On the Role of the Critical Value Polynomial in Algebraic Optimization

Bernard Hanzon\textsuperscript{1}
Joint Work With
Andrei Mustata\textsuperscript{2}

\textsuperscript{1}Edgeworth Centre for Financial Mathematics, Department of Mathematics, University College Cork, Ireland
\textsuperscript{2}Department of Mathematics, University College Cork, Ireland

7 July 2014
Introduction

- Optimization is omnipresent in the mathematical sciences, including financial mathematics, control theory etc.

- Often criterion is a real polynomial or rational function in several variables.

- Then: First Order Conditions for Optimality described by a system of polynomial equations: hence Algebraic optimization problem:

- Algebraic techniques can be used to solve such a problem in principle, but there are computational and theoretical challenges: we have to proceed carefully!
Critical Value Polynomial

- Stress importance of "Critical Value Polynomial" that can be associated with an algebraic optimization problem.
- Can be used in conjunction with Numerical Optimization methods to determine the Global Optimum, and to analyze the Algebraic Complexity of the problem.
- Example: \( f(x) = x^4 - 3x^2 + 2 \).
- Its derivative is \( f'(x) = 4x^3 - 6x \).
- Its critical points (points at which derivative function is zero) are \( x = 0 \lor x = \pm \sqrt[3]{\frac{3}{2}} \).
- Its critical values are the corresponding function values: \( f(0) = 2; f(\pm \sqrt[3]{\frac{3}{2}}) = -\frac{1}{4} \).
Example continued

- Now let $\phi = f(x) = x^4 - 3x^2 + 2$ and $f'(x) = 4x^3 - 6x = 0$.
- We can derive a relation between the powers of $\phi$ using $x^3 = \frac{3}{2}x$.
- $\phi = x^4 - 3x^2 + 2 = -\frac{3}{2}x^2 + 2$; $\phi^2 = (-\frac{3}{2}x^2 + 2)^2 = -\frac{21}{8}x^2 + 4$.
- So $\phi^2 - \frac{7}{4}\phi - \frac{1}{2} = (\phi + \frac{1}{4})(\phi - 2) = 0$
- This will be called Critical Value Polynomial (CVP) as it has all its critical values as its zeros.
- Note that the CVP can be calculated without calculating the critical points! The same holds for multivariable polynomials!
Let \( f(x_1, x_2, \ldots, x_n) \) denote a multivariate polynomial.

The number of critical values of \( f \) is finite.

The ideal \( \langle \frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}, \phi - f(x_1, x_2, \ldots, x_n) \rangle \) contains a univariate polynomial in the variable \( \phi \).

The (monic) minimal degree polynomial which is in the ideal and univariate in \( \phi \) will be called the minimal critical value polynomial in the ideal. Its zero set coincides with the set of critical values.

The number of critical values is not more than the product of the degrees of \( \frac{\partial f}{\partial x_i}, i = 1, 2, \ldots, n \). Same for the degree of the minimal critical value polynomial in the ideal!

Note that this even holds in case the number of critical points is infinite!
Example

Let \( f(x_1, x_2) = (x_1x_2 - 1)^2 \).

Then \( \frac{\partial f}{\partial x_i} = 2(x_1x_2 - 1)x_i, \ i = 1, 2. \)

Set of critical points is the union of the origin with the hyperbola \( \{(x_1, x_2) | x_1x_2 = 1\} \)

Let \( \phi = f(x_1, x_2) \). Using \( x_1x_2^2 = x_2, \ x_1^2x_2 = x_1 \), we obtain that \( \phi^2 = x_1^2x_2^2 - 2x_1x_2 + 1 = -x_1x_2 + 1 \) hence \( \phi^2 = \phi. \)

The critical values are \( \phi = 0 \lor \phi = 1. \)

Clearly the number (=2) of critical values is not more (in this case much less) than the product (=9) of the degrees of the partial derivatives \( \frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2} \).
A Further Example

