

L. Slot

Department of Computer Science

ETH Zürich

lucas.slot@inf.ethz.ch

Onderzoek Research

Polynomiale optimalisatie en sommen van kwadraten

Lucas Slot ontving op het Nederlands Mathematisch Congres 2024 de Stieltjesprijs voor zijn proefschrift “Asymptotic Analysis of Semi-definite Bounds for Polynomial Optimization and Independent Sets in Geometric Hypergraphs”, verdedigd in 2022 aan de Universiteit van Tilburg. In dit artikel legt hij uit waar zijn proefschrift over ging.

Het doel van wiskundige optimalisatie is het minimaliseren (of maximaliseren) van een gegeven *doelfunctie* onder een set *voorwaarden*. Denk bijvoorbeeld aan het ontwerpen van een dienstregeling die de gemiddelde reistijd voor passagiers minimaliseert, maar binnen een voorgeschreven budget blijft. Formeel gesproken krijgen we een functie $f: \mathbb{R}^n \rightarrow \mathbb{R}$ (de doelfunctie) en functies $g_1, g_2, \dots, g_m: \mathbb{R}^n \rightarrow \mathbb{R}$ (de voorwaarden), en moeten we het minimum van f berekenen over de verzameling

$$\mathcal{X} := \{x \in \mathbb{R}^n : g_1(x) \geq 0, \dots, g_m(x) \geq 0\}$$

van punten die aan de voorwaarden voldoen. Met andere woorden, we moeten de waarde bepalen van

$$\text{opt} := \inf_{x \in \mathcal{X}} f(x) \quad (1)$$

Omdat optimalisatieproblemen in het algemeen te complex zijn om met de hand op te lossen, zijn we geïnteresseerd in het ontwikkelen van *efficiënte algoritmen* waarmee het optimum met behulp van een computer kan worden berekend.

Polynomiale optimalisatie

In mijn proefschrift [8] bestudeer ik zogenaamde *polynomiale optimalisatieproblemen* (POPs), waarin de doelfunctie en de voorwaarden polynomen zijn. Deze optimalisatieproblemen hebben brede toepassingen, zowel binnen als buiten de wiskunde. Daar staat tegenover dat het moeilijk (NP-hard) is om ze *exact* op te lossen. Dit motiveert de zoektocht naar efficiënte *benaderingsmethoden*, die het optimum berekenen met een (hopelijk) kleine foutmarge. De

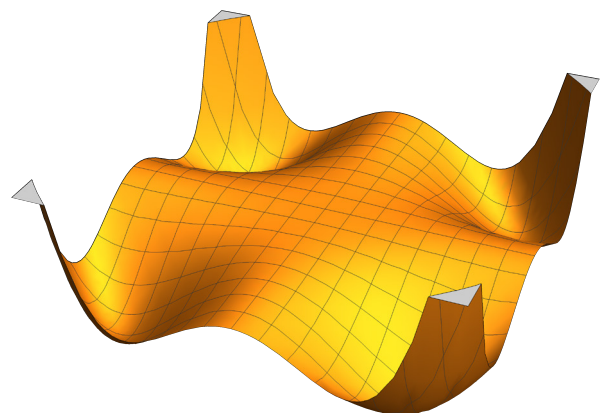
meest succesvolle aanpak is gebaseerd op een link tussen polynomiale optimalisatie en algebraïsche meetkunde. Om deze link te kunnen leggen, hebben we eerst een alternatieve formulatie nodig van het probleem (1). We zeggen dat een polynoom p *niet-negatief* is op $\mathcal{X}$ wanneer $p(x) \geq 0$ voor alle $x \in \mathcal{X}$. We schrijven dan $p \in \mathcal{P}(\mathcal{X})$. Nu geldt dat

$$\text{opt} = \inf_{x \in \mathcal{X}} f(x) = \sup_{\lambda \in \mathbb{R}} \{\lambda : f - \lambda \in \mathcal{P}(\mathcal{X})\}. \quad (2)$$

