Polynomial optimisation, LMI and dynamical systems

Didier HENRION
henrion@laas.fr

Process Control Conference
Slovakia June 2011

Outline

1.1. Measures, moments and LMI
1.2. Polynomial optimisation
1.3. Examples and software

2.1. Occupation measures and dynamical systems
2.2. Stability analysis
2.3. Polynomial optimal control
2.4. Examples and software

Measures

Measure = function assigning a number to a set

$$\mu \,:\, K \subset \mathbb{R}^n \mapsto \mathbb{R} \quad \mu(K) = \int_K d\mu = \int_K d\mu(x) = \int_K \mu(dx)$$

Examples:

- Lebesgue measure $d\mu(x) = dx$, $\mu(K) = \text{vol}(K)$
- Hermite measure $d\mu(x) = e^{-x^T x} dx$
- probability measure $\mu(K) = 1$
- Dirac measure $d\mu(x) = \delta_{x^*}$, $\mu(\{x^*\}) = 1$

Measures as distributions or linear functionals

Measures are a particular class of distributions, continuous linear functionals acting on test functions (infinitely differentiable functions with compact support)

Riesz representation theorems identify measures with continuous linear functionals acting on continuous functions with compact support or vanishing at infinity

So a measure can indifferently act on sets or functions

Examples:

- Lebesgue measure $f \mapsto \int_K f(x)dx$
- Hermite measure $f \mapsto \int f(x)e^{-x^T x}dx$
- probability measure $f \mapsto \int_K f(x)d\mu(x) = E[f(x)]$
- Dirac measure $f \mapsto \int f(x)\delta_{x^*} = f(x^*)$

Some terminology

Support = smallest closed set $K \subset \mathbb{R}^n$ for which $\mu(\mathbb{R}^n / K) = 0$

Examples:

- Dirac measure supp$(\delta_x) = \{x\}$
- atomic measure supp$(\mu) = \{x_1, \ldots, x_r\}$
- Hermite measure supp$(\mu) = \mathbb{R}^n$
- Lesbegue measure on $K = [-1, 1]$, vol$(K) = 2$
- Lesbegue measure on $K = \{x \in \mathbb{R}^2 \ : \ x^T x \leq 1\}$, vol$(K) = \pi$

Indicator, or characteristic function of a set K

$$\begin{aligned} I_K(x) &= 1 \quad \text{if } x \in K \\ &= 0 \quad \text{otherwise} \end{aligned}$$

Classical analysis: from functions to measures

A univariate real function

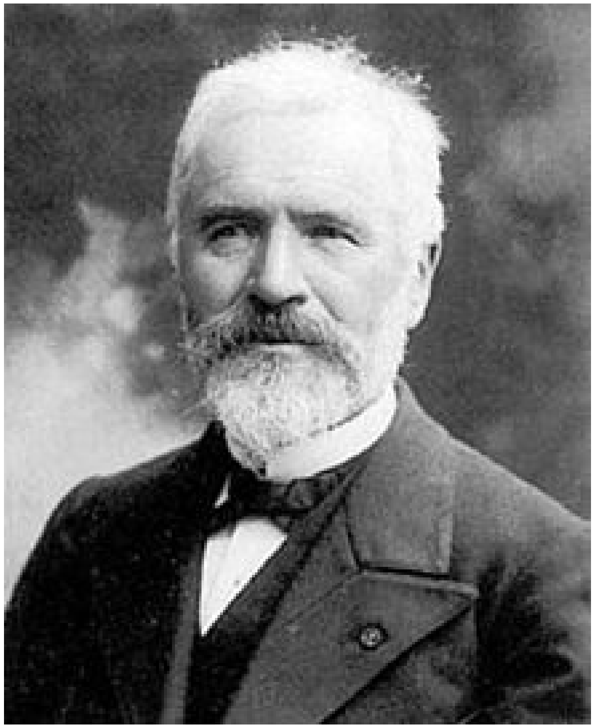
$$f : [a, b] \subset \mathbb{R} \to \mathbb{R}$$

is of bounded variation whenever

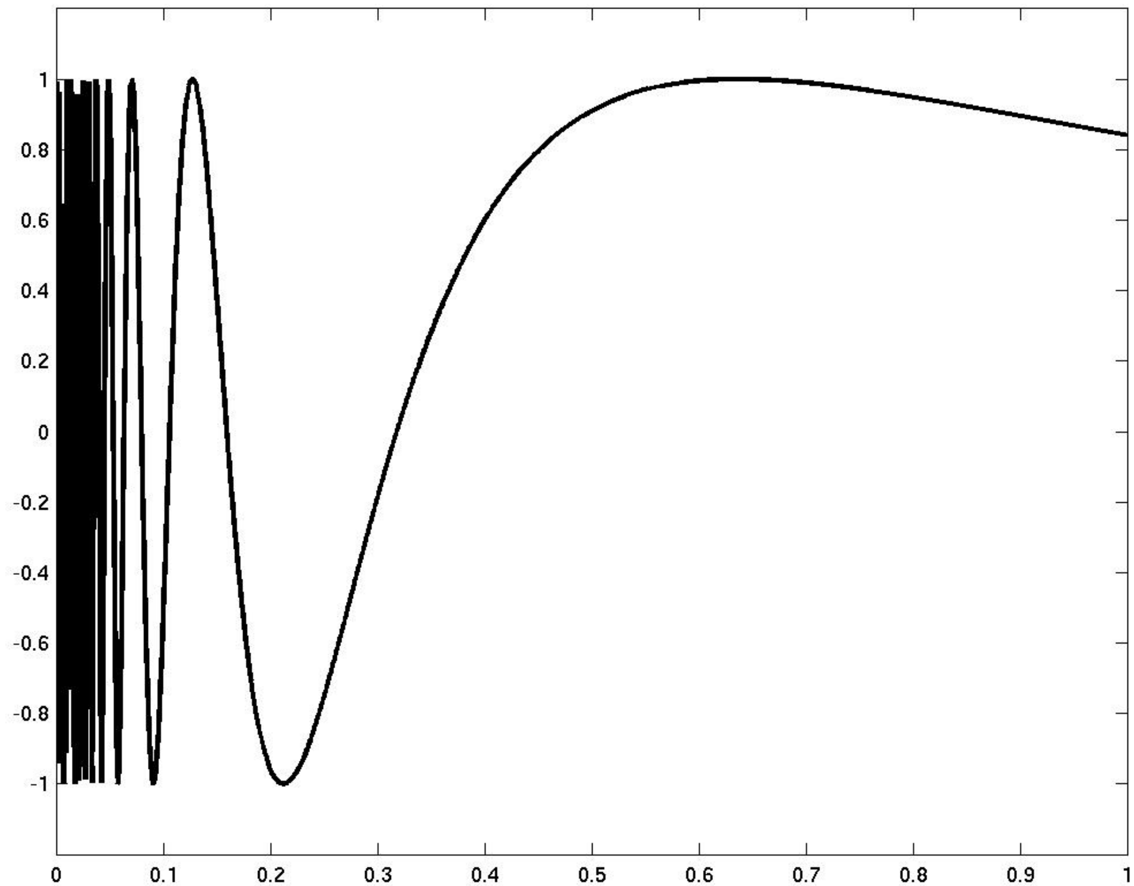
$$\sum_{k=1}^{n} |f(x_k) - f(x_{k-1})|$$

is finite over all possible partitions

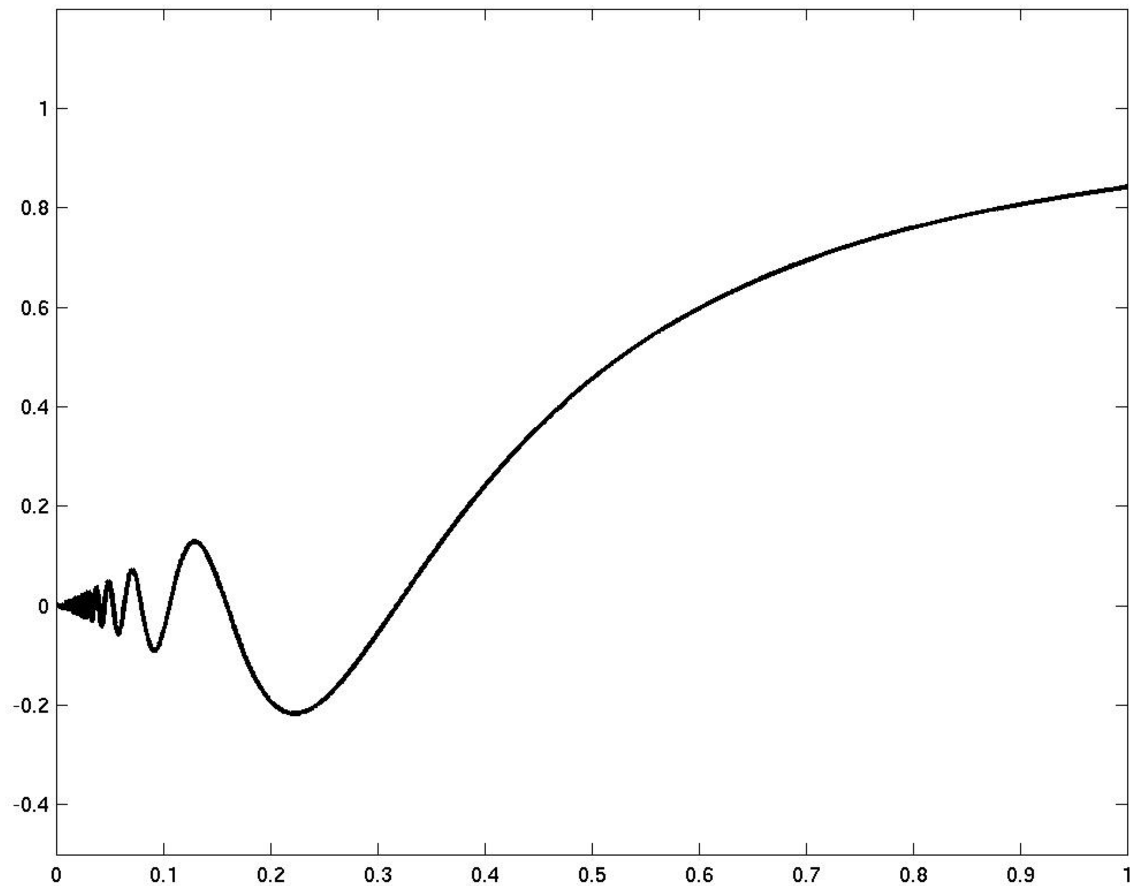
$$a = x_0 < x_1 < x_2 \cdots < x_{n-1} < x_n = b$$



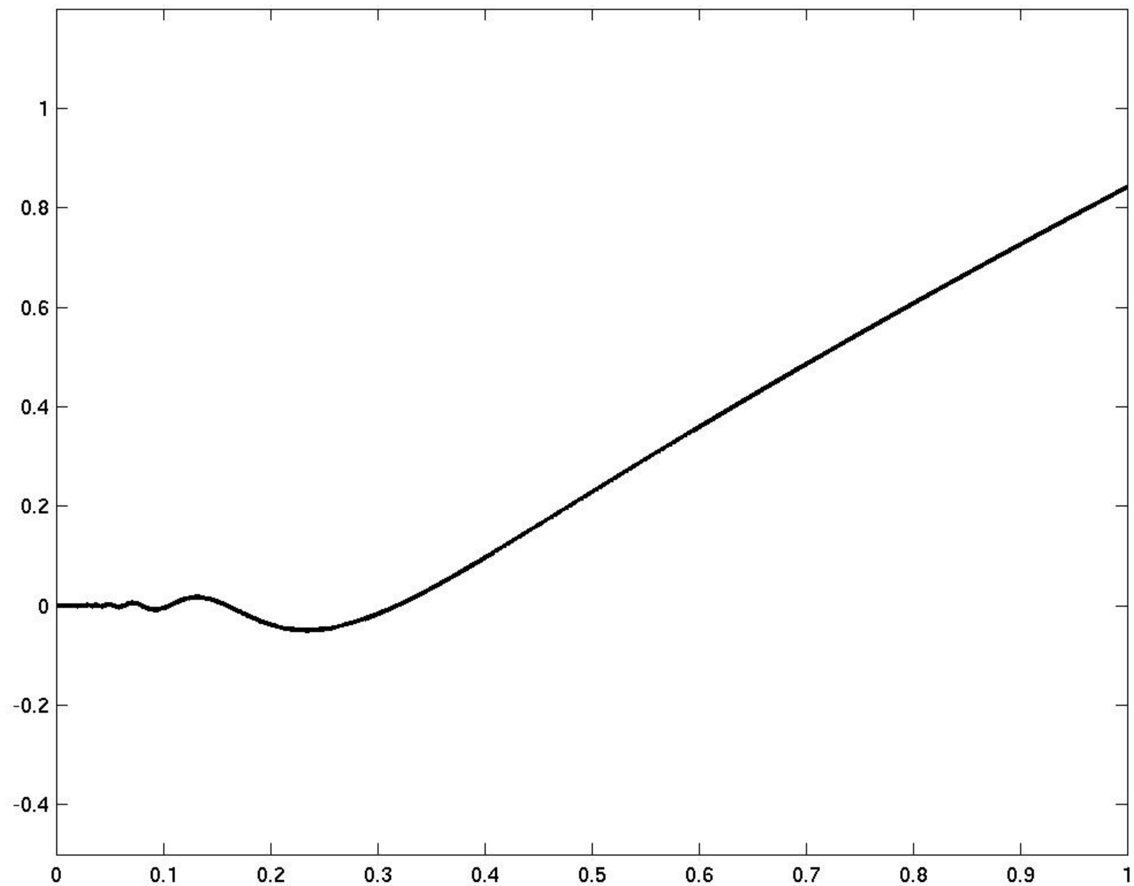
Camille Jordan
(1838-1922)



$\sin x^{-1}$ is not of bounded variation on $[0, 1]$



$x \sin x^{-1}$ is not of bounded variation on $[0, 1]$



$x^2 \sin x^{-1}$ is of bounded variation on $[0, 1]$

When the fundamental theorem of calculus fails

When is a function the (indefinite) integral of an other function ?

$$\int_a^b f'(x)dx = f(b) - f(a)$$

This fundamental identity can fail e.g. for the jump function

$$\begin{aligned} f(x) &= 0 \text{ if } 0 \le x < \tfrac{1}{2} \\ &= 1 \text{ if } \tfrac{1}{2} < x \le 1 \end{aligned}$$

for which

$$\int_0^1 f'(x)dx = 0 < f(1) - f(0) = 1$$

What is $f'(x)$ in this case ?