- The minimal critical value polynomial in the ideal can have a zero with multiplicity larger than one.
- An example is given by \( f(x_1, x_2) = x_1^5 - 35x_1^2x_2^2 + x_2^7 \). The minimal critical value polynomial is now of the form \( \phi^{13} + \alpha\phi^2 \) for some nonzero value of \( \alpha \). Hence it has a zero of multiplicity two at \( \phi = 0 \).
- The reduction process is quite straightforward in this case and the analysis can be done by hand.
- Actually \( \alpha = 11^{11}14^{14}10^{10} \) and the only critical values are zero (with multiplicity two) and \(-11.14^{14/11}.10^{10/11}\). The critical value zero corresponds to a local maximum at the origin and the negative critical value corresponds to a local minimum (actually the minimum over the positive orthant).
Consider $f(x) = \frac{p(x)}{q(x)}$, $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, $p, q$ real polynomials.

Want to find the infimum of $f$ over the set $\mathbb{R}^n$.

Can we replace "infimum" by "minimum"?

Example: $f(x_1, x_2) = p(x_1, x_2) = x_1^2 + (x_1x_2 - 1)^2$, $q = 1$.

Clearly the infimum is zero, but this value is NOT attained, it would require $x_1 = 0$, but $f(0, x_2) = 1 \neq 0$.

In general, the infimum can also be minus infinity, in which case again it is not equal to the minimum.

How to deal with this?
Reformulation of the Optimization Problem

(i) For the moment assume that the infimum is finite. Later on we will discuss how to test this assumption.

(ii) Without loss of generality we will assume that $p, q$ do not have a common factor.

(iii) Under these conditions $q$ will have a constant sign (cf. D. Jibetean[1]). Assume this sign is positive (otherwise multiply $p, q$ both by $−1$).

Under the assumptions $(i), (ii), (iii)$ we have

$$\inf(f) = \max\{\gamma \in \mathbb{R} | \forall x \in \mathbb{R}^n : p(x) − \gamma q(x) \geq 0\}.$$
Reformulation of the Optimization Problem

- Homogenization: let \(2d \geq \max(\deg(p), \deg(q))\), \(d \in \mathbb{N}\).
- Let \(P(x_0, x_1, \ldots, x_n) := x_0^{2d}p\left(\frac{x_1}{x_0}, \ldots, \frac{x_n}{x_0}\right)\), \(Q(x_0, x_1, \ldots, x_n) := x_0^{2d}q\left(\frac{x_1}{x_0}, \ldots, \frac{x_n}{x_0}\right)\), then \(P, Q\) are both homogeneous polynomials of degree \(2d\).
- Because \(2d\) is even, \((\forall x \in \mathbb{R}^n) : p(x) - \gamma q(x) \geq 0 \iff (\forall x \in \mathbb{R}^{n+1}) : P(x) - \gamma Q(x) \geq 0\)
- So \(\inf(f) = \max\{\gamma \in \mathbb{R} | \forall x \in \mathbb{R}^{n+1} : P(x) - \gamma Q(x) \geq 0\}\).
- Use \((P - \gamma Q)(tx) = t^{2d}(P - \gamma Q)(x)\) for any real number \(t\).
- \(\implies\) Can restrict to a (compact!) Minkowski ball \(S_{2r} = \{x \in \mathbb{R}^{n+1} | x_0^{2r} + \ldots + x_n^{2r} = 1\}\):
- \(\inf(f) = \max\{\gamma \in \mathbb{R} | \forall x \in S_{2r} : P(x) - \gamma Q(x) \geq 0\}\).
Reformulation of the Optimization Problem

- How can we find the minimum of $P(x) - \gamma Q(x)$ on $S_{2r}$ for given $\gamma$?
- Lagrangian $L = (P - \gamma Q) - \lambda (x_0^{2r} + x_1^{2r} + \ldots + x_n^{2r} - 1)$
- First order conditions: 
  \[
  \frac{\partial (P - \gamma Q)}{\partial x_i} - 2d.\lambda.x_i^{2r-1} = 0, \quad i = 0, 1, \ldots, n \text{ and } \\
  x_0^{2r} + x_1^{2r} + \ldots + x_n^{2r} = 1
  \]
- Using homogeneity of $P - \gamma Q$ and 'Euler' one obtains $P - \gamma Q = 2r\lambda$
- For $r > d$ and for fixed values of $\gamma, \lambda$ the first order conditions (without the Minkowski ball constraint) form a Groebner basis with finitely many common zeros.
- Form the corresponding 'Stetter' matrix corresponding to multiplication by $P - \gamma Q$ and let its determinant be denoted by $F(\gamma, \frac{1}{\lambda})$, $F$ polynomial.
Reformulation of the Optimization Problem

