ABSTRACT

HUTTON, SHARON ELIZABETH. Exact Sums-of-Squares Certificates in Numeric Algebraic Geometry. (Under the direction of Erich L. Kaltofen.)

We consider the problem of finding the nearest polynomial/system with either a fixed or arbitrary root. Our distance measure to the nearest polynomial/system is the weighted Euclidean, one, or infinity coefficient vector norm. Although much work has already been done on this problem, we offer a new proof in the Euclidean norm case, which uses parameterized Lagrangian multipliers. We present formulas for when the root is real or complex, and when the function has real or complex coefficients. Our formulas also allow fixing selected coefficients of f to their input values and only deforming the other coefficients in $\tilde{f}$, thus preserving sparsity or monicity, for instance. We present an algorithm for computing the nearest polynomial with given linear equality and inequality coefficient constraints. Linear inequality constraints on the coefficients of $\tilde{f}$, for instance non-negativity ($c_i \geq 0$), can now be imposed via Karush-Kuhn-Tucker (KKT) conditions and the arising systems solved via linear programming, at least for a fixed real root. We further extend our algorithms to systems.

Furthermore, we consider the weighted infinity norm and one norm as the distance measure. We give explicit solutions for finding the nearest polynomial with a given root. The resulting functions are optimizable over the root in the unconstrained case. We also consider finding the nearest polynomial with linear inequality constraints on the coefficients. For a given root, this results in solving a linear program, due to Tchebycheff, and for an arbitrary root, this results in conducting a grid search.

In addition, we explore using sums-of-squares certificates to certify a lower bound for the distance to the nearest polynomial with a real root. Some polynomials that cannot be written as a sum-of-squares, such as a modified Motzkin polynomial, have a positive distance to the nearest polynomial with a real root and a sum-of-squares certificate for a positive lower bound on that distance. These sums-of-squares certificates offer an alternative proof that a polynomial has no real root and a deformation analysis for Seidenberg's problem.

Our last result is on a somewhat separate area of research than the rest of our results, approximate GCD. We generalize the univariate resultant to several polynomials.

Exact Sums-of-Squares Certificates in Numeric Algebraic Geometry

by
Sharon Elizabeth Hutton

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Applied Mathematics

Raleigh, North Carolina

2011

APPROVED BY:

Hoon Hong	Agnes Szanto
Pierre Gremaud	Lihong Zhi

Erich L. Kaltofen
Chair of Advisory Committee

DEDICATION

To my parents, Shane and Sandra Kay Hutton for their love and support throughout my life.

BIOGRAPHY

Sharon Elizabeth Hutton was born in Kettering, Ohio to Shane and Sandra Kay Hutton. By the age of four, her mom knew that she would always love mathematics. They say a mother's intuition is never wrong and this time the saying held true. Sharon thrived on solving her first mathematical proofs and later graduated Summa Cum Laude from Baylor University in Mathematics with a minor in Computer Science. Her love for mathematics did not stop there, though, she went on to study Applied Mathematics at North Carolina State University. Her dissertation is in the field of computer algebra under the direction of Dr. Erich L. Kaltofen.

ACKNOWLEDGEMENTS

This accomplishment and the entire experience would not have been possible without the support, prayers, and influence of so many in my life. I want to thank all of my friends and family members for the part that each of you played in helping me to achieve my goals. I could not have done this without you. I would like to acknowledge a few of the most influential people:

- My advisor, Erich L. Kaltofen, for teaching me how to conduct meaningful research and for the many international educational opportunities you provided.
- My Chinese co-author, Lihong Zhi, for her research guidance in Beijing for a summer.
- Lihong Zhi's students, especially Zhenfeng Yang, Xiaoli Wu, Yue Ma, and Feng Guo for your friendship and help during my time in China.
- My brother, Shaun Hutton, for his prayers and encouragement. Thank you for all of your help along the way and your encouragement to keep working even when things were difficult.
- My dad, Shane Hutton, for your love and help. I could not have done this without your never ceasing motivation and encouragement.
- My mom, Sandra Kay Hutton, for your love and help. Thank you for helping me keep my sanity through your many hours of prayer.
- My math professors at Baylor University, especially Dr. Henderson, Dr. Cannon, Dr. Arnold, Dr. Raines, Dr. Sepanski, and Dr. Ryden for encouraging me in my pursuit of mathematics. Thank you for making me aware of various opportunities, further instilling a love of mathematics in me, and for investing in me.
- My committee members, Hoon Hong, Pierre Gremaud, and Agnes Szanto for your help and suggestions.

TABLE OF CONTENTS

List of Tables . . . . **vii**

List of Figures . . . . **viii**

Chapter 1 Overview and Problem Statements **1**
1.1 Introduction 1
1.1.1 Hybrid Symbolic-Numeric Computing 1
1.2 Problem Statements 2
1.2.1 Problem Formulation 1: Coefficient Perturbation 2
1.2.2 Problem Formulation 2: Nearest Polynomial with A Real Root . . 4
1.2.3 Problem Formulation 3: Global Minimum 4
1.3 Sums-Of-Squares Approach 5
1.4 Semidefinite Programming 7
1.5 Related Previous Results 8
1.6 Our Contributions - Results 11

Chapter 2 Nearest Polynomial and Coefficient Constraints **13**
2.1 History 13
2.2 Stetter's Results 14
2.3 Nearest Polynomial with a Real Root in Weighted Euclidean Norm . . . 15
2.4 Deformation Analysis 19
2.5 Nearest Polynomial with Complex Coefficients and/or Roots in Weighted Euclidean Norm 22
2.6 Coefficient Constraints 24
2.6.1 Linear Constraints 24
2.6.2 Inequality Constraints 27

Chapter 3 The Case of Several Polynomials **31**
3.1 System of Polynomials with a Real Root in Weighted Euclidean Norm . . 31

Chapter 4 Infinity Norm and One Norm **34**
4.1 Nearest Polynomial with a Real Root 34
4.1.1 Coefficient Constraints with a Given Root 37
4.1.2 Nearest Polynomial with an Arbitrary Root 40
4.1.3 Nearest Polynomial without Constraints and an Arbitrary Root . 41
4.1.4 Coefficient Constraints with an Arbitrary Root 45
4.2 Systems 47

Chapter 5 Exact Sums-of-Squares Distance Certificates **49**
5.1 History 49
5.2 Motzkin 50
5.3 Lower Bound Certificates 52
5.4 Examples 54
5.5 Future Work 57

Chapter 6 Approximate GCD **59**
6.1 Background 59
6.2 Univariate multi-polynomial Sylvester matrices 59

Chapter 7 Conclusion **62**

References **63**

LIST OF TABLES

Table 4.1 Norm Comparison Table 44
Table 4.2 Norm Comparison Table With Corresponding Roots 44

LIST OF FIGURES

Figure 1.1 Root sensitivity. This figure is from [13] 3
Figure 1.2 Polynomial with relative minimum that is not the global minimum 5
Figure 5.1 Rationalizing a Sum-of-Squares 53

Chapter 1

Overview and Problem Statements

1.1 Introduction

1.1.1 Hybrid Symbolic-Numeric Computing

Hybrid symbolic-numeric computations should involve both a symbolic computation component and a numerical computation component. Numerical computation is the study of approximate techniques for numerically solving mathematical problems. Some examples are solving linear programs, numerical PDE's, or optimization problems. The computations often involve floating point numbers of fixed size. In numerical computing, sometimes complex operations are solved using table lookup. Symbolic computing is different from numerical computing in two ways, namely that the results are symbolic and exact, instead of numerical and approximate [16]. In symbolic computation, the computations are exact in the sense that they guarantee a certain precision or the algorithms guarantee a real root lies within a given interval. Furthermore, the numerical arithmetic performed cannot have digit overflow, since the computation is exact. Algorithms within symbolic computation can also have symbolic variables as input variables versus only numerical data. Hybrid symbolic-numeric computation includes topics such as [5]:

- Polynomial problems which are converted into eigenvalue problems
- The use of numerical coefficients in polynomial arithmetic
- Problems where the pre-computations utilize symbolic computation and later computations use efficient and/or stable numerical algorithms

- Problems where exact computations are combined with faster intermediate floating-point arithmetic computations

Within hybrid symbolic-numeric computation, our focus is on exact computations that utilize prior numerical approximation techniques. Our algorithm first utilizes the speed of numerical sums-of-squares algorithms within Matlab's `solvesos` package and then utilizes the exactness of symbolic sums-of-squares algorithms. This hybrid approach is faster than performing solely symbolic computations. It also achieves an exact solution instead of a numerical approximate solution. We consider the problem of computing the nearest polynomial, using some distance norm, which satisfies a property which the original input polynomial does not. More specifically, given a polynomial that is positive definite, we consider finding the nearest polynomial with a real root, with the distance measure being the distance in coefficient vector Euclidean, one, and infinity norms. We further extend this to finding the nearest consistent system to an inconsistent one. Using semidefinite programming and exact sums-of-squares certificates, we certify a lower bound on the distance to the nearest polynomial with a real root.

1.2 Problem Statements

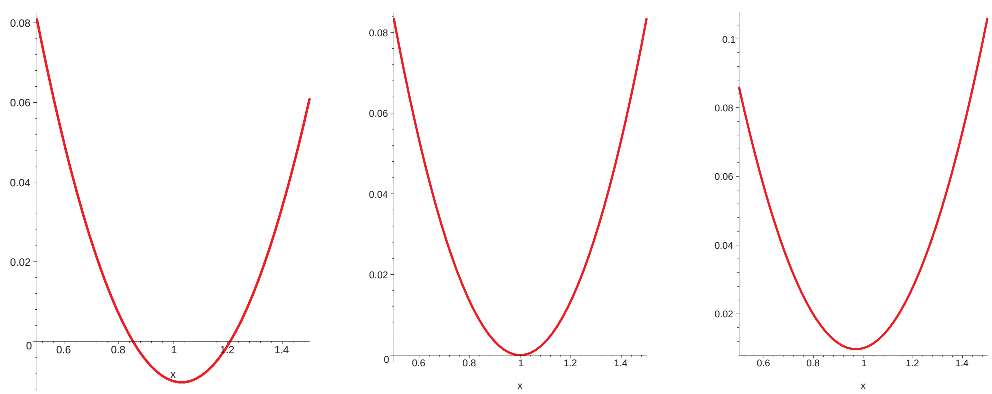
Our problem of finding the nearest polynomial with a real root to a given positive semidefinite polynomial (see Section 1.2.2, Problem Formulation 2) can be formulated in a number of equivalent problem statements. We discuss three different problem formulations below.

1.2.1 Problem Formulation 1: Coefficient Perturbation

Imprecision of empirical data can create ill-conditioned polynomial inequalities. For example, for a given near singular matrix A and solution vector x, a slight deformation in the matrix A can radically change the solution vector. This principle is also true for inequalities: a tiny deformation in the coefficients of the polynomial f can cause the inequality $f \geq 0$ to become invalid. For example, consider the inequality

$$0.33\,x^2 - 0.66\,x + 0.33 \geq 0.$$

Is this inequality true for $\forall x \in \mathbb{R}$? In Figure 1.1, we see that this is an ill-conditioned inequality. The middle polynomial $f(x) = \frac{1}{3}(x-1)^2$ has a double real root at $x = 1$. In fact, it is the nearest polynomial with a real root to the polynomial $x^2 + 1$ under the infinity norm [11]: $\|f(x) - (x^2+1)\|_\infty = \frac{2}{3}$, where for a polynomial g the norm $\|g\|_\infty$ is the maximum of the absolute values of the coefficients of g (i.e. the height of g). Small perturbations in the leading coefficient (one could also perturb the constant coefficient) make the polynomial f either indefinite (left polynomial in Figure 1.1, the polynomial changes sign) or positive definite (right polynomial). Therefore, the right polynomial, although positive definite, as an approximate polynomial is not numerically positive because a small change in its coefficients can make the polynomial indefinite. So, computations of approximate data can give rise to problems of ill-conditionness.



$\left(\frac{1}{3} - \frac{1}{100}\right)x^2 - \frac{2}{3}x + \frac{1}{3}$ $\quad$ $\frac{1}{3}x^2 - \frac{2}{3}x + \frac{1}{3}$ $\quad$ $\left(\frac{1}{3} + \frac{1}{100}\right)x^2 - \frac{2}{3}x + \frac{1}{3}$

Figure 1.1: Root sensitivity. This figure is from [13]

Therefore, one possible problem formulation is for the infimum $h \in \mathbb{R}$ of a polynomial $f \in \mathbb{R}[x_1, \ldots, x_n]$, we want to find the largest perturbation allowed in the coefficients of f such that $f(\xi_1, \ldots, \xi_n) - h \geq 0$, for all $x_i = \xi_i \in \mathbb{R}$. In other words, given a positive semidefinite polynomial f (denoted by $f \succeq 0$) with imprecise coefficients, then what is the most the coefficients can be altered in order for the polynomial to remain positive semidefinite?

1.2.2 Problem Formulation 2: Nearest Polynomial with A Real Root

This problem is equivalent to finding the nearest polynomial with a real root. As stated in [18] deciding if a polynomial is not positive semidefinite is equivalent to deciding if $f(x_1, \ldots, x_n)x_{n+1}^2 + 1$ has a real root. Seidenberg's Problem is: Determine whether a multivariate polynomial has a real root. Therefore, all polynomial inequalities can be reduced to solving Seidenberg's problem. Given a polynomial f, what is the nearest polynomial with a real root? The solution to an approximate polynomial (i.e., the nearest polynomial with a real root) may be as useful as the exact solution and may be more efficient. This problem can be extended to systems and reformulated as finding the nearest system with a non-trivial GCD. We want to find the nearest polynomials $\tilde{f}$ and $\tilde{g}$, where $\tilde{f} = f(x) + \Delta f(x)$ and $\tilde{g} = g(x) + \Delta g(x)$ to the given polynomials f and g that have a common root or non-trivial GCD.

1.2.3 Problem Formulation 3: Global Minimum

As stated in [18] any polynomial inequality $f \geq h$, where h is the infimum of all values of f, is equivalent to $f - h$ being positive semidefinite. Thus, if we can write $f - h$ as a sum-of-squares then we have proven the inequality. We use hybrid symbolic-numeric computing to answer this problem. First, we use numerical optimization algorithms for semidefinite programming. Then we convert the imprecise sum-of-squares to an exact identity over the rational numbers [[29],[19],[20]]. Our method utilizes the speed of numerical algorithms within Matlab's `solvesos` to obtain an approximate sum-of-squares, while utilizing the exactness of symbolic computation by converting this imprecise sum-of-squares to an exact sum-of-squares certificate with exact rational scalars and polynomials. Thereby we solve a truly hybrid symbolic-numeric problem.

If you have an equation resulting from imprecise measurement, how do you prove that h is the global minimum? Many numerical algorithms will converge to a local minimum, but not necessarily the global minimum. Certifying that $f - h$ is a sum-of-squares guarantees that we have obtained a lower bound for the global minimum. Consider the polynomial $f = 1/4x^4 - 5/3x^3 + x^2 + 8x + 6$, which has a local minimum at $(4, 34/3)$ and a global minimum at $(-1, 11/12)$ (see Figure 1.2). If we are able to certify a local minimum, i.e. $f - 11/12 = \text{SOS}$, then we know that it is the global minimum.

1.3 Sums-Of-Squares Approach

Consider the polynomial

$$f = 1/4x^4 - 5/3x^3 + x^2 + 8x + 6$$

and its derivative

$$f' = (x-2)(x-4)(x+1).$$

Using methods from Calculus, we know that there are relative minima at $x = -1$ and $x = 4$. See Figure 1.2.

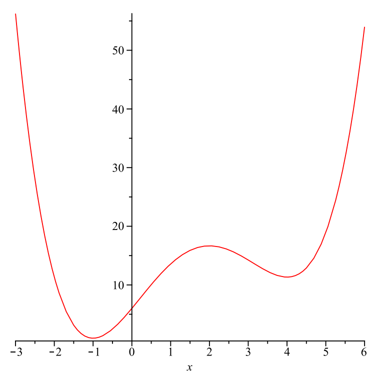


Figure 1.2: Polynomial with relative minimum that is not the global minimum

If we are able to write f as a sum-of-squares then we know that f is positive semi-definite. In order to find a sum-of-squares certificate for f, first we write f as

$$f = [1, x, x^2] \begin{bmatrix} q_{11} & q_{12} & q_{13} \\ q_{12} & q_{22} & q_{23} \\ q_{13} & q_{23} & q_{33} \end{bmatrix} [1, x, x^2]^T.$$

When we multiply-out the right-hand side of the equation and combine coefficients, we obtain the equation

$$1/4x^4 - 5/3x^3 + x^2 + 8x + 6 = q_{11} + (2q_{12})x + (2q_{13} + q_{22})x^2 + (2q_{23})x^3 + q_{33}x^4. \quad (1.1)$$

Next, we equate the coefficients in Equation (1.1) and solve for the unknown $q_{i,j}$ values. We then substitute these values into our matrix, which we call W.

$$W = \begin{bmatrix} 6 & 4 & -1 \\ 4 & 3 & -5/6 \\ -1 & -5/6 & 1/4 \end{bmatrix}$$

After we have our matrix W, we perform a Cholesky factorization on the matrix W. Note that by construction W is symmetric. A symmetric matrix W is positive semidefinite if $x^T W x \geq 0$, for all $x \neq 0$, or equivalently, if all of its eigenvalues are non-negative. The values for W were determined in such a way that W is symmetric positive semidefinite. Recall that a symmetric positive definite matrix W can be factored as $W = R^T R$, where R is the Cholesky factor of W, an unique upper triangular matrix with positive entries on the diagonal.

The Cholesky decomposition of W is:

$$W = \begin{bmatrix} \sqrt{6} & 0 & 0 \\ 2/3\sqrt{6} & 1/3\sqrt{3} & 0 \\ -1/6\sqrt{6} & -1/6\sqrt{3} & 0 \end{bmatrix} \begin{bmatrix} \sqrt{6} & 2/3\sqrt{6} & -1/6\sqrt{6} \\ 0 & 1/3\sqrt{3} & -1/6\sqrt{3} \\ 0 & 0 & 0 \end{bmatrix} \tag{1.2}$$

So we have that:

$$\begin{aligned} f &= [1, x, x^2] W [1, x, x^2]^T \\ &= (\sqrt{6} + (2/3)\sqrt{6}x - (1/6)\sqrt{6}x^2)^2 + ((1/3)\sqrt{3}x - (1/6)\sqrt{3}x^2)^2 \\ &= 6\left(1 + 2/3x - 1/6x^2\right)^2 + 3\left(1/3x - 1/6x^2\right)^2. \end{aligned}$$

Remark 1 As shown in the previous example, we identify 1×1 matrices by their entries.

This problem becomes challenging when you increase the number of variables in f from 1 to n variables. Bezout's inequality states that when you increase the number of variables in a polynomial, f, then the increase in the number of critical values of f is exponential in the number of variables of f. Consider the polynomial

$$f(x_1, \ldots, x_n) = \sum_i f_i(x_i) = x_1^d - x_1 + x_2^d - x_2 + \ldots + x_n^d - x_n.$$

Then

$$\frac{\partial f}{\partial x_i} = 0 \Leftrightarrow dx_i^{d-1} - 1 = 0 \Leftrightarrow x_i^{d-1} - \frac{1}{d} = 0$$

There are $d-1$ solutions for each x_i where $i = 1 \dots n$. Thus, we have $(d-1)^n$ critical values of f of the form

$$(\alpha_1, \dots, \alpha_n) = \left(\dots, \exp\left(\frac{2\pi j_i}{d-1}\right) \sqrt[d-1]{\frac{1}{d}}, \dots \right), \quad \text{where } 1 \leq j_i \leq d-1$$

Note that there are $\prod_i (d_i - 1)$ real points if all f_i have only real roots. Therefore, an increase in the number of variables is an exponential increase in the number of critical values. We thank Mohab Safey El Din for this example. We further note that the computation is even more difficult because the matrix W is singular at the minimizer.

1.4 Semidefinite Programming

We use a fixed-precision semidefinite programming solver in Matlab to obtain a numerical positive semidefinite matrix W as described in section 1.3. Semidefinite programming can be solved using interior-point methods of linear programming. Theoretically, a semidefinite program (SDP) can be solved in polynomial time [19]. In semidefinite programming we minimize a linear objective function over the intersection of the cone of positive semidefinite matrices with an affine space, which is a convex optimization problem.

Semidefinite programming solves global optimization problems of the form [19]

$$\left.\begin{aligned} r^* := &\min_{x \in \mathbb{R}^n} \; p(x) \\ &\text{s. t. } \; q_1(x) \geq 0, \dots, q_l(x) \geq 0 \\ &\qquad\quad \text{where } p, q_1, q_l \in \mathbb{R}[\mathbf{X}_1, \dots, \mathbf{X}_n] \end{aligned}\right\} \tag{1.3}$$

where the q_j are the constraints.

