Polynomial Optimization

Sums of Squares and Moment Generation

CHEATSHEET: DECEMBER 14, 2022

1 INTRODUCTION AND NOTATION

Let n be a fixed positive integer.

Let $(u_i)_{i=1}^{\infty}$ be a sequence containing all monomials of n variables, ordered in such a way that the degree of u_i is nondecreasing with i. For every $\nu \in \mathbb{N}$, let $s(\nu)$ be the number of monomials of n variables with degree smaller than or equal to ν, and let $\mathbf{v}_\nu = [u_1, \dots, u_{s(\nu)}]^\top$.

For every $\nu \in \mathbb{N}$ and $i \in \{1, \dots, s(2\nu)\}$, let $B_i^\nu \in \mathbb{R}^{s(\nu)\times s(n)}$ be defined by $[B_i^\nu]_{j_1,j_2} = 1$ if $u_{j_1}u_{j_2} = u_i$ and $[B_i^\nu]_{j_1,j_2} = 0$ otherwise.

With the above definitions, it holds that if $p = \sum_{i=1}^{s(2\nu)} c_i u_i$, then $p = \mathbf{v}_\nu{}^\top X \mathbf{v}_\nu$ for any $X \in \mathbb{R}^{s(\nu)\times s(\nu)}$ satisfying $\langle B_i^\nu, X\rangle = c_i$ for all $i \in \{1, \dots, s(2\nu)\}$.

Definition 1.1 (Sum-of-squares polynomial). A polynomial $p = \sum_{i=1}^{s(2\nu)} c_i u_i$ is said to be *sum of squares* if there is $X \succeq 0$ such that $\langle B_i^\nu, X\rangle = c_i$ for all $i \in \{1, \dots, s(2\nu)\}$.

2 UNCONSTRAINED OPTIMIZATION

Fix $\nu \in \mathbb{N}$ and $p = \sum_{i=1}^{s(2\nu)} c_i u_i$, where for each $i \in \{1, \dots, s(2\nu)\}$, $c_i \in \mathbb{R}$.

The goal is to find the minimum of p: $p^* := \min_{x\in\mathbb{R}^n} p(x)$. Computing p^* is in general difficult. A relaxation consists in finding a value $t \in \mathbb{R}$ such that $p - t$ is sum of squares. Any such value t will then be a lower bound on p^*.

Hence, we can formulate the following optimization problem to optimize on such t:

$$\begin{array}{llll} \mathbb{P}(p) & : & \sup_{X,t} & t \\ & & \text{s.t.} & X \succeq 0, \\ & & & \langle B_1^\nu, X\rangle = c_1 - t, \\ & & & \langle B_i^\nu, X\rangle = c_i, \quad \forall\, i \in \{2, \dots, s(2\nu)\}, \end{array} \tag{1}$$

with variables $X = X^\top \in \mathbb{R}^{s(m)\times s(m)}$ and $t \in \mathbb{R}$. The dual of (1) is the problem:

$$\begin{array}{llll} \mathbb{D}(p) & : & \inf_{(y_i)_{i=1}^{s(2\nu)}} & \sum_{i=1}^{s(2\nu)} c_i y_i \\ & & \text{s.t.} & \sum_{i=1}^{s(2\nu)} B_i^\nu y_i \succeq 0, \\ & & & y_1 = 1, \end{array} \tag{2}$$

with variables $\{y_i\}_{i=1}^{s(2\nu)} \subseteq \mathbb{R}$.

Proposition 2.1 (see, e.g., 1, Proposition 3.1). *$\mathbb{D}(p)$ is always strictly feasible. Consequently, by strong duality, if $\mathbb{P}(p)$ is feasible, then it has an optimal solution, and there is no duality gap, i.e.,* $\max \mathbb{P}(p) = \inf \mathbb{D}(p)$.

Proof. Take any finite measure μ on $\mathbb{R}^n$, and for each $i \in \{1, \dots, s(2\nu)\}$, let $y_i = \int_{\mathbb{R}^n} u_i \,\mathrm{d}\mu$. Then, for any $z \in \mathbb{R}^{s(2\nu)} \setminus \{0\}$, it holds that $z^\top(\sum_{i=1}^{s(2\nu)} B_i^\nu y_i)z = \int_{\mathbb{R}^n} (z^\top \mathbf{v}_\nu)^2 \,\mathrm{d}\mu > 0$. □

Theorem 2.2 (see, e.g., 1, Theorem 3.2).