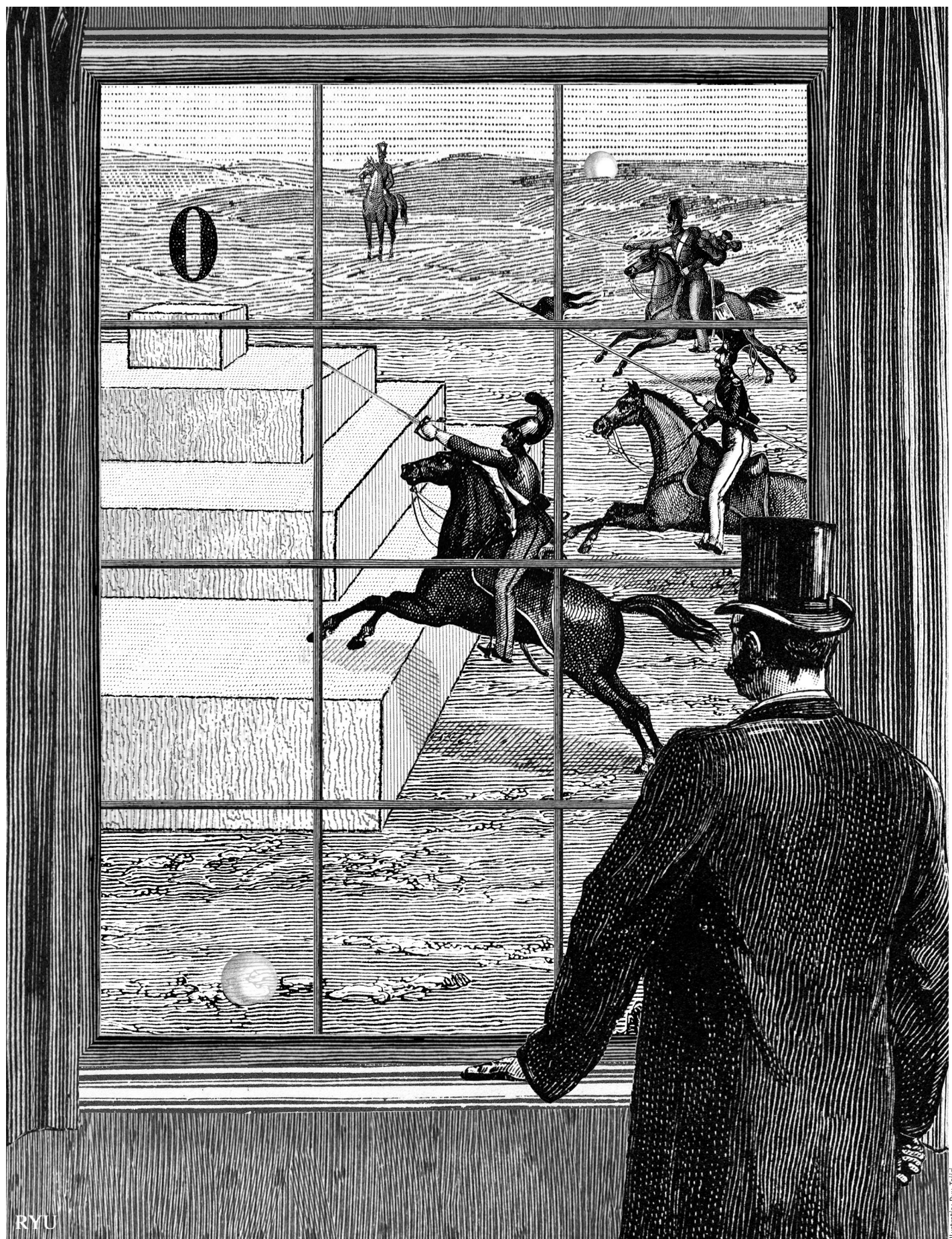
Dat wil zeggen, het minimum van f op $\mathcal{X}$ is gelijk aan het grootste getal λ zodanig dat $f - \lambda$ niet-negatief is op $\mathcal{X}$. Deze observatie laat zien dat het oplossen van polynomiale optimalisatieproblemen in feite equivalent is aan het bepalen van niet-negativiteit van polynomen.

Sommen van kwadraten

Hoe bepalen we of een polynoom niet-negatief is? Neem bijvoorbeeld



Figuur 1 Het *Motzkin-polynoom* $p(x_1, x_2) = x_1^4 x_2^2 + x_1^2 x_2^4 - 3x_1^2 x_2^2 + 1$ is niet-negatief op geheel $\mathbb{R}^2$ (als gevolg van de AM-GM-ongelijkheid). Het is echter geen som van kwadraten.



$$p(x_1, x_2) = x_1^6 - 2x_1^3x_2^2 + x_1^2 - 4x_1x_2 + x_2^4 + 4x_2^2.$$

Aan de hand van de definitie is het niet onmiddellijk duidelijk of p negatieve waarden aan kan nemen. Een korte berekening leert echter dat p kan worden geschreven als een *som van kwadraten*, namelijk

$$p(x_1, x_2) = (x_1 - 2x_2)^2 + (x_1^3 - x_2^2)^2.$$

Dit laat meteen zien dat $p(x_1, x_2) \geq 0$ voor alle $x_1, x_2 \in \mathbb{R}$.