**Theorem**

Let \( m(\gamma) \) denote the coefficient of the highest power of \( \frac{1}{\lambda} \) in \( F(\gamma, \frac{1}{\lambda}) \) when viewed as a polynomial in \( \frac{1}{\lambda} \). Then \( m(\gamma) \) is a polynomial which has the (finite) infimum of \( f \) as one of its zeros.

- The polynomial \( m(\gamma) \) will be called the generalized critical value polynomial and its zeros will be called the generalized critical values of the optimization problem.
- There are similarities with the results and methods of [Hanzon and Jibetean(2003)]
- A similar result holds for the case \( r \leq d \).
From Critical Value Polynomial to the Infimum

- Is smallest real critical value the infimum of the criterion function?
- NO!, not always (despite statements to the contrary in the literature.)

Example: $f(x) = (x^2 + 1)^2$ univariate polynomial. Clearly $\min_{x \in \mathbb{R}} f(x) = 1$. Computation of critical value polynomial: $f'(x) = 0 \iff x^3 = -x$, therefore $\phi = f(x) = x^4 + 2x^2 + 1 = x^2 + 1 \implies \phi^2 = \phi$. Hence the critical values are $\phi = 0 \lor \phi = 1$, $0 < 1$. The critical value zero corresponds to the complex critical points $x = \pm \sqrt{-1}$ and does NOT correspond to a real critical point!
How can we determine which critical value is the infimum?

Answer: Find sufficiently tight upper and lower bounds to the infimum (or show no lower bound exists, in case we allow the function to have an infinite infimum). Because there are only finitely many real critical values, each of which can be determined with any desired precision (e.g. using Sturm chain etc) this will isolate the infimum from the other critical values.
An upper bound is obtained by minimizing the function over a finite grid of points, e.g. using Monte Carlo or Quasi-Monte Carlo techniques. Increasing the sample size should make it more probable that a sufficiently tight upper bound is obtained.

A further refinement of the Quasi-Monte Carlo/low discrepancy sequence technique is possible in case of the generalized critical values and the associated reformulation of the optimization problem (as optimization problem over the compact set $S_{2r}$ for which a Lipschitz constant can be derived). However the decision procedure obtained in this way may be computationally involved. Further investigation along these lines would be of interest.
On the Role of the Critical Value Polynomial in Algebraic Optimization

Lower bound determination

Here we use a classical result of Polya:

**Theorem**

Consider the \( n+1 \)-simplex

\[ \Delta_{n+1} = \{(x_0, x_1, \ldots, x_n) | x_i \geq 0, i = 0, \ldots, n; \sum_{k=1}^{n} x_k = 1\} \].

If the homogeneous polynomial \( R(x_0, x_1, \ldots, x_n) \) is (strictly!) positive on the simplex \( \Delta_{n+1} \) then there exists a positive integer \( N \) such that

\[ R(x_0, \ldots, x_n)(x_0 + x_1 + \ldots + x_n)^N \]

is a homogeneous polynomial with all coefficients non-negative.

- Clearly a polynomial with non-negative coefficients is non-negative on \( \Delta_{n+1} \).
- Note that on the simplex \( \Delta_{n+1} \) one has the equality:

\[ R(x_0, x_1, \ldots, x_n)(x_0 + x_1 + \ldots + x_n)^N = R(x_0, x_1, \ldots, x_n) \]
Note that, for example, \((x_0 - x_1)^2 \geq 0\) but \((x_0 - x_1)^2(x_0 + x_1)^N\) has at least one negative coefficient, for any positive integer \(N\), because the value of the expression at the point \((\frac{1}{2}, \frac{1}{2})\) is zero.