We compute a lower bound $\tilde{r} \leq r$ of our computed function $\frac{f(\mathbf{X})^2}{g(\mathbf{X})} \geq r$, where $\frac{f(\mathbf{X})^2}{g(\mathbf{X})}$ is the formula for the distance to the nearest polynomial to f with a real root from Theorem 2. We can reformulate the inequality $\frac{f(\mathbf{X})^2}{g(\mathbf{X})} \geq r$ as $f(\mathbf{X})^2 - rg(\mathbf{X}) \geq 0$. If we can write $f(\mathbf{X})^2 - rg(\mathbf{X})$ as a sum-of-squares (i.e. $f(\mathbf{X})^2 - \tilde{r}g(\mathbf{X}) = m_d(\mathbf{X})^T W \, m_d(\mathbf{X})$), then we have shown that $f(\mathbf{X})^2 - rg(\mathbf{X}) \geq 0$. We compute the matrix W by solving the sums-of-squares program [15, 28, 19, 20]:

$$\left.\begin{array}{rl} r^* := \sup\limits_{r\in\mathbb{R},W} & r \\ \text{s. t.} & f(\mathbf{X})^2 - \tilde{r}g(\mathbf{X}) = m_d(\mathbf{X})^T\, W\, m_d(\mathbf{X}) \\ & W \succeq 0,\ W^T = W \end{array}\right\} \tag{1.4}$$

where $m_d(\mathbf{X})$ is the column vector of all terms in $\mathbf{X}_1, \ldots, \mathbf{X}_n$ up to degree d. The dimension of $m_d(\mathbf{X})$ is $\binom{n+d}{d}$. Using the SDP solver in Matlab, we can only obtain a numerical positive semidefinite matrix W and floating point number r^* which satisfy approximately

$$f(\mathbf{X})^2 - r^* g(\mathbf{X}) \approx m_d(\mathbf{X})^T \cdot W \cdot m_d(\mathbf{X}) \succeq 0,\ W \succapprox 0. \tag{1.5}$$

See section 5.3 for more details.

1.5 Related Previous Results

Real polynomial or rational function global optimization is equivalent to establishing a polynomial inequality: the infimum $h \in \mathbb{R}$ of a polynomial $f \in \mathbb{R}[x_1, \ldots, x_n]$ satisfies $f(\xi_1, \ldots, \xi_n) - h \geq 0$ for all $x_i = \xi_i \in \mathbb{R}$. In other words, the polynomial $f - h$ is positive semidefinite. For univariate f (i.e., $n = 1$) Sturm sequences [14] yield an efficient algorithm for deciding semidefiniteness. The bivariate case $n = 2$ can be solved by Seidenberg's [37] algorithm (see also [14] and [17]), which is generalized to arbitrarily many variables via Lagrangian multipliers in [1, 36] or used in nonstandard decision methods [41]. Alternatively, one can use Artin's theorem of sum-of-squares and semidefinite programming (see, e.g., [19, 20]).

Here we consider the more difficult situation when the coefficients of f are not exactly known, which is the case when f is the result of an empirical measurement or a computation with floating point numbers. As a simple example consider Figure 1.1.

As in Kharitonov's [25] interval polynomial stability criterion, we seek to compute by how much the coefficients in a polynomial can be deformed while still preserving nonnegativity. This distance is the coefficient vector norm distance to the nearest polynomial with a real root, which we shall call the *radius of positive semidefiniteness*. Note that there may not exist an affine optimizer—hence radius of positive semidefiniteness rather than distance to the nearest polynomial with a real root.

We follow the approach by Karmarkar and Lakshman [24] (see also [4]) which first fixes a real root $(\alpha_1, \ldots, \alpha_n) \in \mathbb{R}^n$ and gives a rational function $\mathcal{N}(\alpha_1, \ldots, \alpha_n)$ in the indeterminate α's for the minimal distance from the given f to the nearest polynomial $\tilde{f}$ with $\tilde{f}(\alpha_1, \ldots, \alpha_n) = 0$. One then can compute the infimum of $\mathcal{N}(\alpha_1, \ldots, \alpha_n)$ over all real α's. The case $n \geq 2$ is from [38].

We rederive the multivariate formula for $\mathcal{N}(\alpha_1, \ldots, \alpha_n)$ in [38], for weighted ℓ^2 distance norms, by the method of Lagrangian multipliers. The weighted norms subsume the fixing of coefficients in [34] (see [5, Section 2.12.3.2.6] and Remark 7 below). Our approach also allows us to introduce linear constraints on the coefficients of $\tilde{f}$, as is done in [22] for the approximate GCD problem. Linear equality constraints on the coefficients of $\tilde{f}$ generalize sparsity, which are equations of the form $c_i = 0$. Because the Jacobian of the Lagrange function remains linear in the control variables and multipliers, determinantal formulas parametric in the real root coordinates can be computed. Linear inequality constraints on the coefficients of $\tilde{f}$, for instance non-negativity ($c_i \geq 0$), can now be imposed via Karush-Kuhn-Tucker (KKT) conditions (see, e.g., [6]) and the arising systems solved via linear programming, at least for a fixed real root. Parametric root coordinates or non-linear constraints necessitate non-linear techniques on the Lagrange and KKT conditions and are therefore in general of much higher computational complexity. Our approach allows multiple simultaneous f's and complex coefficients without modification.

Seidenberg's algorithm (and Safey El Din's generalization) computes to a given real point in $\mathbb{R}^n$ the nearest real point on f in terms of Euclidean distance. If f has no real solution the tangent equations have no real solutions. Our algorithm computes the nearest surface (in terms of coefficient norm) that has a real point. If f has a real point, the nearest surface is f itself. However, if a lower bound on the radius of semidefiniteness for any weight vector is greater than zero, f has no real point, even when the coefficients of f are approximate. The latter can be certified by a sum-of-squares of rational functions, which leads to an entirely new verification that f is definite, i.e., has no real point, with possibly a very short certificate.

Polynomials with a radius of positive semidefiniteness greater than zero are quite special. Our Example 27 below demonstrates that a positive polynomial that is not a sum-of-squares of polynomials can have a lower bound certificate for the radius of positive semidefiniteness that is in fact a sum-of-squares of polynomials, which implies positive semidefiniteness of the polynomial itself. For such polynomials, sum-of-squares denominators in Artin-style certificates may never become necessary (see our conjecture at the end of Chapter 5).

The computation of the nearest polynomial with a real root can be interpreted as a dual of Seidenberg's method that decides if a real hypersurface contains a real point. Sums-of-squares rational lower bound certificates for the radius of semidefiniteness provide a new approach to solving Seidenberg's problem, especially when the coefficients are numeric. They also offer a surprising alternative sum-of-squares proof for those polynomials that themselves cannot be represented by a polynomial sum-of-squares but that have a positive distance to the nearest indefinite polynomial.

Our method is conceptually that of hybrid symbolic-numeric computation, such as computing approximate polynomial greatest common divisors and factorization.

Hitz and Kaltofen [10] derive Lakshman's and Karmarkar's formula for univariate f by a least square fit for the cofactor $f(x)/(x-\alpha)$ and introduce linear equality constraints on the deformed coefficients. Zhi, Wu, Noda, Kai, Rezvan and Corless [43, 42, 33] generalize the formula to roots with given multiplicities. In [11] ℓ^∞-norm distances are introduced and Markus Hitz in the Summer of 1999 considered dual ℓ^p-norms. Stetter [38] then generalized Lakshman and Karmarkar's formula to an arbitrary number of variables and dual ℓ^p-norm distances via Hölders inequality. We generalize the infinity norm results for the multivariate case (see section 4). Results from [11] are generalized.

In [34, 30] Stetter's multivariate (complex) formula is applied to the important problem of computing the nearest consistent polynomial system, with zeros of a minimum given multiplicity, and a different proof via generalized Lagrangian interpolation is given. We observe that the ℓ^∞-norm formulas apply to the problem of consistent systems as well (see Theorem 4 below). In our setting, we determine the smallest deformation where all inequalities are simultaneously violated.

A related result [11] computes the nearest matrix in Frobenius norm that has a real eigenvalue. Sum-of-squares rational lower bound certificates were introduced in [19] to overcome the high algebraic degree in the exact real algebraic minima.

This section was adapted from [13].

1.6 Our Contributions - Results

We present five new results. The first four results are within the area of hybrid symbolic-numeric computation. The last result is in the somewhat separate field of approximate GCD.

Result 1:
Stetter [38] proved an inequality for the distance to the nearest polynomial with a real or complex root. Szanto and Pope [30] generalized the result for systems with roots with higher multiplicities. In Chapter 2, we present a new proof of [38] and [34] based on Lagrangian multipliers for finding the nearest polynomial in weighted Euclidean coefficient vector norm with a constrained root (see Theorem 2). We present formulas for when the root is real or complex, and when the function has real or complex coefficients (see Theorem 2, 3). Our formulas also allow the keeping of selected coefficients of f as their input values and only deforming the others in $\tilde{f}$, thus preserving sparsity or monicity, for instance. Since our formulas allow weighted norms, we can just set the weights of those coefficients to infinity in the limit (see Remark 7). Sometimes there may not exist an affine optimizer. We explore some examples where the radius of positive semidefiniteness is 0 or when the infimum is not attainable (see Example 4).

Result 2:
We present an algorithm for computing the nearest polynomial with given linear equality and inequality coefficient constraints. Because the Jacobian of the Lagrange function remains linear in the control variables and multipliers, determinantal formulas parametric in the real root coordinates can be computed (see Equation (2.11)). Linear inequality constraints on the coefficients of $\tilde{f}$, for instance non-negativity ($c_i \geq 0$), can now be imposed via Karush-Kuhn-Tucker (KKT) conditions (see, e.g., [6]) and the arising systems solved via linear programming, at least for a fixed real root (see Equation (2.15), Equation (2.18)). In Chapter 3 we extend our algorithms to systems (see Equation (3.1)). We are able to compute the nearest consistent system to an inconsistent one with additional equality and inequality coefficient constraints.

Result 3:
We expand upon Stetter's result [38] for finding the nearest polynomial with the weighted infinity norm and one norm as the distance measure. In Chapter 4, we give explicit

solutions for a given root (see Theorem 8 and Theorem 9). The functions are optimizable over the root in the unconstrained case (see Example 21). We also consider finding the nearest polynomial with linear equality constraints on the coefficients. For a given root, this results in solving a linear program due to Tchebycheff [11] (see Linear Program 4.3 and 4.4). If we want to find the nearest polynomial with coefficient constraints on $\tilde{f}$ with an arbitrary root, we conduct a grid search (see Example 24).

Result 4:
In Chapter 5, we use sums-of-squares certificates, as were introduced in [19], to certify a lower bound for the radius of positive semidefiniteness. Given a polynomial f, then in Theorem 2, we give a formula for the distance to the nearest polynomial to f with root α, denoted here by $\mathcal{N}(\alpha) = \dfrac{f(\alpha)^2}{g(\alpha)}$, a rational function in α. If the radius of positive semidefiniteness, denoted by $\inf\limits_{\alpha \in \mathbb{R}^n} \mathcal{N}(\alpha)$, is greater than or equal to zero, call it r, then minimizing the rational function $\mathcal{N}(\alpha)$ is equivalent to maximizing r such that $f(\alpha)^2 - rg(\alpha) \geq 0$. We compute a lower bound of the radius of positive semidefiniteness, $\inf\limits_{\alpha \in \mathbb{R}^n} \mathcal{N}(\alpha)$, by solving a a sums-of-squares program to write $f(\mathbf{X})^2 - rg(\mathbf{X}) = \text{SOS}$ (see SOS Program 5.1). We thereby prove $\dfrac{f(\mathbf{X})^2}{g(\mathbf{X})} \succeq r$. As a result, we are now able to show polynomials, which themselves are not a sum-of-squares, are positive semidefinite via a sums-of-squares certificate (see Example 27).

Result 5:
Our last result is on a somewhat separate area of research than the rest of our results. We wanted to include a result in the research area of approximate GCD. In Chapter 6, we generalize the univariate resultant to several polynomials (see Theorem 14).

Chapter 2

Nearest Polynomial and Coefficient Constraints

2.1 History

We give a stability criterion for real polynomial inequalities with floating point or inexact scalars by estimating from below or computing the radius of positive semidefiniteness. That radius is the maximum deformation of the polynomial coefficient vector measured in a weighted Euclidean vector norm within which the inequality remains true. A large radius means that the inequalities may be considered numerically valid.

The radius of positive (or negative) semidefiniteness is the distance to the nearest polynomial with a real root, which has been thoroughly studied before.

Hitz and Kaltofen [10] derive Lakshman's and Karmarkar's formula for univariate f by a least square fit for the cofactor $f(x)/(x-\alpha)$ and introduce linear equality constraints on the deformed coefficients. Zhi, Wu, Noda, Kai, Rezvan and Corless [43, 42, 33] generalize the formula to roots with given multiplicities. In [11] ℓ^∞-norm distances are introduced and Markus Hitz in the Summer of 1999 considered dual ℓ^p-norms. Stetter [38] then generalized Lakshman and Karmarkar's formula to an arbitrary number of variables and dual ℓ^p-norm distances via Hölders inequality.

In [34, 30] Stetter's multivariate (complex) formula is applied to the important problem of computing the nearest consistent polynomial system, with zeros of a minimum given multiplicity, and a different proof via generalized Lagrangian interpolation is given.

We solve this problem for the Euclidean norm by parameterized Lagrangian multipliers and Karush-Kuhn-Tucker conditions. Our algorithms can compute the radius for several simultaneous inequalities including possibly additional linear coefficient constraints.

2.2 Stetter's Results

Stetter proved an inequality for the distance to the nearest polynomial with a real root. We discuss in more detail the results presented in [38, 5].

Definition 1 *We consider $\mathbb{C}^n$ equipped with some norm $\|\ldots\|$. The associated dual norm or operator norm $\|\ldots\|^*$ for the column vector $v \in \mathbb{C}^n$ is defined by*

$$\|v^T\|^* = \sup_{u \neq 0} \frac{|v^T u|}{\|u\|} = \sup_{\|u\|=1} |v^T u|.$$

Since we are taking the supremum over a compact domain, the maximum value is attained.

Proposition 1 (Proposition 1 in [38]) *For each $u \in \mathbb{C}^n$, with $\|u\| = 1$, there exist vectors $v \in \mathbb{C}^n$, with $\|v^T\|^* = 1$, such that $|v^T u| = 1$.* $\square$

It is well known that with $\frac{1}{p} + \frac{1}{q} = 1$, $1 \leq p, q \leq \infty$,

$$\|\ldots\| = \ell^p\text{-norm} \Leftrightarrow \|\ldots\|^* = \ell^q\text{-norm}.$$

Theorem 1 (see [38]) *Let the vector of possible term values of f and $\tilde{f}$ be given by $\vec{f}$, $\vec{\tilde{f}} \in \mathbb{C}^n$ respectively. Let the given root be denoted by $\alpha = [\alpha_1, \ldots, \alpha_n] \in \mathbb{C}^n$. Let $\tau = [1, \alpha_1, \ldots, \alpha_n, \ldots, \alpha_1^{i_1} \cdots \alpha_n^{i_n}, \ldots]$ the term vector evaluated at the root. Let $\|\ldots\|$ be the given norm and $\|\ldots\|^*$ the associated dual norm. The nearest polynomial with a real root, i.e. $\tilde{f}(\alpha) = \tau^T \vec{\tilde{f}} = 0$ requires*

$$\|\vec{f} - \vec{\tilde{f}}\|^* \geq \frac{|f(\alpha)|}{\|\tau\|}.$$

2.3 Nearest Polynomial with a Real Root in Weighted Euclidean Norm

We present a new proof (see Theorem 2) using Lagrangian multipliers for finding the nearest polynomial with a real/complex root. This new approach allows us to consider linear inequality and equality constraints on the coefficients of the polynomial.

Definition 2 *Let $w \in \mathbb{R}^n_{>0}$ be a vector of positive weights. For $x = [x_1, \ldots, x_n]^T \in \mathbb{R}^n$ the weighted ℓ^2-norm is*

$$\|x\|_{2,w} = \sqrt{w_1 x_1^2 + \ldots + w_n x_n^2}.$$

Definition 3 *Let $\alpha = [\alpha_1, \ldots, \alpha_n] \in \mathbb{R}^n$ be a prescribed real root and $w \in \mathbb{R}^n_{>0}$ a weight vector. The distance to the nearest polynomial with a real root α is defined as*

$$\left.\begin{aligned} \mathcal{N}^{[f]}_{2,w}(\alpha) = \inf_{\tilde{f} \in \mathbb{R}[x_1,\ldots,x_n]} \quad & \|f - \tilde{f}\|^2_{2,w} \\ \text{s. t.} \quad & \tilde{f}(\alpha) = 0, \\ & \deg(\tilde{f}) \leq \deg(f). \end{aligned}\right\} \tag{2.1}$$

If f and the used norm are clear from the context, we may write $\mathcal{N}(\alpha)$ for the above infimum, which is actually a minimum (see Theorem 2 below).

Theorem 2 *Let $f \in \mathbb{R}[x_1, \ldots, x_n]$,*

$$f(x_1, \ldots, x_n) = \sum_{i_1+\cdots+i_n=0}^{d} f_{i_1,\ldots,i_n} x_1^{i_1} \cdots x_n^{i_n}.$$

For $\tau = [1, \alpha_1, \ldots, \alpha_n, \ldots, \alpha_1^{i_1} \cdots \alpha_n^{i_n}, \ldots]^T$, the vector of possible term values in $\tilde{f}$, the distance to the nearest polynomial with a real root α is

$$\mathcal{N}^{[f]}_{2,w}(\alpha) = \frac{f(\alpha)^2}{\tau^T D_w^{-1} \tau}. \tag{2.2}$$

Furthermore, the coefficient vector $\vec{\tilde{f}}$, for the polynomial $\tilde{f}$ as in Equation (2.1), is

$$\vec{\tilde{f}} = \vec{f} - \frac{\tau^T \vec{f}}{\tau^T D_w^{-1} \tau} D_w^{-1} \tau, \tag{2.3}$$

where $\vec{f}$ is the coefficient vector of f and D_w is a diagonal matrix of the weights. The polynomial $\tilde{f}$ is the only polynomial that attains the infimum given in Equation (2.2).

Remark 2 The infimum

$$\rho_{2,w}(f) = \inf_{\alpha \in \mathbb{R}^n} \mathcal{N}_{2,w}^{[f]}(\alpha) \tag{2.4}$$

is the unconstrained radius of positive semidefiniteness.

Remark 3 If the weighted norm is the Euclidean norm then the formula becomes

$$\mathcal{N}_2(\alpha) = \frac{f(\alpha)^2}{\sum_{i_1+\cdots+i_n=0}^{d} \alpha_1^{2i_1} \cdots \alpha_n^{2i_n}}. \tag{2.5}$$

Remark 4 The formulas in [10] and [24] use the weights w_i in the denominator of Equation (2.2), not correctly their reciprocals $1/w_i$.

Remark 5 A degree constraint is required. If we are allowed to grow the degree of the polynomial then the radius of positive semidefiniteness is always 0. To illustrate this point consider a given polynomial $f(x)$ which does not have a real root, and construct the polynomial $g(x) = (\epsilon\, x + 1) f(x) = f(x) + \epsilon\, x\, f(x)$. Notice that $g(x)$ has the root $(-1/\epsilon)$. Therefore, $\rho_2(f) = 0$.

Remark 6 Different degree conditions in Equation (2.5) give different rational functions. For example, if the individual variable degrees are bounded by d, where $\deg_{x_j}(f) \le d$ for all j with $1 \le j \le n$, then for the ℓ^2-norm,

$$\mathcal{N}_2(\alpha) = \frac{f(\alpha)^2}{\sum_{i_1=0}^{d} \cdots \sum_{i_n=0}^{d} \alpha_1^{2i_1} \cdots \alpha_n^{2i_n}}.$$

Comparing the denominators, we have

$$\sum_{i_1=0}^{d} \cdots \sum_{i_n=0}^{d} \alpha_1^{2i_1} \cdots \alpha_n^{2i_n} \ge \sum_{i_1+\cdots+i_n=0}^{d} \alpha_1^{2i_1} \cdots \alpha_n^{2i_n},$$

so

$$\inf \frac{f(\alpha)^2}{\sum_{i_1=0}^{d} \cdots \sum_{i_n=0}^{d} \alpha_1^{2i_1} \cdots \alpha_n^{2i_n}} \le \inf \frac{f(\alpha)^2}{\sum_{i_1+\cdots+i_n=0}^{d} \alpha_1^{2i_1} \cdots \alpha_n^{2i_n}},$$

which must be, since we optimize over a larger set of $\tilde{f}$.

Remark 7 Theorem 2 is the real case of the theorems in [34] and [30] for complex roots. However, they use generalized Lagrangian interpolation for their proof. They also allow keeping selected coefficients of f as their input values and only deform the others in $\tilde{f}$, thus preserving sparsity or monicity, for instance. Our Theorem 2 has theirs as an immediate corollary by setting the weights of those coefficients to infinity in the limit.

However, the problem may have no solution. Consider the case that f has a nonzero constant coefficient which is fixed, and $\alpha = 0$. Then the set of $\tilde{f}$ is empty. Notice that $\lim_{w \to \infty} \mathcal{N}_2^{[f]}(0) = 1/0$.

If a weight $w_i \to 0$ in the limit then the corresponding coefficient in $\tilde{f}$ becomes a "don't care" deformation, i.e., any change in that coefficient is not taken into account in the distance measure. The "nearest" polynomial $\tilde{f}$ with $\alpha \in (\mathbb{R} \setminus \{0\})^n$ as a root then has distance 0, namely $\tilde{f}(x) = f(x) - (f(\alpha)/\alpha^i)\, x^i$, unless there are additional constraints on the coefficients of $\tilde{f}$ in effect.

In Section 2.6 we generalize our approach to handle arbitrary linear constraints on the coefficients of $\tilde{f}$.

