Lecture 11 : Sum of Squares Programming

goals:
- intro to sum of squares programming
- connection between polynomial nonnegativity and SDP
- useful in many areas of engineering + applied math, especially systems + control

Multivariate Polynomials:

Let's start with some definitions:

Definition: A polynomial f in variables $x_1, \ldots, x_n$ is a finite linear combination of monomials

$$f = \sum_\alpha c_\alpha x^\alpha = \sum_\alpha c_\alpha x_1^{\alpha_1} \cdots x_n^{\alpha_n}, \quad c_\alpha \in \mathbb{R}$$

where the sum is over a finite number of n-tuples $\alpha = [\alpha_1, \ldots, \alpha_n]$ with $\alpha_i \in \mathbb{N}_0 = \{0, 1, 2, \ldots\}$.

- set of all polynomials w/ real coefficients and total degree d denoted $P_{n,d}$
- total degree of a monomial x^α : $\sum_i \alpha_i$
- total degree of a polynomial is max degree of its monomials

$\underline{\underline{Ex}}$ $f(x) = x_1^4 - 6x_1^3x_2 + 2x_1^3x_3 + 6x_1^2x_3^2 + 9x_1^2x_2^2$

$- 6x_1^2x_2x_3 - 14x_1x_2x_3^2 + 4x_1x_3^3 + 5x_3^4 - 7x_2^2x_3^2$

$+ 16x_2^4$

Definition: A form is a polynomial where all monomials have the same total degree d

$$\Rightarrow f(\lambda x_1, \dots, \lambda x_n) = \lambda^d f(x_1, \dots, x_n)$$ "homogeneous of degree d"

Connecting polynomials and forms

- Every form in n variables and degree d can be "dehomogenized" to a polynomial in $n-1$ variables of degree $\leq d$ by fixing any variable to 1
- Conversely, every polynomial in n variables can be "homogeaized" by multiplying each monomial by powers of a new variable to obtain a form in $n+1$ variables
  - $\underline{\underline{Ex}}$ $x^2 + 2x + 1 \longrightarrow x^2 + 2xy + y^2$
- These transformations preserve some important properties

- $P_{n,d}$ can be identified with a vector space $\mathbb{R}^N$ with $N = \binom{n+d}{d}$, with each vector entry corresponding to a monomial coefficient
    - dimension comes from counting combinations of variables with repetition (aka multicombo/multisubset)
- Likewise, set of forms of degree d in n variables corresponds w/ a vector space of dimension $\binom{n+d-1}{d}$

Polynomial Nonnegativity

- A polynomial f in n variables is called nonnegative if

$$f(x) \geq 0 \quad \forall x \in \mathbb{R}^n$$

- Many important questions across engineering and applied math ultimately boil down to questions about polynomial nonnegativity
    - e.g. Lyapunov theory in systems + control (more soon!)
- The set $P^+_{n,d} = \{ f \in P_{n,d} \mid f \geq 0 \}$ of nonnegative polynomials forms a convex cone in the coefficient vector space $\mathbb{R}^{\binom{n+d}{d}}$
    - Why?

- Imagine we are given $f \in P_{n,d}$ and asked if $f \in P^+_{n,d}$, or that we are asked if we can choose certain coefficients on f such that $f \in P^+_{n,d}$
    - It's a convex feasibility problem if constraints on coefficients are convex!

However...

FACT: Given a polynomial f (even one of degree 4), it is NP-hard to decide if it's nonnegative!

- Yet more fine print that computational tractability involves more than just convexity!

Sum of squares and SDP

- Could we try to write a polynomial in a way that its nonnegativity becomes obvious?

Definition: A polynomial f is called a sum of squares (SOS) if it can be written

$$f(x) = \sum_i q_i^2(x)$$

for some polynomials q_i.

- clearly f SOS $\Rightarrow$ f nonnegative
- existence of SOS decomposition an algebraic certificate of nonnegativity