Henri Lebesgue
(1875-1941)

Lebesgue decomposition

Any $f(x)$ of bounded variation can be decomposed as

$$f(x) = f_+(x) - f_-(x)$$

where f_+ and f_- are both monotone (e.g. nondecreasing)

Any $f(x)$ of bounded variation can be decomposed as

$$f(x) = f_{AC}(x) + f_{SC}(x) + f_{SD}(x)$$

where

- f_{AC} is an absolutely continuous function
- f_{SC} is a singularly continuous, or singular function
- f_{SD} is a singularly discrete, or jump function

Absolutely continuous functions

Function $f(x)$ is absolutely continuous when it is continuous

$$|f(x_k) - f(x_{k-1})| \to 0 \quad \text{when} \quad |x_k - x_{k-1}| \to 0$$

and in addition

$$\sum_{k=1}^{n} |f(x_k) - f(x_{k-1})| \to 0 \quad \text{when} \quad \sum_{k=1}^{n} |x_k - x_{k-1}| \to 0$$

for all possible partitions

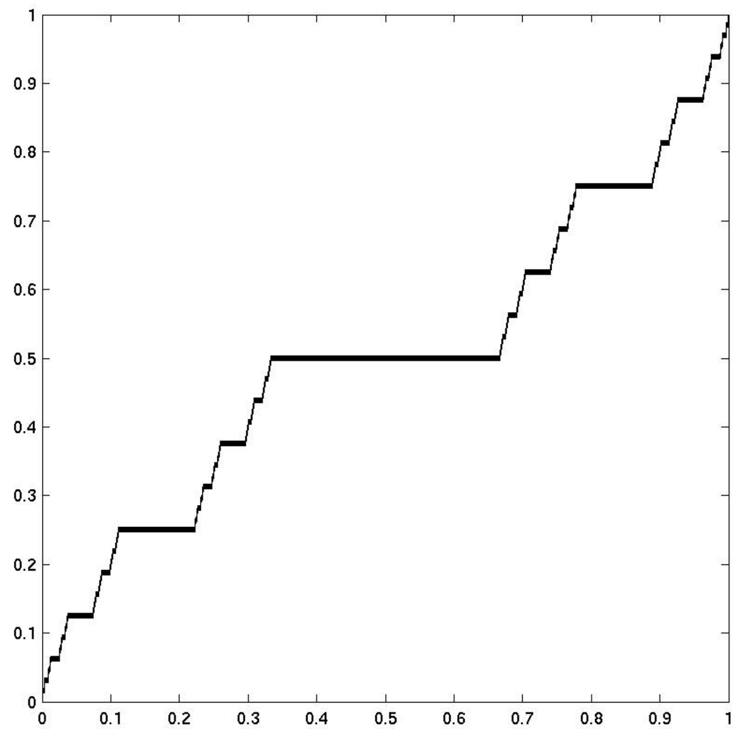
$$a = x_0 < x_1 < x_2 \cdots < x_{n-1} < x_n = b$$

In particular, Lipschitz functions

$$|f(x_k) - f(x_{k-1})| \leq L|x_k - x_{k-1}|$$

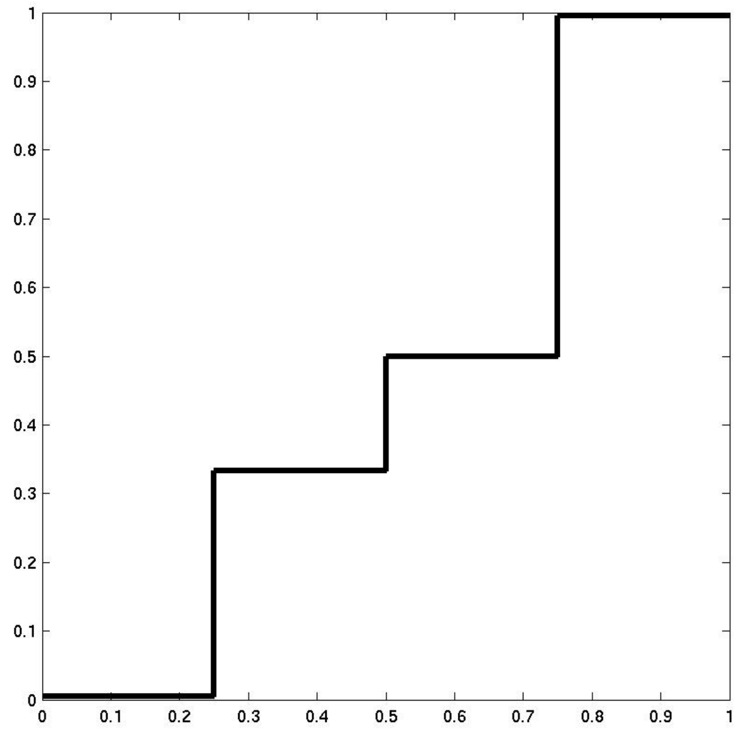
are absolutely continuous

Singularly continuous functions



Cantor's devil staircase function is continuous and monotone but not absolutely continuous

Singularly discrete or jump functions



Piecewise constant with discontinuities

Stieljtes integral

For $f(x)$ of bounded variation
and $v(x)$ continuous

$$\int_a^b v(x)df(x) = \lim_{|x_k - x_{k-1}| \to 0} \sum_k v(z_k)(f(x_k) - f(x_{k-1}))$$

over all possible partitions

$$a = x_0 < x_1 < x_2 \cdots < x_{n-1} < x_n = b$$

with $x_{k-1} \le z_k \le x_k$

Thomas J Stieltjes
(1856-1894)

From functions to measures

Every continuous linear functional acting on continuous functions on $[a, b]$ can be expressed as

$$v \mapsto \int_a^b v(x) df(x)$$

with $f(x)$ a function of bounded variation, or equivalently as

$$v \mapsto \int_a^b v(x) d\mu(x)$$

with μ a measure

Frigyes Riesz
(1880-1956)

Lebesgue decomposition

Any measure μ can be decomposed as

$$\mu = \mu_+ - \mu_-$$

where μ_+ and μ_- are both nonnegative measures

Any measure μ can be decomposed as

$$\mu = \mu_{AC} + \mu_{SC} + \mu_{SD}$$

where

- μ_{AC} is an absolutely continuous measure
- μ_{SC} is a singularly continuous, or singular measure
- μ_{SD} is a singularly discrete, or atomic measure

Moments

Multi-index notation $x^\alpha = \prod_{i=1}^n x_i^{\alpha_i}$ with $x \in \mathbb{R}^n$, $\alpha \in \mathbb{N}^n$

The α-th moment of measure μ is the real number

$$\boxed{y_\alpha = \int_K x^\alpha d\mu(x)}$$

μ is a representing measure for sequence $y = (y_\alpha)_{\alpha \in \mathbb{N}^n}$

Classical problem of moments (Hausdorff, Markov, Stieltjes): characterise sequence y having representing measure μ supported on a (given) set K

Conditions on y_α ? Construction of μ and K, given y ?

LMI conditions

Given a sequence y, define the moment matrix $M_d(y)$ of order d with entries indexed by multi-indices β (rows) and γ (columns)

$$[M_d(y)]_{\beta,\gamma} = y_{\beta+\gamma}, \quad |\beta| + |\gamma| \leq 2d$$

which are linear in y

Necessary condition: if y has a representing measure μ then $M_d(y) \succeq 0$ $\forall d$

Sufficient condition (Berg 1987): if $|y_\alpha| \leq 1$ $\forall \alpha$ and $M_d(y) \succeq 0$ $\forall d$, then y has a representing measure μ with supp$(\mu) \subset [-1, 1]$

LMI conditions

Given a sequence y and a polynomial $p(x) = \sum_\alpha p_\alpha x^\alpha$, define the localising matrix $M_d(py)$ of order d with entries

$$[M_d(py)]_{\beta,\gamma} = \sum_\alpha p_\alpha y_{\alpha+\beta+\gamma}, \quad |\alpha| + |\beta| + |\gamma| \leq 2d$$

Let $K = \{x \in \mathbb{R}^n : p_k(x) \geq 0, \forall k\}$ be compact basic semialgebraic with $\{x : p_k(x) \geq 0\}$ compact for some k

Necessary condition: if y has a representing measure μ with support in K, then $M_d(y) \succeq 0$, $M_d(p_k y) \succeq 0$ $\forall k$ $\forall d$

Sufficient condition (Putinar 1993): if $M_d(y) \succeq 0$, $M_d(p_k y) \succeq 0$ $\forall k$ $\forall d$ then y has a representing measure with $\mathrm{supp}(\mu) \subset K$

Outline

1.1. Measures, moments and LMI
1.2. Polynomial optimisation
1.3. Examples and software

2.1. Occupation measures and dynamical systems
2.2. Stability analysis
2.3. Polynomial optimal control
2.4. Examples and software

Polynomial optimisation

Consider the problem

$$\begin{array}{rcl} p^* & = & \min \;\; p(x) = \sum_\alpha p_\alpha x^\alpha \\ & & \text{s.t.} \;\; x \in K = \{x \in \mathbb{R}^n \,:\, p_k(x) \geq 0,\;\; k = 1,\ldots,m\} \end{array}$$

where unknowns are entries of vector x and K is a given basic semialgebraic set possibly nonconvex and/or nonconnected

For example K can be the union of a finite number of points, i.e. a zero-dimensional variety

Includes polynomial matrix inequalities

$$K = \{x \,:\, P(x) = \sum_\alpha x^\alpha P_\alpha \succeq 0\}$$

and in particular bilinear matrix inequalities (BMIs)

Primal formulation

Linearisation

$$\begin{array}{rcl} p^* & = & \min_\mu \; \int_K p(x)d\mu(x) \\ & = & \min_\mu \; \sum_\alpha p_\alpha \int_K x^\alpha d\mu(x) \\ & = & \min_y \; \sum_\alpha p_\alpha y_\alpha \end{array}$$

unknowns are moments of a probability measure supported on K

Proof (lower bound):
$p(x) \geq p^*$ for all $x \in K$ so $\int_K p(x)d\mu(x) \geq \int_K p^* d\mu(x) = p^*$

Proof (upper bound):
choose a particular probability measure $\mu^* = \delta_{x^*}$ where
x^* is a global minimizer, then $p^* \geq \int_K p(x)d\mu^*(x) = p(x^*)$

Hierarchy of relaxations

Use Putinar's condition to generate hierarchy of LMI relaxations

$$
\begin{array}{rcl}
p_d^* & = & \min_y \ \sum_\alpha p_\alpha y_\alpha \\
 & & \text{s.t.} \ \ M_d(y) \succeq 0, \ \ M_d(p_k y) \succeq 0, \ k = 1, \ldots, m
\end{array}
$$

and monotonically increasing asymptotically converging sequence of lower bounds

$$p_0^* \leq p_1^* \leq \cdots p_\infty^* = p^*$$

Dual formulation

Maximize lower bound on epigraph

$$\begin{array}{rcl} p^* & = & \max \ \underline{p} \\ & & \text{s.t.} \ \ p(x) - \underline{p} \geq 0 \quad \forall x \in K \end{array}$$

involves a polynomial positivity condition which is relaxed as

$$\begin{array}{rcl} p^* & = & \max_q \ \underline{p} \\ & & \text{s.t.} \ \ p(x) - \underline{p} = (\sum_j q_{j0}^2(x)) + \sum_k (\sum_j q_{jk}^2(x)) p_k(x) \end{array}$$

with unknown polynomial sum-of-squares (SOS) multipliers

Lagrangian with polynomial multipliers

Can be formulated as a dual hierarchy of LMI problems by fixing the degree of SOS multipliers to $d = 0, 2, 4 \ldots$

S-procedure

Frequently used in robust control, e.g. to prove positive-real lemma (absolute stability) or bounded-real lemma (H_∞ control)