- *If $p - p^*$ is sum of squares, then $\mathbb{P}(p)$ and $\mathbb{D}(p)$ have an optimal solution and* $p^* = \max \mathbb{P}(p) = \min \mathbb{D}(p)$. *Moreover, for any $x \in \mathbb{R}^n$ such that $p(x) = p^*$, $(y_i)_{i=1}^{s(2\nu)} = (u_i(x))_{i=1}^{s(2\nu)}$ is an optimal solution of* $\mathbb{D}(p)$.

- *Conversely, if $\mathbb{P}(p)$ is feasible and $\mathbb{D}(p)$ has an optimal solution $(y_i)_{i=1}^{s(2\nu)} = (u_i(x))_{i=1}^{s(2\nu)}$ for some $x \in \mathbb{R}^n$, then $p - p^*$ is sum of squares and $p^* = \max \mathbb{P}(p) = \min \mathbb{D}(p)$.*

Proof.

- The first part is obvious by definition of sum-of-squares polynomials and $\mathbb{P}(p)$. The second part follows directly from the fact that $\sum_{i=1}^{s(2\nu)} c_i y_i = p(x) = p^*$. Since, by weak duality $\inf \mathbb{D}(p) \geq \max \mathbb{P}(p) \geq p^*$, it follows that $(y_i)_{i=1}^{s(2\nu)}$ is optimal for $\mathbb{D}(p)$.
- It holds that $p^* \leq p(x) = \sum_{i=1}^{s(2\nu)} c_i y_i = \min \mathbb{D}(p) = \max \mathbb{P}(p) \leq p^*$, where the third equality follows from Proposition 2.1. Hence, $p^* = \max \mathbb{P}(p)$, and thus $p - p^*$ is sum of squares.

□

Definition 2.3. A sequence $(y_i)_{i=1}^{s(2\nu)} \in \mathbb{R}^{s(2\nu)}$ is said to be *atomic* if there is $x \in \mathbb{R}^n$ such that for all $i \in \{1, \ldots, s(2\nu)\}$, $y_i = u_i(x)$.

Corollary 2.4. *Assume $\mathbb{P}(p)$ is feasible. Then $p - p^*$ is sum of squares if and only if $\mathbb{D}(p)$ has an atomic optimal solution.*

3 CONSTRAINED OPTIMIZATION

Fix $\nu \in \mathbb{N}$, $H \in \mathbb{N}_{>0}$ and $\{\nu_h\}_{h=1}^{H} \subseteq \mathbb{N}_{>0}$. Let $p = \sum_{i=1}^{s(2\nu)} c_i u_i$, where for each $i \in \{1, \ldots, s(2\nu)\}$, $c_i \in \mathbb{R}$, and for each $h \in \{1, \ldots, H\}$, let $q_h = \sum_{i=1}^{s(2\nu_h)} c_{h,i} u_i$, where for each $i \in \{1, \ldots, s(2\nu_h)\}$, $c_{h,i} \in \mathbb{R}$. Denote $Q = (q_h)_{h=1}^{H}$.

The goal is to find the minimum of p over the set $S(Q) = \bigcap_{h=1}^{H} \{x \in \mathbb{R}^n : q_h(x) \geq 0\}$: $p_Q^* := \min_{x \in S(Q)} p(x)$. Computing p_Q^* is in general difficult. A relaxation consists in finding a value $t \in \mathbb{R}$ such that $p - t = s_0 + \sum_{h=1}^{H} q_h s_h$ for some sum-of-squares polynomials $s_0, \ldots, s_H$, of respective degrees $2\pi_0, \ldots, 2\pi_H$, with $\{\pi_h\}_{h=0}^{H} \subseteq \mathbb{N}$. Any such value t will then be a lower bound on p_Q^*.

Hence, we can formulate the following optimization problem to optimize on such t:

$$
\begin{array}{llll}
\mathbb{P}(p, Q) : & \sup_{(X_h)_{h=0}^{H}, t} & t & \\
& \text{s.t.} & X_h \succeq 0, \quad \forall\, h \in \{0, \ldots, H\} & \\
& & \langle B_1^{\pi_0}, X_0 \rangle + \sum_{h=1}^{H} c_{h,1} \langle B_1^{\pi_h}, X_h \rangle = c_1 - t, & \\
& & \langle B_i^{\pi_0}, X_0 \rangle + \sum_{h=1}^{H} \sum_{(i_1,i_2)\sim(i,\nu_h,\pi_h)} c_{h,i_1} \langle B_{i_2}^{\pi_h}, X_h \rangle = c_i, & \\
& & \qquad \forall\, i \in \{2, \ldots, s(2\nu)\}, &
\end{array} \tag{3}
$$