- Remarkable observation:

$$\boxed{\text{Deciding if } f \text{ is SOS is an SDP!}}$$

Theorem: A multivariate polynomial f in n variables of degree $2d$ is a sum of squares (SOS) iff $\exists\, Q = Q^T \succeq 0$ such that

$$f(x) = z^T Q z \qquad (*)$$

where z is the vector of monomials of degree up to d

$$z = [1, x_1, x_2, \dots, x_n, x_1^2, x_1 x_2, \dots, x_n^d]^T$$

Proof: ($\Rightarrow$) If $(*)$ holds with $Q \succeq 0$, we can factorize as $Q = V^T V$ and obtain an SOS decomposition as

rows of V

$$f(x) = z^T V^T V z = \| V z \|_2^2 = \sum_i (v_i^T z)^2$$

($\Leftarrow$) If f is SOS, there are coefficients a_i such that

$$f = \sum_i q_i^2(x) = \sum_i (a_i^T z)^2 = z^T \Big(\sum_i a_i a_i^T \Big) z$$

so $Q = \sum_i a_i a_i^T \succeq 0$ can be extracted. ∎

- Corollary: # squares in SOS decomp. = rank Q = # rows V
- Finding $Q \succeq 0$ for a given f is an SDP feasibility problem with affine constraints from matching coefficients in $(*)$

Ex Consider $f = 4x_1^4 + 4x_1^3x_2 - 7x_1^2x_2^2 - 2x_1x_2^3 + 10x_2^4$

$$= \begin{bmatrix} x_1^2 \\ x_1x_2 \\ x_2^2 \end{bmatrix}^T \begin{bmatrix} q_{11} & q_{12} & q_{13} \\ q_{12} & q_{22} & q_{23} \\ q_{13} & q_{23} & q_{33} \end{bmatrix} \begin{bmatrix} x_1^2 \\ x_1x_2 \\ x_2^2 \end{bmatrix}$$

$$= q_{11}x_1^4 + 2q_{12}x_1^3x_2 + (q_{22} + 2q_{13})x_1^2x_2^2$$
$$+ 2q_{23}x_1x_2^3 + q_{33}x_2^4$$

Matching coefficients:

$q_{11} = 4$ $\quad$ $q_{23} = -1$

$q_{12} = 2$ $\quad$ $q_{22} + 2q_{13} = -7 \Rightarrow q_{22} = -7 - 2q_{13}$

$q_{33} = 10$

Simplified SDP feasibility problem:

find q_{13} such that $Q = \begin{bmatrix} 4 & 2 & q_{13} \\ 2 & -7-2q_{13} & -1 \\ q_{13} & -1 & 10 \end{bmatrix} \succeq 0$

With $q_{13} = -6$ we have

$$Q = \begin{bmatrix} 4 & 2 & -6 \\ 2 & 5 & -1 \\ -6 & -1 & 10 \end{bmatrix} = \begin{bmatrix} 0 & 2 \\ 2 & 1 \\ 1 & -3 \end{bmatrix} \begin{bmatrix} 0 & 2 & 1 \\ 2 & 1 & -3 \end{bmatrix}$$

yielding an SOS decomposition

$$f = \begin{bmatrix} x_1^2 \\ x_1x_2 \\ x_2^2 \end{bmatrix}^T \begin{bmatrix} 0 & 2 \\ 2 & 1 \\ 1 & -3 \end{bmatrix} \begin{bmatrix} 0 & 2 & 1 \\ 2 & 1 & -3 \end{bmatrix} \begin{bmatrix} x_1^2 \\ x_1x_2 \\ x_2^2 \end{bmatrix} = \left\| \begin{bmatrix} 2x_1x_2 + x_2^2 \\ 2x_1^2 + x_1x_2 - 3x_2^2 \end{bmatrix} \right\|_2^2$$

$$= (2x_1x_2 + x_2^2)^2 + (2x_1^2 + x_1x_2 - 3x_2^2)^2$$

- Similarly, the 3 variable polynomial from page 2 has SOS decomposition

$$f(x) = \left(x_1^2 - 3x_1x_2 + x_1x_3 + 2x_3^2\right)^2 + \left(x_1x_3 - x_2x_3\right)^2 + \left(4x_2^2 - x_3^2\right)^2 \geq 0$$

  - <u>not</u> obvious, but can be automated computationally!

Are <u>all</u> nonnegative polynomials SOS?

- No, in general $SOS_{n,d} \subset P^+_{n,d}$

<u>Theorem</u> (Hilbert 1888)

All nonnegative polynomials in n variables of degree d are SOS iff

- $n = 1$. OR $d = 2$ OR $n = 2,\ d = 4$

Equivalently, all nonneg. forms in n variables of degree d are SOS iff

- $n = 2$ OR $d = 2$ OR $n = 3,\ d = 4$

<u>Ex</u> The Motzkin Polynomial (nonnegative but <u>NOT</u> SOS)

$$M(x) = x_1^2 x_2^4 + x_1^4 x_2^2 + 1 - 3x_1^2 x_2^2$$

- nonnegativity via arithmetic-geometric inequality w/ $[x_1^2x_2^4, x_1^4x_2^2, 1]$
- SOS SDP is infeasible