Proof of Theorem 2. Let $\vec{f}$, τ, and f be as above. Denote the coefficients of $\tilde{f}$ in Equation (2.1) by

$$\tilde{f}(x_1, \ldots, x_n) = \sum_{i_1 + \cdots + i_n = 0}^{d} \tilde{f}_{i_1, \ldots, i_n} x_1^{i_1} \cdots x_n^{i_n}.$$

Let $\vec{\tilde{f}}$ be the coefficient vector of $\tilde{f}$. Also, $\tilde{f}(\alpha_1, \ldots, \alpha_n) = \tau^T \vec{\tilde{f}} = 0$. We have

$$\|f - \tilde{f}\|_{2,w} = (\vec{f} - \vec{\tilde{f}})^T D_w (\vec{f} - \vec{\tilde{f}}),$$

the weighted ℓ^2-norm, where D_w is a diagonal matrix of the weights. Let λ be the Lagrange multiplier and

$$L(\alpha_1, \ldots, \alpha_n, \lambda) = (\vec{f} - \vec{\tilde{f}})^T D_w (\vec{f} - \vec{\tilde{f}}) + \lambda \tau^T \vec{\tilde{f}}$$

the Lagrange function of our constrained optimization problem. We must check that α is a regular point (i.e., the gradient of the constraint is not 0 at α). Since $\nabla \tilde{f}(\alpha) \neq 0$ if $\tau \neq 0$, then α is a regular point as long as $\alpha \neq 0$. In the case $\alpha = 0$ the constant coefficient of f is deformed to 0 and the formulas hold. The Jacobian of L w.r.t. $\vec{\tilde{f}}$ and

λ is

$$J_L = \begin{bmatrix} \vdots \\ \frac{\partial L}{\partial (\vec{\tilde{f}})_i} \\ \vdots \\ \frac{\partial L}{\partial \lambda} \end{bmatrix} = \begin{bmatrix} -2D_w(\vec{f} - \vec{\tilde{f}}) + \lambda\tau \\ \\ \tau^T \vec{\tilde{f}} \end{bmatrix}.$$

Looking at the first block of the vector we have

$$-2D_w(\vec{f} - \vec{\tilde{f}}) + \tau\lambda = -2D_w\vec{f} + 2D_w\vec{\tilde{f}} + \tau\lambda = 0. \tag{2.6}$$

Multiplying by $\tau^T D_w^{-1}$ we have

$$-2\tau^T D_w^{-1} D_w \vec{f} + 2\tau^T D_w^{-1} D_w \vec{\tilde{f}} + \tau^T D_w^{-1}\tau\lambda = 0.$$

Recalling that $\tilde{f}(\alpha_1, \ldots, \alpha_n) = 0$ which means that $\tau^T \vec{\tilde{f}} = 0$, then we have

$$-2\tau^T I \vec{f} + 2\tau^T I \vec{\tilde{f}} + \tau^T D_w^{-1}\tau\lambda = -2\tau^T \vec{f} + \tau^T D_w^{-1}\tau\lambda = 0.$$

Solving for λ we get $\lambda = \dfrac{2\tau^T \vec{f}}{\tau^T D_w^{-1}\tau}$. Looking at Equation (2.6), we have $\vec{f} - \vec{\tilde{f}} = \dfrac{D_w^{-1}\tau\lambda}{2}$. Substituting in for λ we obtain as the only solution $\vec{f} - \vec{\tilde{f}} = \dfrac{\tau^T \vec{f}}{\tau^T D_w^{-1}\tau} D_w^{-1}\tau$. Finally,

$$\mathcal{N}_{2,w}^{[f]}(\alpha) = (\frac{\tau^T \vec{f}}{\tau^T D_w^{-1}\tau} D_w^{-1}\tau)^T D_w (\frac{\tau^T \vec{f}}{\tau^T D_w^{-1}\tau} D_w^{-1}\tau) = \frac{\vec{f}^T \tau\tau^T D_w^{-1}\tau\tau^T \vec{f}}{\tau^T D_w^{-1}\tau\tau^T D_w^{-1}\tau}.$$

Therefore, $\mathcal{N}_{2,w}^{[f]}(\alpha) = \dfrac{f(\alpha)^2}{\tau^T D_w^{-1}\tau}$. □

Example 1 Consider the polynomial $f = x^2 + y^2 + 1$. We want to find the nearest polynomial $\tilde{f} = a_{2,0}x^2 + a_{1,1}xy + a_{0,2}y^2 + a_{1,0}x + a_{0,1}y + a_{0,0}$ with root $(\alpha, \beta) = (0, 0)$. In other words, we are trying to find

$$\mathcal{N}_2^{[f]}(0,0) = \inf_{\tilde{f} \in \mathbb{R}[x_1,\ldots,x_n]} (1-a_{2,0})^2 + (0-a_{1,1})^2 + (1-a_{0,2})^2 + (0-a_{1,0})^2 + (0-a_{0,1})^2 + (1-a_{0,0})^2.$$

Using Theorem 2 we get that

$$\mathcal{N}_2^{[f]}(0,0) = \frac{(0^2+0^2+1)^2}{[0,0,0,0,0,1][0,0,0,0,0,1]^T} = 1$$

and

$$\tilde{f} = x^2 + y^2.$$

The nearest polynomial with a real root is:

$$\rho_2(f) = \inf_{(\alpha,\beta)} \mathcal{N}_2^{[f]}(\alpha,\beta) = 1. \ \square$$

Example 2 Given the polynomial $f(x) = x^2 + 1$, what is the relationship between $\rho(f)$ and $\rho(g)$ where $g(x) = f(x)(y^2+1)$? According to [10], the nearest polynomial to f occurs at the root $x = 0$. Theorem 2 gives us that $\mathcal{N}_2^f(0) = 1$. However, $\rho_2(g) = 0$. Consider $g(x,1) = 2x^2 + 2$. If we subtract ϵx^4, we have $g(x,1) - \epsilon x^4 = -\epsilon x^4 + 2x^2 + 2$, which has the root $x = \dfrac{-2 \pm \sqrt{4+8\epsilon}}{-2\epsilon}$ for all ϵ. This is similar to Remark 5. We have increased the degree of the terms allowed in the root vector τ; therefore, we will always have $\rho_2(g) = 0$. $\square$

The majority of this section is from [13].

2.4 Deformation Analysis

As mentioned in Section 1.5, there may not exist an affine optimizer.

Example 3 Consider the polynomial

$$f = x^2 + y^2 - 2xy + 4.$$

The deformed polynomial $(1-\epsilon^2)x^2 + y^2 - 2xy + 4$ attains for any $\epsilon > 0$ negative values at $x = y > 2/\epsilon$. Thus the polynomial $x^2 + y^2 - 2xy + 4$ has a radius of positive semidefiniteness $= 0$ although its global minimum is 4. $\square$

Example 4 Similar to Example 1, if we consider the polynomial $f = x^2 + 1$ then

$$\rho_2(\, x^2 + 1\,) = 1, \quad \text{and} \quad \tilde{f} = x^2.$$

Notice that $f \succeq 1$, with $\rho_2(f) = 1$. Consider the polynomial

$$g = x^4y^2 + x^2y^4 + z^6 - 3x^2y^2z^2 + 1 \succeq 1.$$

However, the nearest polynomial with a real root has distance $\rho_2(g) = 0$ because $(\frac{1}{\epsilon}, \frac{1}{\epsilon}, \frac{1}{\epsilon})$ is a root of $g - \epsilon^6 x^2y^2z^2$. $\square$

Remark 8 Within any ϵ of the radius, as in Equation (2.4), there is a polynomial that attains negative values: for any $\epsilon > 0$ there is an $\tilde{f}_\epsilon$ with a real root α and

$$\|f - \tilde{f}_\epsilon\|_{2,w}^2 < \rho_{2,w}(f) + \epsilon/2.$$

Then $(\tilde{f}_\epsilon - \delta)(\alpha) < 0$ for all $\delta > 0$, and in particular if $w_1\delta^2 < \epsilon/2$ we have

$$\|f - (\tilde{f}_\epsilon - \delta)\|_{2,w}^2 < \rho_{2,w}(f) + \epsilon.$$

In Section 2.6 we permit constraints for the coefficients of $\tilde{f}$. Then a negative evaluation may be impossible: e.g., $\tilde{f}(x,y) = \tilde{f}_{2,0}x^2 + \tilde{f}_{0,2}y^2$ and $\tilde{f}_{2,0} \geq 0$, $\tilde{f}_{0,2} \geq 0$. However, within less of the distance to the nearest polynomial with a real root, a deformed $\tilde{f}$ remains positive definite.

Example 5 Here we give another example for the case that the infimum in Equation (2.4) is not always attainable. Consider the polynomial

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

We have that

$$\mathcal{N}_2(\alpha,\beta) = \frac{((1-\alpha\beta)^2 + \alpha^2)^2}{\sum_{i+j=0}^{4} \alpha^{2i}\beta^{2j}}.$$

Then $\inf\limits_{\alpha,\beta} \mathcal{N}_2(\alpha,\beta) = 0$. Suppose now that there exists α, β such that the numerator is 0. Then $(1-\alpha\beta) = 0$ and $\alpha = 0$. But if $\alpha = 0$ then $\alpha\beta = 0$. Then $1-\alpha\beta \neq 0$, contradiction.

Thus f does not have a real root and the infimum is not attainable.

We have

$$\mathcal{N}_2(\epsilon, \frac{1}{\epsilon}) = \frac{\epsilon^4}{\delta}, \delta = 3 + 2\epsilon^2 + \frac{2}{\epsilon^2} + 2\epsilon^4 + \frac{2}{\epsilon^4} + \epsilon^6 + \frac{1}{\epsilon^6} + \epsilon^8 + \frac{1}{\epsilon^8},$$

and the nearest polynomial to f with $(\alpha, \beta) = (\epsilon, 1/\epsilon)$ as its root is

$$\begin{aligned}\tilde{f}(x, y) =& -\frac{\epsilon^6}{\delta}x^4 - \frac{\epsilon^4}{\delta}x^3y + (1 - \frac{\epsilon^2}{\delta})x^2y^2 - \frac{1}{\delta}xy^3 - \frac{1}{\epsilon^2\delta}y^4 \\ & -\frac{\epsilon^5}{\delta}x^3 - \frac{\epsilon^3}{\delta}x^2y - \frac{\epsilon}{\delta}xy^2 - \frac{1}{\epsilon\delta}y^3 + (1 - \frac{\epsilon^4}{\delta})x^2 \\ & -(2 + \frac{\epsilon^2}{\delta})xy - \frac{1}{\delta}y^2 - \frac{\epsilon^3}{\delta}x - \frac{\epsilon}{\delta}y + 1 - \frac{\epsilon^2}{\delta}.\end{aligned}$$

Note that $f(\epsilon, 1/\epsilon) - \epsilon^2 = 0$ has squared distance ϵ^4 from f, which is larger than $\epsilon^4/3 > \epsilon^4/\delta$ for all $\epsilon \neq 0$. $\square$

Example 6 Given a polynomial

$$\begin{aligned} f &= x^2y^2 + x^2 - xy + y^4 - y^2 + 1 \\ &= (xy - 1/2)^2 + (y^2 - 1/2)^2 + x^2 + 1/2, \end{aligned}$$

decide the minimum perturbation such that the perturbed polynomial has a real root. We can perturb f by any monomial term with degree bounded by 4, so we consider $f(x, y) - \epsilon x^4$. For $f(x, y) - \epsilon x^4$ one has for $x = y^2$ that $g(y) = f(y^2, y) - \epsilon y^8$. Notice that $g(y)$ always has a real root, because $g(0) = 1$ and $g(\infty) = -\infty$. We see that f has a radius of positive semidefiniteness that is 0. $\square$

Example 7 Consider the polynomial $f(x, y) = (xy)^2 + 1$. Then $f(x, y) - \epsilon x$ has the root $(2/\epsilon, \epsilon/2)$. Hence $\rho(f) = 0$. $\square$

Another example of when the infimum is not attainable is given in Example 23. Most of this section is from [13].

2.5 Nearest Polynomial with Complex Coefficients and/or Roots in Weighted Euclidean Norm

Theorem 2 can be generalized to a complex root α and real/complex coefficients for f, which is the original setting of [4, 24, 38, 34].

Below we present a proof for the Euclidean norm case. This can be easily generalized to include weighted norms.

Theorem 3 *Let $f \in \mathbb{C}[x_1, \ldots, x_n]$,*

$$f(x_1, \ldots, x_n) = \sum_{i_1+\cdots+i_n=0}^{d} f_{i_1,\ldots,i_n} x_1^{i_1} \cdots x_n^{i_n}.$$

For $\tau = [1, \alpha_1, \ldots, \alpha_n, \ldots, \alpha_1^{i_1} \cdots \alpha_n^{i_n}, \ldots]$, the vector of possible term values in $\tilde{f}$, the distance to the nearest polynomial with root $\alpha \in \mathbb{C}^n$ is

$$\mathcal{N}_2^{[f]}(\alpha) = \frac{(\bar{f}(\bar{\alpha}))(f(\alpha))}{\tau^H \tau}. \tag{2.7}$$

Here ^{H} denotes the Hermitian transpose and $^-$ complex conjugation. Furthermore, the coefficient vector $\vec{\tilde{f}}$, for the polynomial $\tilde{f}$ as in Equation (2.1), is

$$\vec{\tilde{f}} = \vec{f} - \frac{\tau^T \vec{f}}{\tau^H \tau} \bar{\tau}, \tag{2.8}$$

where $\vec{f}$ is the coefficient vector of f. The polynomial $\tilde{f}$ is the only polynomial that attains the infimum (Equation (2.7)).

Proof of Theorem 3. Let $\vec{f}$, τ, and f be as above. Denote the coefficients of $\tilde{f}$ by

$$\tilde{f}(x_1, \ldots, x_n) = \sum_{i_1+\cdots+i_n=0}^{d} \tilde{f}_{i_1,\ldots,i_n} x_1^{i_1} \cdots x_n^{i_n}.$$

Let $\vec{\tilde{f}}$ be the coefficient vector of $\tilde{f}$. Also, $\tilde{f}(\alpha_1, \ldots, \alpha_n) = \tau^T \vec{\tilde{f}} = 0$. We have

$$\|f - \tilde{f}\|_2 = (\vec{f} - \vec{\tilde{f}})^H (\vec{f} - \vec{\tilde{f}}),$$

the ℓ^2-norm. Let $\vec{f} = R_{\vec{f}} + I_{\vec{f}}\boldsymbol{i}$ and $\vec{\bar{f}} = R_{\vec{\bar{f}}} + I_{\vec{\bar{f}}}\boldsymbol{i}$ where $R_{\vec{f}}$, $R_{\vec{\bar{f}}}$ are vectors of the real parts of $\vec{f}$ and $\vec{\bar{f}}$ and $I_{\vec{f}}$, $I_{\vec{\bar{f}}}$ are vectors of the imaginary parts of $\vec{f}$ and $\vec{\bar{f}}$ respectively. Let λ be the Lagrange multiplier and

$$\begin{aligned} L(\alpha_1, \dots, \alpha_n, \lambda) =& (R_{\vec{f}}^T + I_{\vec{f}}^T\boldsymbol{i} - (R_{\vec{\bar{f}}} - I_{\vec{\bar{f}}}\boldsymbol{i}))(R_{\vec{f}}^T + I_{\vec{f}}^T\boldsymbol{i} - (R_{\vec{\bar{f}}} + I_{\vec{\bar{f}}}\boldsymbol{i})) \\ &+ \bar{\lambda}(R_\tau^T + I_\tau^T\boldsymbol{i})(R_{\vec{\bar{f}}} + I_{\vec{\bar{f}}}\boldsymbol{i}) + \lambda^T(R_\tau^T - I_\tau^T\boldsymbol{i})(R_{\vec{\bar{f}}} - I_{\vec{\bar{f}}}\boldsymbol{i}) \end{aligned} \tag{2.9}$$

the Lagrange function of our constrained optimization problem.

The Jacobian of L with respect to $R_{\vec{\bar{f}}}$ and $I_{\vec{\bar{f}}}$ is

$$J_L = \begin{bmatrix} \vdots \\ \frac{\partial L}{\partial (R_{\vec{\bar{f}}})_i} \\ \vdots \\ \frac{\partial L}{\partial I_{\vec{\bar{f}}}} \end{bmatrix} = \begin{bmatrix} 2R_{\vec{\bar{f}}} - 2R_{\vec{f}} + \bar{\lambda}R_\tau + \bar{\lambda}I_\tau\boldsymbol{i} + \lambda R_\tau - \lambda I_\tau\boldsymbol{i} \\ \\ 2I_{\vec{\bar{f}}} - 2I_{\vec{f}} + \bar{\lambda}R_\tau\boldsymbol{i} - \bar{\lambda}I_\tau - \lambda R_\tau\boldsymbol{i} - \lambda I_\tau \end{bmatrix}.$$

Adding the first block of the vector with $-\boldsymbol{i}$ times the second block of the vector we see that

$$-2(R_{\vec{f}} - I_{\vec{f}}\boldsymbol{i} - R_{\vec{\bar{f}}} + I_{\vec{\bar{f}}}\boldsymbol{i} -) + 2\bar{\lambda}(R_\tau + I_\tau\boldsymbol{i}) = 0.$$

Multiplying by $(R_\tau - I_\tau\boldsymbol{i})^T$ and recalling $\tau^T\vec{\bar{f}} = 0$ we get

$$-2(R_\tau - I_\tau\boldsymbol{i})^T(R_{\vec{f}} - I_{\vec{f}}\boldsymbol{i}) + 2\bar{\lambda}(R_\tau + I_\tau\boldsymbol{i})^T(R_\tau - I_\tau\boldsymbol{i}) = 0.$$

Solving for $\bar{\lambda}$ we get $\bar{\lambda} = \dfrac{\bar{\tau}^T\,\overline{\vec{\bar{f}}}}{\tau^H\tau}$.

Adding the first block of J_L and $\boldsymbol{i}$ times the second block of J_L we obtain

$$2(R_{\vec{\bar{f}}} + I_{\vec{\bar{f}}}\boldsymbol{i} - R_{\vec{f}} - I_{\vec{f}}\boldsymbol{i}) + 2\lambda R_\tau - 2\lambda I_\tau\boldsymbol{i} = -2(\vec{f} - \vec{\bar{f}}) + 2\lambda\bar{\tau}.$$

Multiplying by $(R_\tau + I_\tau\boldsymbol{i})^T$ and recalling $\tau^T\vec{\bar{f}} = 0$ we get

$$-2(R_\tau^T R_{\vec{f}} + R_\tau^T I_{\vec{f}}\boldsymbol{i} + I_\tau^T R_{\vec{f}}\boldsymbol{i} - I_\tau^T I_{\vec{f}}) + 2\lambda(R_\tau^T + I_\tau^T\boldsymbol{i})(R_\tau - I_\tau\boldsymbol{i}) = -2\tau^T\vec{f} + 2\lambda\tau^T\bar{\tau}.$$

Solving for λ we get $\lambda = \dfrac{\tau^T \vec{f}}{\tau^H \tau}$. Solving Equation (2.5) and substituting in for λ we obtain

$$\vec{f} - \vec{\tilde{f}} = \lambda \bar{\tau} = \frac{\tau^T \vec{f}}{\tau^H \tau} \bar{\tau}.$$

Finally, $\mathcal{N}_{2,w}^{[f]}(\alpha) = (\dfrac{\tau^T \vec{f}}{\tau^H \tau} \bar{\tau})^H (\dfrac{\tau^T \vec{f}}{\tau^H \tau} \bar{\tau}) = \dfrac{\bar{\vec{f}}^T \bar{\tau} \tau^T \bar{\tau} \tau^T \vec{f}}{\tau^T \bar{\tau} \tau^T \bar{\tau}} = \dfrac{(\bar{f}(\bar{\alpha})) f(\alpha)}{\tau^H \tau}$. $\square$

Remark 9 For $\alpha \in \mathbb{R}$ then

$$\left.\begin{aligned} \inf_{\tilde{f} \in \mathbb{C}[x_1,\ldots,x_n]} \|f - \tilde{f}\|_{2,w}^2 = \inf_{\tilde{f} \in \mathbb{R}[x_1,\ldots,x_n]} & \|f - \tilde{f}\|_{2,w}^2 \\ \text{s. t. } & \tilde{f}(\alpha) = 0, \\ & \deg(\tilde{f}) \leq \deg(f). \end{aligned}\right\} \tag{2.10}$$

Notice that $\|\vec{f} - \vec{\tilde{f}}\|_2 = (R_{\vec{f}} - (R_{\vec{\tilde{f}}} - I_{\vec{\tilde{f}}}))(R_{\vec{f}} - (R_{\vec{\tilde{f}}} + I_{\vec{\tilde{f}}})) = \|R_{\vec{f}} - R_{\vec{\tilde{f}}}\|_2 - I_{\vec{\tilde{f}}}^2$. Thus, $I_{\vec{\tilde{f}}} = 0$. Therefore if we want to find the nearest polynomial to $f \in \mathbb{R}[x_1, \ldots, x_n]$ with a real root α then we only need to consider polynomials with real coefficients.

