HONORS PROJECT

YIXUAN ZHOU AND MAREIKE DRESSLER

ABSTRACT. Short description

1. Introduction

Polynomial with real coefficients is a very powerful tool in terms of expressing problems. It has a vast range of applications, in both theory side and engineering side. Deciding the nonegativity of a multivariate polynomial, natually, become a cenctral prblem for many optimizations and feasibility problems. However, it is known that deciding the nonegativity of an arbitrary multivariate polynomial is a NP-hard problem when the degree of the polynomial is greater than or equals to 2. Therefore, we restrict our focus to decide whether the polynomial can be expressed in *sum of squares (SOS)* form.

Definition 1.1. Sum of Suqares (SOS)

Let $\mathbb{R}[x]_{n,d}$ denotes the set of polynomials with n real variables with degree at most d. A polynomial $p(x) \in \mathbb{R}[x]_{n,2d}$, is a *sum of squares* if there exists $q_1(x), ..., q_n(x) \in \mathbb{R}[x]_{n,d}$ such that

$$p(x) = \sum_{i=1}^{n} q_i^2(x) \tag{1.1}$$

It is clear that if a polynomial $p(x)$ is a SOS, then it is nonegative, Thus, SOS is a (proper) subset of the set of nonegative polynomials.

We are interested in SOS beacuse given a multivariate polynomial, the decision problem of whether it can be decomposed into SOS is a polynomial time problem, by using *Semidefinite Programing (SDP).*

Theorem 1.2. *A multivariate polynomial $p(x) \in \mathbb{R}[x]_{n,2d}$ is a sum of squares if and only if there exists $\mathcal{Q} \in \mathcal{S}^{\binom{n+d}{d}}$ such that*

$$p(x) = [x]_d^T \mathcal{Q} [x]_d \quad \mathcal{Q} \succcurlyeq 0 \tag{1.2}$$

Where $[x]_d$ denotes the vector of monomials with degree at most d.

[BPT13]
[Lau09]

As a consequence of the above theorem, whether a multivariate polynomial is SOS can be determined by a *Semidefinite Programming.* Notice that, the size of the semidefinite

Key words and phrases. some keywords go here.

matrix, $\binom{n+d}{d}$, when fixing d, grows in polynomial time with respect to n, and when fixing n, it grows in polynomial time with respect to d. Thus, though imposed some limitation, we have "redueced" a NP-Problem to a P problem.

In the above theorem, there is nothing special with monomials other than being a basis of the vector space $\mathbb{R}[x]_{n,2d}$. Instead of using $[x]_d$ to proceed the calculation, we can choose any basis of the vector space of polynomials $\mathbb{R}[x]_{n,d}$. For example, we can choose Lagrange basis, Chebychev basis, ... And in the most cases, the monomial basis is not the right choice due to its instability of monomials bases.

In this paper, we will use python program to analyze the computational effeciency and numerical stability of the usage of different bases. In particular, the analysis will be focusing on the condition number of the linear system generated when solving the semidefinite program, that is generated by $p(x) = [x]_d^T \mathcal{Q}[x]_d$. We will use the *condition number* of the matrix in the linear system to identify how stable the it is under noises that are introduced in practical problems. We will further attempt to identify some most effective bases for some particular type of polynomials, and will try to justify the reasons.

2. Preliminaries

Here background definitions etc will be put. You can do it in several subsections (like notation, bla, blabla)

2.1. Notations and Definiations.

Because we are going to use a lot of tools from lienar algebra, we first introduce some key definiations that we will use in this thesis.

Definition 2.1. Given a matrix $A \in \mathbb{R}^{n\times n}$, we say it is *symmetric* if $A^T = A$.

Theorem 2.2. *Spectrum Theorem*
Given a symmetric matrix$A \in \mathbb{R}^{n\times n}$, it can be diagnoalized as

$$A = P^{-1}DP \tag{2.1}$$

where D is the diagnoal matrix with all real values, and P is an orthonomal matrix.
In other word, A has all real eigenvalues, and their correspdoning eigenvectors form an orthonomal basis of $\mathbb{R}^n$.