Nonnegativity over Basic Semialgebraic Sets

- Often we are interested in a polynomial being nonnegative over a certain <u>subset</u>
- There's an important special case involving quadratic functions known as the <u>S-Procedure</u>

<u>S-Procedure</u> (aka S-lemma)

- When is it true that $\forall x \in \mathbb{R}^n$

$$\underbrace{\begin{bmatrix} x \\ 1 \end{bmatrix}^T \begin{bmatrix} P_1 & q_1 \\ q_1^T & r_1 \end{bmatrix} \begin{bmatrix} x \\ 1 \end{bmatrix}}_{Q_1(x)} \geq 0 \implies \underbrace{\begin{bmatrix} x \\ 1 \end{bmatrix}^T \begin{bmatrix} P_0 & q_0 \\ q_0^T & r_0 \end{bmatrix} \begin{bmatrix} x \\ 1 \end{bmatrix}}_{Q_0(x)} \geq 0$$

i.e., when is Q_0 nonnegative on the subset defined by $Q_1 \geq 0$

- these functions need not be convex!

<u>Theorem</u>: Assume $\exists z$ such that $Q_1(z) > 0$. We have $Q_1(x) \geq 0 \implies Q_0(x) \geq 0$ iff

$$\exists \tau \geq 0 \text{ such that } \begin{bmatrix} P_0 & q_0 \\ q_0^T & r_0 \end{bmatrix} \succeq \tau \begin{bmatrix} P_1 & q_1 \\ q_1^T & r_1 \end{bmatrix}$$

- an LMI (in variable τ) even when Q_0 not convex and $Q_1 \geq 0$ not convex!

- sufficiency is obvious but necessity is <u>NOT</u> obvious, hard to prove (Yakubovich 1971)

- can be used to solve non-convex problems with exactly one quadratic constraint and quadratic objective
    - even more fine print that strictly speaking convexity $\neq$ tractability

- the sufficient condition extends to more general polynomial settings

<u>Definition</u> A <u>basic</u> semi<u>algebraic</u> <u>set</u> is a set of the form

$$K = \{ x \in \mathbb{R}^n \mid g_i(x) \geq 0, \ i=1,\dots,m, \ h_i(x) = 0, \ i=1,\dots,p \}.$$

where g_i and h_i are multivariate polynomials.

- K not necessarily convex!

SOS S-Procedure

- Consider a multivariate polynomial $f(x)$ (not necessarily convex)

- When is it true that

$$x \in K \implies f(x) \geq 0$$

i.e., when is f nonnegative on K?

Theorem We have $x \in K \implies f(x) \geq 0$ if there exist polynomials $t_1(x), \dots, t_p(x)$ and nonnegative polynomials $s_1(x), \dots, s_m(x)$ such that

$$f(x) \geq \sum_{i=1}^{m} s_i(x) g_i(x) + \sum_{i=1}^{p} t_i(x) h_i(x)$$

- Thus if we could solve the feasibility problem

$$\begin{aligned} &\text{find} \quad t_1(x), \dots, t_p(x),\ s_1(x), \dots, s_m(x) \\ &\text{subject to} \quad f(x) - \sum_{i=1}^{m} s_i(x) g_i(x) - \sum_{i=1}^{p} t_i(x) h_i(x) \quad \text{SOS} \\ &\qquad\qquad\quad s_i(x) \quad \text{SOS} \quad \forall i = 1, \dots, p \end{aligned}$$

we could certify nonnegativity of f on K (variables are coefficients of t_is and s_is)

- This is an SDP feasibility problem!

$\implies$ we can automate search for nonnegativity certificates

- There's a stronger version of this called the Positivstellensatz that exploits full power of real algebraic geometry and SDP to certify infeasibility of semialgebraic problems.

- Final Comments on SOS Programming
  - Beautiful, deep topic w/ surprisingly many applications at boundaries of research
  - There are software tools to automatically parse SOS problems into SDPs
    - SOSTOOLS
    - YALMIP
    - Gloptipoly
  - extensions to some non-polynomials (e.g. trig functions)
  - However, numerically very challenging to scale methods to large, realistic problems
    - SDPs get very large fast as # of variables + polynomial degree grows
    - active research on improving scalability + efficiency by exploiting structure of many practical problems (sparsity, symmetries, etc.), first-order methods for SDPs, etc.
      - Ahmadi et al. CDC 2017