2.6 Coefficient Constraints

2.6.1 Linear Constraints

Our method can be further generalized to include problems with linear constraints of the form $H\vec{\tilde{f}} = p$, where $H \in \mathbb{R}^{t \times s}$, $p \in \mathbb{R}^t$, on the coefficient vector $\vec{\tilde{f}}$ of $\tilde{f}$. We define

$$\left.\begin{aligned} \mathcal{N}_{2,w}^{[f;H]}(\alpha) = \inf_{\tilde{f} \in \mathbb{R}[x_1,\ldots,x_n]} & \|f - \tilde{f}\|_{2,w}^2 \\ \text{s. t. } & \tilde{f}(\alpha) = 0, H\vec{\tilde{f}} = p, \\ & \deg(\tilde{f}) \leq \deg(f). \end{aligned}\right\} \tag{2.11}$$

We note that the Jacobian of the Lagrange function corresponding to Equation (2.11) constitutes a linear system in the unknown coefficients of $\tilde{f}$ and the multipliers, hence a determinantal formula parameterized by the real root for the solution can be computed, which one can minimize.

Example 8 Given a polynomial $f(x, y) = x^2 + y^2 + 1$ find the nearest polynomial $\tilde{f}(x, y) = \tilde{f}_{2,0}x^2 + \tilde{f}_{0,2}y^2 + \tilde{f}_{1,1}xy + \tilde{f}_{1,0}x + \tilde{f}_{0,1}y + \tilde{f}_{0,0}$ with $\tilde{f}_{1,1} = \tilde{f}_{0,0}$ and $\tilde{f}_{0,1} = 0$ and with the root (α, β). The Lagrangian function is

$$L(\alpha, \beta, \lambda) = (\vec{f} - \vec{\tilde{f}})^T (\vec{f} - \vec{\tilde{f}}) + \lambda_0 \tau^T \vec{\tilde{f}} + \lambda_1 (\tilde{f}_{1,1} - \tilde{f}_{0,0}) + \lambda_2 \tilde{f}_{0,1},$$

with term vector $\tau = [\alpha^2, \beta^2, \alpha\beta, \alpha, \beta, 1]$. The Jacobian of L in $\vec{\tilde{f}}$ and λ is zero for

$$\vec{\tilde{f}} = \begin{bmatrix} \tilde{f}_{2,0} \\ \tilde{f}_{0,2} \\ \tilde{f}_{1,1} \\ \tilde{f}_{1,0} \\ \tilde{f}_{0,1} \\ \tilde{f}_{0,0} \end{bmatrix} = \begin{bmatrix} -\frac{-\alpha^2-2\beta^4+\alpha^2\beta^2-2\alpha\beta-1+\alpha^3\beta}{2\alpha^2+2\beta^4+2\alpha^4+\alpha^2\beta^2+2\alpha\beta+1} \\ \frac{2\alpha^2+2\alpha^4-\alpha^2\beta^2+2\alpha\beta+1-\beta^2-\alpha\beta^3}{2\alpha^2+2\beta^4+2\alpha^4+\alpha^2\beta^2+2\alpha\beta+1} \\ \frac{\beta^4+\alpha^4-\alpha\beta^3-\alpha^3\beta-\beta^2}{2\alpha^2+2\beta^4+2\alpha^4+\alpha^2\beta^2+2\alpha\beta+1} \\ -\frac{\alpha(1+2\beta^2+\alpha\beta+2\alpha^2)}{2\alpha^2+2\beta^4+2\alpha^4+\alpha^2\beta^2+2\alpha\beta+1} \\ 0 \\ \frac{\beta^4+\alpha^4-\alpha\beta^3-\alpha^3\beta-\beta^2}{2\alpha^2+2\beta^4+2\alpha^4+\alpha^2\beta^2+2\alpha\beta+1} \end{bmatrix},$$

$$\lambda_0 = \frac{2(2\alpha^2 + \alpha\beta + 2\beta^2 + 1)}{2\alpha^4 + 2\alpha^2 + \alpha^2\beta^2 + 2\alpha\beta + 1 + 2\beta^4},$$

$$\lambda_1 = \frac{-2(\alpha^4 + \alpha^3\,\beta + \alpha^2\,\beta^2 + \alpha\beta + \beta^3\alpha + \beta^4 - \beta^2)}{2\alpha^4 + 2\alpha^2 + \alpha^2\beta^2 + 2\alpha\beta + 1 + 2\beta^4},$$

$$\lambda_2 = \frac{-2(2\alpha^2 + \alpha\beta + 2\beta^2 + 1)\beta}{(2\alpha^4 + 2\alpha^2 + \alpha^2\beta^2 + 2\alpha\beta + 1 + 2\beta^4)}.$$

The minimum perturbation is

$$\mathcal{N}_2 = \frac{3\alpha^4+2\alpha^3\beta+5\alpha^2\beta^2+3\alpha^2+2\alpha\beta+2\alpha\beta^3+1+3\beta^4+2\beta^2}{2\alpha^2+2\alpha^4+2\beta^4+\alpha^2\beta^2+2\alpha\beta+1}. \tag{2.12}$$

Running the `Minimize` procedure in Maple 14 we obtain $\min_{(\alpha,\beta)} \mathcal{N}_2 = 1$ at the root $(0, 0)$ and $\tilde{f} = x^2 + y^2$. That is the same deformed polynomial as for the unconstrained problem but derived from a different norm expression (Equation (2.12)).

Note that before minimizing Equation (2.12) one could restrict (α, β) to lie on a parametric curve, thus constraining the variables rather than the coefficients, as is done in [10]. $\square$

Example 9 Given the polynomial $f = x^2 + y^2 - 2xy + 4$, find the nearest polynomial $\tilde{f}(x, y) = \tilde{f}_{2,0}x^2 + \tilde{f}_{0,2}y^2 + \tilde{f}_{1,1}xy + \tilde{f}_{1,0}x + \tilde{f}_{0,1}y + \tilde{f}_{0,0}$ with $\tilde{f}_{1,1} = -3$ and with the root (α, β).

The Lagrangian function is

$$L(\alpha,\beta,\lambda) = (\vec{f} - \vec{\tilde{f}})^T(\vec{f} - \vec{\tilde{f}}) + \lambda_0 \tau^T \vec{\tilde{f}} + \lambda_1(\tilde{f}_{1,1} + 3),$$

where the term vector $\tau = [\alpha^2, \beta^2, \alpha\beta, \alpha, \beta, 1]$.

The Jacobian of L in $\vec{\tilde{f}}$ and λ is zero for

$$\vec{\tilde{f}} = \begin{bmatrix} \tilde{f}_{2,0} \\ \tilde{f}_{0,2} \\ \tilde{f}_{1,1} \\ \tilde{f}_{1,0} \\ \tilde{f}_{0,1} \\ \tilde{f}_{0,0} \end{bmatrix} = \begin{bmatrix} \frac{1-3\,\alpha^2+\beta^2+\beta^4-\alpha^2\beta^2+3\,\alpha^3\beta}{1+\alpha^2+\beta^2+\alpha^4+\beta^4} \\ \frac{1+\alpha^2-3\,\beta^2+\alpha^4-\alpha^2\beta^2+3\,\alpha\,\beta^3}{1+\alpha^2+\beta^2+\alpha^4+\beta^4} \\ -3 \\ -\frac{\alpha\left(4+\alpha^2+\beta^2-3\,\alpha\,\beta\right)}{1+\alpha^2+\beta^2+\alpha^4+\beta^4} \\ -\frac{\beta\left(4+\alpha^2+\beta^2-3\,\alpha\,\beta\right)}{1+\alpha^2+\beta^2+\alpha^4+\beta^4} \\ \frac{3\,\alpha\,\beta+3\,\alpha^2+3\,\beta^2+4\,\alpha^4+4\,\beta^4}{1+\alpha^2+\beta^2+\alpha^4+\beta^4} \end{bmatrix}$$

$$\lambda_0 = \frac{2(4+\alpha^2+\beta^2-3\,\alpha\,\beta)}{1+\alpha^2+\beta^2+\alpha^4+\beta^4}$$

$$\lambda_1 = \frac{2(1+\alpha^2+\beta^2+\alpha^4+\beta^4-4\,\alpha\,\beta-\alpha^3\beta-\alpha\,\beta^3+3\,\alpha^2\beta^2)}{1+\alpha^2+\beta^2+\alpha^4+\beta^4}$$

The minimum perturbation is

$$\mathcal{N}_2 = \frac{2\,\alpha^4-6\,\alpha^3\beta+9\,\alpha^2+11\,\alpha^2\beta^2-6\,\alpha\,\beta^3-24\,\alpha\,\beta+17+9\,\beta^2+2\,\beta^4}{1+\alpha^2+\beta^2+\alpha^4+\beta^4}. \tag{2.13}$$

Running the `Minimize` procedure in Maple 14 we obtain $\min\limits_{(\alpha,\beta)} \mathcal{N}_2 = 1$ at the root $(-2,-2)$ and $\tilde{f} = x^2 + y^2 - 3xy + 4$. $\square$

Example 10 Looking at the same polynomial as in the last example $f = x^2+y^2-2xy+4$ but with the constraint $\tilde{f}_{1,1} = -4\tilde{f}_{2,0}$ we obtain

$$\mathcal{N}_2 = \frac{5\,\alpha^4-40\,\alpha^3\beta+76\,\alpha^2+98\,\alpha^2\beta^2-72\,\alpha\,\beta^3-288\,\alpha\,\beta+21\,\beta^4+140\,\beta^2+276}{17+17\,\alpha^2+17\,\beta^2+17\,\beta^4+\alpha^4-8\,\alpha^3\beta+16\,\alpha^2\beta^2}. \tag{2.14}$$

The nearest polynomial is

$$\tilde{f} = 0.5294117644\,x^2 + 0.9999999996\,y^2 - 2.117647058\,xy + 3.999999998$$

at

$$(3.58092825040923168, 5.68559836091241166)$$

with

$$\min_{(\alpha,\beta)} \mathcal{N}_2 = 0.235294117647058626. \quad \square$$

This section has been adapted from [13].

2.6.2 Inequality Constraints

Our method can be generalized even further to include inequalities, $G\vec{\tilde{f}} \leq q$ with $G \in \mathbb{R}^{m \times s}$. Then

$$\left.\begin{array}{rl} \mathcal{N}_{2,w}^{[f;H;G]}(\alpha) = & \inf\limits_{\tilde{f} \in \mathbb{R}[x_1,\ldots,x_n]} \|f - \tilde{f}\|_{2,w}^2 \\ \text{s. t.} & \tilde{f}(\alpha) = 0, G\vec{\tilde{f}} \leq q, H\vec{\tilde{f}} = p, \\ & \deg(\tilde{f}) \leq \deg(f). \end{array}\right\} \tag{2.15}$$

Note that our constraint functions, being linear, are always convex. We can use the Karush-Kuhn-Tucker (KKT) conditions and the quantities as defined in Theorem 2.

The KKT conditions (for a regular point) are then (see [6]):
Given the canonical form nonlinear programming problem

$$\begin{array}{lll} \min\limits_{x \in \mathbb{R}^n} & f_0(x) & \\ \text{subject to} & f_i(x) \leq 0 & i = 1, \ldots, m, \end{array}$$

if the Lagrangian function is given by

$$L(x, u) = f_0(x) + \sum_{i=1}^{m} u_i f_i(x)$$

then the KKT conditions for the problem are

$$\left.\begin{array}{lll} \dfrac{\partial L}{\partial(x_j)} = 0, & j = 1, \ldots, n, & \text{gradient condition} \\ u_i f_i(x) = 0 & i = 1, \ldots, m & \text{orthogonality} \\ f_i(x) \leq 0 & i = 1, \ldots, m & \text{feasibility} \\ u_i \geq 0 & i = 1, \ldots, m & \text{non-negativity} \end{array}\right\} \tag{2.16}$$

A point $(\bar{x}, \bar{u})$ satisfying the KKT conditions is called a KKT point. If the functions $f_i(x)$ are all convex, the points $\bar{x}$ are global minimizing points.

Using the conditions in Equation (2.15) with the KKT conditions as defined in Equation (2.16) with the Lagrange function:

$$L = (\vec{f} - \vec{\tilde{f}})^T D_w(\vec{f} - \vec{\tilde{f}}) + \lambda_0 \tau^T \vec{\tilde{f}} + \lambda^T (H\vec{\tilde{f}} - p) + \mu^T (G\vec{\tilde{f}} - q), \tag{2.17}$$

the KKT conditions (for a regular point) become

$$\left.\begin{array}{l} \dfrac{\partial L}{\partial (\vec{\tilde{f}})_i} = 0, \quad i = 1, \ldots, s, \\ \tau^T \vec{\tilde{f}} = 0, \\ H\vec{\tilde{f}} = p, \\ G\vec{\tilde{f}} \leq q, \\ \mu_i \geq 0, \quad i = 1, \ldots, m, \\ \mu^T (G\vec{\tilde{f}} - q) = 0. \end{array}\right\} \tag{2.18}$$

The last orthogonality conditions constitute branching: $\mu_i = 0$ or $(G\vec{\tilde{f}} - q)_i = 0$, and (2.18) form linear programs.

Example 11 Given a polynomial $f(x, y) = x^2 + y^2 - 2y + 1$ and constraint $\tilde{f}_{0,1} \geq 0$, we determine the nearest polynomial $\tilde{f}(x, y) = \tilde{f}_{2,0}x^2 + \tilde{f}_{0,2}y^2 + \tilde{f}_{1,1}xy + \tilde{f}_{1,0}x + \tilde{f}_{0,1}y + \tilde{f}_{0,0}$ with real root $(0, 0)$. The term vector for the root is $\tau = [0, 0, 0, 0, 0, 1]$. The Lagrangian function is $L(\alpha, \beta, \lambda, \mu) = (\vec{f} - \vec{\tilde{f}})^T (\vec{f} - \vec{\tilde{f}}) + \lambda \tau^T \vec{\tilde{f}} + \mu(-\tilde{f}_{0,1})$. We can formulate the KKT conditions as solving two linear programs:

$$\begin{array}{ll} \text{Minimize} & 1 \\ \text{subject to} & \partial L / \partial (\vec{\tilde{f}})_i = 0, \quad i = 1, \ldots, 6 \\ & \tilde{f}_{0,0} = 0, \\ & -\tilde{f}_{0,1} \leq 0, \\ & \mu = 0, \end{array}$$

and

$$\begin{array}{ll} \text{Minimize} & 1 \\ \text{subject to} & \partial L / \partial (\vec{\tilde{f}})_i = 0, \quad i = 1, \ldots, 6, \\ & \tilde{f}_{0,0} = 0, \\ & -\tilde{f}_{0,1} = 0, \\ & \mu \geq 0. \end{array}$$

The first linear program is infeasible; for the second linear program we obtain:

$$\tilde{f} = x^2 + y^2,\ \lambda = 2,\ \mu = 4,\ \text{and}\ \mathcal{N}_2^{[f;G]} = 5.$$

The minimum perturbation can also be obtained by running the `Minimize` procedure in Maple 14 on the original optimization problem Equation (2.15). $\square$

Example 12 Given a polynomial $f(x,y) = 1 + x + x^2 - 2x^3y + x^2y^2$ and constraint $\tilde{f}_{3,1} \geq 0$, we determine the nearest polynomial

$$\begin{aligned}\tilde{f}(x,y) = \ & \tilde{f}_{4,0}x^4 + \tilde{f}_{0,4}y^4 + \tilde{f}_{3,1}x^3y + \tilde{f}_{1,3}xy^3 + \tilde{f}_{2,2}x^2y^2 + \tilde{f}_{3,0}x^3 + \tilde{f}_{0,3}y^3 \\ & + \tilde{f}_{2,1}x^2y + \tilde{f}_{1,2}xy^2 + \tilde{f}_{2,0}x^2 + \tilde{f}_{0,2}y^2 + \tilde{f}_{1,1}xy + \tilde{f}_{1,0}x + \tilde{f}_{0,1}y + \tilde{f}_{0,0},\end{aligned}$$

with real root $(2,0)$. The Lagrangian function is

$$L(\alpha,\beta,\lambda,\mu) = (\vec{f} - \vec{\tilde{f}})^T(\vec{f} - \vec{\tilde{f}}) + \lambda\tau^T\vec{\tilde{f}} + \mu(-\tilde{f}_{3,1}).$$

We can formulate the KKT conditions as solving two linear programs:

$$\begin{array}{ll} \text{Minimize} & 1 \\ \text{subject to} & \partial L/\partial(\vec{\tilde{f}})_i = 0, \quad i = 1,\ldots,15 \\ & \tau^T\vec{\tilde{f}} = 0 \\ & -\tilde{f}_{3,1} \leq 0, \\ & \mu = 0, \end{array}$$

and

$$\begin{array}{ll} \text{Minimize} & 1 \\ \text{subject to} & \partial L/\partial(\vec{\tilde{f}})_i = 0, \quad i = 1,\ldots,15, \\ & \tau^T\vec{\tilde{f}} = 0, \\ & -\tilde{f}_{3,1} = 0, \\ & \mu \geq 0. \end{array}$$

The first linear program is infeasible; for the second linear program we obtain:

$$\begin{aligned}\tilde{f} = \ & -0.3284457478005864440x^4 - 0.164222873900293276x^3 + x^2y^2 \\ & +0.917888563049853224x^2 + 0.958944281524926944x + 0.979472140762463250,\end{aligned}$$

$$\lambda = 0.0410557184750733190, \quad \mu = 4$$

and

$$\mathcal{N}_2^{[f;G]} = 4.143695016.$$

For f without constraints, $\mathcal{N}_2^f(2,0) = \frac{49}{341} \approx 0.1436950147$. $\square$

Example 13 We consider the polynomial $f(x,y) = x^2 + y^2 + 1$. We have seen that the nearest polynomial with the root $(0,0)$ is $\tilde{f}(x,y) = x^2 + y^2$, but what is the nearest polynomial with the root $(1,1)$ and the constraint that $\tilde{f}_{0,0} \geq 1$? The Lagrangian function is $L(\alpha,\beta,\lambda,\mu) = (\vec{f} - \vec{\tilde{f}})^T(\vec{f} - \vec{\tilde{f}}) + \lambda\tau^T\vec{\tilde{f}} + \mu(1 - \tilde{f}_{0,0})$.

We can formulate the KKT conditions as solving two linear programs:

$$\begin{array}{ll} \text{Minimize} & 1 \\ \text{subject to} & \partial L/\partial(\vec{\tilde{f}})_i = 0, \quad i = 1,\ldots,6 \\ & \tilde{f}(1,1) = 0, \\ & -\tilde{f}_{0,0} \leq -1, \\ & \mu = 0, \end{array}$$

and

$$\begin{array}{ll} \text{Minimize} & 1 \\ \text{subject to} & \partial L/\partial(\vec{\tilde{f}})_i = 0, \quad i = 1,\ldots,6, \\ & \tilde{f}(1,1) = 0, \\ & \tilde{f}_{0,0} = 1, \\ & \mu \geq 0. \end{array}$$

The first linear program is infeasible; for the second linear program we obtain:

$$\tilde{f} = 0.4x^2 + 0.4y^2 - 0.6x - 0.6y - 0.6xy + 1$$

$$\lambda = 1.2, \quad \mu = 1.2$$

and

$$\mathcal{N}_2^{[f;G]} = 1.8. \ \square$$

This section has been adapted from [13].

Chapter 3

The Case of Several Polynomials

3.1 System of Polynomials with a Real Root in Weighted Euclidean Norm

Theorem 2 can be extended to systems. The distance to the nearest system with k equations and common root α is defined as

$$\left.\begin{array}{rl} \inf\limits_{\tilde{f}^{[1]},\ldots,\tilde{f}^{[k]}} & \|f^{[1]}-\tilde{f}^{[1]}\|_2^2+\cdots+\|f^{[k]}-\tilde{f}^{[k]}\|_2^2 \\ \text{s. t.} & \tilde{f}^{[i]}(\alpha)=0,\ i=1,\ldots,k \\ & f^{[i]}\in\mathbb{R}[x_1,\ldots,x_n],\ i=1,\ldots,k \\ & \deg(\tilde{f}^{[i]})\leq\deg(f^{[i]}),\ i=1,\ldots,k \end{array}\right\} \tag{3.1}$$

Applying Theorem 2 to each individual $\tilde{f}^{[i]}$ easily yields the following. Since we are summing up the smallest deformation of each individual $f^{[i]}$ for $i=1,\ldots,k$ then we will have the nearest system with a common root α.

Theorem 4 *Let $f^{[1]},\ldots,f^{[k]}\in\mathbb{R}[x_1,\ldots,x_n]$, with $d_i=\deg(f^{[i]})$, The distance to the nearest system with a common root $\alpha\in\mathbb{R}^n$ is (in ℓ^2-norm)*

$$\mathcal{N}_2^{\{f^{[1]},\ldots,f^{[k]}\}}(\alpha)=\frac{f^{[1]}(\alpha)^2}{\sum_{i_1+\cdots+i_n=0}^{d_1}\alpha_1^{2i_1}\cdots\alpha_n^{2i_n}}+\cdots+\frac{f^{[k]}(\alpha)^2}{\sum_{i_1+\cdots+i_n=0}^{d_k}\alpha_1^{2i_1}\cdots\alpha_n^{2i_n}}. \tag{3.2}$$

The nearest polynomials, if they exist (see Example 5) are again determined by Equation (2.3). Theorem 4 easily generalizes to include weighted norms. Linear equality and inequality constraints on the coefficients as described in Equation (2.11) and Equation (2.15) can also be applied.

Example 14 Given polynomials

$$f^{[1]}(x,y) = x^4 + y^4 + 1 \text{ and } f^{[2]}(x,y) = x^2 + x^2y^2 - 2xy + 1$$

we shall determine the minimum perturbation such that the deformed system of 2 equations has a real root.

