Applied algebraic geometry: polynomial optimisation

Jan Draisma, Universität Bern

New topic: Optimisation

Typical goal: maximise $f(x)$ subject to $x \in X$ (1)

Here $X \subseteq \mathbb{R}^n$ is a compact set given by polynomial weak inequalities (and equalities) and f is a polynomial function.

We'll discuss an approach due to Lasserre.

Write $R = \mathbb{C}[x_1, \ldots, x_n]$ and write R_d for poly of degree $\leq d$.

Simple idea: think of points in X as linear functions on R. This linear function $\mathrm{ev}_x : R \to \mathbb{R}, h \mapsto h(x)$ satisfies:

1. $\mathrm{ev}(1) = 1$
2. $\mathrm{ev}_x(h) \geq 0$ for all $h \in R$ that are nonnegative everywhere on X.

We find that the maximum we're trying to find has the *upper bound* the following maximum:

Relaxation

$\sup \ell(f)$ where $\ell : R \to \mathbb{R}$ is linear and satisfies $\ell(1) = 1$, $\ell(h) \geq 0$ for all h nonnegative on X. (2)

Proposition

In fact, this sup is attained at ev_x for some $x \in X$.

Proof: First, note that ℓ is automatically continuous in the ∞-norm: if $|h - h'|_\infty < \epsilon$ then $\epsilon - l(h - h') = l(\epsilon - (h - h')) \geq 0$, etc. By Stone-Weierstrass, since R separates points of X, R is dense in $C^0(X, R)$. Hence ℓ extends uniquely to a continuous linear functional $\ell : C^0(X, R) \to \mathbb{R}$ with the same properties. (If h is continous and nonnegative on X, approximate $k := h + \epsilon 1$ by a polynomial p with $|p - k|_\infty \leq \epsilon$ and so that $|\ell(p - k)| \leq \epsilon$. Then $\ell(p) \geq 0$ implies $\ell(h) \geq -2\epsilon$.)

By the Riesz representation theorem, there is a probability measure μ on X such that $\ell(h) = \int_X h(x)\mathrm{d}\mu(x)$ for all $h \in C^0(X, \mathbb{R})$. Then, $\ell(f) = \int_X f(x)d\mu(x) \leq \max_{x\in X} f(x)$. □

For each $d \geq \deg(f)$ we get the following further:

Relaxation

$\max \ell(f)$ such that $\ell : R_d \to \mathbb{R}$ linear, $\ell(1) = 1$ and $\ell(h) \geq 0$ for all $h \in R_d$ nonnegative on X. $(3)_d$

This is a finite-dimensional linear program, but still with infinitely many, hard-to-handle constraints. Note that the maximum is attained, since for every monomial $m \in R_d$ the functions $c1 - m$ and $c1 + m$ are nonnegative on X for $c \gg 0$, so the domain over which ℓ runs is compact. Using Tychonov, find that the max in $(3)_d$ converges to the sup/max in (2).

How to certify positivity?

There are polynomials that are clearly positive everywhere, namely *sums of squares*: $p_1^2 + \ldots + p_k^2$ where the p_i are polynomials. Let $S \subseteq R$ be the set of these. It is closed under addition.

Now assume $X = \{x \in \mathbb{R}^n \mid g_1(x), \ldots, g_r(x) \geq 0\}$. Set $g_0 := 1$. Define the *quadratic module* generated by the g_i as $M := \sum_{i=0}^{m} S\, g_r$. These polynomials are clearly nonnegative on X.

So it is natural to replace the condition "$\ell(h) \geq 0$ for $h \in R_d$ nonnegative on X" in $(3)_d$ by "$\ell(h) \geq 0$ for $h \in M_d$".

The condition $\ell(h) \geq 0$ for each $h \in M_d$ is equivalent to $\ell(p^2 g_i) \geq 0$ for each $p \in R_{\lfloor (d - \deg(g_i))/2 \rfloor} =: P_{i,d}$ and each $i = 1, \ldots, r$.

Let β_i be the symmetric bilinear form on $P_{i,d}$ defined by $\beta_i(p,q) := \ell(pg_iq)$. Then the above is equivalent to β_i being *positive semidefinite*.

This leads to the following optimisation problem:

Lasserre's hierarchy

(For any $d \geq 2\deg f$): $\max \beta_0(f,1)$ where $\beta_0,\ldots,\beta_r$ are symmetric bilinear forms on $P_{0,d},\ldots,P_{r,d}$ satisfying

1. $\beta_0(1,1) = 1$ and $\quad (4)_d$
2. $\beta_i(p,q) = \beta_j(p',q')$ for all i,j,p,q such that $pg_iq = p'g_jq'$ and
3. β_i is positive semidefinite.

Conditions 2,3 imply that there is a linear $\ell : R_d \to \mathbb{R}$ with $\beta_i(p,q) = \ell(pg_iq)$ for all i and $p,q \in P_{i,d}$ and with $\ell(M_d) \subseteq \mathbb{R}_{\geq 0}$.