Definition 2.3. Given a matrix $A \in \mathbb{R}^{n\times n}$, it is *positive semidifinate* if A is symmetric and

$$x^T A x \geq 0 \quad \forall x \in \mathbb{R}^n \tag{2.2}$$

And we denote it as $A \succcurlyeq 0$.

Proposition 2.4. *A matrix is* positive semidifinate *if and only if all its eigenvalues are greater than or equals to 0*

Proof. If a matrix A has eigenvalue $\lambda < 0$, then if x is the correspdoning eigenvector, we have $x^T Ax = \lambda x^T x < 0$. On the other hand, if A has all positive eignevalues, because A being a symmetric matrix, its eigenvalue form a basis. Thus for any $x \in \mathbb{R}^n$, we have $x = \sum_{i=1}^n c_i v_i$ where v_i are the eigenvectors of A, that are also orthnormal to each other. Hence, $x^T Ax = \sum_{i=1}^n \lambda_i$. Since all $\lambda_i \geq 0$, we have that $x^T Ax \geq 0$.

□

Definition 2.5. Given a matrix $A \in \mathbb{R}^{n\times m}$, the *pesudo-inverse*, which is also knows as the *Moore-Penrose* inverse of A, is the matrix A^+ satisfying:

- $AA^+A = A$
- $A^+AA^+ = A^+$
- $(AA^+)^T = AA^+$
- $(A^+A)^T = AA^+$

Every matrix has its pesudo-inverse, and when $A \in \mathbb{R}^{n\times m}$ is *full rank*, that is $rank(A) = min\{n, m\}$, A can be expressed in simple algebric form.

In particular, when A has linearly independent columns, A^+ can be computed as

$$A^+ = (A^T A)^{-1} A^T \tag{2.3}$$

In this case, the pesudo-inverse is called the *left inverse* since $A^+A = I$.

And when A has linearly independent rows, A^+ can be be computed as

$$A^+ = A^T (AA^T)^{-1} \tag{2.4}$$

In this case, the pesudo-inverse is called the *right inverse* since $AA^+ = I$.

Definition 2.6. Given a matrix $A \in \mathbb{R}^{n\times m}$, the condition number of A, $\kappa(A)$ is defined as

$$\kappa(A) = ||A|| \cdot ||A^+|| \tag{2.5}$$

for any norm imposed on A, for instance *Frobenius norm*.

Remark 2.7. *The condition number measure how stable the system is. It can be alternatively defined as*

$$\kappa(A) = \frac{\sigma_{max}(A)}{\sigma_{min}(A)} \tag{2.6}$$

where the σ denotes the singular values of A.

Thus, it can be understand as how stale our system is. Intuiatively, when the condition number is large, some error in the input along the max direction of the singular value, our result would largly flacuate because the error, maginified by the singular value, will dominate the input that is along the direction of the minimum singular value. Therefore, the smaller the condition number is, the more stable our system is under flcuations caused by noises. The rigirous explaination of the condition numebr can be found in [CK07]

Now with the tools from linear algebra, we are ready to proceed to the realm of real coefficients polynomials.

Definition 2.8. Let $\mathbb{R}[x]_{n,d}$ denotes the set of real coefficient polynmials with n variables and at most d degree.

Definition 2.9. Let $P_{n,2d}$ denotes the set of nonegative polynomials with n variables and at most $2d$ degree, that is

$$P_{n,2d} = \{p \in \mathbb{R}[x]_{n,2d} : p(x) \geq 0, \forall x \in \mathbb{R}^d\} \tag{2.7}$$

Remark 2.10. *When trying to determine the nonnegativity of a polynomial, there is no reason to consider the set $P_{n,d}$ when d is odd, since if the degree of a polynomial is odd, then it will always be nonegative at some point.*