Definitie

Een polynoom p is een *som van kwadraten* (svk) wanneer het geschreven kan worden als

$$p(x) = p_1(x)^2 + p_2(x)^2 + \dots + p_\ell(x)^2 \quad (p_i \in \mathbb{R}[x]).$$

We zeggen dan dat $p \in \Sigma$, en merken op dat $\Sigma \subseteq \mathcal{P}(\mathbb{R}^n)$.

We kunnen sommen van kwadraten ook toepassen om niet-negativiteit op $X \subseteq \mathbb{R}^n$ te bepalen. Stel dat een polynoom p geschreven kan worden als

$$p(x) = \sigma_0(x) + \sigma_1(x)g_1(x) + \dots + \sigma_m(x)g_m(x), \quad (3)$$

waar elk van de $\sigma_i \in \Sigma$ een som van kwadraten is. Dan volgt direct dat $p(x) \geq 0$ voor alle $x \in X$. (Immers, per definitie geldt $g_i(x) \geq 0$ voor alle $x \in X$.) We noemen de representatie (3) een *certificaat* van niet-negativiteit voor p op X . De verzameling van alle polynomen met zulk een certificaat heet het *kwadratische moduul* $Q(X)$ van X . Er geldt dus dat $Q(X) \subseteq \mathcal{P}(X)$. Met het oog op formulatie (2) geeft deze inclusie ons een *ondergrens* voor het optimalisatieprobleem (3), namelijk

$$\text{svk} := \sup_{\lambda \in \mathbb{R}} \{\lambda : f - \lambda \in Q(X)\} \leq \text{opt}. \quad (4)$$

Immers, we nemen het supremum hierboven over een kleinere verzameling dan in (2).

Hilbert, Artin en Positivstellensätze

Is de ondergrens $\text{svk} \leq \text{opt}$ in het algemeen scherp? Anders gezegd, heeft elk niet-negatief polynoom een som-van-kwadraten-certificaat? Deze vraag werd al in de late 19e eeuw gesteld door Hilbert. Hij liet zien dat dit niet het geval is: voor $n \geq 2$ is de inclusie $\Sigma \subsetneq \mathcal{P}(\mathbb{R}^n)$ strict, en bijgevolg bestaan er polynomiale optimalisatieproblemen (met $X = \mathbb{R}^n$) waar $\text{svk} < \text{opt}$. Hilbert's bewijs is niet constructief, en het eerste expliciete voorbeeld van een polynoom $p \in \mathcal{P}(\mathbb{R}^n) \setminus \Sigma$ werd pas in 1957 gevonden door Motzkin, zie Figuur 1.

Gelukkig verbetert de situatie wanneer we voorwaarden toevoegen: Putinar [7] bewees in 1993 dat, wanneer X compact is, elk *positief* polynoom op X een som-van-kwadraten certificaat heeft.

Stelling (Putinar's Positivstellensatz)

Neem aan dat $X \subseteq \mathbb{R}^n$ compact [10] is. Laat $\mathcal{P}_{>0}(X)$ de verzameling van (strikt) positieve polynomen op X . Dan geldt dat

$$\mathcal{P}_{>0}(X) \subseteq Q(X).$$

Hieruit volgt dat $\text{svk} = \text{opt}$ voor polynomiale optimalisatieproble-

men over X .

Stellingen van dit type worden *Positivstellensätze* genoemd, een verwijzing naar Hilbert's bekende *Nullstellensatz*. Ze laten net als de Nullstellensatz een correspondentie zien tussen een *meetkundige* eigenschap van polynomen (positiviteit) en een *algebraïsch* object (het kwadratisch moduul). Overigens bewees Artin [1] in 1927 al dat elk niet-negatieve polynoom op $\mathbb{R}^n$ geschreven kan worden als som van *rationale* kwadraten, dat wil zeggen

$$p(x) = \frac{p_1(x)^2}{q_1(x)^2} + \frac{p_2(x)^2}{q_2(x)^2} + \dots + \frac{p_\ell(x)^2}{q_\ell(x)^2}.$$

Hiermee loste hij destijds het 17e van Hilbert's lijst van 23 problemen op. De Positivstellensatz van Krivine [4] en Stengle [9] veralgemeniseert dit resultaat naar niet-negatieve polynomen over verzamelingen $X \subseteq \mathbb{R}^n$. Rationale certificaten zijn echter computationeel lastig om mee te werken, en leiden derhalve niet tot efficiënte algoritmen. We laten ze daarom verder buiten beschouwing.

