s,q)}{x(s,q)}z(s)$$

also depending on q (scheduled in q) and ensuring
closed-loop stability and performance (H_2 or H_∞ norm)

Gain scheduling

Controller design can be reformuled as a parametrized PMI problem

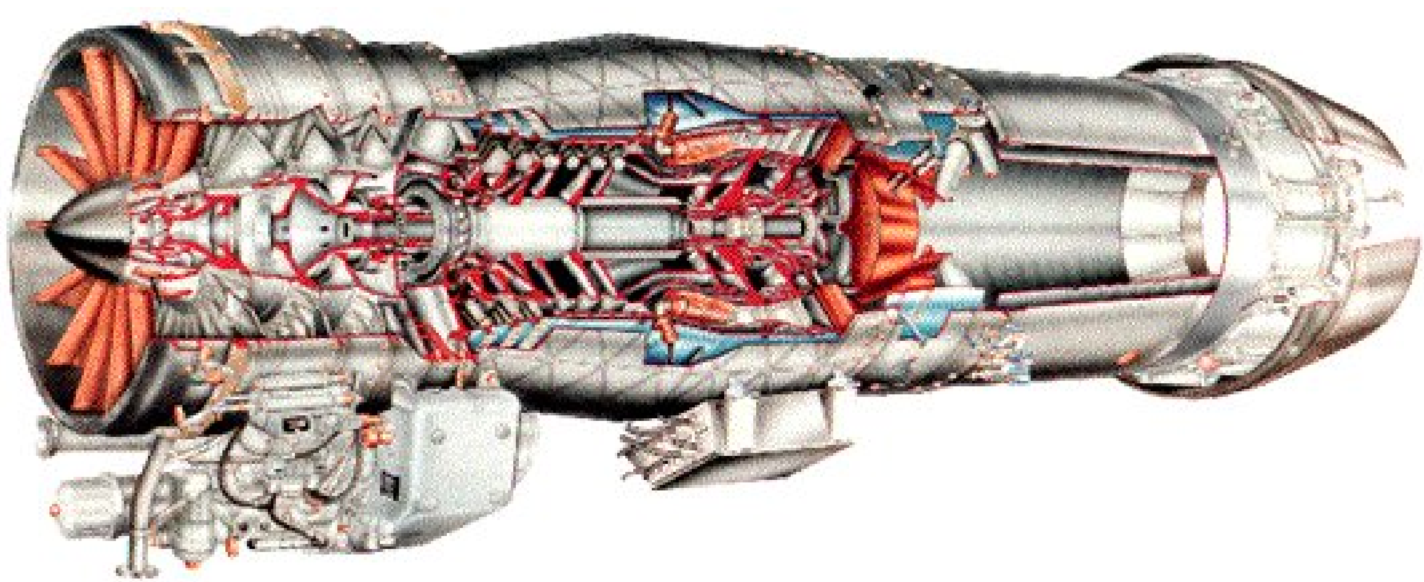
$$F(k, q) \succeq 0$$

where k is a variable to be found (parametrizing the controller) and q can be anywhere in Q

Difficult semi-infinite non-convex optimisation problem for which a hierarchy of finite-dimensional LMI problems can be devised

Aircraft turbofan engines

Research contract with Snecma since 2000



Inputs: fuel flow WF32, nozzle area A8
Outputs: fan speed XN2, compressor speed XN25
Scheduling parameter q: pressure in combustion chamber

Main points

Convex optimisation can be used to solve numerically non-convex optimisation problems, specifically polynomial matrix inequalities

Numerical linear algebra is a key ingredient

Applications in many branches of engineering and applied maths, in particular systems control

Current limitations

In general, no bounds available on the size of the LMI problem to be solved - we can construct worst-case exponential instances

No working implementation of SDP solver exploting the sparse quasi-Hankel structure

Numerical stability of solution extraction mechanism remains unclear = unknown sensitivity w.r.t. problem formulation