There exists no $x \neq 0$ such that $p_k(x) = x^T A_k x \geq 0$, $k = 1, \ldots, m$ if there exists $q \geq 0$ such that $\sum_{k=1}^m q_k A_k \prec 0$, an LMI

Corresponds to $p(x) = 1$ and constant multipliers $q_k(x)$ used as infeasibility Farkas certificates

When does converse statement hold ?
Issue of conservatism of robust control LMIs

Exactness

Unconstrained polynomial optimisation

$$\min p(x) \text{ s.t. } x \in \mathbb{R}^n, \;\; \deg p(x) = 2\delta$$

First LMI relaxation exact for $n = 1$ (univariate polynomials), $\delta = 1$ (conics), $n = \delta = 2$ (bivariate quartics)

Already known to Hilbert (1900), but first explicit counter-example of nonexactness given by Motzkin (1965) as a bivariate sextic which is nonnegative but not polynomial SOS

Artin (1927) proved however that every nonnegative polynomial is rational SOS

Used by Packard and Doyle (1993) for μ-analysis, and then extended by Parrilo (2000)

Exactness

Constrained polynomial optimisation

$$p^* = \min_x p(x) \text{ s.t. } x \in K = \{x : p_k(x) \geq 0,\ k = 1, \ldots, m\}$$

with LMI relaxations

$$\begin{array}{rcl} p_d^* & = & \min_y \ \sum_\alpha p_\alpha y_\alpha \\ & & \text{s.t. } \ M_d(y) \succeq 0,\ \ M_d(p_k y) \succeq 0,\ k = 1, \ldots, m \end{array}$$

Exactness certificate $p_d^* = p^*$ whenever

$$r = \operatorname{rank} M_d(y^*) = \operatorname{rank} M_{d-\delta}(y^*),\ \ \delta = \max_k \deg p_k(x)/2$$

moment matrix with flat extension (Curto and Fialkow 1993)

Corresponds to an r-atomic optimal measure
and we can extract minimizers x^* from $M_d(y)$

Imperial College Press Optimization Series Vol. 1

Moments, Positive Polynomials and Their Applications

Many important applications in global optimization, algebra, probability and statistics, applied mathematics, control theory, financial mathematics, inverse problems, etc. can be modeled as a particular instance of the *Generalized Moment Problem* (GMP).

This book introduces a new general methodology to solve the GMP when its data are polynomials and basic semi-algebraic sets. This methodology combines semidefinite programming with recent results from real algebraic geometry to provide a hierarchy of semidefinite relaxations converging to the desired optimal value. Applied on appropriate cones, standard duality in convex optimization nicely expresses the duality between moments and positive polynomials.

In the second part, the methodology is particularized and described in detail for various applications, including global optimization, probability, optimal control, mathematical finance, multivariate integration, etc., and examples are provided for each particular application.

P665 hc
ISBN-13 978-1-84816-445-1
ISBN-10 1-84816-445-9
9 781848 164451

Imperial College Press
www.icpress.co.uk

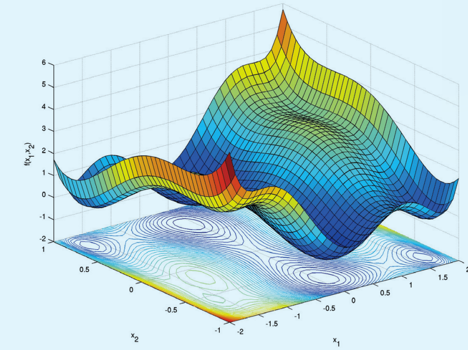
Vol. 1

Moments, Positive Polynomials and Their Applications

Lasserre

ICP

Imperial College Press Optimization Series Vol. 1



Moments, Positive Polynomials and Their Applications

Jean Bernard Lasserre

Imperial College Press

Outline

1.1. Measures, moments and LMI
1.2. Polynomial optimisation
1.3. Examples and software

2.1. Occupation measures and dynamical systems
2.2. Stability analysis
2.3. Polynomial optimal control
2.4. Examples and software

Univariate polynomials

Global minimisation of univariate polynomial

$$\min_x p(x) = \sum_{\alpha=0}^{d} p_\alpha x^\alpha, \quad x \in \mathbb{R}$$

Primal moment problem

$$\begin{array}{ll} \min_y & p_0 + \sum_{\alpha=1}^{d} p_\alpha y_\alpha \\ \text{s.t.} & H_0 + \sum_{\alpha=1}^{d} H_d y_k \succeq 0 \end{array}$$

where H_α are unit Hankel matrices

Dual SOS problem

$$\begin{array}{ll} \max_X & p_0 - \text{trace}\, H_0 X \\ \text{s.t.} & \text{trace}\, H_\alpha X = p_\alpha,\ \alpha = 1, \ldots, d \\ & X \succeq 0 \end{array}$$

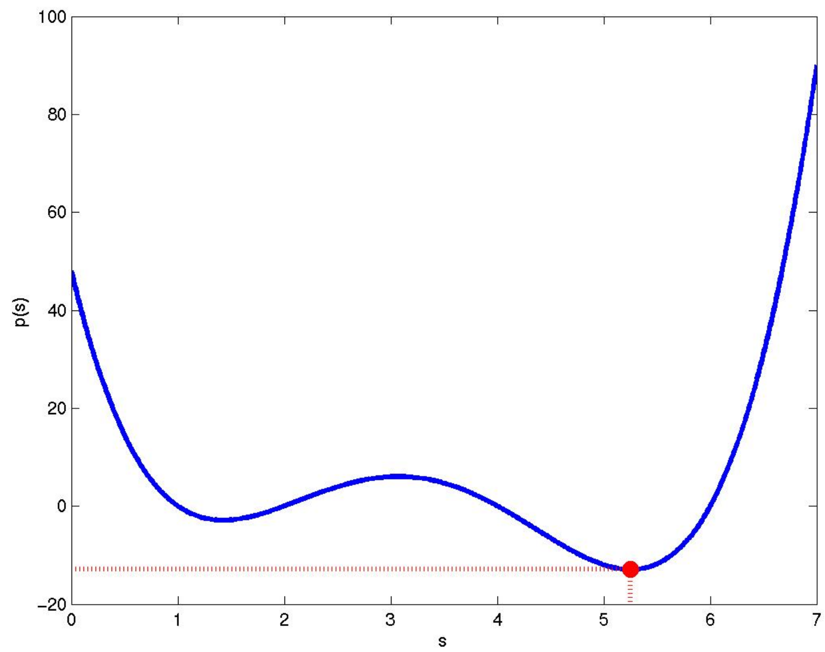
Univariate polynomials

Example: quartic polynomial

$$p(x) = 48 - 92x + 56x^2 - 13x^3 + s^4$$

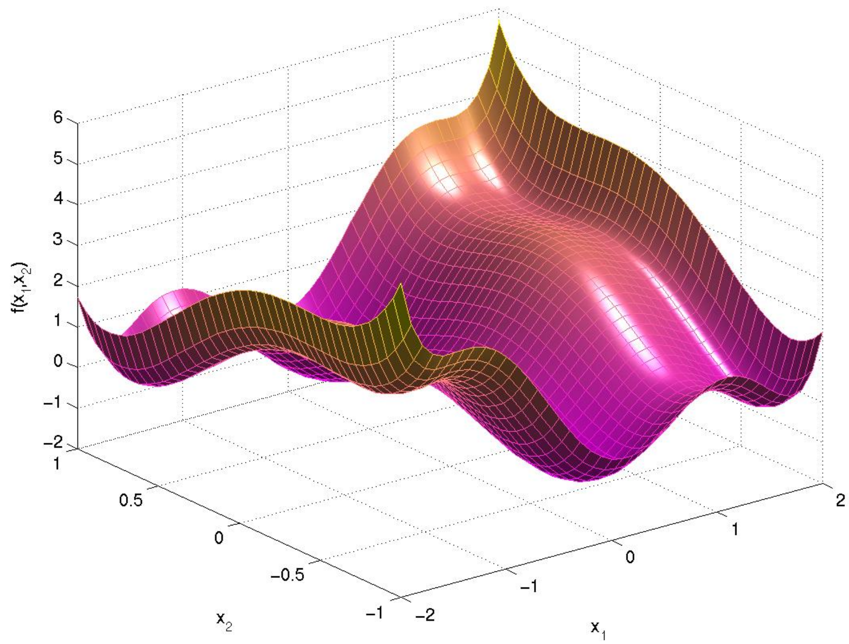
Solving the moment LMI problem yields $p^* = p(5.25) = -12.89$

$$\begin{array}{ll} \min & 48 - 92y_1 + 56y_2 - 13y_3 + y_4 \\ \text{s.t.} & \begin{bmatrix} 1 & y_1 & y_2 \\ y_1 & y_2 & y_3 \\ y_2 & y_3 & y_4 \end{bmatrix} \succeq 0 \end{array}$$



Camelback function

For the six-hump camelback function



with two global optima and six local optima, the global optimum is reached at the first LMI relaxation ($d = 1$) without any problem splitting

LMI hierarchy

Quadratic problem

$$
\begin{array}{ll}
\text{min} & -2x_1 + x_2 - x_3 \\
\text{s.t.} & x_1(4x_1 - 4x_2 + 4x_3 - 20) + x_2(2x_2 - 2x_3 + 9) \\
& \qquad +x_3(2x_3 - 13) + 24 \geq 0 \\
& x_1 + x_2 + x_3 \leq 4, \quad 3x_2 + x_3 \leq 6 \\
& 0 \leq x_1 \leq 2, \quad 0 \leq x_2, \quad 0 \leq x_3 \leq 3
\end{array}
$$

Computational burden increases quickly with relaxation order

order d	1	2	3	4	5	6
bound p_d^*	-6.0000	-5.6923	-4.0685	-4.0000	-4.0000	-4.0000
`size(y)`	9	34	83	164	285	454

..yet fourth LMI relaxation solves globally the problem

Robust stability analysis

Linear system of order n

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

with polytopic uncertainty

$$A(q) = A_0 + \sum_{i=1}^{m} q_i A_i \quad \forall q \in K \subset \mathbb{R}^m$$

Particular "easy" cases:

- $K =$ box and rank $A_i = 1$: Kharitonov's Theorem
- rank $A_i = 1$: Edge Theorem

Otherwise no general polynomial-time algorithm for checking robust stability of this system

Hermite stability criterion

Uncertain linear system robustly stable iff

$$H(q) \succ 0 \quad \forall q \in K$$

Charles Hermite (1822-1901)

where $H(q)$ is Hermite matrix (1854) of det $(sI_n - A(q))$

Quadratic Lyapunov matrix depending polynomially in q

Assessing robust stability

Defining $p(q) = \det H(q)$ and $K = \{q : p_k(q) \geq 0\}$, solve $p^\star = \min_{q \in K} p(q)$ and check whether $p^\star > 0$

Hierarchy of LMI relaxations

$$\begin{array}{rcll} p_d^* & = & \min_y & p^T y \\ & & \text{s.t.} & M_d(y) \succeq 0 \\ & & & M_d(p_k y) \succeq 0 \end{array}$$

with $M_d(y)$ truncated moment matrix and $M_d(p_k y)$ truncated localizing matrices

Converging sequence $p_1^* \leq p_2^* \leq \cdots p_\infty^* = p^\star$

Interval matrix stability

Consider the interval matrix

$$A(q) = \begin{bmatrix} q_1 & 0 & 0 \\ 0 & q_2 & q_3 \\ 0 & -0.7115 & q_4 \end{bmatrix}$$

where

$$\begin{aligned} q \quad \in \quad & Q = [-2.4780,\ -1.4471] \times [-0.0518,\ -0.0194] \\ & \times [2.0000,\ 3.4347] \times [-0.0026,\ -0.0012] \end{aligned}$$

LMI relaxation of order 3 inconclusive with $p_3^* = -171$

LMI relaxation of order 4 yields certified global optimum $p_4^* = p^* = 0.1505$ attained at

$$q^* = [-1.4471,\ -0.0194,\ 2.0000,\ -0.0012]$$

hence proving robust stability

Software for polynomial optimisation, moments and LMI

Matlab interfaces

- GloptiPoly (Henrion/Lasserre 2002)
- SOSTOOLS (Parrilo at al. 2002)
- YALMIP (Löfberg 2005)
- SparsePOP (Kojima et al. 2005)

Semidefinite programming solvers

- SeDuMi (Sturm 1999 and Terlaky 2005)
- SDPT3 (Toh et al. 1999)
- CSDP (Borchers 1999)
- SDPA (Kojima et al. 1996)
- PENSDP (Kočvara and Stingl 2004)

Outline

Occupation measures

Dynamical system described by ODE

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

with Lipschitz vector field f and (unique) solution, or flow ϕ_t starting from initial condition x_0

Occupation measure of trajectory from $t = 0$ to $t = T$

$$\mu(X) = \int_0^T I_X(\phi_t)dt$$

where X is a subset of $\mathbb{R}^n$

It is the time spent in X by the solution of the ODE

Moments of occupation measure

By definition, in any subset $X \subset \mathbb{R}^n$,
the α-th moment of occupation measure μ is given by

$$\begin{aligned} y_\alpha &= \int_X x^\alpha d\mu(x) \\ &= \int_X \int_0^T x^\alpha \delta_x(\phi_t) dt dx \\ &= \int_0^T \int_X x^\alpha \delta_x(\phi_t) dx dt \\ &= \int_0^T \phi_t^\alpha dt \end{aligned}$$

So if sequence y is given we can find system trajectories
by solving the corresponding problem of moments

Given dynamics $f(x)$, how can we find sequence y ?