with variables $X_h = X_h^\top \in \mathbb{R}^{s(\pi_h) \times s(\pi_h)}$ for all $h \in \{0, \ldots, H\}$, and $t \in \mathbb{R}$, and where for each $h \in \{0, \ldots, H\}$, $(i_1, i_2) \sim (i, \nu_h, \pi_h)$ means that $i_1 \in \{1, \ldots, \nu_h\}$, $i_2 \in \{1, \ldots, \pi_h\}$ and $u_{i_1} u_{i_2} = u_i$. The dual of (1) is the problem:

$$
\begin{array}{llll}
\mathbb{D}(p, Q) : & \inf_{(y_i)_{i=1}^{s(2\nu)}} & \sum_{i=1}^{s(2\nu)} c_i y_i & \\
& \text{s.t.} & \sum_{i=1}^{s(2\nu)} B_i^{\pi_0} y_i \succeq 0, & \\
& & \sum_{i=1}^{s(2\nu)} \sum_{(i_1,i_2)\sim(i,\nu_h,\pi_h)} c_{h,i_1} B_{i_2}^{\pi_h} y_i \succeq 0, \quad \forall\, h \in \{1, \ldots, H\}, & \\
& & y_1 = 1, &
\end{array} \tag{4}
$$

with variables $\{y_i\}_{i=1}^{s(2\nu)} \subseteq \mathbb{R}$.

Assumption 3.1. *$S(Q)$ has nonempty interior, and for all $h \in \{1, \ldots, H\}$, $q_h \neq 0$.*

Proposition 3.2. *Under Assumption 3.1, $\mathbb{D}(p, Q)$ is always strictly feasible. Consequently, by strong duality, if $\mathbb{P}(p, Q)$ is feasible, then it has an optimal solution, and there is no duality gap, i.e., $\max \mathbb{P}(p, Q) = \inf \mathbb{D}(p, Q)$.*

Proof. Take any finite measure μ on $S(Q)$, and for each $i \in \{1, \ldots, s(2\nu)\}$, let $y_i = \int_{\mathbb{R}^n} u_i \, \mathrm{d}\mu$. Then, for any $z \in \mathbb{R}^{s(2\nu)} \setminus \{0\}$, it holds that $z^\top(\sum_{i=1}^{s(2\nu)} B_i^\nu y_i)z = \int_{\mathbb{R}^n} (z^\top \mathbf{v}_\nu)^2 \, \mathrm{d}\mu > 0$, and for every $h \in \{1, \ldots, H\}$, $z^\top(\sum_{i=1}^{s(2\nu)} \sum_{(i_1,i_2)\sim(i,\nu_h,\pi_h)} c_{h,i_1} B_{i_2}^{\pi_h} y_i)z = \int_{\mathbb{R}^n} q_h \cdot (z^\top \mathbf{v}_\nu)^2 \, \mathrm{d}\mu > 0$. □

Theorem 3.3. *Let Assumption 3.1 hold.*

- *If $p - p_Q^* = s_0 + \sum_{h=1}^H q_h s_h$ for some sum-of-squares polynomials $s_0, \ldots, s_H$, of respective degrees $2\pi_0, \ldots, 2\pi_H$, then $\mathbb{P}(p)$ and $\mathbb{D}(p)$ have an optimal solution and $p_Q^* = \max \mathbb{P}(p, Q) = \min \mathbb{D}(p, Q)$. Moreover, for any $x \in S(Q)$ such that $p(x) = p_Q^*$, $(y_i)_{i=1}^{s(2\nu)} = (u_i(x))_{i=1}^{s(2\nu)}$ is an optimal solution of $\mathbb{D}(p, Q)$.*
- *Conversely, if $\mathbb{P}(p, Q)$ is feasible and $\mathbb{D}(p, Q)$ has an optimal solution $(y_i)_{i=1}^{s(2\nu)} = (u_i(x))_{i=1}^{s(2\nu)}$ for some $x \in S(Q)$, then $p - p_Q^* = s_0 + \sum_{h=1}^H q_h s_h$ for some sum-of-squares polynomials $s_0, \ldots, s_H$, of respective degrees $2\pi_0, \ldots, 2\pi_H$, and $p_Q^* = \max \mathbb{P}(p, Q) = \min \mathbb{D}(p, Q)$.*

Proof. Same as the proof of Theorem 2.2. □

REFERENCES

[1] Jean B Lasserre. 2001. Global optimization with polynomials and the problem of moments. *SIAM Journal on Optimization* 11, 3 (2001), 796–817. https://doi.org/10.1137/S1052623400366802