Sum of Squares: Part 1

Apoorv Vikram Singh

February 2, 2021

Introduction

Reason two basic problems about polynomial inequalities

1. Feasibility of polynomial system.

$$\begin{aligned} p_1(\boldsymbol{x}) &\geq 0 \\ &\vdots \\ p_m(\boldsymbol{x}) &\geq 0 \end{aligned} \tag{1}$$

   Is there an $\boldsymbol{x}$ such that (1) is satisfied?

2. Checking non-negativity: Is $q(\boldsymbol{x}) \geq 0$, $\forall \boldsymbol{x}$ satisfying (1) ?

Justification: Feasibility

Feasibility checking is highly expressive.

- Example: MaxCut.
  - Input: $G = (V, E)$, $|V| = n$.
  - Goal: Find $S \subseteq V$ such that $\left|E(S, \overline{S})\right|$ is maximized.
  - Polynomial Feasibility: For some $\beta \in \mathbb{Z}^+$,

$$\frac{1}{4} \sum_{(i,j) \in E} (\boldsymbol{x}_i - \boldsymbol{x}_j)^2 = \beta \, |E|$$

$$\boldsymbol{x}_i^2 = 1, \qquad \forall 1 \leq i \leq n.$$

  - $n + 1$ degree-2 polynomials. Enumerate over polynomial number of values of β and solve MaxCut.

Other examples: MaxClique, Max 3-SAT, Knapsack. Therefore, polynomial feasibility checking problem is *NP*-Hard.

Goal

- Analyze a relaxation for the feasibility problem, and try to find interesting situations where one can get a poly-time algorithms.

Justification: Non-Negativity

- Checking positivity: Given $f : \{-1, 1\}^n \to \mathbb{R}$ with rational coefficients, decide if:
  - $f \geq 0$, $\forall \boldsymbol{x} \in \{-1, 1\}^n$, or,
  - find an $\boldsymbol{x} \in \{-1, 1\}^n$ such that $f(\boldsymbol{x}) \leq 0$.
- Example MaxCut: Decide if MaxCut $\leq c$.
  - Let $f_G(\boldsymbol{x}) \stackrel{\text{def}}{=} \frac{1}{4} \sum_{(i,j) \in E} (\boldsymbol{x}_i - \boldsymbol{x}_j)^2$.
  - Decide if $c - f_G(\boldsymbol{x}) \geq 0$, $\forall \boldsymbol{x} \in \{-1, 1\}^n$.

Certifying Non-Negativity

Given $f : \{-1, 1\}^n \to \mathbb{R}$, find an "efficiently verifiable" certificate of non-negativity.

SoS Certificates

Definition (SoS cert of non-neg (or) SoS proof of non-neg)

A degree-d SoS certificate of non-negativity of $f : \{-1, 1\}^n \to \mathbb{R}$ is a list of polynomials $g_1, \ldots, g_r : \{-1, 1\}^n \to \mathbb{R}$, such such that

- $\deg(g_i) \leq d/2$, and
- $f(\boldsymbol{x}) = \sum_{i \leq r} g_i^2(\boldsymbol{x})$, $\forall \boldsymbol{x} \in \{-1, 1\}^n$.

Efficiently Verifiable (?)

Polynomials $f, g_1, \ldots, g_r$ are represented as a vector of coefficients.

1. How large if r? ($\leq n^d$, see later)
2. How large are coefficients of g_i?

Efficiently Verifiable

Proposition (Efficiently Verifiable)

Suppose $r \leq n^d$, all coefficients of g_i are bounded in magnitude by $2^{\mathrm{poly}(n^d)}$. Then the identity $f = \sum_{i \leq r} g_i^2$ over all $\boldsymbol{x} \in \{-1, 1\}^n$ can be checked in poly(n^d) time.

Proof.