Test functions

Consider a continuously differentiable test function $v(x)$ whose time-derivative along system trajectories is given by

$$\dot{v} = \nabla v \cdot \dot{x} = \nabla v \cdot f$$

Since flow ϕ_t is absolutely continuous, from the fundamental theorem of calculus

$$\begin{array}{rcl} \int_0^T dv & = & \int_0^T \nabla v \cdot f dt \\ & = & \int_X \nabla v \cdot f d\mu \\ & = & v(x_T) - v(x_0) \end{array}$$

it follows that occupation measure μ satisfies (infinitely many) linear equations

$$\int_X \nabla v \cdot f d\mu = v(x_T) - v(x_0) \quad \forall v$$

Variational formulation

Now assume x_0 and x_T are not known exactly, they are modeled by probability measures μ_0 and μ_T with supports X_0 and X_T, respectively

Our three measures satisfy the following constraints

$$\int_X \nabla v \cdot f d\mu = \int_{X_T} v d\mu_T - \int_{X_0} v d\mu_0, \quad \forall v$$

(compare with previous slide where $\mu_0 = \delta_{x_0}$, $\mu_T = \delta_{x_T}$)

This is an infinite-dimensional linear problem in measure space

Compare with weak or variational formulations of PDE problems

Duality between measures and functions

More formally, let X be a compact topological space
Let $M(X)$ be the Banach space of finite measures
Let $C(X)$ be the Banach space of bounded continuous functions
Then $M(X)$ can be identified with the dual $C(X)^*$, in the sense that $C(X), M(X)$ form a dual pair with duality bracket

$$< v, \mu > = \int_X v d\mu$$

Let $L : C(X) \to C(Y)$ be a linear mapping
Let $L^* : M(Y) \to M(X)$ be its adjoint

$$< L(v), \mu > = < v, L^*(\mu) >$$

Duality along dynamics

Let $v(x) \in C^1(X)$ be continuously differentiable
Define linear mapping $F : C^1(X) \to C^1(X)$ such that

$$\frac{\partial v}{\partial t} = \frac{\partial v}{\partial x}\frac{dx}{dt} = \nabla v \cdot f = -F(v)$$

Once again, integration along system trajectories yields
linear relation linking measures μ, μ_0 and μ_T

$$\begin{array}{rcl}
-\int_X \nabla v \cdot f d\mu & = & \int_{X_0} v d\mu_0 - \int_{X_T} v d\mu_T \\
\int_X F(v) d\mu & = & \int_{X_0} v d\mu_0 - \int_{X_T} v d\mu_T \\
< F(v), \mu > & = & < v, d\mu_0 > - < v, d\mu_T > \\
< v, F^*(\mu) > & = & < v, d\mu_0 > - < v, d\mu_T > \\
F^*(\mu) & = & \mu_0 - \mu_T \\
\nabla \cdot (f\mu) & = & \mu_0 - \mu_T
\end{array}$$

Linear measure problem

Nonlinear Cauchy problem

$$\dot{x} = f(x), \;\; x(0) \in X_0, \;\; x(T) \in X_T$$

replaced by linear problem

$$\boxed{\nabla \cdot (f\mu) = \mu_0 - \mu_T}$$

where differentiation should be understood
in the sense of distributions (Schwartz 1950)

Fluid dynamics and linear transport equation

Now suppose that $f(x)$ describes the velocity field of a compressible fluid with density $\rho(t,x)$

Time evolution of fluid density described by continuity equation from fluid dynamics, i.e. principle of mass conservation

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (f\rho) = 0, \quad \rho(0,x) = \rho_0(x)$$

Transportation of density of flow along trajectories

Liouville's theorem and advection PDE

See e.g. Cédric Villani's 2003 book on optimal transport

Linear transport equation

Flow density $\rho(t, dx)$ is a measure with boundary conditions

$$\rho(0, dx) = \mu_0(dx), \quad \rho(T, dx) = \mu_T(dx)$$

and time integration

$$\begin{aligned}
\mu_T - \mu_0 &= \int_0^T d\rho(t) \\
&= \int_0^T \frac{\partial \rho}{\partial t} dt \\
&= -\int_0^T \nabla \cdot (f\rho) dt \\
&= -\nabla \cdot (f \int_0^T \rho dt) \\
&= -\nabla \cdot (f\mu)
\end{aligned}$$

allows to recover our LP measure problem

Outline

Invariant measures

Define the Frobenius-Perron operator P, also called the push-forward of measure μ along flow ϕ_t, as

$$P\mu(X) = \mu(\phi_t^{-1}(X))$$

Measure μ is called invariant if

$$P\mu(X) = \mu(X)$$

for all subsets X, i.e. if it is a fixed point of P

Invariant measures satisfy

$$\nabla \cdot (f\mu) = 0$$

and they characterize stable, unstable or periodic trajectories

Equilibrium points

Let x^* satisfy $f(x^*) = 0$. Then $\mu = \delta_{x^*}$ is such that

$$\int \nabla \cdot (f\mu)v = -\int \nabla v \cdot f\delta_{x^*} = -\nabla v(x^*) \cdot f(x^*) = 0$$

so it is invariant

Periodic solutions

Let $T > 0$ satisfy $\phi_{t+T}(x) = \phi_t(x)$ for all t and x
Then $\mu(X) = \frac{1}{T}\int_0^T I_X(\phi_t)dt$ is such that

$$\int \nabla \cdot (f\mu)v = \frac{1}{T}(v(x) - v(\phi_T(x))) = 0$$

so it is invariant

Stability

Assuming $f(0) = 0$, Rantzer (2000) observes that the existence of a density ρ such that

$$\boxed{\nabla \cdot (f\rho) > 0}$$

implies that $x(t) \to 0$ when $t \to \infty$ for almost all $x(0)$

Dual to existence of a Lyapunov function $v(x)$ such that

$$v > 0, \;\; \nabla v \cdot f < 0$$

Outline

1.1. Measures, moments and LMI
1.2. Polynomial optimisation
1.3. Examples and software

2.1. Occupation measures and dynamical systems
2.2. Stability analysis
2.3. Polynomial optimal control
2.4. Examples and software

Optimal control and value function

Consider the optimal control problem (OCP)
with fixed initial condition and fixed terminal time

$$\begin{array}{rcll} v(x_0) & = & \min_{u \in U} & g(x(T)) + \int_0^T h(x,u)dt \\ & & \text{s.t.} & \dot{x} = f(x,u), \quad x(0) = x_0 \end{array}$$

Cost $x \mapsto v(x)$ is called the value function

Using calculus of variations, it can be shown that
the value function satisfies a nonlinear first order PDE..

HJB PDE

Defining the Hamiltonian

$$H(x,p) = \min_{u \in U}\{h(x,u) + p \cdot f(x,u)\}$$

the value function solves the Hamilton-Jacobi-Bellman PDE

$$H(x, \nabla v) = 0, \quad v(x(T)) = g(x(T))$$

Under standard assumptions, the HJB PDE has a unique viscosity solution $v^* = \lim_{\varepsilon \to 0} v_\varepsilon$ with

$$H(x, \nabla v_\varepsilon) = \varepsilon \Delta v_\varepsilon, \quad v_\varepsilon(x(T)) = g(x(T))$$

cf. P.-L. Lions (1983)

Feedback control from solution of HJB PDE

At time t for a given state $x(t)$ we let

$$u^*(x(t)) = \arg\min_u \{h(x,u) + \nabla v^* \cdot f(x,u)\}$$

so that the Hamiltonian is minimized, i.e.

$$h(x,u^*) + \nabla v^* \cdot f(x,u^*) = H(x,\nabla v^*)$$

This is an optimal feedback control policy

Polynomial optimal control

Consider now the OCP

$$\begin{array}{ll} \min_u & g(x(T)) + \int_0^T h(x,u)dt \\ \text{s.t.} & \dot{x} = f(x,u) \\ & x(0) \in X_0,\ x(T) \in X_T \\ & x \in X,\ u \in U \end{array}$$

with f, g, h polynomials and X_0, X_T, X, U compact basic semialgebraic sets (intersections of polynomial sublevel sets)

Weak formulation = LP on measures

Suppose μ_0 is given, the OCP can be written as a linear but infinite-dimensional problem on measures μ and μ_T

$$\begin{array}{ll} \min_{\mu,\mu_T} & \int_{X_T} g d\mu_T + \int_X h d\mu \\ \text{s.t.} & \int_X \nabla v \cdot f(x,u) d\mu = \int_{X_T} v d\mu_T - \int_{X_0} v d\mu_0, \quad \forall v \in C \end{array}$$

Without test functions it can be written

$$\begin{array}{ll} \min_{\mu,\mu_T} & < g, \mu_T > + < h, \mu > \\ \text{s.t.} & \mu_T + \nabla \cdot (f\mu) = \mu_0 \end{array}$$

or more abstractly

$$\begin{array}{ll} \min_{\nu} & < c, \nu > \\ \text{s.t.} & < A, \nu > = b, \quad \nu \in M_+(X) \times M_+(X_T) \end{array}$$

as a primal LP on the Banach space of nonnegative measures

Dual LP

Using duality on compact Banach spaces, we obtain

$$\begin{array}{ll} \max_v & < b, v > \\ \text{s.t.} & c - < A^*, v > \in \; C_+(X) \times C_+(X_T) \end{array}$$

a dual LP on the space of nonnegative continuous functions that can be written explicitly as

$$\begin{array}{ll} \max_v & < \mu_0, v > = \int_{X_0} v d\mu_0 \\ \text{s.t.} & < \mu, h + \nabla v \cdot f > = \int_X (h + \nabla v \cdot f) d\mu \geq 0 \\ & < \mu_T, g - v > = \int_{X_T} (g - v) d\mu_T \geq 0 \end{array}$$

Conic complementarity

By conic complementarity, along optimal trajectories (x^*, u^*) and for optimal dual function v^* it holds

$$< h + \nabla v^* \cdot f, \mu^* >=< g - v^*, \mu_T^* >= 0$$

or equivalently

$$\begin{array}{l} H(x^*, \nabla v^*) = h(x^*, u^*) + \nabla v^* \cdot f(x^*, u^*) = 0 \\ v^*(x^*(T)) = g(x^*(T)) \end{array}$$

which means that v^* solves the HJB PDE

Solving primal problem on measures = solving OCP

Solving dual problem on functions = solving HJB PDE

How do we proceed numerically ?

Generalized problem of moments

We face linear problems involving several measures μ_i respectively supported on semialgebraic sets X_i

All the data are polynomials, so we can replace measures by their moments (e.g. $\int_{X_i} h_i(x)d\mu_i = \int_{X_i} \sum_\alpha h_{i\alpha} x^\alpha d\mu_i = \sum_\alpha h_{i\alpha} \int_{X_i} x^\alpha d\mu_i$)

$$\begin{array}{ll} \min_\mu & \sum_i \int_{X_i} h_i d\mu_i \\ \text{s.t.} & \sum_i \int_{X_i} a_{ij} d\mu_i = b_j \\ & \text{measures } \mu_i \end{array} \qquad \begin{array}{ll} \min_y & \sum_i \sum_\alpha h_{i\alpha} y_{i\alpha} \\ \text{s.t.} & \sum_i \sum_\alpha a_{ij_\alpha} y_{i\alpha} = b_j \\ & \text{moments } y_i \end{array}$$

provided we can handle the representation condition

$$y_{i\alpha} = \int_{X_i} x^\alpha d\mu_i(x)$$

Moment LP as LMI

Using Putinar's representation conditions we obtain

$$\begin{array}{ll} \min_y & c^T y \\ \text{s.t.} & Ay = b \\ & y_\alpha = \int_X x^\alpha d\mu \\ & X = \{x \,:\, p_k(x) \geq 0, \forall k\} \end{array}$$

infinite-dimensional

LP problem

$$\begin{array}{ll} \min_y & c^T y \\ \text{s.t.} & Ay = b \\ & M_d(y) \succeq 0 \\ & M_d(p_k y) \succeq 0, \forall k \end{array}$$

finite-dim. LMI

relaxation of order d

our familiar hierarchy of LMI relaxations

Compare with static polynomial optimisation: dynamics are now taken into account by introducing several measures whose moments are linearly constrained

Dual function LP as LMI

Dual to LMI moment problem yields polynomial supersolution of HJB PDE with polynomial sign conditions enforced by polynomial SOS conditions