For that, we compute the Gröbner basis of the numerators of the partial derivatives of Equation (3.2) (cf. [2]). The Gröbner basis computation took 150.54 seconds[1]. In Section 5 we present an alternative approach based on sums-of-squares certificates. The first equation in the obtained Gröbner basis is a polynomial in terms of β of degree 195. Next, we find all real roots of this polynomial and plug all 9 choices into a second polynomial in the Gröbner basis. We compute the norm of each possible point and select the minimum value. The minimum perturbation obtained by solving the Gröbner basis of Equation (3.2) in Maple is

$$\widetilde{\mathcal{N}}_2 = 0.645973069980782776667 \tag{3.3}$$

for

$$(\alpha, \beta) = (-0.9138289555225176138,\\ -1.1947071766554875688).$$

Note that for this example at least 25 mantissa digits must be used in Maple 14 in order to obtain the correct minimum.

We can then find the nearest polynomial system by plugging the root into Equation (2.3) for each of the two polynomials:

$$\tilde{f}^{[1]} = 0.83448994938 + 0.15028000318\,x + 0.19773604528\,y - 0.17954059831\,xy - \\ 0.13645140747\,x^2 - 0.23623667238\,y^2 + 0.12389530347\,x^3 + 0.21449844130\,xy^2 + \\ 0.16301947576\,x^2y + 0.28223364788\,y^3 + 0.88750540206\,x^4 - 0.14801860821\,x^3y - \\ 0.19476053763\,x^2y^2 - 0.25626282720\,xy^3 + 0.66281343538\,y^4,$$

[1]All reported timings were run on an 8 CPU (2.8GHz Xeon) MacPro with 8GB of memory under Linux 2.6.32-31 (Ubuntu) using Maple 14 and/or Matlab R2010a.

$\tilde{f}^{[2]} = 0.96296934167 + 0.03362313909\,x + 0.04424079327\,y - 2.04016980557\,xy + 0.96947082410\,x^2 - 0.05285479322\,y^2 + 0.02771991571\,x^3 + 0.04799115499\,xy^2 + 0.03647342555\,x^2y + 0.06314600078\,y^3 - 0.02516916045\,x^4 - 0.03311718223\,x^3y + 0.95642493674\,x^2y^2 - 0.05733537729\,xy^3 - 0.07544098031\,y^4$. □

Example 15 We consider the overdetermined system

$$f^{[1]}(x,y) = x^2 + y^2 - 1, \quad f^{[2]}(x,y) = x^2 - y^2 - 1, \quad f^{[3]}(x) = x^2 - 2.$$

Using Theorem 4 we obtain

$$\begin{aligned}\mathcal{N}_2^{\{f^{[1]},f^{[2]},f^{[3]}\}}(\alpha,\beta) = \; & \frac{\left(\alpha^2+\beta^2-1\right)^2}{1+\alpha^2+\beta^2+\alpha^4+\beta^4+\alpha^2\beta^2} + \frac{\left(\alpha^2-\beta^2-1\right)^2}{1+\alpha^2+\beta^2+\alpha^4+\beta^4+\alpha^2\beta^2} \\ & + \frac{\left(\alpha^2-2\right)^2}{1+\alpha^2+\beta^2+\alpha^4+\beta^4+\alpha^2\beta^2}\end{aligned}$$

Running `Minimize` on $\mathcal{N}_2^{\{f^{[1]},f^{[2]},f^{[3]}\}}(\alpha,\beta)$ we obtain

$$\rho_2(f^{[1]}, f^{[2]}, f^{[3]}) = 0.151159953434232108$$

with

$$(\alpha,\beta) = (1.19718389644442, 0.3153893315661)$$

and

$$\begin{aligned}\tilde{f}^{[1]} = \; & -1.11240274409999995 - 0.134566755152683015\,\alpha \\ & -0.0354506263278943912\,\beta - 0.0424409189517329297\,\alpha\,\beta \\ & +0.838898847836248018\,\alpha^2 + 0.988819250656115645\,\beta^2 \\ \tilde{f}^{[2]} = \; & -1.07042663899999990 - 0.0843136380915045502\,\alpha \\ & -0.0222118105986570287\,\beta - 0.0265916219552682041\,\alpha\,\beta \\ & +0.899061070290003483\,\alpha^2 - 1.00700536809929386\,\beta^2 \\ \tilde{f}^{[3]} = \; & -1.88041678439999993 + 0.143163100001361149\,\alpha \\ & +0.0377152704346088233\,\beta + 0.0451521144070290079\,\alpha\,\beta \\ & +1.17139255777836793\,\alpha^2 + 0.0118949939351104834\,\beta^2.\end{aligned}$$ □

This section has been adapted from [13].

Chapter 4

Infinity Norm and One Norm

4.1 Nearest Polynomial with a Real Root

In this chapter we further explore finding the nearest polynomial or system of polynomials with the weighted infinity norm and one norm as the distance measure. Furthermore, we consider the problem of linear constraints on the coefficients of the polynomial.

As mentioned in Section 2.2, Stetter proved the following theorem.

Theorem 5 (see Theorem 1) *Let the vector of possible term values of f and $\tilde{f}$ be given by $\vec{f}$, $\vec{\tilde{f}} \in \mathbb{C}^n$ respectively. Let the given root be denoted by $\alpha = [\alpha_1, \ldots, \alpha_n] \in \mathbb{C}^n$. Let $\tau = [1, \alpha_1, \ldots, \alpha_n, \ldots, \alpha_1^{i_1} \cdots \alpha_n^{i_n}, \ldots]$, the term vector evaluated at the root. Let $\|\ldots\|$ be the given norm and $\|\ldots\|^*$ the associated dual norm. The nearest polynomial with a real root, i.e. $\tilde{f}(\alpha) = \tau^T \vec{\tilde{f}} = 0$ requires*

$$\|\vec{f} - \vec{\tilde{f}}\|^* \geq \frac{|f(\alpha)|}{\|\tau\|}.$$

Theorem 1 shows that Theorem 2 can be extended to any ℓ^p-norm. We extend the results from Theorem 1 to the weighted ℓ^1 and ℓ^∞-norms. We prove Hölder's inequality for weighted ℓ^1, ℓ^2 and ℓ^∞-norms, which allows us to then follow the same proof as in [38] for Theorem 1. We further give an explicit formula for $\vec{\tilde{f}}$.

Theorem 6 *Let u, $v \in \mathbb{C}^n$ and weights w_i. Then $|v^T u| \leq \|u\|_{\infty,w} \|v\|_{1,1/w}$, where $1/w$ is the vector of reciprocals of entries of w.*

Proof. Looking at $|v^T u|$:

$$|v^T u| = \Big|\sum_i v_i u_i\Big| = \Big|\sum_i w_i v_i \frac{1}{w_i} u_i\Big| \leq (\max_j w_j |v_j|) \sum_i \frac{1}{w_i}|u_i| = \|v\|_{\infty,w}\|u\|_{1,1/w}. \quad \square$$

Corollary 1 *Let u, $v \in \mathbb{C}^n$, and let w_i be the weights. Then*

$$|v^T u| \leq \|u\|_{1,w}\|v\|_{\infty,1/w}.$$

Proof. Looking at $|v^T u|$:

$$|v^T u| = \Big|\sum_i v_i u_i\Big| = \Big|\sum_i \frac{1}{w_i} v_i w_i u_i\Big| \leq (\max_j \frac{1}{w_j}|v_j|) \sum_i w_i|u_i| = \|v\|_{\infty,1/w}\|u\|_{1,w}. \quad \square$$

Theorem 7 *Let u, $v \in \mathbb{C}^n$ and let w_i be the weights. Then $|v^T u| \leq \|u\|_{2,w}\|v\|_{2,1/w}$.*

Proof of Theorem 7. Let $\widehat{u_i} = \sqrt{w_i}u_i$, $\widehat{v_i} = \dfrac{v_i}{\sqrt{w_i}}$. Using the Cauchy-Schwartz inequality, we have:

$$|v^T u| = |\widehat{v}^T \widehat{u}| \leq \Big(\sum_i (\sqrt{w_i}u_i)^2\Big)^{1/2}\Big(\sum_i (\frac{v_i}{\sqrt{w_i}})^2\Big)^{1/2}.$$

Therefore, $|v^T u| \leq \|u\|_{2,w}\|v\|_{2,1/w}$. $\square$

Now that we have proven Hölder's inequality for the weighted ℓ^1, ℓ^2 and ℓ^∞-norms, we can follow the proof of Theorem 1 to extend Theorem 2 to the weighted ℓ^1 and ℓ^∞-norms. Theorem 7 would also yield an alternative proof of Theorem 2.

Theorem 8 *For τ, f, and $\tilde{f}$ as described in Theorem 1, and for $v = [1, \mathrm{sgn}(\tau_i), \ldots]$, where $v \in \mathbb{R}^\kappa$, κ is the dimension of f and* $\mathrm{sgn}(\tau_i) = \begin{cases} 1 & \text{for } \tau_i > 0 \\ 0 & \text{for } \tau_i = 0 \\ -1 & \text{for } \tau_i < 0 \end{cases}$, *with the distance measure being the weighted ℓ^∞-norm (with weights w_i) then*

$$\mathcal{N}^{[f]}_{\infty,w}(\alpha) = \frac{|f(\alpha)|}{\|\tau\|_{1,1/w}} \tag{4.1}$$

and

$$\vec{\tilde{f}} = \vec{f} - \frac{f(\alpha)}{\|\tau\|_{1,1/w}} D_w^{-1} v.$$

Proof of Theorem 8. From [38] and Theorem 6 we know that

$$|f(\alpha)| = |(\vec{\tilde{f}} - \vec{f})^T \tau| \leq \|\vec{\tilde{f}} - \vec{f}\|_{\infty,w}\|\tau\|_{1,1/w}.$$

Therefore, $\dfrac{|f(\alpha)|}{\|\tau\|_{1,1/w}} \leq \|\vec{f} - \vec{\tilde{f}}\|_{\infty,w}$. For all j construct $(\vec{\tilde{f}})_j$ such that

$$\frac{f(\alpha)}{\|\tau\|_{1,1/w}} = w_j(\vec{f} - \vec{\tilde{f}})_j.$$

From this we get that $w_j(\vec{\tilde{f}})_j = w_j(\vec{f})_j - \dfrac{f(\alpha)}{\|\tau\|_{1,1/w}}$. Therefore,

$$\vec{\tilde{f}} = \vec{f} - \frac{f(\alpha)}{\|\tau\|_{1,1/w}} D_w^{-1} v,$$

which yields equality in the above inequality. This gives

$$\tilde{f}(\alpha) = \tau^T \vec{f} - \frac{f(\alpha)}{\|\tau\|_{1,1/w}} \tau^T D_w^{-1} v = 0.$$

Thus the constraint that $\tilde{f}(\alpha) = 0$ is satisfied. $\square$

In the same way we obtain the following theorem.

Theorem 9 *For τ, f, $\tilde{f}$, and $\operatorname{sgn}(\tau_i)$ as described in Theorem 8 with weighted ℓ^1-norm and weights $w_i \geq 0$ we have*

$$\mathcal{N}_{1,w}^{[f]}(\alpha) = \frac{|f(\alpha)|}{\|\tau\|_{\infty,1/w}}$$

and

$$\vec{\tilde{f}}_i = \begin{cases} \vec{f}_i & \text{for } i \neq i_{max} \\ \vec{f}_i - \operatorname{sgn}(\tau_i) \dfrac{f(\alpha)}{\|\tau\|_{\infty,1/w}} \dfrac{1}{w_i} & \text{for } i = i_{max} \end{cases}$$

where $i_{max} = \underset{i}{\operatorname{argmax}} \left\{ \dfrac{|\tau_i|}{w_i} \right\}$.

Proof. From [38] and Theorem 6 we know that

$$|f(\alpha)| = |(\vec{\tilde{f}} - \vec{f})^T \tau| \leq \|\vec{\tilde{f}} - \vec{f}\|_{1,w} \|\tau\|_{\infty,1/w}.$$

Therefore,

$$\frac{|f(\alpha)|}{\|\tau\|_{\infty,1/w}} \leq \|\vec{f} - \vec{\tilde{f}}\|_{1,w}. \tag{4.2}$$

Let

$$\vec{\tilde{f}}_i = \begin{cases} \vec{f}_i & \text{for } i \neq i_{max} \\ \vec{f}_i - \operatorname{sgn}(\tau_i)\dfrac{f(\alpha)}{\|\tau\|_{\infty,1/w}}\dfrac{1}{w_i} & \text{for } i = i_{max} \end{cases}$$

where $i_{max} = \underset{i}{\operatorname{argmax}}\left\{\dfrac{|\tau_i|}{w_i}\right\}$. This gives equality in inequality 4.2. Substituting into the above inequality, the inequality becomes

$$\begin{aligned} \frac{|f(\alpha)|}{\|\tau\|_{\infty,1/w}} &\leq \|\vec{f} - \vec{\tilde{f}}\|_{1,w} \\ &= 0 + |w_{i_{max}}(\vec{f}_{i_{max}} - \vec{f}_{i_{max}} + \operatorname{sgn}(\tau_{i_{max}})\frac{f(\alpha)}{\|\tau\|_{\infty,1/w}}\frac{1}{w_{i_{max}}})| \\ &= \frac{|f(\alpha)|}{\|\tau\|_{\infty,1/w}}. \end{aligned}$$

We also need to check that $0 = \tilde{f}(\alpha)$.

$$0 = \tilde{f}(\alpha) = \tau^T \vec{f} - \operatorname{sgn}(\tau_{i_{max}})\frac{f(\alpha)}{\|\tau\|_{\infty,1/w}}\frac{1}{w_{i_{max}}}\tau_{i_{max}}.$$

Therefore, $\tilde{f}(\alpha) = 0$. $\square$

4.1.1 Coefficient Constraints with a Given Root

Next, we consider finding the nearest polynomial with equality constraints on the coefficients. For the ℓ^∞ and ℓ^1-norms, we cannot follow the same method as in Section 2.6. In order to compute the nearest polynomial with real coefficients using the ℓ^∞-norm as the distance measure, one reformulates the problem as a linear program [10]. For a given polynomial, f of dimension κ, we want to minimize

$$\max_{0 \leq k \leq \kappa} |\vec{f}_k - \vec{\tilde{f}}_k|$$

with constraints $\vec{\tilde{f}}^T\tau = 0$ and $\vec{\tilde{f}}_i = \vec{\tilde{f}}_j$. This is equivalent to solving the linear program

$$
\begin{array}{ll}
\text{Minimize} & w \\
\text{subject to} & \vec{\tilde{f}}^{\,T}\tau = 0, \\
& \vec{\tilde{f}}_i = \vec{\tilde{f}}_j, \\
& \vec{f}_k - \vec{\tilde{f}}_k \leq w, \quad k = 0, 1, \ldots \kappa \\
& -(\vec{f}_k - \vec{\tilde{f}}_k) \leq w, \quad k = 0, 1, \ldots \kappa
\end{array}
\tag{4.3}
$$

Similarly, for the ℓ^1 - norm we want to minimize

$$
\sum_{k=1}^{\kappa} |\vec{f}_k - \vec{\tilde{f}}_k|.
$$

We obtain the linear program:

$$
\begin{array}{ll}
\text{Minimize} & \displaystyle\sum_{k=1}^{\kappa}(d_k^+ + d_k^-) \\
\text{subject to} & \vec{\tilde{f}}^{\,T}\tau = 0, \\
& \vec{\tilde{f}}_i = \vec{\tilde{f}}_j, \\
& d_k^+ - d_k^- = \vec{f}_k - \vec{\tilde{f}}_k, \quad k = 0, 1, \ldots, \kappa \\
& d_i^+ \geq 0, \quad k = 0, 1, \ldots \kappa \\
& d_i^- \geq 0, \quad k = 0, 1, \ldots \kappa
\end{array}
\tag{4.4}
$$

For several linear constraints on the coefficients of $\tilde{f}$, we can embed the constraints into a matrix H and the problem becomes

$$
\left.
\begin{array}{r}
\mathcal{N}_{\infty,w}^{[f;H]}(\alpha) = \displaystyle\inf_{\tilde{f}\in\mathbb{R}[x_1,\ldots,x_n]} \|f - \tilde{f}\|_{\infty,w} \\
\text{s. t. } \tilde{f}(\alpha) = 0, H\vec{\tilde{f}} = p, \\
\deg(\tilde{f}) \leq \deg(f).
\end{array}
\right\}
\tag{4.5}
$$

Note that we can impose inequality constraints on the coefficients of $\tilde{f}$ in the same way.

Example 16 We want to find the nearest polynomial, $\tilde{f} = \tilde{f}_{2,0}x^2 + \tilde{f}_{0,2}y^2 + \tilde{f}_{1,1}xy + \tilde{f}_{1,0}x + \tilde{f}_{0,1}y + \tilde{f}_{0,0}$ in ℓ^∞-norm distance to the polynomial $f(x, y) = 4x^2 + 2y^2 + 5x + 3$ with coefficient constraint $\tilde{f}_{0,0} = \tilde{f}_{2,0}$ and given root $(\alpha, \beta) = (0, 0)$. The linear program in 4.3 becomes

$$
\begin{array}{ll}
\text{Minimize} & w \\
\text{subject to} & \vec{\tilde{f}}^{\,T}\tau = 0, \\
& \tilde{f}_{0,0} = \tilde{f}_{2,0}, \\
& f_k - \tilde{f}_k \leq w, \quad k = (0,0),(1,0),(0,1),(1,1),(2,0),(0,2) \\
& -(f_k - \tilde{f}_k) \leq w, \quad k = (0,0),(1,0),(0,1),(1,1),(2,0),(0,2)
\end{array}
$$

Running `LPSolve` in Maple 14 we obtain

$$\mathcal{N}_\infty^f(0,0) = 4, \quad \tilde{f} = 1.0718y^2 + 4xy + x + 4y.$$

The distance to the nearest polynomial in ℓ^∞-norm without constraints is $\mathcal{N}_\infty^f(0,0) = 3$. One can deform the constant coefficient to 0 to obtain this result, however the polynomial $\tilde{f}$ is not unique. Also, $\mathcal{N}_2^f(0,0) = 3$. If we consider a different root, $(\alpha, \beta) = (1,1)$, then

$$\mathcal{N}_\infty^f(1,1) = 2.5, \quad \tilde{f} = 1.5 - 2.5y + 2.5x - 2.5xy - 0.5y^2 + 1.5x^2$$

and if the linear constraint is removed then

$$\mathcal{N}_2^f(1,1) = 98/3, \text{ and } \tilde{f} = 2/3 - 7/3y + 8/3x - 7/3xy - 1/3y^2 + 5/3x^2$$

and

$$\mathcal{N}_\infty^f(1,1) = 2.333 \text{ and } \tilde{f} = 2/3 - 7/3y + 8/3x - 7/3xy - 1/3y^2 + 5/3x^2. \quad \square$$

Example 17 Next, we consider a polynomial where the nearest polynomial in ℓ^∞-norm is different from the nearest polynomial in ℓ^2- norm. We consider the bivariate polynomial $f(x,y) = 3x^2 + y^2 + 2$ with the constraint $\tilde{f}_{2,0} = \tilde{f}_{0,0}$ and root $(\alpha, \beta) = (1,1)$.

$$
\begin{array}{ll}
\text{Minimize} & w \\
\text{subject to} & \vec{\tilde{f}}^{\,T}\tau = 0, \\
& \vec{\tilde{f}}_{0,0} = \vec{\tilde{f}}_{2,0}, \\
& f_k - \tilde{f}_k \leq w, \quad k = (0,0),(1,0),(0,1),(1,1),(2,0),(0,2) \\
& -(f_k - \tilde{f}_k) \leq w, \quad k = (0,0),(1,0),(0,1),(1,1),(2,0),(0,2)
\end{array}
$$

Using Theorem 2, the distance to the nearest polynomial in ℓ^2-norm is

$$\mathcal{N}_2^f(1,1) = 26/4 = 6.5$$

and

$$\tilde{f}^{[\ell^2]} = 3/2x^2 - xy - x - y + 3/2.$$

Then, the distance in ℓ^∞-norm between f and $\tilde{f}$ is $\|\vec{f} - \vec{\tilde{f}}^{[\ell^2]}\|_\infty = 1.5$. However, if we calculate $N_\infty^f(1,1)$ by running `LPSolve` in Maple 14 on the above linear program, we obtain

$$\mathcal{N}_\infty^f(1,1) = 7/6$$

and

$$\tilde{f}^{[\ell^\infty]} = 11/6 - 7/6y - 7/6x - 7/6xy - 1/6y^2 + 11/6x^2$$

for the infinity norm. Notice that

$$\mathcal{N}_\infty^f(1,1) = 7/6 < 3/2 = \|\vec{f} - \vec{\tilde{f}}^{[\ell^2]}\|_\infty \text{ at } (1,1).$$

Similarly, for the root $(\alpha, \beta) = (0,0)$ and constraint $v_{2,0} = v_{0,0}$ we obtain:

$$\mathcal{N}_\infty^f(0,0) = 3 \text{ and } \tilde{f} = 3y + 3x + 3xy$$

and

$$\mathcal{N}_2^f(0,0) = 13 \text{ and } \tilde{f} = y^2.$$

If we want to find the nearest polynomial to f without any coefficient constraints, then we use Theorem 2 and obtain

$$N_2^f(0,0) = 4 \text{ and } \tilde{f} = 3x^2 + y^2.$$

This example will be continued on page 45. □

4.1.2 Nearest Polynomial with an Arbitrary Root

We have explored how to find the nearest polynomial to a given polynomial f with a given root α. Namely, we can use the formulas in Theorem 8 or Theorem 9 or we can use the linear programs described above (see 4.3 and 4.4). In the unconstrained case, we have a formula in Theorem 8 and 9, which we can minimize over the root α. In the constrained case, we do not have a formula, so we construct a grid and sample at every point on the grid. We consider several examples below.