- Given g_i, can compute g_i^2, and $\sum_{i \leq r} g_i^2$ in polynomial time.
- Check if $(f - \sum_{i \leq r} g_i^2)(\boldsymbol{x}) = 0$, $\forall \boldsymbol{x} \in \{-1, 1\}^n$.
- Using the fact that coefficient vector representation is unique, just check if $f - \sum_{i \leq r} g_i^2 = \mathbf{0}$

□

Fact (Unique Representation)

$\forall f : \{-1,1\}^n \to \mathbb{R}$, there exists a <u>*unique*</u> *representation of f: The multi-linear representation of f*

$$f = \sum_{S \subseteq [n]} \hat{f}(S) \prod_{i \in S} \boldsymbol{x}_i \,.$$

(this representation is its Fourier transform)

$\implies$ coefficient vector representation is unique.
(multilinear representation exists because $\boldsymbol{x}_i^2 = 1$)

Are non-negative functions always certifiable?

Proposition (Certifiablity of non-negative functions)
Let $f : \{-1, 1\}^n \to \mathbb{R}$ be non-negative over $\{-1, 1\}^n$. Then, there exists a $\deg(2n)$*-SoS certificate of non-negativity.*

Proof.

- Consider $g : \{-1, 1\}^n \to \mathbb{R}$, and $g(\boldsymbol{x}) = \sqrt{f(\boldsymbol{x})}$.
- Every function on $\{-1, 1\}^n$ is a polynomial of deg $\leq n$.
- $f = g^2 \implies \deg(2n)$-SoS Certificate.

$\square$

Tensor Notation

- Suppose vector $\boldsymbol{v} \in \mathbb{R}^n$.
- $\boldsymbol{v}^{\otimes 2} \in \mathbb{R}^{n^2}$, where $\boldsymbol{v}(i, j) = \boldsymbol{v}_i \boldsymbol{v}_j$.
- $\boldsymbol{v}^{\otimes k} \in \mathbb{R}^{n^k}$.

Proving Efficient Verifiability

Theorem (PSD Matrices and SoS Certificates)

$f : \{-1,1\}^n \to \mathbb{R}$ has a $\deg(d)$-SoS certificate of non-negativity iff there exists a matrix A such that $A \succcurlyeq 0$, and

$$f(\boldsymbol{x}) = \left\langle (1,\boldsymbol{x})^{\otimes \frac{d}{2}}, A \cdot (1,\boldsymbol{x})^{\otimes \frac{d}{2}} \right\rangle .$$

- Parsing Notation:
  - $(1,\boldsymbol{x}) \in \mathbb{R}^{(n+1)}$.
  - $(1,\boldsymbol{x})^{\otimes \frac{d}{2}}$: populate in a vector all possible monomials in the variable $\boldsymbol{x}$ of degree at most $d/2$.
  - $A \in \mathbb{R}^{(n+1)^{d/2} \times (n+1)^{d/2}}$.

Proof

- If Part:

$$f(\boldsymbol{x}) = \left\langle (1,\boldsymbol{x})^{\otimes \frac{d}{2}}, A\cdot(1,\boldsymbol{x})^{\otimes \frac{d}{2}} \right\rangle$$
$$= \left\langle (1,\boldsymbol{x})^{\otimes \frac{d}{2}}, (B^\top B)\cdot(1,\boldsymbol{x})^{\otimes \frac{d}{2}} \right\rangle$$
$$= \left\langle B\cdot(1,\boldsymbol{x})^{\otimes \frac{d}{2}}, B\cdot(1,\boldsymbol{x})^{\otimes \frac{d}{2}} \right\rangle$$
$$= \left\| B\cdot(1,\boldsymbol{x})^{\otimes \frac{d}{2}} \right\|^2 . \tag{2}$$

- Let $g_i(\boldsymbol{x}) = \left\langle e_i, B\cdot(1,\boldsymbol{x})^{\otimes \frac{d}{2}} \right\rangle$, i.e., i-th entry of the vector.
- B is a matrix of constants, applied to monomials of degree at most $d/2$, therefore, $deg(g_i) \le d/2$.
- Therefore,

$$f(\boldsymbol{x}) = \sum_{i=1}^{(n+1)^{\otimes d/2}} g_i^2(\boldsymbol{x}) .$$

Proof Cont...