Good approximation of value function along optimal trajectories

For example, if $f(x,u) = f_1(x) + f_2(u)$ and $h(x,u) = h_1(x) + u^T u$ use first-order optimality condition

$$\partial_u(h(x,u) + \nabla v^* \cdot f(x,u)) = 2u + \nabla v^* \cdot f_2 = 0$$

to derive state-feedback control law

$$u^*(x) = -\frac{1}{2}\nabla v^*(x) \cdot f_2(x)$$

Outline

Example of linear ODE analysis

Consider the scalar linear ODE

$$\dot{x} = -x$$

with initial measure μ_0 in $X_0 = \{x \, : \, p_0(x) = \frac{1}{4} - (x - \frac{3}{2})^2 \geq 0\}$
with terminal measure μ_T in $X_T = \{x \, : \, p_T(x) = \frac{1}{4} - x^2 \geq 0\}$
with occupation measure μ in $X = \{x \, : \, p(x) = 4 - x^2 \geq 0\}$

We want to find trajectories minimising the energy $\int_0^T x^2 dt$

Linear measure problem

$$\begin{array}{rl} \text{min} & \int_0^T x^2 d\mu(x) \\ \text{s.t.} & \int_X \nabla v(x)(-x) d\mu(x) = \int_{X_T} v d\mu_T - \int_{X_0} v d\mu_0, \quad \forall v \end{array}$$

Example of linear ODE analysis

Setting $v = x^\alpha$ we introduce sequences y_0, y_T, y representing measures μ_0, μ_T, μ, and we obtain the linear moment problem

$$\begin{array}{rl} \min & y_2 \\ \text{s.t.} & -\alpha y_\alpha = y_{T\alpha} - y_{0\alpha}, \quad \forall \alpha \end{array}$$

and the corresponding LMI relaxation of order d

$$\begin{array}{rl} \min & y_2 \\ \text{s.t.} & -\alpha y_\alpha = y_{T\alpha} - y_{0\alpha}, \quad \forall \alpha,\ |\alpha| \leq 2d \\ & M_d(y_0) \succeq 0,\ M_d(y_T) \succeq 0,\ M_d(y) \succeq 0 \\ & M_d(p_0 y_0) \succeq 0,\ M_d(p_T y_T) \succeq 0,\ M_d(py) \succeq 0 \end{array}$$

Solving LMI relaxations of increasing orders d yields a sequence of monotonically increasing lower bounds on the optimum

Example of linear ODE analysis

This problem can be solved analytically, with optimal trajectory $x(t) = e^{-t}$ leaving X_0 at $x(0) = 1$ and reaching X_T at $x(T) = \frac{1}{2}$ for $T = \log 2 \approx 0.6931$

Moment matrix $M(y)$ has entries $y_\alpha = \int_0^{\log 2} e^{-\alpha t} dt = \frac{1-2^{-\alpha}}{\alpha}$

We get with SeDuMi 1.1R3 the following sequence of valid significant digits on T: 0, 2, 4, 7, 10, 13 (fast convergence)

Convergence at a finite relaxation order is impossible since the optimum is transcendental, whereas the solution of an integer coefficient LMI is algebraic

Linear ODE analysis

More generally, for the first-order linear Cauchy problem

$$\dot{x} = Ax, \;\; x(0) = x_0, \;\; x(\infty) = 0$$

the moment LMI problem reads

$$AQ + (AQ)^T = Q_0 \succeq 0, \;\; Q \succeq 0$$

with Q_0, Q nonzero covariance matrices of μ_0 (initial measure) and μ (occupation measure) respectively

Infeasible if and only if dual Lyapunov LMI problem

$$A^T P + PA \prec 0, \;\; P \succ 0$$

is feasible

Example of LQR design

Consider the linear quadratic regulator design problem

$$\begin{array}{ll} \min_{u,T} & \int_0^T (x^2+u^2)dt \\ \text{s.t.} & \dot{x} = u \\ & x(0) = x_0,\ x(T) = x_T \end{array}$$

with given initial and terminal conditions

Measure μ_0 is the Dirac δ_{x_0} and measure μ_T is the Dirac δ_{x_T} so that only measure μ must be found

Linear measure problem

$$\begin{array}{rl} \min_{\mu} & \int_0^T (x^2+u^2)d\mu(x,u) \\ \text{s.t.} & \int_{\mathbb{R}^2} \nabla v(x) u d\mu(x,u) = v(x_T) - v(x_0), \quad \forall v \end{array}$$

Example of LQR design

Moment LMI problem

$$\begin{array}{rl} \min & y_{20} + y_{02} \\ \text{s.t.} & y_{01} = 2y_{11} = 3y_{21} = \cdots = -1 \\ & M_d(y) \succeq 0 \end{array}$$

The moment matrix has the following quasi-Hankel structure

$$M_d(y) = \begin{bmatrix} y_{00} & y_{10} & y_{01} & \\ y_{10} & y_{20} & y_{11} & \\ y_{01} & y_{11} & y_{02} & \\ & & & \ddots \end{bmatrix}$$

with

$$y_\alpha = \int x^{\alpha_1} u^{\alpha_2} d\mu(x,u)$$

Example of LQR design

Solving with SeDuMi 1.1R3 the LMI relaxation of order $d = 1$ yields

$$M_1(y) = \begin{bmatrix} 3.66 & 1.00 & -1.00 \\ 1.00 & 0.50 & -0.50 \\ -1.00 & -0.50 & 0.50 \end{bmatrix} \quad X = \begin{bmatrix} 0.00 & 0.00 & 0.00 \\ 0.00 & 1.00 & 1.00 \\ 0.00 & 1.00 & 1.00 \end{bmatrix}$$

where X is the multiplier matrix such that trace $M_1(y)X = 0$

From entries $M_1(y)$ we can read the optimal trajectory, with $T \approx 3.66$ (exact value $= \infty$, but objective function almost equal), $\int x = 1$, $\int u = -1$, $\int x^2 = \frac{1}{2}$ etc

Example of LQR design

From multipliers corresponding to equality constraints we retrieve $v^*(x) = x^2$ as a polynomial subsolution to the HJB PDE, which is here a standard algebraic Riccati equation

From X we notice that the sum of the second (indexed by x) and third (indexed by u) row/column in $M(y)$ vanishes, so that the optimal control policy $u^*(x)$ satisfies the equation

$$x + u^*(x) = 0$$

We can also use the optimality condition

$$u^*(x) = -\frac{1}{2}\nabla v^*(x) = -x$$

LQR design

More generally, consider the LQR design problem

$$\begin{array}{ll} \min_u & \int_0^\infty (x^T R x + u^T u) dt \\ \text{s.t.} & \dot{x} = Ax + Bu, \;\; x(0) = x_0 \end{array}$$

Moment LMI problem restricted to covariance matrices

$$\begin{array}{ll} \min_Q & \operatorname{trace} RQ_{11} + \operatorname{trace} Q_{22} \\ \text{s.t.} & AQ_{11} + BQ_{21} + (AQ_{11} + BQ_{21})^T + Q_0 = 0 \\ & Q = \left[\begin{array}{cc} Q_{11} & Q_{21}^T \\ Q_{21} & Q_{22} \end{array} \right] \succeq 0 \end{array}$$

Optimal static state-feedback $u = Kx$ with

$$K = Q_{21} Q_{11}^{-1}$$

Convex design LMI, cf. Bernussou, Geromel, Peres (1988)

LQR design

Dual LMI problem

$$\begin{array}{ll} \max & x_0^T P x_0 \\ \text{s.t.} & \begin{bmatrix} A^T P + PA + R & PB \\ B^T P & -I \end{bmatrix} \preceq 0, \;\; P \succeq 0 \end{array}$$

Complementarity conditions

$$\begin{bmatrix} Q_{11} & Q_{21}^T \\ Q_{21} & Q_{22} \end{bmatrix} \begin{bmatrix} A^T P + PA + R & PB \\ B^T P & -I \end{bmatrix} =$$

$$\begin{bmatrix} I \\ K \end{bmatrix} \begin{bmatrix} I & K^T \end{bmatrix} \begin{bmatrix} A^T P + PA + R & PB \\ B^T P & -I \end{bmatrix} = 0$$

Null-space of covariance matrix of occupation measure provides optimal state-feedback

Nonlinear stabilization

Consider nonlinear polynomial system

$$\dot{x} = f(x) + g(x)u$$

Apply Rantzer's density condition

$$\nabla \cdot ((f + gu)\rho) > 0$$

and choose

$$\rho(x) = \frac{p_1(x)}{p_0(x)}, \quad u(x)\rho(x) = \frac{p_2(x)}{p_0(x)}$$

with $p_0(x)$ given positive polynomial ensuring integrability and $p_1(x)$, $p_2(x)$ polynomials to be found

Positivity relaxed to SOS

Convex design LMI, rational stabilizing feedback $u(x) = \frac{p_2(x)}{p_1(x)}$

Software

GloptiPoly 3 (DH, JB. Lasserre, J. Löfberg) for Matlab models generalised problems of moments as LMI problems

POCP (C. Savorgnan) models polynomial optimal control problems as generalised problems of moments

`homepages.laas.fr/henrion/software`

Can explicitly address state constraints, impulsive controls, discontinuous trajectories..

Example

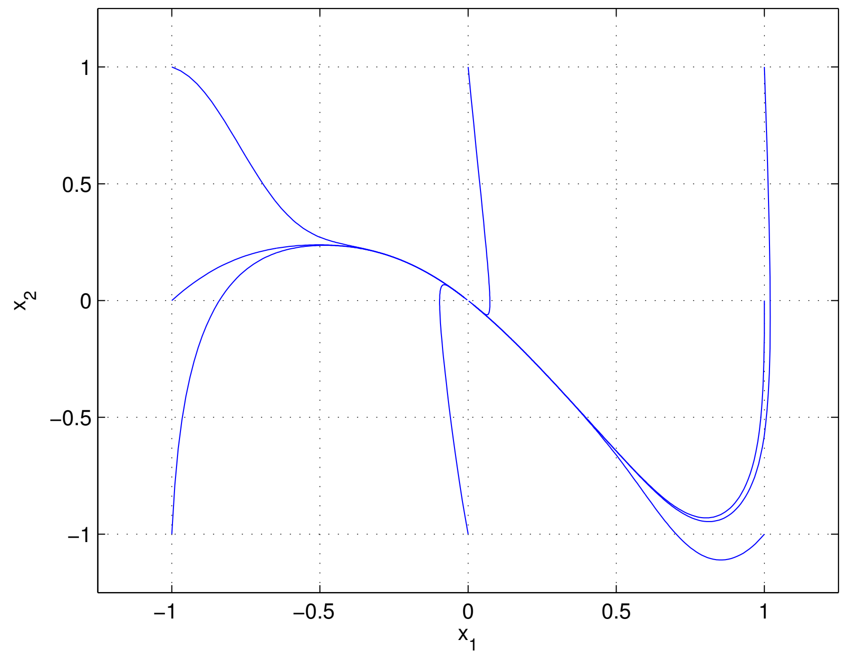
POCP translates polynomial optimal control problem into generalised problem of moments in GloptiPoly 3 format

LMI relaxations then solved with SDP solver (e.g. SeDuMi)

$$\begin{array}{ll} \min_{u,T} & \int_0^T x_1^2 + x_2^2 + \frac{u^2}{100} \\ \text{s.t.} & \dot{x}_1 = x_2 - x_1^3 + x_1^2, \quad \dot{x}_2 = u \\ & x(0) \in [-1,1]^2,\ x(T) = 0 \end{array}$$

```
mpol x 2; mpol u
% problem definition
P = pocp('state', x, 'input', u, ...
 'dynamics', [x(2)+x(1)^2-x(1)^3; u], ...
 'horizon', 0, 'iuniform', x, [-1 1; -1 1], ...
 'fdirac', x, [0;0], 'scost', x'*x+u^2/100);
```

```
% problem solved with test function V of degree 8
[status, J, mu, v] = solvepocp(prob, 'tf', 8);
% gradient control law
ux = -50 * diff(v, x)*[0; 1];
```



Piecewise affine optimal control

Optimal control problem

$$\begin{array}{ll} \min_{u,T} & \int_0^T h(x,u)dt + h_T(x(T)) \\ \text{s.t.} & \dot{x} = A_i x + a_i + B_i u \quad \text{when } x \in X_i \\ & x(0) = x_0, \quad x(T) = x_T \end{array}$$

where state-space is partitioned

$$X = \cup_i X_i$$

into compact basic semialgebraic sets X_i (e.g. polytopes)