4.1.3 Nearest Polynomial without Constraints and an Arbitrary Root

First, we consider finding $\rho_\infty(f)$ when there are no constraints on the coefficients of $\tilde{f}$. To solve this problem we consider minimizing $\dfrac{|f(\alpha,\beta)|}{\|\tau\|_1}$ (see Theorem 8) in the four quadrants: $\alpha \geq 0$, $\beta \geq 0$; $\alpha \geq 0$, $\beta \leq 0$; $\alpha \leq 0$, $\beta \geq 0$; $\alpha \leq 0$, $\beta \leq 0$. We make the substitution $\alpha = u^2$ and $\beta = v^2$, which increases the degree of the problem, but allows us to drop the absolute value of the denominator. Since there are no absolute values in the expression in the Euclidean norm case, this substitution is not necessary.

Example 18 We consider the polynomial in Example 17 but without coefficient constraints on $\tilde{f}$. The polynomial $f = 3x^2 + y^2 + 2$ has the property that $f \geq 0$ for all $x, y \in \mathbb{R}$. The value of the polynomial is invariant depending on which quadrant we are in. Therefore, we only need to consider the case $\alpha \geq 0$ and $\beta \geq 0$ and therefore do not need to make the substitution $\alpha = u^2$ and $\beta = v^2$. Using `Minimize` in Maple 14, for Example 17 we obtain

$$\rho_\infty(f) = 0.7639320225,\ \alpha = 0.61803383, \text{ and } \beta = 2.618033705.$$

and

$$\tilde{f} = 1.236 - 0.764\,x - 0.764\,y - 0.764\,xy + 2.236\,x^2 + 0.236\,y^2$$

This example will be continued on page 43. □

Example 19 We consider the polynomial in Example 16 but without a constraint on the coefficients of $\tilde{f}$. Since the polynomial $f = 4x^2 + 2y^2 + 5x + 3$ is symmetric in y, we compute the minimum for the two quadrants, $\alpha \leq 0$ and $\beta \geq 0$, and $\alpha \geq 0$ and $\beta \geq 0$. For $\alpha \geq 0$ and $\beta \geq 0$ we obtain $\rho_\infty(f) = 1.56949912595693952$ and for $\alpha \leq 0$ and $\beta \geq 0$ we get $\rho_\infty(f) = 0.563757167$. Taking the minimum of the two solutions we achieve the value $\rho_\infty(f) = 0.563757167$. We obtain this solution using `Minimize` in Maple 14 and substituting $\alpha \leq 0$ and $\beta \geq 0$ into our polynomial f. The solution is

$$\rho_\infty(f) = 0.563757167,\ \alpha = -0.8391797248 \text{ and } \beta = 0.3609594177.$$

and

$$\tilde{f} = 2.436 + 5.564\,x - 0.564\,y + 0.564\,xy + 3.436\,x^2 + 1.436\,y^2. \quad \square$$

Example 20 Given $f = x^2 + y^2 + 1$, what is the nearest polynomial? In Theorem 8 we have a formula for finding the nearest polynomial in ℓ^∞-norm for a given root. We now follow the procedure described above and minimize Equation (4.1). The solution is

$$\rho_\infty(f) = 0.5 \text{ and } (\alpha, \beta) = (1, 1).$$

This gives us that

$$\tilde{f} = 0.5 - 0.5x - 0.5y - 0.5xy + 0.5x^2 + 0.5y^2. \quad \square$$

Example 21 Consider the polynomial $f = (x-y)^2 + (x+2y)^2 + 1 = 2x^2 + 2xy + 5y^2 + 1$. We want to find the nearest polynomial $\tilde{f}$ in ℓ^∞-norm. We use Equation (4.1) and compute the minimum of $\dfrac{|f(\alpha,\beta)|}{\|\tau\|_1}$ in the 2 cases:

Case 1: $\alpha \geq 0$, $\beta \geq 0$

$\min \mathcal{N}_\infty^f(\alpha, \beta) = 0.837722339831620699$

Case 2: and $\alpha \leq 0$, $\beta \geq 0$

$\min \mathcal{N}_\infty^f(\alpha, \beta) = 0.713254935043394989.$

So $\rho_\infty(f) = 0.713254935043394989$ and $(\alpha, \beta) = (-0.5475652821, 0.2564812120)$

$$\tilde{f} = 0.2867 + 0.7133\,x - 0.7133\,y + 2.7133\,x\,y + 1.2867\,x^2 + 4.2867\,y^2. \quad \square$$

Example 22 Given the polynomial $f = 1 - 2xy + x^2y^2 + x^2 = (1 - xy)^2 + x^2$ (see Example 5), find the nearest polynomial with root $(\alpha, \beta) = (\epsilon, 1/\epsilon)$. We have that

$$N_\infty^f(\alpha, \beta) = \frac{|(1 - \alpha\beta)^2 + \alpha^2|}{\sum_{i+j=0}^{4} |\alpha^i \beta^j|}.$$

Then

$$N_\infty^f(\epsilon, 1/\epsilon) = \frac{\epsilon^6}{3\,\epsilon^4 + 2\,\epsilon^3 + 2\,\epsilon^2 + \epsilon + 1 + 2\,\epsilon^5 + 2\,\epsilon^6 + \epsilon^7 + \epsilon^8}$$

and the nearest polynomial to f with $(\alpha, \beta) = (\epsilon, 1/\epsilon)$ as a root is

$$\begin{aligned}\tilde{f}(x,y) =& \frac{3\,\epsilon^4 + 2\,\epsilon^3 + 2\,\epsilon^2 + \epsilon + 1 + 2\,\epsilon^5 + \epsilon^6 + \epsilon^7 + \epsilon^8 - \epsilon^6 y - \epsilon^6 y^2 - \epsilon^6 y^3 - \epsilon^6 y^4 - \epsilon^6 x}{\delta} \\ &+ \frac{(-6\,\epsilon^4 - 4\,\epsilon^3 - 4\,\epsilon^2 - 2\,\epsilon - 2 - 4\,\epsilon^5 - 5\,\epsilon^6 - 2\,\epsilon^7 - 2\,\epsilon^8)\,xy - \epsilon^6 xy^2 - \epsilon^6 xy^3}{\delta} \\ &+ \frac{(3\,\epsilon^4 + 2\,\epsilon^3 + 2\,\epsilon^2 + \epsilon + 1 + 2\,\epsilon^5 + \epsilon^6 + \epsilon^7 + \epsilon^8)\,x^2 - \epsilon^6 x^2 y}{\delta} \\ &+ \frac{(3\,\epsilon^4 + 2\,\epsilon^3 + 2\,\epsilon^2 + \epsilon + 1 + 2\,\epsilon^5 + \epsilon^6 + \epsilon^7 + \epsilon^8)\,x^2y^2 - \epsilon^6 x^3 - \epsilon^6 x^3 y - \epsilon^6 x^4}{\delta}\end{aligned}$$

$$\begin{aligned}\tilde{f}(\epsilon, 1/\epsilon) =& \frac{-2\,\epsilon^4 - \epsilon^3 + (3\,\epsilon^4 + 2\,\epsilon^3 + 2\,\epsilon^2 + \epsilon + 1 + 2\,\epsilon^5 + \epsilon^6 + \epsilon^7 + \epsilon^8)\,\epsilon^2}{\delta} \\ &+ \frac{-\epsilon^2 - 2\,\epsilon^5 - 3\,\epsilon^6 - 2\,\epsilon^7 - \epsilon^8 - \epsilon^9 - \epsilon^{10}}{\delta},\end{aligned}$$

where

$$\delta = 3\,\epsilon^4 + 2\,\epsilon^3 + 2\,\epsilon^2 + \epsilon + 1 + 2\,\epsilon^5 + 2\,\epsilon^6 + \epsilon^7 + \epsilon^8.$$

For $\epsilon = 1/10$,

$$\mathcal{N}^f_\infty(1/10, 10) = 100/112232211. \quad \square$$

Remark 10 The distance to the nearest polynomial in ℓ^∞ -norm is zero if and only if the distance to the nearest polynomial in ℓ^2 -norm is zero. i.e.

$$0 = \rho_\infty(f) = \inf_\alpha \frac{|f(\alpha)|}{\|\tau\|_1} \Leftrightarrow \inf_\alpha \frac{f(\alpha)^2}{\|\tau\|_2} = \rho_2(f) = 0.$$

This can be seen using the fact that $\|v\|_\infty \leq \|v\|_2 \leq \sqrt{n}\|v\|_\infty$. Note that in the same way, $\rho_1(f) = 0$ if and only if $\rho_2(f) = 0$.

Example 23 In this example we consider the nearest polynomial $\tilde{f}^{[\infty]}$, $\tilde{f}^{[2]}$ and $\tilde{f}^{[1]}$ to $f = 3x^2 + y^2 + 2$ in ℓ^∞, ℓ^2, and ℓ^1 -norms respectively. We also compute $\|f - \tilde{f}^{[\cdot]}\|_\infty$, $\|f - \tilde{f}^{[\cdot]}\|_2^2$, and $\|f - \tilde{f}^{[\cdot]}\|_1$ for each $\tilde{f}$.

Table 4.1: Norm Comparison Table

$f =$	$2+3x^2+y^2$	$\|f-\tilde{f}^{[\cdot]}\|_\infty$	$\|f-\tilde{f}^{[\cdot]}\|_2^2$	$\|f-\tilde{f}^{[\cdot]}\|_1$
$\tilde{f}^{[\ell^\infty]} =$	$1.236 - 0.764\,x - 0.764\,y$ $-0.764\,xy + 2.236\,x^2 + 0.236\,y^2$	0.764	3.502	4.584
$\tilde{f}^{[\ell^2]} =$	$1.999 - 0.001\,y + 3.0\,x^2$ $-0.00000099\,y^2$	1.000001	1.000003	1.001
$\tilde{f}^{[\ell^1]} =$	$2+3x^2-0.000002y^2$	1.000002	1.000004	1.000002

The last line would have all entries of 1.000002 if it was $\|f-\tilde{f}^{[\cdot]}\|_2$ instead of the squared distance $\|f-\tilde{f}^{[\cdot]}\|_2^2$. It is already known that $\|f-\tilde{f}^{[\cdot]}\|_\infty \leq \|f-\tilde{f}^{[\cdot]}\|_2 \leq \|f-\tilde{f}^{[\cdot]}\|_1$. We also know that the diagonal entries are the minimum for each column. What is interesting is comparing the other two entries in each column.

Notice that the optimum for $\tilde{f}^{[\ell^1]}$ is $\tilde{f}^{[\ell^1]} = 2 + 3\,x^2 - \epsilon\,y^2 = 0$. Rewriting the equation, we have $3\,x^2 = \epsilon\,y^2 - 2$. This gives us that $y = \sqrt{\dfrac{2}{\epsilon}}$. Thus, $\|f-\tilde{f}^{[\ell^1]}\|_1 = 1$, $\|f-\tilde{f}^{[\ell^1]}\|_\infty = 1$, and there is no real root. This is another example of when the infimum is not attainable.

Table 4.2: Norm Comparison Table With Corresponding Roots

$f =$	$2+3x^2+y^2$	Root $= (\alpha, \beta)$
$\tilde{f}^{[\ell^\infty]} =$	$1.236 - 0.764\,x - 0.764\,y$ $-0.764\,xy + 2.236\,x^2 + 0.236\,y^2$	$(0.618, 2.618)$
$\tilde{f}^{[\ell^2]} =$	$1.999 - 0.001\,y + 3.0\,x^2$ $-0.00000099\,y^2$	$(0, 1000)$
$\tilde{f}^{[\ell^1]} =$	$2+3x^2-0.000002y^2$	$(0, 1000)$

In Table 4.2 the nearest polynomial to $f = 2+3x^2+y^2$ is given for ℓ^∞, ℓ^2 and ℓ^1 -norms and the corresponding root. $\square$

4.1.4 Coefficient Constraints with an Arbitrary Root

We have seen how to find the nearest polynomial with an arbitrary root using the coefficient vector infinity norm distance measure if there are no constraints. Now we consider the case of when there are coefficient constraints on $\tilde{f}$. We cannot follow the same method as we did for the Euclidean norm. When our distance measure is the Euclidean norm then the distance measure is differentiable. For the ℓ^1 and ℓ^∞-norm this is not the case, since we have absolute values in the formulas we are trying to minimize. Therefore, we cannot use the KKT conditions as with the Euclidean norm. Unlike in the unconstrained case, we do not know a norm expression to minimize for the constrained case. Therefore, we cannot minimize over the four quadrants for the root like we did in the unconstrained case. Instead, we lay a fine grid for the root and run the linear program described in 4.3 at each grid point. In the following examples we have polynomials of two variables, α and β. We use the grid $-100 \leq \alpha \leq 100$ and $-100 \leq \beta \leq 100$ in steps of 0.1 and obtain $(\alpha_{\min,1}, \beta_{\min,1})$, the root that corresponds to $\hat{\rho}_\infty(f)$ on the grid. We want an extra decimal place in the solution for the root, so we search in steps of 0.01 in our refined grid. We refine the grid around the solution and search in a 4×4 grid around $(\alpha_{\min,1}, \beta_{\min,1})$ in steps of 0.01 to obtain $(\alpha_{\min,2}, \beta_{\min,2})$. The number of times the solution has been refined is given as the second coordinate of the subscript of the root.

Example 24 Example 17 continued. For $f(x) = 3x^2 + y^2 + 2$ with the constraint $\tilde{f}_{2,0} = \tilde{f}_{0,0}$ we compute

$$\begin{array}{ll} \text{Minimize} & w \\ \text{subject to} & \vec{\tilde{f}}^T\tau = 0, \\ & \tilde{f}_{0,0} = \tilde{f}_{2,0}, \\ & f_k - \tilde{f}_k \leq w, \quad k = (0,0),(1,0),(0,1),(1,1),(2,0),(0,2) \\ & -(f_k - \tilde{f}_k) \leq w, \quad k = (0,0),(1,0),(0,1),(1,1),(2,0),(0,2) \end{array} \tag{4.6}$$

at the roots $-100 \leq \alpha \leq 100$ and $-100 \leq \beta \leq 100$ in increments of 0.1.

We obtain the minimum $\alpha_{\min,1} = -1.0$, $\beta_{\min,1} = -4.4$, $\hat{\rho}_{\infty,1}(f) = 0.813863928112965$. The maximum value on the grid is $\max \hat{\rho}_{\infty,1}(f) = 3$ at $(\alpha_{\max,1}, \beta_{\max,1}) = (0,0)$. The average value is $\text{avg}(\rho_\infty(f)) = 1.481924505$. The average time for one iteration is 0.02 seconds and the total time to run all of the linear programs was 9612.940 seconds[1].

Next, we can further refine the grid around the solution $(\alpha_{\min,1}, \beta_{\min,1}) = (-1.0. -4.4)$ in order to obtain an extra decimal place in accuracy in the solution. We compute the minimum on the grid $-3 \leq \alpha \leq 1$ and $-6 \leq \beta \leq -2$ in steps of 0.01. Our refined root $(\alpha_{\min,2}, \beta_{\min,2})$ is $(-1.00, -4.37)$ which gives us

$$\hat{\rho}_{\infty,2}(f) = 0.813859369781010.$$

The distance to the nearest polynomial with a real root has improved by

$$\hat{\rho}_{\infty,1}(f) - \hat{\rho}_{\infty,2}(f) = 0.0000045583.$$

So by obtaining an extra decimal place in accuracy in the root, we get an improvement in $\hat{\rho}_\infty(f)$ of 0.00000456. Running the linear program 4.6 for $(\alpha, \beta) = (-1, -4.37)$ we obtain

$$\begin{aligned} \tilde{f}(x,y) = \;& 2.186140630 + 0.813859369781010\, x + 0.813859369781010\, y \\ & -0.813859369781009\, xy + 2.18614063021899\, x^2 + 0.186140630218991\, y^2. \end{aligned}$$

The maximum value on our refined grid was

$$\max \hat{\rho}_\infty(f) = 1.36$$

at the root $(\alpha_{\max,2}, \beta_{\max,2}) = (-3, -2)$.

The average value for $\rho_\infty(f)$ on the grid $-3 \leq \alpha \leq 1$ and $-6 \leq \beta \leq -2$ was

$$\text{avg}(\rho_\infty(f)) = 0.8881288462.$$

We don't use the bisection method when computing the minimum on the refined grid because we don't know how the values for $\rho_\infty(f)$ will change. The average time for one linear program to run was about 0.217 seconds and the total time to compute all of the linear programs for the refined grid was 346.971 seconds. $\square$

4.2 Systems

As in Chapter 3, we now consider the problem of finding the nearest system to an inconsistent one. Applying Theorem 8 to each individual $f^{[j]}$ yields the following.

Theorem 10 *Let $f^{[1]}, \ldots, f^{[k]} \in \mathbb{R}[x_1, \ldots, x_n]$, with $d_i = \deg(f^{[i]})$ and w_j the weight vector for the j-th polynomial. The distance to the nearest system with a common root $\alpha \in \mathbb{R}^n$ in ℓ^∞-norm is*

$$\mathcal{N}_{\infty, w_{j_{max}}}^{\{f^{[1]}, \ldots, f^{[k]}\}}(\alpha) = \max_{1 \le j \le k} \frac{|f^{[j]}(\alpha)|}{\|\tau\|_{1,1/w_j}}, \tag{4.7}$$

where $j_{max} = \underset{j}{\operatorname{argmax}} \left\{ \frac{|f^{[j]}(\alpha)|}{\|\tau\|_{1,1/w_j}} \right\}$

Proof. For each $f^{[j]}$, $1 \le j \le k$ we have that

$$\mathcal{N}_{\infty, w_j}^{[f^{[j]}]}(\alpha) = \frac{|f^{[j]}(\alpha)|}{\|\tau\|_{1,1/w_j}}.$$

We want to find what the maximum change is in a single coefficient. So we take the maximum over all polynomials. Therefore, $\mathcal{N}_{\infty, w_{j_{max}}}^{\{f^{[1]}, \ldots, f^{[k]}\}}(\alpha) = \max_{1 \le j \le k} \frac{|f^{[j]}(\alpha)|}{\|\tau\|_{1,1/w_j}}$. $\square$

Remark 11 We cannot have a system of equations that has an equation with weight vector $w = [\infty, \ldots, \infty]$. For this reason, we cannot formulate root constraints as a system of equations. For example, for the root (α, β) if we want to impose the constraint $g(\alpha, \beta) = 2\alpha^2 - 3\beta^2 + 1 = 0$, then we would also have the corresponding weight vector $w = [\infty, \ldots, \infty]$ for g. Now we are dividing by zero in Equation (4.7) in Theorem 10. For the same reason we cannot formulate root constraints as a system of equations in the ℓ^1 or ℓ^p -norms.

We can impose linear constraints on the coefficients of a system of polynomials. These constraints can be coupled, where the constraint applies to the coefficients of multiple polynomials, or uncoupled, where the constraint is on the coefficients of just one of the polynomials. The distance to the nearest system with k equations, a common root α, and linear coefficient constraints of the form $\mathcal{H}[\vec{\tilde{f}}^{[1]}, \ldots, \vec{\tilde{f}}^{[k]}]^T = b$ is defined as

$$
\left.
\begin{aligned}
\mathcal{N}_{\infty}^{[f^{[1]},\ldots,f^{[k]};\mathcal{H}]}(\alpha) = \max_{1\le j\le k} & \frac{|f^{[j]}(\alpha)|}{\|\tau\|_1} \\
\text{s. t. } & \tilde{f}^{[i]}(\alpha) = 0,\ i = 1,\ldots,k \\
& \mathcal{H}[\vec{\tilde{f}}^{\,[1]},\ldots,\vec{\tilde{f}}^{\,[k]}]^T = b \\
& f^{[i]} \in \mathbb{R}[x_1,\ldots,x_n],\ i = 1,\ldots,k \\
& \deg(\tilde{f}^{[i]}) \le \deg(f^{[i]}),\ i = 1,\ldots,k.
\end{aligned}
\right\} \tag{4.8}
$$

Chapter 5

Exact Sums-of-Squares Distance Certificates

5.1 History

In Hilbert's famous 1900 lecture, Hilbert posed the problem: For any $f \in \mathbb{R}[x_1, \ldots, x_n]$, is it true that $f \succeq 0$ on $\mathbb{R}^n$ implies that f is a sum-of-squares of rational functions?