Definition 2.11. Let $\Sigma_{n,2d}$ denotes the set of polynomials with n variables and at most d degree that are *Sum of Suqares*, that is

$$\Sigma_{n,2d} = \{p \in \mathbb{R}[x]_{n,2d} : \exists\ q_1(x), ..., q_k(x) \in \mathbb{R}[x]_{n,d}\ s.t.\ p(x) = \sum_{i=1}^{k} q_i^2(x)\} \tag{2.8}$$

Remark 2.12. *It is easy to check that $P_{n,2d}$ form an vector space. There are a lot of choices of basis, and the most canonical one is the monomial basis.*

Example 2.13. $P(2, 2)$ is a vector space, then $\mathcal{B}_{2,2} = x^2, xy, y^2, x, y, 1$ is a basis of the vector space.

Remark 2.14. *Given $\mathcal{B}_{n,d}$ be a basis of $P_{n,d}$. If we list the elements of $\mathcal{B}_{n,d}$ in a column vector, write as b, then bb^T form a matrix whose upper triangle entries can be collected to form a basis of $P_{n,2d}$.*

Example 2.15. Let $\mathcal{B}_{2,1} = x, y, 1$, be a basis of $P_{2,1}$, then

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \begin{bmatrix} x & y & 1 \end{bmatrix} = \begin{bmatrix} x^2 & xy & x \\ xy & y^2 & y \\ x & y & 1 \end{bmatrix} \tag{2.9}$$

Then the upper triangle of the result form a basis of $P_{2,2}$.

Then the question that we are interested in is given any polynomial $p(x) \in P_{n,2d}$, decide whether it is in $\Sigma_{n,2d}$. To help solving this decision problem, the following proposition will become very helpful.

Proposition 2.16. *Given $p(x) \in P_{n,2d}$, if $p(x) \in \Sigma_{n,2d}$, then for any basis $\mathcal{B}_{n,d}$ of $P_n n, d$, there exists a matrix $\mathcal{Q} \succcurlyeq 0$ such that*

$$\mathcal{B}_{n,d}^T \mathcal{Q} \mathcal{B}_{n,d} = p(x) \tag{2.10}$$

Proof. For any $p(x) \in P_{n,2d}$, if $p(x) \in \Sigma_{n,2d}$, then we can write

$$p(x) = \sum_{i=1}^{k} q^2(x) = \begin{bmatrix} q_1(x), ..., q_k(x) \end{bmatrix} \begin{bmatrix} q_1 \\ \vdots \\ q_k \end{bmatrix} \tag{2.11}$$

Notice that $q_j(x) \in P_{n,d}$.

Now given $\mathcal{B}_{n,d} = \{b_1, ..., b_{\binom{n+d}{d}}\}$ be a basis of $P_{n,d}$, we have

$$q_j(x) = \sum_{i=1}^{\binom{n+d}{d}} c_j b_j = \begin{bmatrix} c_1, ..., c_{\binom{n+d}{d}} \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_{\binom{n+d}{d}} \end{bmatrix} \tag{2.12}$$

By subsituting the section equation into the first, we have

$$p(x) = \begin{bmatrix} b_1 & \cdots & b_{\binom{n+d}{d}} \end{bmatrix} \begin{bmatrix} c_{1,1} & \cdots & c_{1,k} \\ \vdots & & \\ c_{\binom{n+d}{d},1} & \cdots & c_{\binom{n+d}{d},k} \end{bmatrix} \begin{bmatrix} c_{1,1} & \cdots & c_{1,\binom{n+d}{d}} \\ \vdots & & \\ c_{k,1} & \cdots & c_{k,\binom{n+d}{d}} \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_{\binom{n+d}{d}} \end{bmatrix} \tag{2.13}$$

Now the matrices in the middle is $C^T C = \mathcal{Q}$ a positive semidefinite matrix, which proofs the foward direction of this theorem.