Multiple open-loop equilibrium points since $a_i \neq 0$

We want a good approximation of the optimal control law $u^*(x)$

Several measures

Introduce local occupation measures μ_i supported in each cell X_i

Global occupation measure $\mu = \sum_i \mu_i$

Linear optimal cost

$$\int_X h d\mu + \int_{X_T} h_T d\mu_T$$

and linear constraints on measures

$$\int_X \nabla v \cdot \sum_i (A_i x + a_i + B_i u) d\mu_i = \int_{X_T} v d\mu_T - \int_{X_0} v d\mu_0, \quad \forall v$$

and hence linear constraints on repective moment sequences y_i

Example

Consider the nonlinear system

$$\dot{x} = \frac{1}{2}(1 - x^2) + u$$

with two equilibrium points, approximated globally by a piecewise affine system

$$\begin{aligned} \dot{x} &= -x + 1 + u \quad \text{if } x \geq 0 \\ &= x + 1 + u \quad \text{if } x \leq 0 \end{aligned}$$

and we would like to solve the optimal control problem

$$\min_{u,T} \int_0^T (2(1-x)^2 + u^2)dt$$

with boundary conditions

$$x(0) = -1, \quad x(T) = +1$$

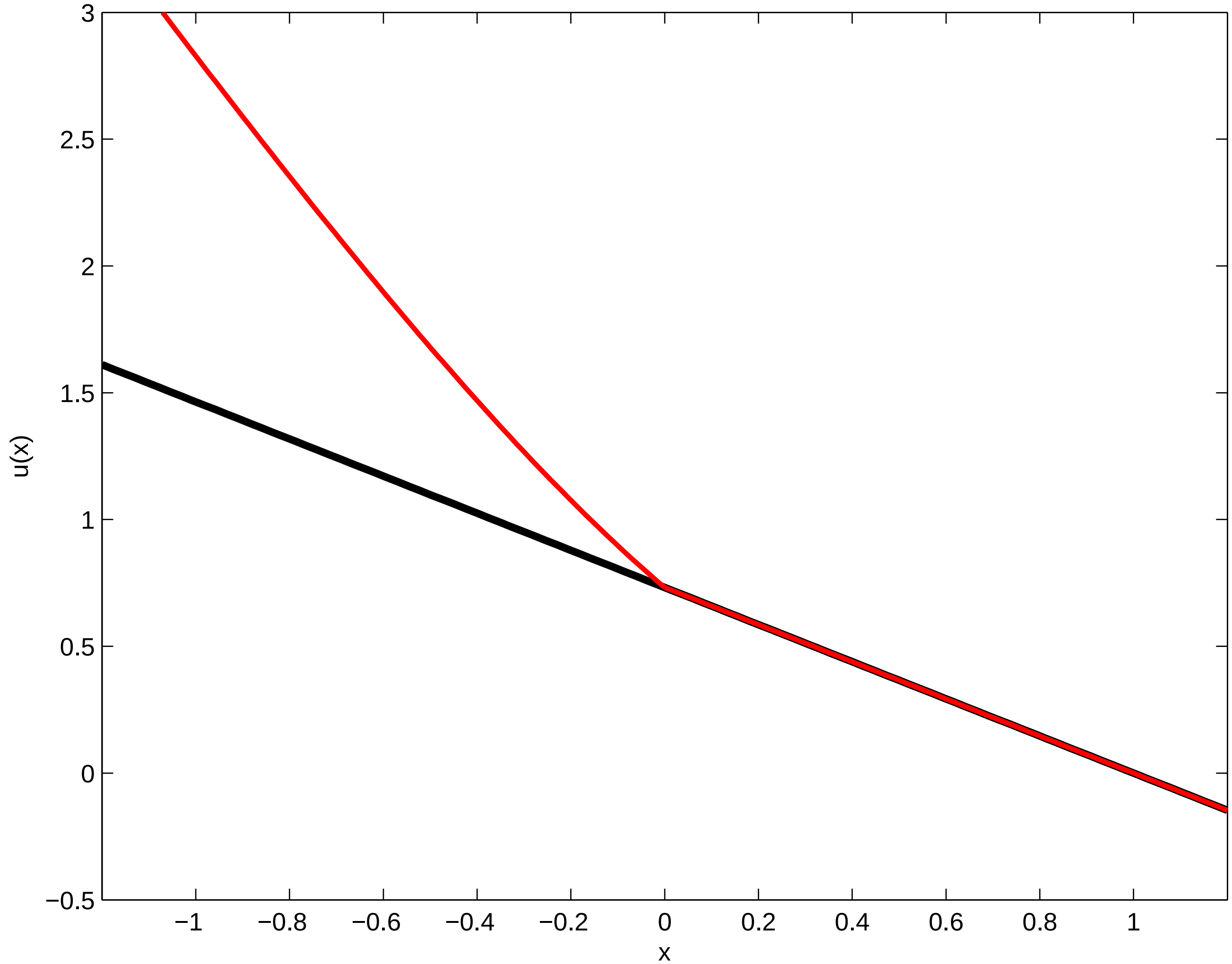
Optimal control

From necessary optimality conditions on the piecewise affine Hamiltonian we obtain the analytic solution

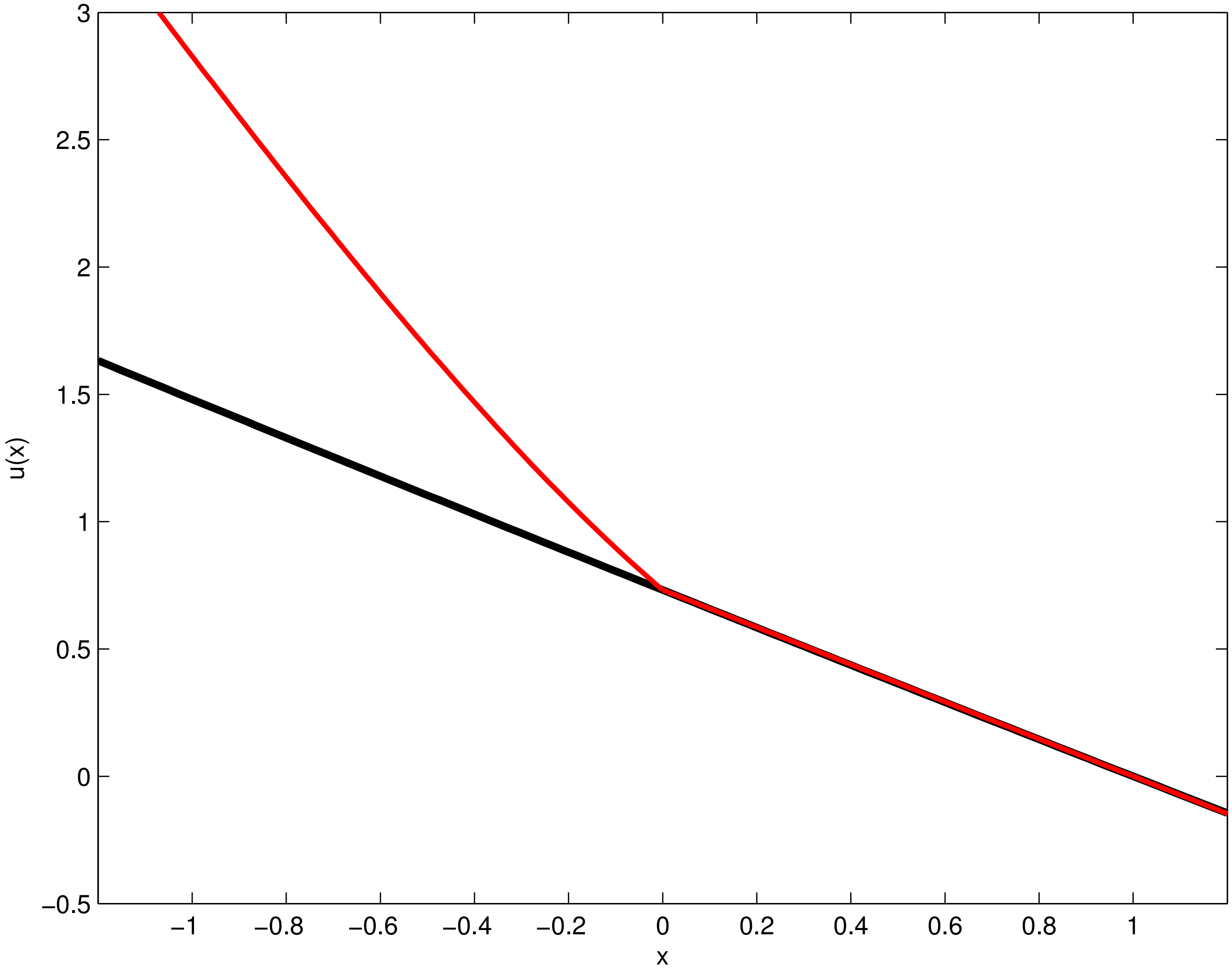
$$
\begin{aligned}
u^*(x) &= (1-\sqrt{3})(x-1) && \text{if } x \geq 0 \\
&= -x-1+\sqrt{2(x-1)^2+(x+1)^2} && \text{if } x \leq 0
\end{aligned}
$$

showing in passing that an optimal controller for a PWA system is not necessarily PWA

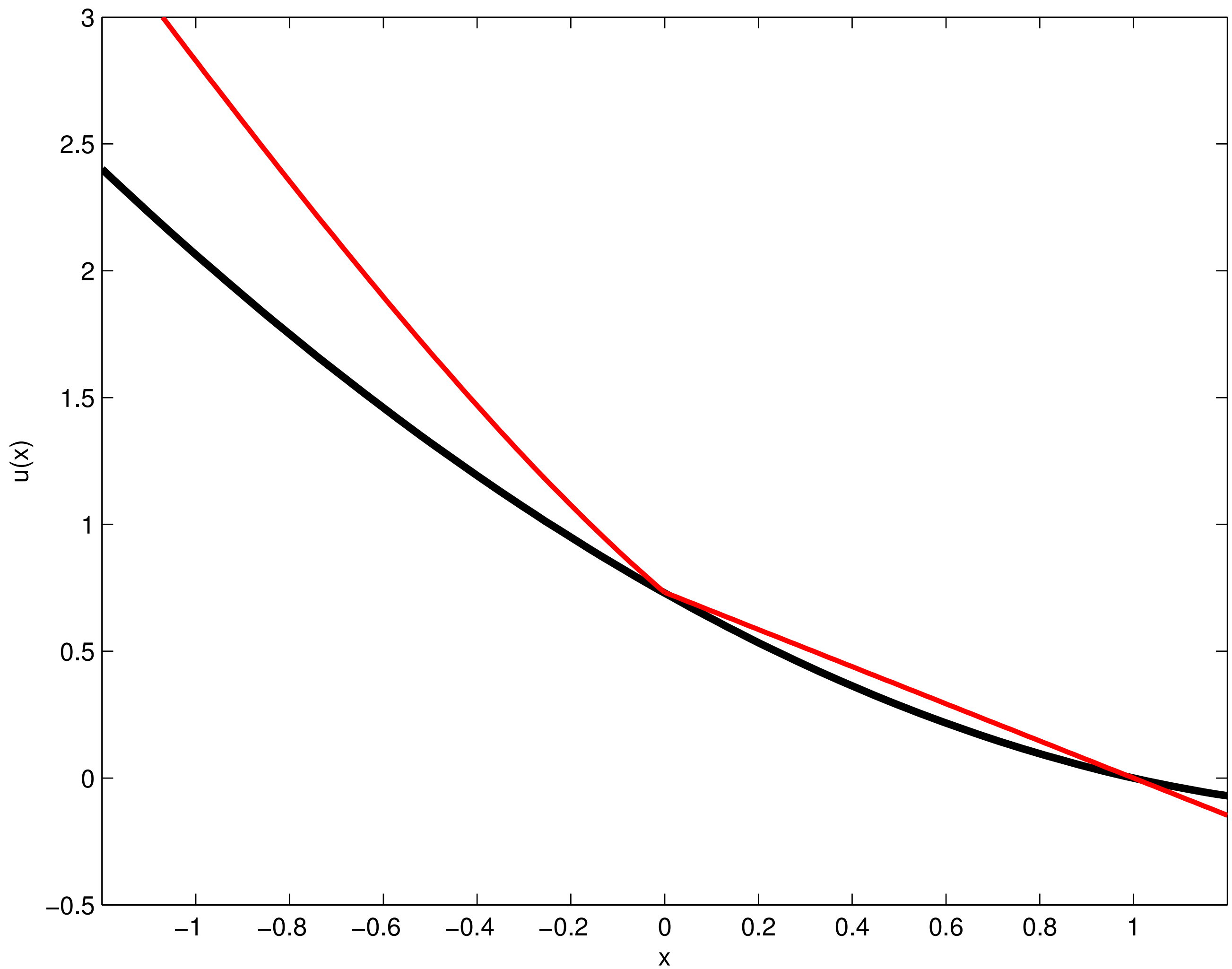
Using LMI relaxations of orders $1, 2, 3, \ldots, 10$
we obtain the following approximations to $u^*(x)$