In 1927, Emil Artin solved this problem. His theorem states:

Theorem 11 *Given a polynomial $f \in \mathbb{Q}[X_1, \ldots, X_n]$ then the following statements are equivalent:*

1. $f \succeq 0$
2. *There exist* $u_i, v_j \in \mathbb{Q}[X_1, \ldots, X_n]$: $f(X_1, \ldots, X_n) = \dfrac{\sum_{i=1}^{m} u_i^2}{\sum_{j=1}^{m} v_j^2}$
3. *There exist rational* $W^{[1]} \succeq 0, W^{[2]} \succeq 0$ *such that* $f = \dfrac{m_d^T\, W^{[1]}\, m_d}{m_e^T\, W^{[2]}\, m_e}$,

   with $m_d(X_1, \ldots, X_n)$ *and* $m_e(X_1, \ldots, X_n)$ *term vectors*

5.2 Motzkin

Artin proved that if $f \succeq 0$, then f can be written as a sum-of-squares of rational functions. Is it true that $f \succeq 0$ on $\mathbb{R}^n$ implies that f is a sum-of-squares? In 1967, T. S. Motzkin discovered a polynomial that cannot be written as a sum-of-squares (without introducing a denominator as Artin proved). Recall that the arithmetic mean (average of n numbers) is $\dfrac{|x_1| + |x_2| + \cdots + |x_n|}{n}$ and the geometric mean is $\sqrt[n]{|x_1| \cdot |x_2| \cdots |x_n|}$. From the AGM inequality,

$$\text{the geometric mean} \leq \text{arithmetic mean.}$$

Multiplying this by 3 and rewriting the inequality we have

$$3(\text{arithmetic mean}) - 3(\text{geometric mean}) \geq 0.$$

For the monomials x^4y^2, x^2y^4, z^6 this equation becomes the Motzkin polynomial

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

We will show that a modified Motzkin polynomial $\widetilde{M} = M(x, y, z) + z^6 = x^4y^2 + x^2y^4 + 2z^6 - 3x^2y^2z^2$ cannot be written as a sum of squares in $\mathbb{R}[x, y, z]$. Later in Example 27 we will prove, using a sum-of-squares, certificate that $\widetilde{M}(x, y, 1) \succeq 0$.

Theorem 12 *For $\widetilde{M} = M(x, y, z) + z^6 = x^4y^2 + x^2y^4 + 2z^6 - 3x^2y^2z^2$ then $\widetilde{M} \geq 0$ and $\widetilde{M}$ is not a sum of squares in $\mathbb{R}[x, y, z]$.*

Proof of Theorem 12. Clearly $\widetilde{M} \geq 0$ since $M(x, y, z) \geq 0$. Following the proof from [32], if $\widetilde{M}$ were a sum-of-squares, then $\widetilde{M} = \sum_k h_k^2(x, y, z)$ for some polynomials $h_k \in \mathbb{R}[x, y, z]$. Writing out $\widetilde{M}$ as a ternary sextic using all potential monomials, we have:

$$0x^6 + 0x^5 + 1x^4y^2 + 0x^3y^3 + 1x^2y^4 + 0x^5 + 0y^6$$

$$+0x^5z + 0x^4yz + 0x^3y^3z + 0x^2y^3z + 0xy^4z + 0y^5z$$

$$+0x^4z^2 + 0x^3yz^2 - 3x^2y^2z^2 + 0xy^3z^2 + 0y^4z^2$$

$$+0x^3z^3 + 0x^2yz^3 + 0xy^2z^3$$

$$+0x^2z^4 + 0xyz^4 + 0y^2z^4$$

$$+0xz^5 + 0yz^5$$

$$+2z^6.$$

Using the same geometric scheme we write out $h_k(x,y,z)$.

$$A_kx^3 + B_kx^2y + C_kxy^2 + D_ky^3$$

$$+E_kx^2z + F_kxyz + G_ky^2z$$

$$+H_kxz^2 + I_kyz^2$$

$$+J_kz^3.$$

Since the coefficient of x^6 in $\widetilde{M}$ is 0, the corresponding coefficient in $\sum_k h_k^2$, $\sum_k A_k^2$, also equals 0. Thus $A_k = 0$ for all k. Next, we look at the coefficient of x^4z^2 in $\sum_k h_k^2$, we have that $\sum_k E_k^2 + 2A_kH_k$. Since $A_k = 0$ and the coefficient of x^4z^2 in $\widetilde{M}$ is 0, it follows that $E_k = 0$ for all k. Continuing down the xz edge, we see that the coefficient of x^2z^4 in $\widetilde{M}$ is 0, so $\sum_k 2E_kJ_k + H_k^2 = 0$. Since $E_k = 0$, then $H_k = 0$. A similar argument applied to the coefficients of y^6, y^4z^2 and y^2z^4, shows that $D_k = G_k = I_k = 0$. We now have that $x^4y^2 + x^2y^4 + 2z^6 - 3x^2y^2z^2 = \sum_k (B_kx^2y + C_kxy^2 + F_kxyz + J_kz^3)^2$. Looking at the coefficient of the $x^2y^2z^2$ term in $\widetilde{M}$, we have $-3 = \sum_k F_k^2$. Contradiction. $\square$

Remark 12 Notice that the polynomial $\widetilde{M}$ in Theorem 12 is just the homogenization of $M(x,y) + 1 = 2 - 3x^2y^2 + x^2y^4 + x^4y^2$.

Remark 13 Notice that the polynomial $\widetilde{M} + g(x)$, where $\deg(g(x)) < \deg(\widetilde{M})$ is also not a sum-of-squares since the highest degree has to be a sum-of-squares in order for the polynomial to be a sum-of-squares.

Remark 14 The polynomial $f(x_1, \ldots, x_n) \in \mathbb{R}[x_1, \ldots, x_n]$ is homogeneous and a sum-of-squares if and only if $f(x_1, \ldots, x_{n-1}, 1)$ is a sum-of-squares.

5.3 Lower Bound Certificates

The minimization of the rational function $\mathcal{N}_{2,w}^{[f]} = \dfrac{f(\alpha)^2}{g(\alpha)}$ where $g = \tau^T D_w^{-1} \tau$ defined in Equation (2.2) can be reformulated as maximizing r such that $f(\alpha)^2 - rg(\alpha)$ is non-negative. We compute a lower bound of $\inf\limits_{\alpha \in \mathbb{R}^n} \mathcal{N}_{2,w}^{[f]}(\alpha)$ by solving the SOS program [15, 28, 19, 20]:

$$\left.\begin{array}{rl} r^* := \sup\limits_{r \in \mathbb{R}, W} & r \\ \text{s. t.} & f(\mathbf{X})^2 - rg(\mathbf{X}) = m_d(\mathbf{X})^T \, W \, m_d(\mathbf{X}) \\ & W \succeq 0, \; W^T = W \end{array}\right\} \tag{5.1}$$

where $m_d(\mathbf{X})$ is the column vector of all terms in $\mathbf{X}_1, \ldots, \mathbf{X}_n$ up to degree d. The dimension of $m_d(\mathbf{X})$ is $\binom{n+d}{d}$.

The SOS program (5.1) can be solved efficiently by algorithms in GloptiPoly [9], SOSTOOLS [31], YALMIP [26] and SeDuMi [39]. One can use GloptiPoly as described in [9] to extract the solutions α which achieve the global minimum. However, since we are running fixed precision SDP solvers in Matlab, we can only obtain a numerical positive semidefinite matrix W and a floating point number r^* which satisfy approximately

$$f(\mathbf{X})^2 - r^* g(\mathbf{X}) \approx m_d(\mathbf{X})^T \cdot W \cdot m_d(\mathbf{X}), \; W \succapprox 0. \tag{5.2}$$

So r^* is a lower bound of $\inf_{\alpha \in \mathbb{R}^n} \mathcal{N}_{2,w}^{[f]}(\alpha)$ approximately. Since it is a fixed precision SDP solver it is subject to numerical error.

We convert the imprecise SOS to an exact SOS certificate with exact rational scalars and polynomials. The lower bound $\tilde{r} \le r^*$ is certified if $\tilde{r}$ and $\widetilde{W}$ hold the following conditions exactly:

$$f(\mathbf{X})^2 - \tilde{r} g(\mathbf{X}) = m_d(\mathbf{X})^T \cdot \widetilde{W} \cdot m_d(\mathbf{X}), \; \widetilde{W} \succeq 0. \tag{5.3}$$

We can use Artin's theorem of sum-of-squares and semidefinite programming (see, e.g., [29, 19, 20]) to certify the computed minimum. This is different from a validated numerical algorithm, since we do not insist on the accuracy of the SDP solution.

The method is described in Figure 5.1, where the affine linear hyperplane, $\mathcal{L}$, is

$$\mathcal{L} = \{A \mid A^T = A, f(\mathbf{X}) = m(\mathbf{X})^T \cdot A \cdot m(\mathbf{X}) = \text{ SOS } .$$

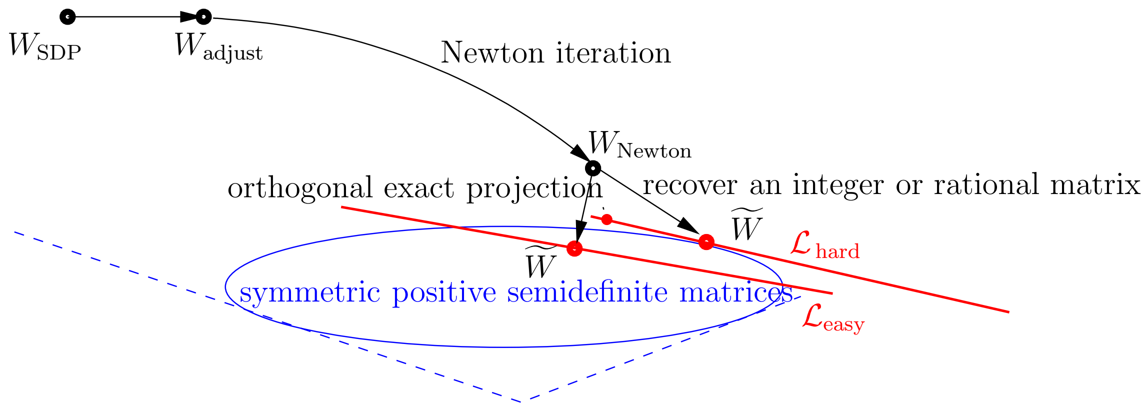


Figure 5.1: Rationalizing a Sum-of-Squares

The numeric semidefinite matrix obtained from `solvesos` in Matlab satisfying Equation (5.2) is represented by W_{SDP}. We compute the rank of W_{SDP}. If W_{SDP} is rank deficient then this may be caused by extra monomials in the SOS decomposition, so we delete the rows and columns of W_{SDP} with small elements (remove the monomials which should not appear in the SOS). More reasons for why W_{SDP} may be rank deficient are discussed in [20]. According to [19] sometimes W_{SDP} is too coarse to allow successful projection to $\widetilde{W}$ using Maple 14's exact linear algebra. Therefore, the W_{SDP} is refined by rank-preserving Gauss-Newton iteration to W_{NEWTON}. After we have refined W_{SDP} to W_{NEWTON} such that

$$\|f(\mathbf{X}) - r^* g(\mathbf{X}) - m_d(\mathbf{X})^T \cdot W_{NEWTON} \cdot m_d(\mathbf{X})\| < \tau,$$

where τ is the given tolerance, we approximate r^* by a nearby rational number $\tilde{r} \leq r^*$. Next, if W_{NEWTON} is of full rank (easy case), then we convert W_{NEWTON} to a rational matrix and perform an exact orthogonal projection, using least squares, to the rational matrix $\widetilde{W}$ on the hyperplane $\mathcal{L}_{\text{easy}}$. If W_{NEWTON} is not of full rank or is near to a singular matrix (hard case), then we use rational vector recovery to compute $\widetilde{W}$. If $\widetilde{W}$ is positive semidefinite we have our solution to Equation (5.3). If $\widetilde{W}$ is not positive semidefinite, we can try a smaller value for $\tilde{r}$ or increase the precision of the Newton iterations and then repeat the previous computation. [19, 20]

We have done so for

$$\min_{(\alpha,\beta)} \mathcal{N}_2 = 1$$

of Equation (2.12) of Example 8 and the rational lower bound

$$\widetilde{\mathcal{N}}_2 = 64597306998078108/100000000000000000$$

of the real algebraic optimum of Equation (3.3) of Example 14.

This section has been adapted from [13].

5.4 Examples

Example 25 ([41]) Given a polynomial

$$f = x^2y^2 + x^2 - xy + y^4 - y^2 + 1 = (xy - 1/2)^2 + (y^2 - 1/2)^2 + x^2 + 1/2$$

decide the minimum perturbation such that the perturbed polynomial has a real root.

If we allow dense perturbations, after running `solvesos` in Matlab, we get the lower bound

$$\widetilde{\mathcal{N}}_2 = 2.453484553428391600 \times 10^{-15}.$$

This is caused by the assumption that we can perturb f by any monomial term with degree bounded by 4. As illustrated in Example 6, we see that f has a radius of positive semidefiniteness that is 0. Hence, it would be more interesting to consider a weighted norm for this polynomial. For instance, if we only allow terms which appear in f to be perturbed, then the lower bound computed by `solvesos` in Matlab is

$$\widetilde{\mathcal{N}}_2 = 0.246916019336920590\text{0}.$$

After applying the certification algorithm in [19, 20], we obtain the certified lower bound

$$\widetilde{\mathcal{N}}_2 = 24691601933692029/100000000000000000.$$

This means that f is positive since $f(0,0) = 1 > 0$. $\square$

Example 26 Consider the Robinson polynomial

$$g(x,y) = 1 + x^6 - x^4 - y^2 - x^4y^2 - x^2 - y^4 - x^2y^4 + y^6 + 3\,x^2y^2.$$

According to [3] $g(x,y) + \beta(x^6 + y^6 + z^6)$ is a sum-of-squares if and only if $\beta \geq 1/8$. Therefore, we consider $f(x,y) = g(x,y) + \frac{1}{10}(x^6 + y^6 + z^6)$. We have that

$$f(x,y) = \frac{11}{10} + \frac{11}{10}\,x^6 - x^4 - y^2 - x^4y^2 - x^2 - y^4 - x^2y^4 + \frac{11}{10}\,y^6 + 3\,x^2y^2.$$

We allow only perturbations in the coefficients of f. The lower bound computed by `solvesos` in Matlab is

$$\mathcal{N}_2 = 0.0089999999999993343200.$$

After applying the certification algorithm in [19, 20], we obtain the certified lower bound

$$\widetilde{\mathcal{N}}_2 = 899999999999334317/100000000000000000000.$$

We have that

$$f(x,y)^2 - \widetilde{\mathcal{N}}_2(1 + x^4 + x^4y^4 + x^4y^8 + y^4 + x^8 + y^8 + x^8y^4 + x^{12} + y^{12}) = \mathrm{SOS}.$$

The sum-of-squares is a sum of 28 polynomial squares. $\square$

Example 27 (see [27]) Consider the polynomial

$$f(x,y) = 2 - 3x^2y^2 + x^2y^4 + x^4y^2 = M(x,y) + 1.$$

Notice that f is the result of adding one to the Motzkin polynomial. It is well-known that f is positive semidefinite but not an SOS, as seen in [27]. In fact $f \geq 1$ for all $x, y \in \mathbb{R}$. First, we consider using a dense perturbation to obtain a lower bound for $\mathcal{N}_2$. We use Matlab to compute the approximate lower bound of $\mathcal{N}_2$ and obtain 0 as the minimum, which is easily proven by considering $f(x, y) - \epsilon x^5$. Hence, we consider a weighted norm. We use infinite weights on the terms that have zero coefficients in f. Thus, we only allow the terms which appear in f to be perturbed (sparse deformation). The lower bound computed by `solvesos` in Matlab is

$$\widetilde{\mathcal{N}}_2 = 0.1285480262594671800.$$

After applying the certification algorithm in [19, 20], we obtain the certified lower bound

$$\widetilde{\mathcal{N}}_2 = 12854802625942833/100000000000000000.$$

We have computed an exact rational certificate (as in Equation (5.3))

$$f(x, y)^2 - 12854802625942833/100000000000000000 \times (1 + x^4y^8 + x^8y^4 + x^4y^4) = \text{SOS}.$$

This is a sum-of-squares of 10 polynomial squares. This means that the non-zero coefficients of f need to be perturbed (by at least 0.128 in ℓ^2-norm squared) for f to have a real root. Since $f(0, 0) = 2$, we have proven that $f(x, y) > 0$ for all real x, y via a polynomial sum-of-squares certificate. $\square$

Example 28 Similar to Example 27 we consider the polynomial

$$f(x, y) = x^4y^2 + x^2y^4 + 1 - 2x^2y^2.$$

Notice that $f(x, y)$ is similar to the inhomogeneous Motzkin polynomial except that the coefficient of the x^2y^2 term is -2 instead of -3. We consider a sparse perturbation to obtain the lower bound for $\mathcal{N}_2$. We follow the algorithm described above and compute an approximate lower bound for $\mathcal{N}_2$ using `solvesos` in Matlab and obtain 0.25 as the minimum. After applying the certification algorithm in [19, 20] we obtain the certified lower bound

$$\widetilde{\mathcal{N}}_2 = 12499999999997733/50000000000000000.$$

We have computed an exact rational certificate that

$$f(x, y)^2 - 12499999999997733/50000000000000000 \times (1 + y^8x^4 + y^4x^8 + y^4x^4)$$

is a sum-of-squares (10 polynomial squares). The sums-of-squares is

$$
\begin{aligned}
&f(x,y)^2 - \tfrac{1249999999999997733}{5000000000000000000}\,(1+y^8x^4+y^4x^8+y^4x^4) = \\
&\tfrac{50000000000000000}{193572598687033141}\,\big(\tfrac{193572598687033141}{50000000000000000} - \tfrac{202647902552528050093}{300000000000000000000}\,x^2y^2 \\
&+\tfrac{257054793517541488391}{1500000000000000000000}\,x^2y^4 + \tfrac{515134358033083171231}{3000000000000000000000}\,x^4y^2\big)^2 \\
&+\tfrac{12195073717283087883000000000000000000000}{53828677492523777636951019568616964202 9}\,\big(\tfrac{538286774925237776369510195686169642029}{12195073717283087883000000000000000000000}\,x^2y^2 \\
&-\tfrac{2890754357101173812702054086343054942 77}{17421533881832982690000000000000000000000}\,x^2y^4 - \tfrac{5639476364153098101522909155515838719 37}{34843067763665965380000000000000000000000}\,x^4y^2\big)^2 \\
&+\tfrac{10092877029848208306928316169115680788043750000000000000000000}{509757872747132365690124695205465293642370941129281099380729} \\
&\big(\tfrac{509757872747132365690124695205465293642370941129281099380729}{10092877029848208306928316169115680788043750000000000000000000}\,x^2y^4 \\
&-\tfrac{6796771648240746880323956135224534279220959579173735167595 99}{13457169373130944409237754892154241050725000000000000000000000}\,x^4y^2\big)^2 \\
&+\tfrac{26929168608411728647084429984452390547254640043196819177172399 0335383}{122341889459311767765629926849311670474169025871027463851374960000000000000000000000}\,x^8y^4 \\
&+\tfrac{75000000000000000000}{42945206482458511609}\,\big(\tfrac{42945206482458511609}{75000000000000000000}\,xy^2 - \tfrac{18603262817549652237}{43750000000000000000}\,x^3y^2\big)^2 \\
&+\tfrac{94671836162476413095314881 7}{631294535292140120652300000000000000000000}\,x^6y^4 \\
&+\tfrac{150000000000000000000}{84865641966916828769}\,\big(\tfrac{84865641966916828769}{150000000000000000000}\,x^2y - \tfrac{58820174232832129311}{140000000000000000000}\,x^2y^3\big)^2 \\
&+\tfrac{37357609989079608434186766607}{249504987382735476580860000000000000000000}\,x^4y^6 \\
&+\tfrac{15000000000000000000}{22647902552528050 93}\,\big(\tfrac{2264790255252805093}{15000000000000000000}\,xy - \tfrac{58284098243642068 2041}{7000000000000000000000}\,x^3y^3\big)^2 \\
&+\tfrac{59418914101946391613553086871}{665848335044324697342000000000000000000000}\,x^6y^6
\end{aligned}
$$

This means that the non-zero coefficients of f need to be perturbed (by at least 0.25 in ℓ^2-norm squared) in order for f to have a real root. $\square$

This section has been adapted from [13].

5.5 Future Work

We conjecture that such polynomial sums-of-squares always exist. More precisely, if for a real polynomial $f(x_1,\ldots,x_n)$ there exists a vector w of positive and infinite weights (excluding an infinite weight for the constant coefficient) such that $\rho_{2,w}(f) > 0$ then in (5.1) $r^* > 0$. We have seen that $\rho_{2,w}(f)$ easily is no larger than 0, provided f has a projective root at infinity, and the condition $\rho_{2,w}(f) > 0$ makes f and w quite special.

Remark 15 Note that if $\rho_2(f) = 0$ then we have $f(\alpha,\beta)^2 - 0 =$ SOS .

Remark 16 We have considered the case $f(\alpha) - \hat{\rho}t(\alpha)^2$ is a sum-of-squares, where $t(\alpha)$

is a single term of the τ-vector. Consider the polynomial

$$f = (x^2 + y^2 - 2xy + 4)^2 = ((x - y)^2 + 4)^2.$$

We subtract ϵ^2, where $\epsilon > 0$ and obtain the sum-of-squares

$$(x^2 + y^2 - 2xy + 4)^2 - \epsilon^2 = (x - y)^4 + 8(x - y)^2 + 16 - \epsilon^2.$$

Notice that f is a sum-of-squares, so this is a trivial example. Can this be proven for polynomials which are not sums-of-squares?