- Only if Part: Suppose f has a degree-d SoS certificate.

$$f = \sum_{i \le r} g_i^2 \,, \text{ and } \underbrace{g_i(\boldsymbol{x})}_{\deg \le d/2} = \left\langle \boldsymbol{v}_i, (1, \boldsymbol{x})^{\otimes \frac{d}{2}} \right\rangle \,.$$

$$\begin{aligned} f(\boldsymbol{x}) &= \sum_{i \le r} \left\langle \boldsymbol{v}_i, (1, \boldsymbol{x})^{\otimes \frac{d}{2}} \right\rangle^2 \\ &= \left\langle (1, \boldsymbol{x})^{\otimes \frac{d}{2}}, \underbrace{\left(\sum_{i \le r} \boldsymbol{v}_i \boldsymbol{v}_i^\top \right)}_{\succcurlyeq 0} \cdot (1, \boldsymbol{x})^{\otimes \frac{d}{2}} \right\rangle \,. \end{aligned}$$

$\square$

Efficient Verifiability

Corollary (Bound on r)

If $f : \{-1,1\}^n \to \mathbb{R}$ has a degree-d SoS certificate, then it has a certificate with $r \le (n+1)^{d/2}$.

Proof.

Follows from (2):

$$f(\mathbf{x}) = \left\| B \cdot (1, \mathbf{x})^{\otimes \frac{d}{2}} \right\|^2 .$$

$\square$

Efficient Verifiability

Lemma (Bit-complexity of SoS proofs)

Suppose f has a degree-d SoS certificate over $\{-1, 1\}^n$. Then, we can find a degree-d SoS certificate for $f + \varepsilon$ in time $\mathrm{poly}(n^d, \log 1/\varepsilon)$.

- $\{-1, 1\}^n$ is important, and doesn't necessarily hold for other domains.

Proof

Since we are given f, we know that it can be efficiently represented. Therefore, we try to bound the entries of A in terms of f.

- $f(\boldsymbol{x}) = \left\langle (1, \boldsymbol{x})^{\otimes \frac{d}{2}}, A \cdot (1, \boldsymbol{x})^{\otimes \frac{d}{2}} \right\rangle$, for some A.
- $f = \sum_u \hat{f}_u \boldsymbol{x}_u$, where $\boldsymbol{x}_u = \prod_{i \in u} \boldsymbol{x}_i$.
- Expanding the inner product, we see $\hat{f}_u = \sum_{S,T} A_{S,T}$, such that $\mathrm{odd}\,(S + T) = u$, and $|S|, |T| \leq d/2$.

$$\hat{f}_{\emptyset} = \sum_S A_{S,S} = \mathrm{tr}\,(A) = \sum_i \underbrace{\lambda_i(A)}_{\geq 0} .$$

$$\|A\|_F^2 = \sum_{S,T} A_{S,T}^2 = \sum_i \lambda_i^2(A) \leq \hat{f}_{\phi}^2 .$$

We do not know if entries of A are rational, therefore, above proof doesn't suffice. We now try to find an A.

Connection between SDP and SoS: Find A

Proof Cont...

Recall

$$f(\boldsymbol{x}) = \Big\langle \underbrace{(1,\boldsymbol{x})^{\otimes \frac{d}{2}}, A \cdot (1,\boldsymbol{x})^{\otimes \frac{d}{2}}}_{\stackrel{\text{def}}{=} g_A(\boldsymbol{x})} \Big\rangle \,,$$

Then, we form the following constraints:

1. $A \succcurlyeq 0$.
2. $\mathrm{mult}\,(g_A(\boldsymbol{x})) = \mathrm{mult}\,(f(\boldsymbol{x}))$.

Therefore, we get the following SDP feasibility problem:

$$\forall u \subseteq [n] : \hat{f}_u = \sum_{\mathrm{odd}(S+T)=u} A_{S,T}; \qquad \Big((n+1)^d \text{constraints}\Big)$$

$$A \succcurlyeq 0\,.$$

It is unknown if we can decide the feasibility of this system.

Therefore, we try to solve it approximately.