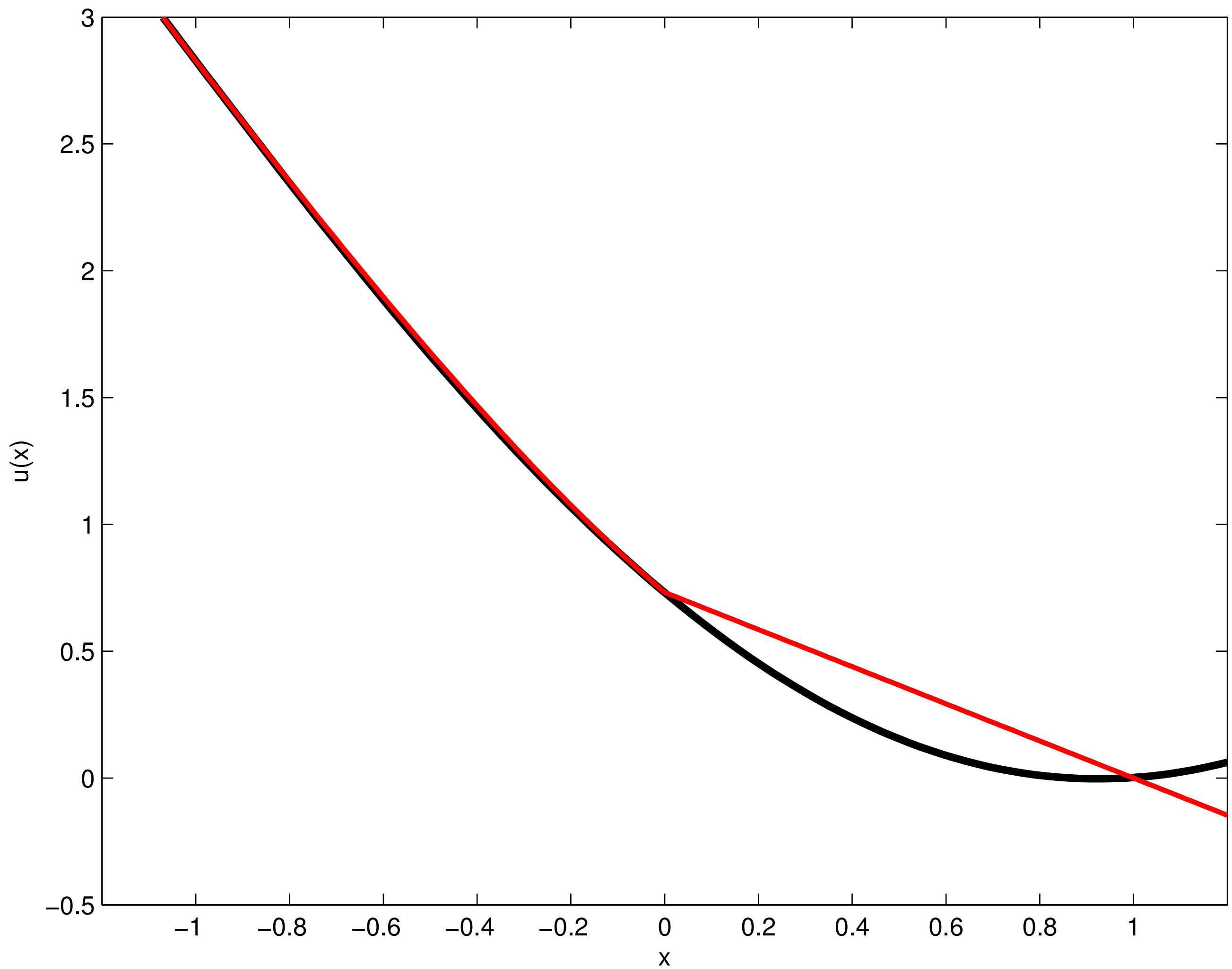
Optimal feedback (red) and degree 2 approximation (black)



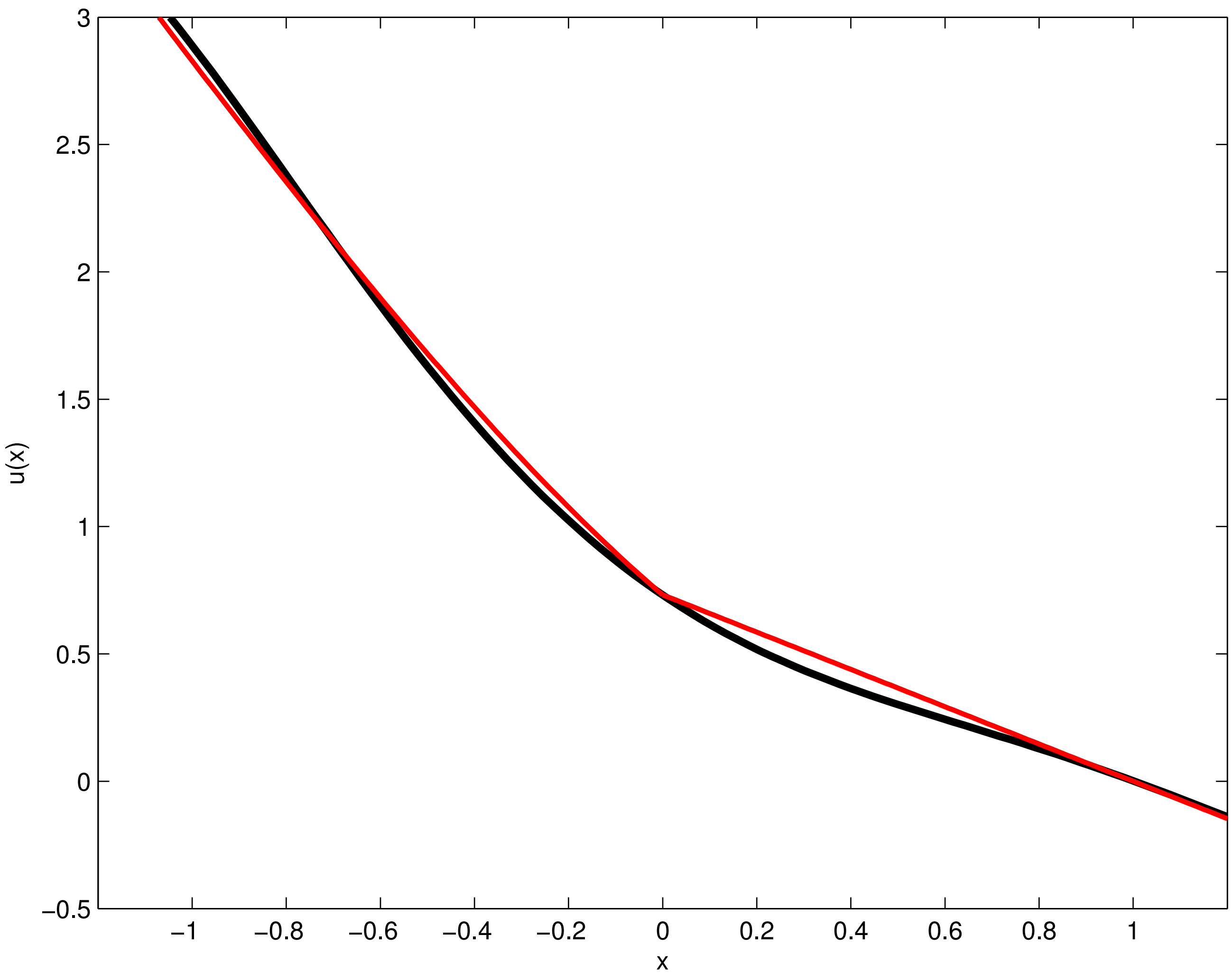
Optimal feedback (red) and degree 4 approximation (black)



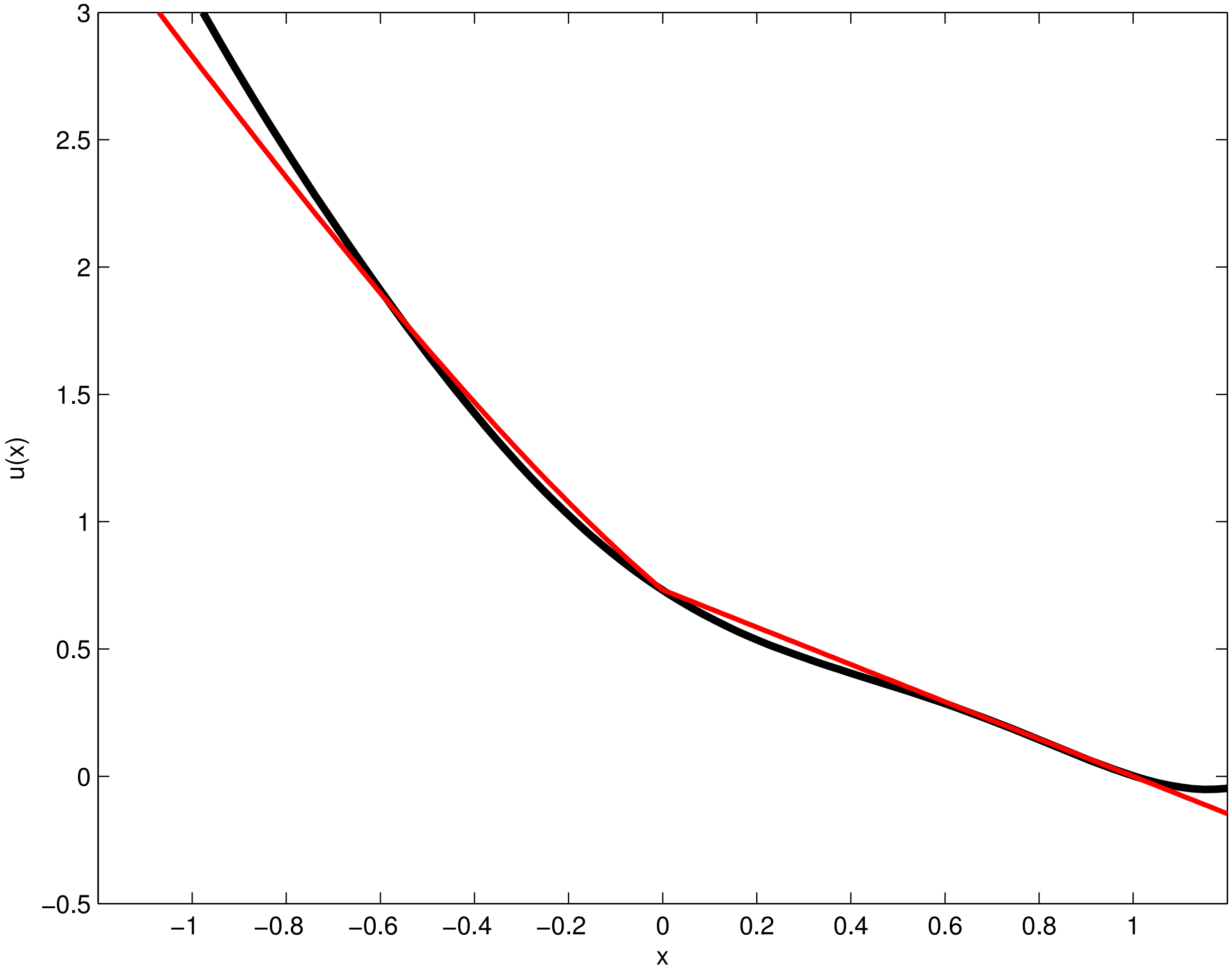
Optimal feedback (red) and degree 6 approximation (black)



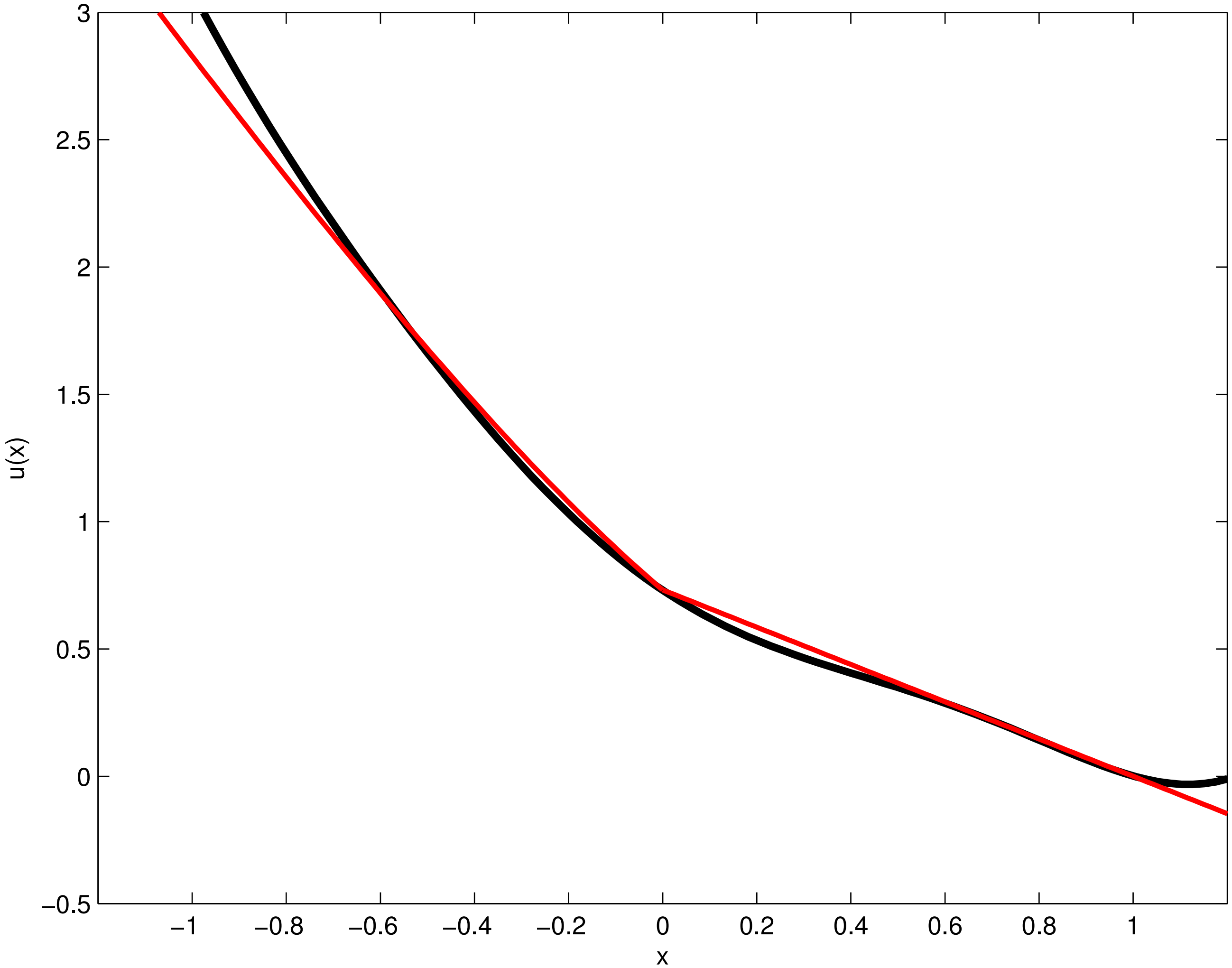
Optimal feedback (red) and degree 8 approximation (black)