Efficiënte benaderingsalgoritmen

We hebben gezien dat polynomiale optimalisatie over compacte verzamelingen X kan worden gereduceerd naar de optimalisatie van een eenvoudige (lineaire) functie over het kwadratisch moduul $Q(X)$, bestaande uit polynomen met een som-van-kwadraten certificaat (3). Op zichzelf is nu nog niet duidelijk hoe dit tot *efficiënte* algoritmen zou leiden. Het is namelijk niet duidelijk of eenvoudig kan worden gecontroleerd of een gegeven polynoom p zo'n certificaat heeft. Sterker nog, het is zeer onwaarschijnlijk dat efficiënt kan worden nagegaan of $p \in Q(X)$! We hebben een laatste inzicht nodig van Lasserre [5] en van Parrilo [6]. Voor $r \in \mathbb{N}$ beschouwen zij het *begrensde* kwadratisch moduul $Q(X)_r \subseteq Q(X)$ dat bestaat uit polynomen met een som-van-kwadraten certificaat van *graad ten hoogste* r . Hiermee wordt bedoeld dat elk van de factoren $\sigma_0, \sigma_1, \dots, \sigma_m \in \Sigma$ in het certificaat (3) graad ten hoogste r heeft. Nu blijkt dat, voor vaste r , efficiënt kan worden bepaald of $p \in Q(X)_r$, middels het oplossen van een zogenaamd *semidefinitief programma*. De benodigde rekenkracht om dit programma op te lossen hangt onder meer af van de gekozen graad r . De conclusie van Lasserre en Parrilo is dat, wederom voor vaste r , de parameter

$$\text{svk}_r := \sup_{\lambda \in \mathbb{R}} \{\lambda : f - \lambda \in Q(X)_r\}$$

efficiënt kan worden berekend. Tevens geldt dat

$$\text{svk}_1 \leq \text{svk}_2 \leq \text{svk}_3 \leq \dots \leq \text{opt}.$$

Tot slot vertelt Putinar's Positivstellensatz ons dat

$$\lim_{r \rightarrow \infty} \text{svk}_r = \text{svk} = \text{opt}$$

De rij $\text{svk}_1, \text{svk}_2, \dots$ vormt daarmee een *convergente hiërarchie van ondergrenzen* op het optimum opt , waarvan iedere individuele term efficiënt kan worden berekend. Deze ondergrenzen staan bekend als de *som-van-kwadraten-hiërarchie* (of ook *Lasserre's hiërarchie*).

Foutmarges en kwantitatieve Positivstellensätze

De vraag die centraal staat in mijn proefschrift is: *Hoe snel convergeert de som-van-kwadraten-hiërarchie naar het ware optimum?* Deze vraag kan zowel vanuit het perspectief van de optimalisatie of van de algebraïsche meetkunde worden geformuleerd: Binnen de optimalisatie is men geïnteresseerd in de relatie tussen de graad r

van de certificaten (en daarmee de benodigde rekenkracht), en de *foutmarge* $\text{opt} - \text{svk}_r$ van de resulterende ondergrenzen. Binnen de algebraïsche meetkunde probeert men *quantitatieve* versies te bewijzen van Positivstellensätze, waarin de benodigde graad van certificaten wordt uitgedrukt in ‘de mate van positiviteit’ van een polynoom. In de afgelopen twee decennia zijn vanuit beide gemeenschappen tal van publicaties verschenen die deze vragen (gedeeltelijk) beantwoorden.

Een raamwerk voor snelle convergentie

In mijn proefschrift bewijs ik nieuwe stellingen over de convergentiesnelheid van Lasserre’s hiërarchie voor optimalisatie over verzamelingen X met bepaalde *symmetrie*. Hieronder vallen de eenheidsbol, de standaardsimplex Δ^n , de kubus $[-1, 1]^n$ en de Booleaanse kubus $\{-1, 1\}^n$. Deze betrekkelijk eenvoudige domeinen omvatten desalniettemin veel van de fundamentele problemen die polynomiale optimalisatie motiveren. In deze bijzondere gevallen is het mogelijk sterke convergentiesnelheden te garanderen. Een voorbeeld:

Stelling (Slot 2022)

Voor optimalisatie over de eenheidsbol geldt dat

$$\text{opt} - \text{svk}_r = O(1/r^2) \quad (r \rightarrow \infty).$$

Wanneer we het best beschikbare resultaat [2] voor *algemene* domeinen X direct op de eenheidsbol zouden toepassen, garandeert dit slechts een foutmarge in $O(1/\sqrt[10]{r})$ (een factor 20 verschil in de exponent). De belangrijkste bijdrage van mijn proefschrift is het ontwikkelen van een theoretisch raamwerk waarbinnen convergentiesnelheden voor elk van de bovengenoemde domeinen kunnen worden bewezen. Dit bouwt voort op eerder werk van Fang en Fawzi [3], die het geval $X = S^{n-1}$ bestudeerden. Ter conclusie beschrijven we dit raamwerk (in zeer grove lijnen).