The coefficients of \(R(x_0, x_1, \ldots, x_n)(x_0 + x_1 + \ldots + x_n)^N\) can be computed recursively for \(N = 1, 2, \ldots\) using a 'Pascal-triangle/Pascal-simplex' like procedure (see e.g. [B. Hanzon, Inaugural lecture 2012]).

\[ R(x_0, x_1, \ldots, x_n) \geq 0, \forall (x_0, x_1, \ldots, x_n) \in \Delta_{n+1} \iff R(x_0, x_1, \ldots, x_n) \geq 0, \forall (x_0, x_1, \ldots, x_n) \in \mathbb{R}_{n+1}^+ \]

due to homogeneity of \(R\). How to check non-negativity on the other orthants? Use:

\[ R(x_0, x_1, \ldots, x_n) \geq 0, \forall (x_0, x_1, \ldots, x_n) \in \mathbb{R}_- \times \mathbb{R}_n^+ \iff R(-x_0, x_1, \ldots, x_n) \geq 0, \forall (x_0, x_1, \ldots, x_n) \in \mathbb{R}_{n+1}^+ \]

etc.
Lower bound determination

- Non-negativity of $R$ on $\mathbb{R}^{n+1}$ can be checked by checking for non-negativity on each orthant.
- Actually in practice one often encounters optimization problems in which some or all of the parameters involved are constrained to be non-negative; if so then the Polya method can still be used for lower bound determination, one just checks on the relevant orthants only. (The earlier problem formulation would have to be adapted to incorporate the non-negativity constraints—we will not go into that here)
Results of [Powers and Reznick] among others, in combination with considerations involving a relevant critical value polynomial, can be used to derive an upper bound for the value $N$ for which the coefficients of $R(x_0, \ldots, x_n)(x_0 + x_1 + \ldots + x_n)^N$ need to be checked before concluding that Polya’s test fails and that the polynomial is NOT strictly positive on the simplex.

This can be used to construct a finite decision procedure to determine the infimum, based on the generalized critical value polynomial and the Polya method. However the decision procedure might be lengthy!
Consider the rational function 
\[
\frac{x_4^2 + 2x_2^4 - x_3^4}{x_1^4 + 4(x_2 + x_3)^4}
\]
This has two real critical values, namely \(1, -\frac{1}{2}(19 + 15\sqrt[3]{2} + 12\sqrt[3]{4})\).

Using the Polya method we find: function is bounded from below, hence infimum is NOT equal to minus infinity.

Using a Monte Carlo method one can determine an upper bound that eliminates the critical value at 1 as a candidate for the infimum.

It follows that the infimum is equal to 
\[-\frac{1}{2}(19 + 15\sqrt[3]{2} + 12\sqrt[3]{4}) \approx -28.4738\] (cf Marketa Adamova, Masters minor dissertation, UCC).

Here the rational function is homogeneous of degree zero, from which it follows that the infimum is actually a minimum. Also it follows that the set of corresponding critical points is infinite.
Final remarks

- Up till now we have only been able to run relatively small examples. Further improvements are called for, in relation to:
  (i) usage of low discrepancy sequences to efficiently search for small values of a polynomial function over a compact set
  (ii) computation of the minimal critical value polynomial
  (iii) computational techniques to carry out the Polya method
- We are presently working on various aspects of these three issues.
Final remarks

Possible usage of this would be to gauge and tune existing minimization algorithms for relatively small cases. To test such an algorithm one would like to be able to run a large number of examples in which the global minimum is known. It is envisioned that the methods presented here could be used for that purpose at least in the case where the examples are relatively small.

Note that the method can also reveal algebraic information about the critical values which is useful to analyse the algebraic optimization problem further. More remarks about this can be found in [B.Hanzon, Leuven lecture 2013]
Thank you!\textsuperscript{1}

For references and internet links please contact b.hanzon@ucc.ie or Andrei.Mustata@ucc.ie

\textsuperscript{1}Research support for B.Hanzon by Science Foundation Ireland, grant nrs RFP2007-MATF802 and 07/MI/2008