This is a *semidefinite program*: except for the positive semidefiniteness condition, the conditions on the β_i are affine-linear, and we're optimising a linear function. The general form of a semidefinite program is as follows: *max* $\ell(t)$ *s.t.* $A_0 + t_1 A_1 + \ldots + A_m t_m$ *is positive semidefinite*. Here ℓ is linear and $A_0, \ldots, A_m$ are real symmetric matrices. The above can be put in this form by parameterising the affine space given by the linear conditions on the β_i.

Archimedean assumption

Assume that there exists a $u \in M$ such that $\{x \in \mathbb{R}^n \mid u(x) \geq 0\}$ is compact.

Remark

One can enforce this assumption by adding the constraint $g_{r+1}(x) := a - \|x\|^2 \geq 0$ for some sufficiently large a so that X is contained in this set.

Theorem
Under the Archimedean assumption, the optimal value of $(4)_d$ converges to the optimal value of (1) as $d \to \infty$.

Given the discussion above, this follows directly from the *Positivstellensätze* on the next slide.

Theorem (Schmüdgen)
$f \in R$ is *strictly positive* on the compact set X iff $f \in a \cdot 1 + \sum_{\epsilon\in\{0,1\}^r} S g_1^{\epsilon_1} \cdots g_r^{\epsilon_r}$ for some $a > 0$.

Theorem (Putinar)
Under the Archimedian assumption, every $f \in R$ *strictly positive* on X lies in M.

Examples (from Nie's lecture notes)
• $f = x_1 x_2 + 1$ is positive on $X = \{x \in \mathbb{R}^2 \mid g := 1 - (x_1^2 + x_2^2) \geq 0\}$, and indeed $f = \frac{1}{2}(x_1 + x_2)^2 + \frac{1}{2} g + \frac{1}{2}$.
• but "nonnegative" does not suffice: $f = 1 - x^2$ is nonnegative on $X = \{x \in \mathbb{R} \mid g(x) := (1 - x^2)^3 \geq 0$

Assume $f = s_1 + s_2 g$ for sums of squares s_1, s_2. Then -1 is a root of s_1, hence a root of all its square terms, hence of multiplicity at least 2 on the rhs, but of mult. 1 on the lhs, contradiction.

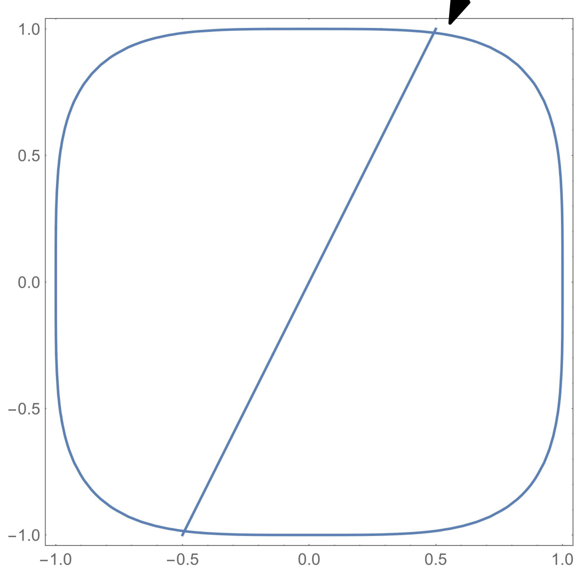
An example

Example

Maximise $f = x + 8y$ subject to $g := 1 - (x^4 + y^4) \geq 0.$

Classical approach: argue that maximum is taken on the boundary where $x^4 + y^4 \leq 1$. Necessary condition is that the derivative of f along the tangent direction to the boundary vanishes. That tangent direction at (x, y) equals $(y^3, -x^3)$, so we want $y^3 - 8x^3 = 0$.

maximum attained here



Then get $y = 2x$ (by reality) and $x = 17^{-1/4}$, so the maximum of f equals $17 \cdot 17^{-1/4} = 17^{3/4} \simeq 8.372144.$

Example, continued

Take $d = 4$.
So $R_d = \langle 1, x, y, x^2, xy, y^2, x^3, x^2y, xy^2, y^3, x^4, \ldots, y^4 \rangle$.

So $P_{0,4} = \langle 1, x, y, x^2, xy, y^2 \rangle$ and $P_{1,4} = \langle 1 \rangle$.

So β_0 has the matrix:

	1	x	y	x^2	xy	y^2
1	1	a	b	c	d	e
x	a	c	d	h	i	j
y	b	d	e	i	j	k
x^2	c	h	i	l	m	n
xy	d	i	j	m	n	o
y^2	e	j	k	n	o	p

Relaxation $(4)_d$
$\max a + 8b$ subject to $\mathrm{diag}(\beta_0, \beta_1)$ PSD.
Optimum using Sage: 8.372144
In hindsight, not surprising:
PSD $\Rightarrow l \geq c^2 \geq a^4$, and $p \geq e^2 \geq b^4$
so $0 \leq 1 - (l + p) \leq 1 - (a^4 + b^4) \ldots$

$\rightsquigarrow$ *we already have equality at $d = 4$!*

And β_1 has the 1×1-matrix $1 - (l + p)$. The condition is that the $(6 + 1) \times (6 + 1)$-block matrix $\mathrm{diag}(\beta_0, \beta_1)$ is positive semidefinite.