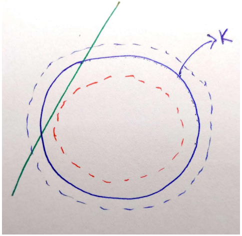
Tools for Approximately Solving SDP

Definition (Weak Separation Oracle)

Let $K \subseteq \mathbb{R}^N$ be a convex set. Weak Separation Oracle:

- Input: Rational vector $\boldsymbol{x} \in \mathbb{R}^N$, and $\varepsilon > 0$.
- Output: Either
  - Correctly asserts that $\boldsymbol{x} \in K + \mathcal{B}(0, \varepsilon)$, or,
  - Returns an "almost separating hyperplane", i.e., returns $\boldsymbol{y} \neq \mathbf{0} \in \mathbb{R}^N$, such that

$$\langle \boldsymbol{y}, \boldsymbol{x} \rangle > \langle \boldsymbol{y}, \boldsymbol{z} \rangle - \varepsilon \|\boldsymbol{y}\|_2, \forall \boldsymbol{z} \in K.$$



Tools for Approximately Solving SDP

Theorem (Grötschel, Lovász, Schrijver '81)

Let K be a closed, convex, and bounded set. Suppose there exists $R > r > 0$, such that $\mathcal{B}(\boldsymbol{p}, r) \subseteq K \subseteq \mathcal{B}(\boldsymbol{0}, R)$. Assume that we have a poly-time weak-separation oracle for K. Then given any rational vector $\boldsymbol{v} \in \mathbb{R}^N$, we can compute a rational vector $\boldsymbol{x} \in \mathbb{R}^N$ such that

1. $\boldsymbol{x} \in K$.
2. $\langle \boldsymbol{v}, \boldsymbol{x} \rangle \geq \langle \boldsymbol{v}, \boldsymbol{z} \rangle - \varepsilon, \forall \boldsymbol{z} \in K$.

Running time: $\operatorname{poly}(\log R/r + \log 1/\varepsilon + N)$.

- Interpreting theorem: If I have a convex set K with non-empty interior, with a weak separation oracle, then I can approximately maximize $\boldsymbol{v}^\top \boldsymbol{x}$ over K.

Proof Cont...

Applying this to our problem. We define the following:

$$S = \left\{ A \,\middle|\, A \succcurlyeq 0, \langle C_i, A \rangle = b_i, \forall i, \|A\|_F^2 \leq \hat{f}_\emptyset^2 \right\} .$$

We note that S is convex, bounded, closed. Now,

$$\mathcal{B}\left(\boldsymbol{p}, r\right) \not\subseteq S .$$

Therefore, relax the equality constraints. And find a point in S',

$$S' = \left\{ A \,\middle|\, A \succcurlyeq 0, \langle C_i, A \rangle = [b_i - \varepsilon, b_i + \varepsilon], \forall i, \|A\|_F^2 \leq \hat{f}_\emptyset^2 \right\} .$$

Now, $\mathcal{B}\left(\boldsymbol{p}, r\right) \subseteq S'$ because for any point $A \in S$, then $A + \delta I \in S'$ for δ small enough.

Applying the Theorem

Proof Cont...

- We can find some f' such that f' has a degree-d SoS certificate and $\left|\hat{f}_u - \hat{f}'_u\right| \leq \varepsilon$.
- Note: $f(\boldsymbol{x}) = f'(\boldsymbol{x}) + (f - f')(\boldsymbol{x})$.
- Small coefficient: $\sum_{|u| \leq d} \left|\hat{f}_u - \hat{f}'_u\right| \leq \varepsilon(n+1)^d$.
- Let $L = \sum_{|u| \leq d} \left|\hat{f}_u - \hat{f}'_u\right|$.
- Then, $L + f - f'$ has a degree d-SoS certificate.
- Then, it implies, $L + f$ has a degree-d SoS certificate, i.e., $\varepsilon(n+1)^d + f$ has a degree-d SoS certificate.
- $\varepsilon = \mathcal{O}\left(n^{-d}\right)$ finishes the proof.

$\square$