Remark 17 Consider Example 5 with $f(x, y) = 1 - 2xy + x^2y^2 + x^2$. We can choose any weighted norm for τ where τ is reduced to one monomial (a single term support), call it τ_i and then $\rho_2(f) = 0$. Take the root $(\sqrt{\epsilon}, 1/\sqrt{\epsilon})$ and the polynomial $\tilde{f}(x, y) = f(x, y) - \frac{\epsilon}{\tau_i}\tau_i$. Then we have that

$$\tilde{f}(\sqrt{\epsilon}, 1/\sqrt{\epsilon}) = f(\sqrt{\epsilon}, 1/\sqrt{\epsilon}) - \frac{\epsilon}{\tau_i}\tau_i = 0.$$

The distance between f and $\tilde{f}$ is ϵ. So as $\epsilon \to 0$ then $\|\vec{f} - \vec{\tilde{f}}\|_2 \to 0$. Notice that this is true for any polynomial g where $\inf_{\alpha} g(\alpha) = 0$ can get arbitrarily close to 0 as long as the root does not cause τ_i to be 0. If the root was $(0, 0)$ instead of $(\sqrt{\epsilon}, 1/\sqrt{\epsilon})$, then we would be dividing by 0 in $\tilde{f}$. Therefore the root must not cause τ_i to be 0.

Chapter 6

Approximate GCD

6.1 Background

The problem of finding the nearest polynomial with a real root to a given polynomial, can be extended to systems and reformulated as finding the nearest system with a non-trivial greatest common divisor. In this chapter we consider finding the approximate GCD of several univariate polynomials.

The problem can be formulated as a minimization problem as follows [23]:
The input for the problem is s univariate polynomials, namely $f^{[1]}, \dots, f^{[s]} \in \mathbb{C}[x]$ with $\deg(f^{[1]}) = m_1, \dots, \deg(f^{[s]}) = m_s$. For a positive integer k with $k \leq \min(m_1, \dots, m_s)$, we wish to compute $\Delta f^{[1]}, \dots, \Delta f^{[s]} \in \mathbb{C}[x]$ such that $\deg(\Delta f^{[1]}) \leq m_1, \dots \deg(\Delta f^{[s]}) \leq m_s$, $\deg(GCD(f^{[1]}+\Delta f^{[1]}, f^{[s]}+\Delta f^{[s]})) \geq k$ and such that $\|\Delta f^{[1]}\|_2^2 + \dots \|\Delta f^{[s]}\|_2^2$ is minimized.

In [40], the authors use Bezout matrices to compute an approximate GCD of several univariate polynomials. We extend some of their results to Sylvester matrices, namely we generalize the univariate resultant to several polynomials. In [7] they prove for several univariate polynomials the degree of the GCD in terms of the rank of the companion matrix.

6.2 Univariate multi-polynomial Sylvester matrices

We introduce the convolution matrix $\mathrm{C}^{[l,m]}(f)$, which for the coefficient vector $\vec{u}$ of a polynomial $u(x)$ of degree $\leq l$ produces the coefficient vector of $(u(x) \cdot f(x)) \bmod x^{m+1}$

as $\mathrm{C}^{[l,m]}(f) \cdot \vec{u}$. If $m = \deg(f) + l$, we can write $\mathrm{C}^{[l]}(f)$ [21, 22]. For instance,

$$\overrightarrow{(a_2x^2 + a_1x + a_0) \cdot (b_2x^2 + b_1x + b_0) \bmod x^6} =$$

$$\mathrm{C}^{[2,5]}(a_2x^2 + a_1x + a_0) \cdot \begin{bmatrix} b_2 \\ b_1 \\ b_0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ a_2 & 0 & 0 \\ a_1 & a_2 & 0 \\ a_0 & a_1 & a_2 \\ 0 & a_0 & a_1 \\ 0 & 0 & a_0 \end{bmatrix} \cdot \begin{bmatrix} b_2 \\ b_1 \\ b_0 \end{bmatrix}.$$

The matrix has dimensions $(m + 1) \times (l + 1)$.

Theorem 13 *Let $f_1, \ldots, f_s \in \mathsf{K}[x] \setminus \mathsf{K}$, where K is a field, $s \geq 2$, and let $d_i = \deg(f_i)$ for $1 \leq i \leq s$. Suppose $d_1 \geq d_2 \geq \cdots \geq d_s$. Let $\delta = d_1 + d_s - 1$ and let*

$$S = \left[\mathrm{C}^{[d_s-1,\delta]}(f_1) \quad \mathrm{C}^{[d_s-1,\delta]}(f_2) \quad \ldots \quad \mathrm{C}^{[d_s-1,\delta]}(f_{s-1}) \quad \mathrm{C}^{[d_1-1,\delta]}(f_s)\right] \in \mathsf{K}^{(\delta+1)\times(d_1+(s-1)d_s)}$$

Then $\operatorname{rank}(S) = \delta + 1 - \deg(\mathrm{GCD}_{1\leq i\leq s}(f_i))$.

Note that for $s = 2$, the matrix S is the classical Sylvester matrix for two polynomials. Theorem 13 is stated in [35, Section 5.2] for the somewhat larger matrix

$$S = \left[\mathrm{C}^{[k-d_1,k]}(f_1) \quad \mathrm{C}^{[k-d_2,k]}(f_2) \quad \ldots \quad \mathrm{C}^{[k-d_s,k]}(f_s)\right] \in \mathsf{K}^{(k+1)\times\nu}$$

where $k = d_1 + d_s$ and $\nu = s(d_1 + d_s + 1) - d_1 - \cdots - d_s$.

Proof. Let $g = \mathrm{GCD}_{1\leq i\leq s}(f_i)$. Then there exist polynomials $u_i \in \mathsf{K}[x]$ such that $\sum_{i=1}^{s} u_i f_i = g$ and therefore for any $h(x) = x^j$ with $j \geq 0$ we have $\sum_{i=1}^{s}(hu_i)f_i = hg$. We can replace

$$(hu_i)f_i = (\underbrace{hu_i - q_i f_s}_{\hat{u}_i})f_i + (q_i f_i)f_s \text{ where } \deg(\hat{u}_i) < d_s \text{ for all } 1 \leq i \leq s-1,$$

where q_i are polynomial quotients. Suppose now that $\deg(h) = j \leq \delta - \deg(g)$. We have

$$(\underbrace{hu_s + q_1 f_1 + \cdots + q_{s-1} f_{s-1}}_{\hat{u}_s})f_s = hg - \hat{u}_1 f_1 - \hat{u}_2 f_2 - \cdots - \hat{u}_{s-1} f_{s-1},$$

where the degree of the right polynomial is $\leq \delta$. Therefore $\deg(\hat{u}_s) \leq \delta - d_s$. We consider the ansatz

$$\left.\begin{array}{l} v_1(x)f_1(x) + \cdots + v_s(x)f_s(x) = x^j g(x), \\ \deg(v_i) \leq d_s - 1 \text{ for } 1 \leq i \leq s-1,\ \deg(v_s) \leq d_1 - 1,\ j \leq \delta - \deg(g), \end{array}\right\} \tag{6.1}$$

where the coefficients of all v_i are the unknowns and the rows correspond to the coefficients of the powers x^0, $x, \ldots, x^\delta$. The coefficient matrix for the left side of (6.1) is S, and by the above considerations, all right side coefficient vectors are in the range of S. Furthermore, the coefficient vectors of $x^j g(x)$ are linearly independent, hence the dimension of the column space of S is at least $\delta - \deg(g) + 1$. But no other linearly independent vector can be in the column space, since all corresponding polynomials are multiples of g, so the dimension cannot be more. □

Theorem 14 *The* $\deg(\mathrm{GCD}_{1\leq i\leq s}(f_i)) \geq k$ *if and only if* $\delta + 2 - k$ *rows are linearly dependent in* S.

Proof. Let S be as defined in Theorem 13. Let $\deg(\mathrm{GCD}_{1\leq i\leq s}(f_i)) = \gamma \geq k$. Let E be the reduced column echelon form of S. From Theorem 13, we know that $\mathrm{rank}(S) = \delta + 1 - \gamma$. Therefore, there are exactly $\delta + 1 - \gamma$ pivots in E. First, we want to show that all of the pivot elements of E will be in the first $\delta + 1 - \gamma$ rows. Suppose this is not the case, that there exists a pivot element in the $\delta + 1 + j - \gamma$ row, where j is a positive integer. Since the last nonzero column of E gives the coefficients of $\mathrm{GCD}_{1\leq i\leq s}(f_i)$, then $\deg(\mathrm{GCD}_{1\leq i\leq s}(f_i)) < \gamma$, which is a contradiction. Therefore, all of the pivot elements of E are in the first $\delta + 1 - \gamma$ rows. Since there are $\delta + 1$ rows in E with the last γ rows containing no pivots, then $\delta + 2 - \gamma$ rows are linearly dependent. Since $k \leq \gamma$, then $\delta + 2 - \gamma \leq \delta + 2 - k$. Therefore $\delta + 2 - k$ rows are linearly dependent in S.

To see the converse, let $\delta + 2 - k$ rows be linearly dependent in S. Since the last nonzero column of E gives the coefficients of $\mathrm{GCD}_{1\leq i\leq s}(f_i)$, then $\deg(\mathrm{GCD}_{1\leq i\leq s}(f_i))$ must be greater than the remaining number of rows. So,

$$\deg\Big(\underset{1\leq i\leq s}{\mathrm{GCD}}(f_i)\Big) > \delta + 1 - (\delta + 2 - k) > k - 1.$$

Therefore, $\deg(\mathrm{GCD}_{1\leq i\leq s}(f_i)) \geq k$. □

Chapter 7

Conclusion

We present a new approach to the old problem of least squares fitting. When finding the nearest consistent system to an inconsistent one (or the nearest system with a non-trivial GCD), we use optimization techniques, such as semidefinite programming and linear programming in symbolic computation, as opposed to exact Gröbner basis methods (see Example 14). How beneficial semidefinite programming and linear programming in symbolic computation would be for our problem was not initially realized. Semidefinite programming allows for more efficient computations.

While conducting research, we were further surprised that the formulas in [38] allow the use of linear and non-linear optimization. This is very fortunate, since it allows us to optimize the distance to the nearest polynomial with a real root in infinity or one norm with given coefficient constraints. We are now able to optimize a non-differentiable function with constraints.

When I first began research in hybrid symbolic-numeric computation, I did not foresee how fertile the use of optimization in symbolic computation would be in both problem applications and in the creation of new problems. In searching for a research topic for my PhD thesis, it was very important to me that my PhD research area have many real-world applications. We did not expect all of the applications that currently exist. For example, there is an application in biochemical systems [12]. Furthermore, this research area is fertile in the creation of new problems. New problems or questions are constantly arising and being studied around the world. As mentioned in Section 5.5, I hope to continue contributing to the area of numerical optimization in symbolic computation in the future.

REFERENCES

[1] Aubry, P., Rouillier, F., and Safey El Din, M. Real solving for positive dimensional systems. *J. Symbolic Comput. 34*, 6 (Dec. 2002), 543–560. URL: http://www-spiral.lip6.fr/~safey/Articles/RR-3992.ps.gz.

[2] Becker, E., Powers, V., and Wörmann, T. Deciding positivity of real polynomials. In *Real Algebraic Geometry and Ordered Structures*, C. N. Delzell and J. J. Madden, Eds., vol. 253 of *Contemporary Math.* AMS, 2000, pp. 251–272.

[3] Choi, M.D., Lam, T. Y., and Reznick, B. Even symmetric sextics, Math. Z., *195* (1987), pp. 559–580 (MR 88j.11019).

[4] Corless, R. M., Gianni, P. M., Trager, B. M., and Watt, S. M. The singular value decomposition for polynomial systems. In *Proc. 1995 Internat. Symp. Symbolic Algebraic Comput. ISSAC'95* (New York, N. Y., 1995), A. H. M. Levelt, Ed., ACM Press, pp. 96–103.

[5] Corless, R. M., Kaltofen, E., and Watt, S. M. Hybrid methods. In *Computer Algebra Handbook*, J. Grabmeier, E. Kaltofen, and V. Weispfenning, Eds. Springer Verlag, Heidelberg, Germany, 2003, Section 2.12.3, pp. 112–125. URL: http://www.math.ncsu.edu/~kaltofen/bibliography/01/symnum.pdf.

[6] Ecker, J. G., and Kupferschmid, M. *Introduction to Operations Research.* John Wiley and Sons, 1988. 509 pp.

[7] Gonzlez-Vega, Laureano An elementary proof of Barnett's theorem about the greatest common divisor of several univariate polynomials. In *Linear Algebra and It's Appl.*, vol. 247, 1996, pp. 185202.

[8] Heintz, J. and Schnorr, C. P. Testing polynomials which are easy to compute (extended abstract). In *STOC*, ACM, 1980, pp. 262-272.

[9] Henrion, D., and Lasserre, J.-B. Detecting global optimality and extracting solutions in GloptiPoly. In *Positive polynomials in control*, D. Henrion and A. Garulli, Eds., vol. 312 of *Lecture Notes on Control and Information Sciences.* Springer Verlag, Heidelberg, Germany, 2005, pp. 293–310. URL: http://homepages.laas.fr/henrion/Papers/extract.pdf.

[10] Hitz, M. A., and Kaltofen, E. Efficient algorithms for computing the nearest polynomial with constrained roots. In *Proc. 1998 Internat. Symp. Symbolic Algebraic Comput. (ISSAC'98)* (New York, N. Y., 1998), O. Gloor, Ed., ACM Press, pp. 236–243. URL: http://www.math.ncsu.edu/~kaltofen/bibliography/98/HiKa98.pdf.

[11] Hitz, M. A., Kaltofen, E., and Lakshman Y. N. Efficient algorithms for computing the nearest polynomial with a real root and related problems. In *Proc. 1999 Internat. Symp. Symbolic Algebraic Comput. (ISSAC'99)* (New York, N. Y., 1999), S. Dooley, Ed., ACM Press, pp. 205–212. URL: http://www.math.ncsu.edu/~kaltofen/bibliography/99/HKL99.pdf.

[12] Horimoto, K., Regensburger, G., Rosenkranz, M., and Yoshida, H., Eds. Algebraic Biology, Third International Conference, AB 2008, Castle of Hagenberg, Austria, July 31-Aug 2,2008, Proceedings. vol. 5147. of *Lecture Notes in Computer Science*, Springer, 2008.

[13] Hutton, S., Kaltofen, E., and Zhi, L. Computing the radius of positive semidefiniteness of a multivariate real polynomial via a dual of Seidenberg's method. In *Proc. 2010 Internat. Symp. Symbolic Algebraic Comput. ISSAC 2010* (New York, N. Y., 2010), S. Watt, ed., ACM Press, pp. 227-234.

[14] Jacobson, N. *Basic Algebra I.* W. H. Freeman & Co., San Francisco, 1974.

[15] Jibetean, D., and de Klerk, E. Global optimization of rational functions: a semidefinite programming approach. *Math. Program. 106*, 1 (2006), 93–109.

[16] Kaltofen, E. Computer algebra algorithms. In J. F. Traub, editor, *Annual Review in Computer Science*, volume 2, pages 91-118. Annual Reviews Inc., Palo Alto, California, 1987.

[17] Kaltofen, E. Computing the irreducible real factors and components of an algebraic curve. *Applic. Algebra Engin. Commun. Comput. 1*, 2 (1990), 135–148. URL: http://www.math.ncsu.edu/~kaltofen/bibliography/90/Ka90_aaecc.pdf.

[18] Kaltofen, E. The "Seven Dwarfs" of symbolic computation, April 2010. Manuscript prepared for the final report of the 1998-2008 Austrian research project SFB F013 "Numerical and Symbolic Scientific Computing," Peter Paule, director.

[19] Kaltofen, E., Li, B., Yang, Z., and Zhi, L. Exact certification of global optimality of approximate factorizations via rationalizing sums-of-squares with floating point scalars. In *ISSAC 2008* (New York, N. Y., 2008), D. Jeffrey, Ed., ACM Press, pp. 155–163. URL: http://www.math.ncsu.edu/~kaltofen/bibliography/08/KLYZ08.pdf.

[20] Kaltofen, E. L., Li, B., Yang, Z., and Zhi, L. Exact certification in global polynomial optimization via sums-of-squares of rational functions with rational coefficients, Jan. 2009. Accepted for publication in J. Symbolic Comput. URL: http://www.math.ncsu.edu/~kaltofen/bibliography/09/KLYZ09.pdf.

[21] Kaltofen, E., May, J., Yang, Z., and Zhi, L.. Approximate factorization of multivariate polynomials using singular value decomposition. Manuscript, 22 pages. Submitted, January 2006a.

[22] Kaltofen, E., Yang, Z., and Zhi, L. Approximate greatest common divisors of several polynomials with linearly constrained coefficients and singular polynomials. In *ISSAC MMVI Proc. 2006 Internat. Symp. Symbolic Algebraic Comput.* (New York, N. Y., 2006), J.-G. Dumas, Ed., ACM Press, pp. 169–176. URL: http://www.math.ncsu.edu/~kaltofen/bibliography/06/KYZ06.pdf.

[23] Kaltofen, E., Yang, Z., Zhi, L. Structured low rank approximation of a Sylvester matrix. Manuscript, 15 pages, Oct. 2005. Preliminary version in SNC 2005 Proceedings, D. Wang and L. Zhi, eds., pp. 188201, distributed at the International Workshop on Symbolic-Numeric Computation in Xian, China, July 1921, 2005.

[24] Karmarkar, N. K., and Lakshman Y. N. On approximate GCDs of univariate polynomials. *J. Symbolic Comput. 26*, 6 (1998), 653–666. Special issue on Symbolic Numeric Algebra for Polynomials S. M. Watt and H. J. Stetter, editors.

[25] Kharitonov, V. L. Asymptotic stability of an equilibrium of a family of systems of linear differential equations. *Differential Equations 14* (1979), 1483–1485.

[26] Löfberg, J. YALMIP : A toolbox for modeling and optimization in MATLAB. In *Proc. IEEE CCA/ISIC/CACSD Conf.* (Taipei, Taiwan, 2004). URL: http://control.ee.ethz.ch/~joloef/yalmip.php.

[27] Marshall, M. *Positive Polynomials and Sums of Squares.* American Math. Soc., 2008. 187 pp.

[28] Nie, J., Demmel, J., and Gu, M. Global minimization of rational functions and the nearest GCDs. *J. of Global Optimization 40*, 4 (2008), 697–718.

[29] Peyrl, H., and Parrilo, P. A. Computing sum of squares decompositions with rational coefficients. *Theoretical Comput. Sci. 409* (2008), 269–281.

[30] Pope, S., and Szanto, A. Nearest multivariate system with given root multiplicities. *Journal of Symbolic Computation 44* (2009), 606–625. URL: http://www4.ncsu.edu/~aszanto/Pope-Szanto-Preprint.pdf.

[31] Prajna, S., Papachristodoulou, A., Seiler, P., and Parrilo, P. A. SOSTOOLS: Sum of squares optimization toolbox for MATLAB. Available from http://www.cds.caltech.edu/sostools and http://www.mit.edu/~parrilo/sostools, 2004.

[32] Reznick, B. Some concrete aspects of Hilbert's 17th problem, Seminaire de Structures Algbriques Ordonnes, (F. Delon, M.A. Dickmann, D. Gondard eds), Publ. quipe de Logique, Univ. Paris VII, Jan. 1996; revised version in *Real Algebraic Geometry and Ordered Structures*, (C. N. Delzell, J.J. Madden eds.) Cont. Math., 253 (2000), 251-272 (MR 2001i:11042).

[33] Rezvani, N., and Corless, R. M. The nearest singular polynomial with a given zero, revisited. *SIGSAM Bulletin 39*, 3 (Sept. 2005), 73–79.

[34] Ruatta, O., Sciabica, M., and Szanto, A. Over-constrained Weierstrass iteration and the nearest consistent system. Rapport de Recherche 5215, INRIA, www.inria.fr, June 2004. Accepted in Journal of Complexity. URL: ftp://ftp-sop.inria.fr/galaad/oruatta/RR-5215.pdf.

[35] David Rupprecht An algorithm for computing certified approximate GCD of n univariate polynomials. *J. Pure Applied Algebra* 139 (1999), 255–284.

[36] Safey El Din, M. *Résolution réelle des systèmes polynomiaux en dimension positive.* Thèse de doctorat, Univ. Paris VI (Univ. Pierre et Marie Curie), Paris, France, 2001. URL: http://www-spiral.lip6.fr/~safey/these_safey.ps.gz.

[37] Seidenberg, A. A new decision method for elementary algebra. *Annals Math. 60* (1954), 365–374.

[38] Stetter, H. J. The nearest polynomial with a given zero, and similar problems. *SIGSAM Bulletin 33*, 4 (Dec. 1999), 2–4.

[39] Sturm, J. F. Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. *Optimization Methods and Software 11/12* (1999), 625–653.

[40] Sun, D. and Zhi, L. Structured Low Rank Approximation of a Bezout Matrix *Mathematics in Computer Science*, vol. 1, no. 2, (2007), pp. 427-437.

[41] Zeng, G. Nonstandard decision methods for the solvability of real polynomial equations. *Science in China Series-A Mathematics 42*, 12 (1999), 1251–1261.

[42] Zhi, L., Noda, M.-T., Kai, H., and Wu, W. Hybrid method for computing the nearest singular polynomials. *Japan J. Industrial and Applied Math. 21*, 2 (June 2004), 149–162. URL: http://www.mmrc.iss.ac.cn/~lzhi/Publications/znkwjjiam04.pdf.

[43] Zhi, L., and Wu, W. Nearest singular polynomial. *J. Symbolic Comput. 26*, 6 (1998), 667–675. Special issue on Symbolic Numeric Algebra for Polynomials S. M. Watt and H. J. Stetter, editors.