1. We definiëren een klasse $\mathcal{K}$ van lineaire operatoren op $\mathbb{R}[x]$ met de eigenschap dat, voor elke $K \in \mathcal{K}$,

$$K(\mathcal{P}(X)) \subseteq \mathcal{Q}(X)_r.$$

Dat wil zeggen, operatoren $K \in \mathcal{K}$ beelden niet-negatieve polynomen af op het (begrensd) kwadratisch moduul. Deze klasse

bestaat uit bijzondere *convolutieoperatoren*, van de vorm

$$K: p \mapsto K \star p, \text{ waar} \\ (K \star p)(x) := \int_X p(y) K(x, y) dy.$$

Hier refereert K abusievelijk zowel naar een lineaire operator, als naar een functie $\mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$.

2. Binnen de klasse $\mathcal{K}$ gaan we op zoek naar een operator K die zich (bijna) gedraagt als de identiteit, dat wil zeggen $K \star p \approx p$. Als we zo een operator kunnen vinden, stelt ons dit in staat een certificaat voor $K \star p$ uit de vorige stap om te vormen tot een certificaat voor $p + \varepsilon$ betrekkelijk klein. Dit laat (ruwweg) zien dat de foutmarge van Lasserre’s ondergrens svk_r ten hoogste ε is.

3. Om onze zoektocht te vereenvoudigen, maken we gebruik van de symmetrie van X om de speciale klasse van operatoren te karakteriseren. Kort gezegd kan de hele klasse worden gevat in termen van polynomen van één variabele. Deze beschrijving is het mooist in het geval van de eenheidssfeer. Daar bestaat de klasse uit alle operatoren van de vorm

$$(K \star p)(x) = \int_{S^{n-1}} p(y) q(x \cdot y) dy,$$

waar q een som van kwadraten in één variabele is.

4. Tot slot laten we zien dat ‘de beste’ operator $K \in \mathcal{K}$ gevonden kan worden door een relatief eenvoudig optimalisatieprobleem op te lossen over de ruimte van polynomen in één variabele. Door de juiste orthogonale basis voor deze ruimte te kiezen, kan dit probleem met de hand worden opgelost. Dit leidt dan uiteindelijk tot een convergentiesnelheid voor Lasserre’s hiërarchie.

Stelling (Slot 2022)

De convergentiesnelheid van Lasserre’s hiërarchie kan, in de aanwezigheid van voldoende symmetrie, worden beschreven in termen van het gedrag van klassieke orthogonale polynomen in één variabele.



Referenties en noten

- 1 E. Artin, Über die Zerlegung definiter Funktionen in Quadrate, *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg* 5, 1927.
- 2 L. Baldi and B. Mourrain, On moment approximation and the effective Putinar’s Positivstellensatz, *Mathematical Programming* 200 (2023) 71–103.
- 3 K. Fang, H. Fawzi, The sum-of-squares hierarchy on the sphere, and applications in quantum information theory, *Mathematical Programming, Series A* 190 (2021) 331–360.
- 4 J.-L. Krivine, Anneaux préordonnés, *Journal d’Analyse Mathématique* 12 (1964) 307–326.
- 5 J.B. Lasserre, Global optimization with polynomials and the problem of moments, *SIAM Journal on Optimization*, 11(3), 796–817, 2001.
- 6 P.A. Parrilo, *Structured Semidefinite Programs and Semialgebraic Geometry Methods in Robustness and Optimization*, Ph.D. thesis, California Institute of Technology, 2000.
- 7 M. Putinar, Positive polynomials on compact semi-algebraic sets, *Indiana University Mathematics Journal* 42 (1993) 969–984.
- 8 L. Slot, *Asymptotic Analysis of Semidefinite Bounds for Polynomial Optimization and Independent Sets in Geometric Hypergraphs*, Ph.D. thesis, Tilburg University, 2022.
- 9 G. Stengle, A Nullstellensatz and a Positivstellensatz in Semialgebraic Geometry, *Mathematische Annalen*. 207(2) (1974) 87–97.
- 10 Om precies te zijn dienen de voorwaarden $g_1, g_2, \dots, g_m$ aan een zekere algebraïsche conditie te voldoen die compactheid van X impliceert, maar deze conditie kan in de context van optimalisatie zonder verlies van algemeenheid worden aangenomen, gegeven dat X compact is.