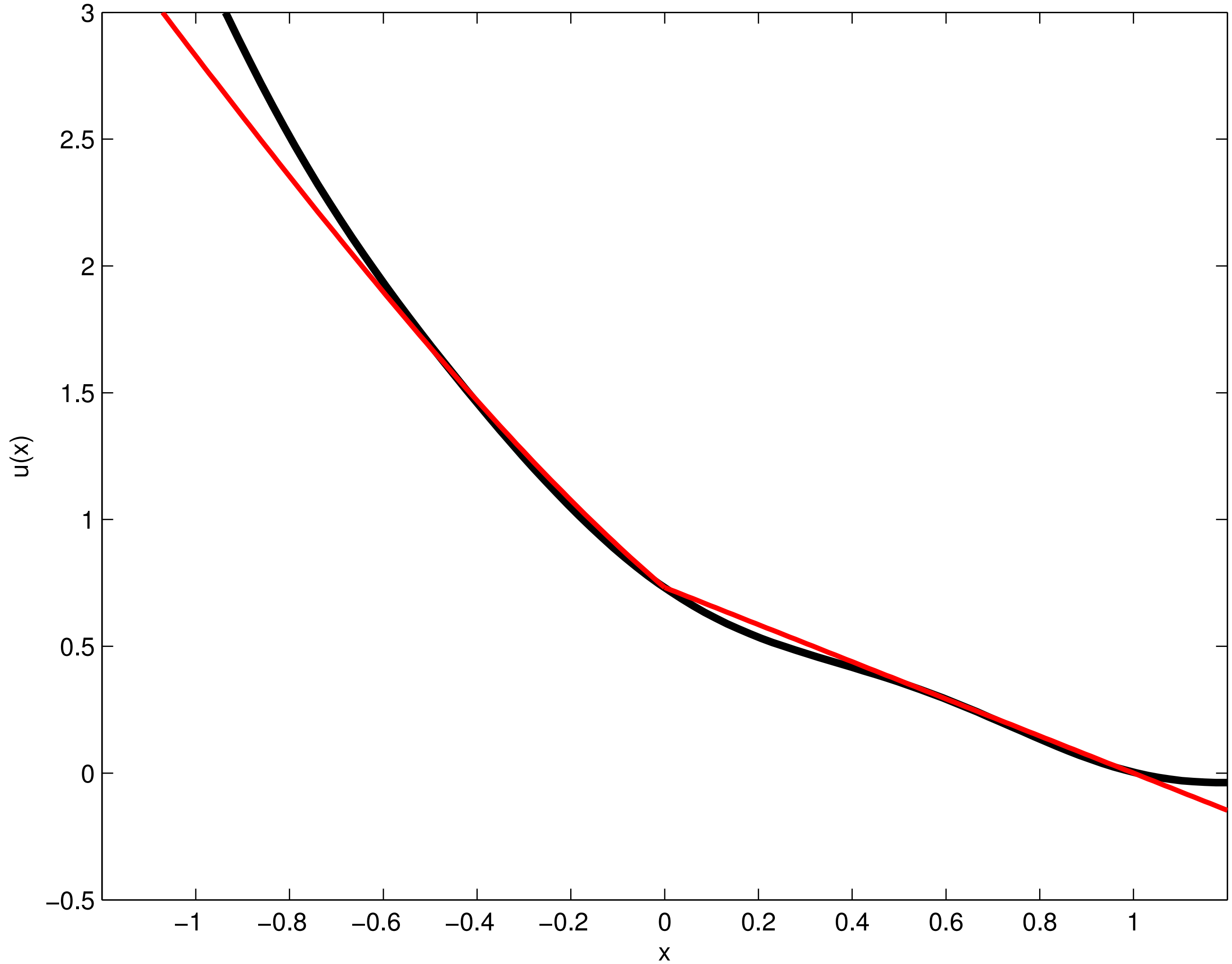
Optimal feedback (red) and degree 10 approximation (black)



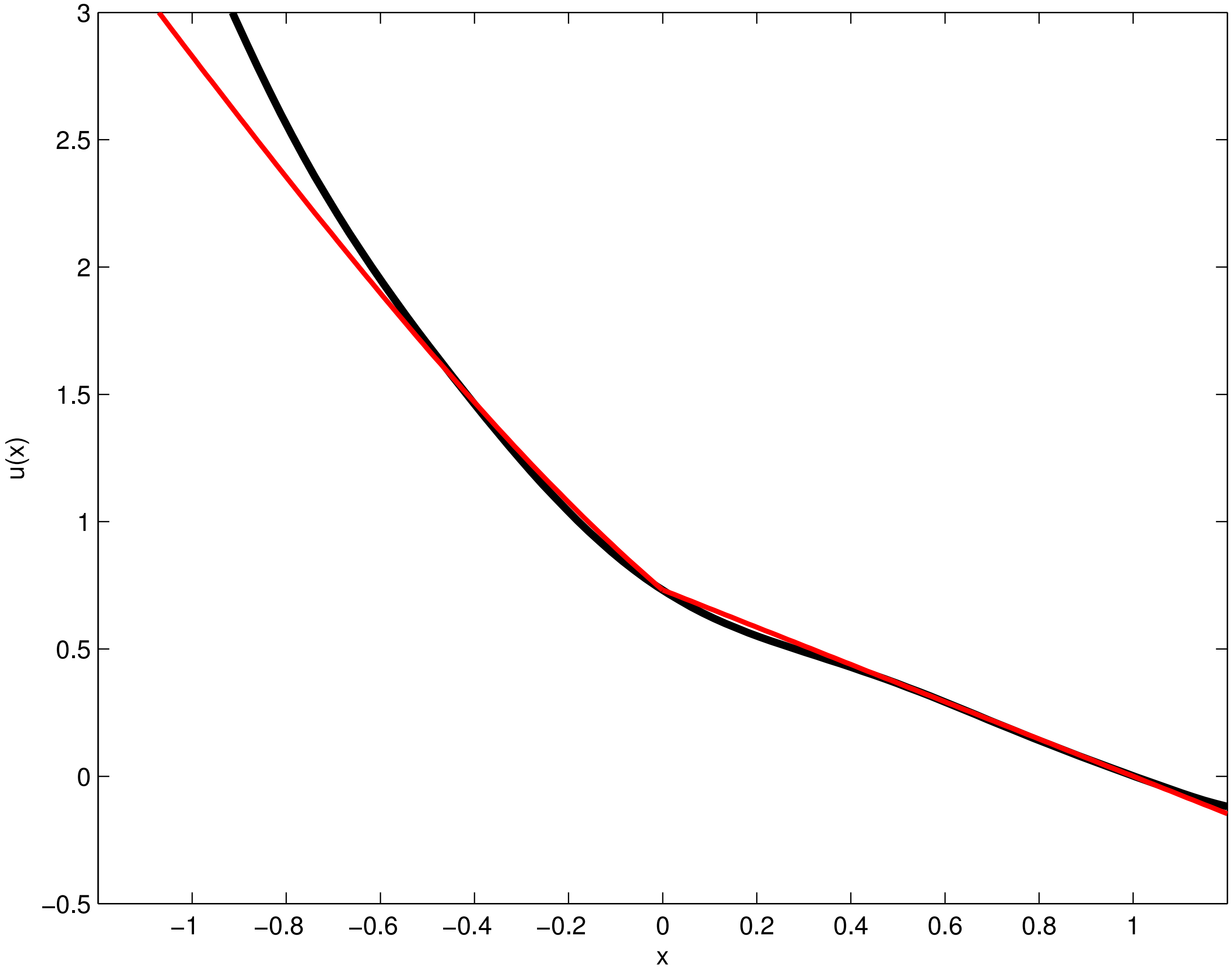
Optimal feedback (red) and degree 12 approximation (black)



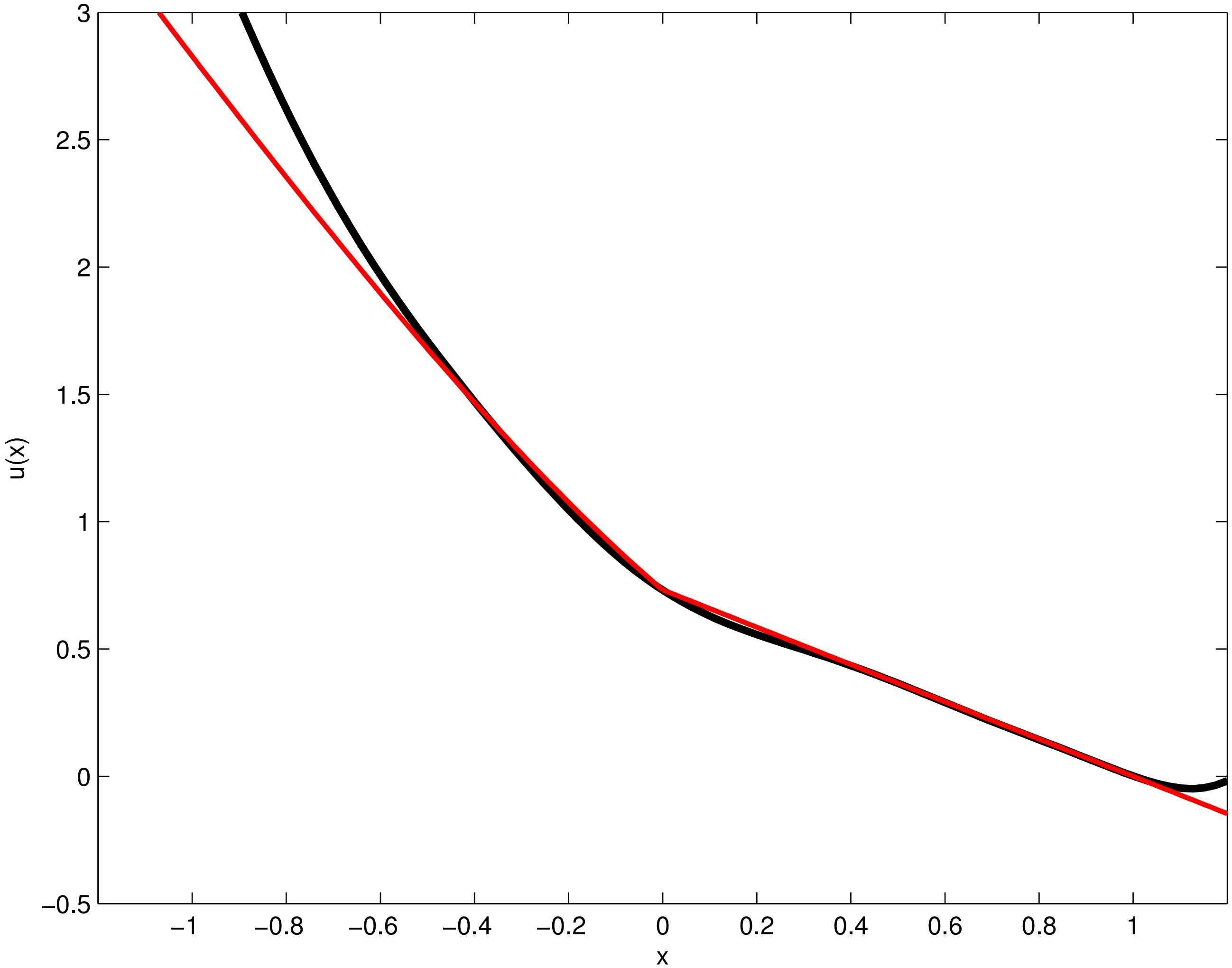
Optimal feedback (red) and degree 14 approximation (black)



Optimal feedback (red) and degree 16 approximation (black)



Optimal feedback (red) and degree 18 approximation (black)



Optimal feedback (red) and degree 20 approximation (black)

Concluding remarks

Hierarchy of LMI relaxations for
- static polynomial optimisation
- (polynomial approximation of) optimal control of dynamical systems (with polynomial dynamics)

Occupation measures can handle piecewise affine models, but also much more (piecewise polynomial models, impulsive controls, hybrid dynamics)

Dynamical systems theory: Liouville's theorem, advection PDEs

Discrete-time piecewise affine systems: chaotic dynamics (e.g. tent map), invariant measures, ergodic theory, Frobenius-Perron and Koopmans operators