On the other hand, if we know $p(x) = \mathcal{B}_{n,d}^T \mathcal{Q} \mathcal{B}_{n,d}$, where $\mathcal{Q}$ is a positive semidefinite matrix, we can just apply the Cholesky decomposition to get $\mathcal{Q} = L^T L$, and recover the SOS form of $p(x)$. [BPT13] □

Therefore, we have reduceed our problem decision problem to finding this positive semidefinite matrix. It turns out that this can be done via *Semidefinite Programming* [BPT13]

Definition 2.17. An *Semidefinite Problem (SDP)* in standard primal form is written as

$$minimize \langle C, X \rangle \quad subject\ to \langle A_i, X \rangle = b_i, i = 1, ..., k \quad X \succcurlyeq 0 \tag{2.14}$$

Returning to our problem, we require $\mathcal{Q} \succcurlyeq 0$, and the set of constraints $\langle A_i, X\rangle = b_i$ is the same as satisfying $p(x) = \mathcal{B}_{n,d}^T \mathcal{Q}\mathcal{B}_{n,d}$, which can be done via comparing the coefficients of each element in the basis that we use. [Rec14] Therefore, when we plug in the constraints into a SDP solver, if the result that we get is not no solution, then we know that the $p(x) \in \Sigma_{n,2d}$

Example 2.18. We can do an example here

Now, we further examine the constraints. Using the inner product of matrices, there is a natural representation of the constraints using the inner products $\langle \cdot, \cdot\rangle$ between matrices. That is $\langle A, B\rangle = tr(A^T B)$

Proposition 2.19. *We pick a basis of $\mathcal{B}_{n,d} = b_1, ..., b_{\binom{n+d}{d}}$ of $P_{n,d}$, and list it in a vector form $b = \begin{bmatrix} b_1 & \cdots & b_{\binom{n+d}{d}} \end{bmatrix}^T$. Then by remark 2.14, we can form a basis $\mathcal{B}_{n,2d} = \{b'_1, ..., b'_{\binom{n+2d}{2d}}\}$ based on the vector b. Suppose the $p(x)$ that we are interested in is written in the form $\sum_{i=1}^{\binom{n+2d}{2d}} c_i b'_i$. Then, we can reformulate $p(x) = b^T \mathcal{Q} b = \langle Q, bb^T\rangle$*

Because the constrants are linear with respect to $\mathcal{Q}$, we can re-write it into a system on linear equations with respect to the entries of the matrix $\mathcal{Q}$. Thus we can reformulate the constrants as $Aq = b$, where $q = [q_{1,1}, ..., q_{1,k}, q_{2,2}, ..., q_{k,k}]$. And we are interested in the numerical stability of A.

Definition 2.20. We define the bb^T in Proposition 2.19 as the *Moment Matrix.*

Definition 2.21.

2.2. **Polynomial Basis.**

2.3. **Solving Semidefinite Program.** Toy examples maybe

3. Numerical Results

Proposition 3.1.

Proof. □

maybe a theorem

Theorem 3.2.

or an example...

Example 3.3.

pictures are always a good idea...

4. Maybe Some Proofs

5. Resume, Outlook, or/and Open Problems

what did you do, what questions are still open, natural next steps etc.

References

[BPT13] G. Blekherman, P.A. Parrilo, and R.R. Thomas, *Semidefinite optimization and convex algebraic geometry*, MOS-SIAM Series on Optimization, vol. 13, SIAM and the Mathematical Optimization Society, Philadelphia, 2013.

[CK07] E. Cheney and D. Kincaid, *Numerical mathematics and computing*, International student edition, Cengage Learning, 2007.

[Lau09] M. Laurent, *Sums of squares, moment matrices and optimization over polynomials*, Emerging applications of algebraic geometry, IMA Vol. Math. Appl., vol. 149, Springer, New York, 2009, pp. 157–270.

[Rec14] M. Recher, *Zur numerischen Auswirkung von Basiswahlen in der polynomiellen Optimierung*, 2014, Master Thesis, Universität zu Köln.

8514 Villa La Jolla Drive # 114, La Jolla, CA, 92037

Email address: yiz044@ucsd.edu

9500 Gilmam Drive # 0112, La Jolla, CA, 92093

Email address: mdressler@ucsd.edu