ECE285: Semidefinite and sum-of-squares optimization **Winter 2023**

Lecture 16: Dual side: moment problems

Lecturer: Yang Zheng *Scribe: Yang Zheng*

Learning goals:

1. Moment interpretation
2. Nonnegative measures on intervals

The sets of nonnegative polynomials and SOS polynomials are convex cones, and they have rich duality structures. In this lecture, we introduce their duals in the univariate case and explain their interpretation in terms of moment problems.

1 Duality and sum-of-squares

Recall that the cone of nonnegative univariate polynomials of degree $2d$ is defined as

$$P_{2d} = \left\{ (p_0, p_1, \dots, p_{2d}) \in \mathbb{R}^{2d+1} \mid \sum_{k=0}^{2d} p_0 x^k \geq 0, \forall x \in \mathbb{R} \right\}. \tag{1}$$

This is equivalent to the cone of SOS univariate polynomials of degree $2d$. We have seen that P_{2d} is a proper cone in $\mathbb{R}^{2d+1}$, and it has a semidefinite representation:

$$p \in P_{2d} \quad \Leftrightarrow \quad \exists Q \in \mathbb{S}_+^{2d+1}, \text{ s.t.} \sum_{i+j=k} Q_{ij} = p_k, k = 0, \dots, 2d. \tag{2}$$

Note that any conic program over P_{2d} is a semidefinite program.

We now discuss the dual of P_{2d}. For any $x \in \mathbb{R}$, we consider the following vector

$$y_x := (1, x, x^2, \dots, x^{2d}) \in \mathbb{R}^{2d+1}. \tag{3}$$

Let M_{2d} be the curve drawn by these vectors in $\mathbb{R}^{2d+1}$, known as *moment curve* of degree $2d$:

$$M_{2d} = \{y_x \in \mathbb{R}^{2d+1} \mid x \in \mathbb{R}\}. \tag{4}$$

Recall the definition of the dual of a set $S \subset \mathbb{R}^n$: $S^* = \{y \in \mathbb{R}^n \mid y^\mathsf{T} x \geq 0, \forall x \in S\}$. Observe that the definition of P_{2d} in (1) simply states that P_{2d} is the dual cone of M_{2d}, i.e.,

$$M_{2d}^* = P_{2d}.$$

Then, we automatically get that the dual of P_{2d} is the closure of the cone of M_{2d}, i.e.,

$$P_{2d}^* = \text{cl cone}(M_{2d}). \tag{5}$$

Remark 16.1 (Closure in (5)). *Note that the cone generated by the moment curve M_{2d} is generally not closed, so we need to keep the closure operation in* (5). *For instance, we can verify that*

$$(0, 0, 1) \in \textit{cl cone}(M_2), \quad \textit{but} \quad (0, 0, 1) \notin \textit{cone}(M_2).$$

Indeed, we cannot write $(0,0,1)$ *as a conic combination of the elements in* $\{y_x \in \mathbb{R}^3 \mid x \in \mathbb{R}\}$ *since the first element of*

$$\sum_{k=1}^{m} \lambda_k y_{x_k} = \sum_{k=1}^{m} \lambda_k (1, x_k, x_k^2), \quad \lambda_k \geq 0$$

is strictly positive unless all $\lambda_k = 0$*. On the other hand, we have*

$$\begin{aligned}\lim_{x\to\infty} \frac{1}{x^2} y_x &= \lim_{x\to\infty} \frac{1}{x^2}(1, x, x^2) \\ &= (0,0,1).\end{aligned}$$

The main reason why $\text{cone}(M_{2d})$ *is not closed is that the value of* x *can be arbitrarily large on the real line in* (4)*. If we restrict the definition* x *in* (4) *to be in a compact interval* $[a,b]$*, then the set* $\text{cone}(M_{2d})$ *is closed.* □

Remark 16.2 (Point evaluation)**.** *The vector* y_x*, defined in* (3)*, can be interpreted as linear functional on the space of polynomials. A linear functional on polynomials is of the form* $\mathbb{R}[x] \to \mathbb{R}$ *that takes a polynomial and returns a real number. Then, given a polynomial* $p(x)$ *of degree* $2d$ *with coefficients* $(p_0, p_1, \ldots, p_{2d})$*, the inner product between* y_x *and* p *is just nothing but*

$$p(x) = \langle p, y_x \rangle,$$

i.e., the point evaluation of $p(x)$ *at* $x \in \mathbb{R}$*. It is clear that point evaluation* $y_x \in P_{2d}^*$ *(since the point evaluation of any nonnegative polynomials at* x *is nonnegative). The relation in* (5) *suggests that any element in* P_{2d}^* *is a conic combination of point evaluations (up to closure).* □

1.1 Moment interpretation of P_{2d}^*

A particular and important interpretation of the dual cone P_{2d}^* is in terms of the *truncated moment sequence* of probability measures. Consider the following question

Definition 16.1 (Truncated moment problem)**.** *Given a vector* $y = (y_0, y_1, y_2, \ldots, y_{2d}) \in \mathbb{R}^{2d+1}$*, does there exist a nonnegative measure* μ *on* $\mathbb{R}$ *such that*

$$y_k = \int_{\mathbb{R}} x^k d\mu, \quad k = 0, \ldots, 2d?$$

If the answer is positive, then we say the vector y is a valid moment vector. It is clear that not any vector $y = (y_0, y_1, y_2, \ldots, y_{2d}) \in \mathbb{R}^{2d+1}$ is a valid moment vector. For example, we must have $y_k \geq 0$ for any k even. Also, we must have

$$y_2 + (y_0 - 2)y_1^2 = \int_{\mathbb{R}} (x - y_1)^2 d\mu \geq 0.$$

Furthermore, given any nonnegative polynomial $p(x)$ on $\mathbb{R}$, we must have

$$\int_{\mathbb{R}} p(x) d\mu \geq 0.$$

In this case, upon denoting $p(x) = \sum_{k=0}^{2d} p_k x^k$, we must have

$$0 \leq \int_{\mathbb{R}} p(x) d\mu = \int_{\mathbb{R}} \sum_{k=0}^{2d} p_k x^k d\mu = \sum_{k=0}^{2d} p_k \int_{\mathbb{R}} x^k d\mu = \sum_{k=0}^{2d} p_k y_k.$$

In other words, if y is a valid moment vector, then we must have

$$\langle p, y \rangle \geq 0, \forall p \in P_{2d}.$$

By the definition of dual cones, we have $y \in P_{2d}^*$. Note that P_{2d}^* is the closure of all valid moment vectors, so not every point on the boundary of P_{2d}^* is a valid moment vector (an explicit counterexample is given in Example 16.1.)

Remark 16.3. *Note that the vector y_x defined in (3) is actually a valid moment vector for any $x \in \mathbb{R}$: y_x is simply the moment vector for the Dirac probability measure δ_x that takes the value of 1 at a single point x. Any conic combination of these vectors is a valid moment vector as well. If $y = \sum_{i=1}^N \lambda_i y_{x_i}, \lambda_1, \dots, \lambda_N \geq 0$, then y is the moment vector of the atomic measure $\mu = \sum_{i=1}^N \lambda_i \delta_{x_i}$. Therefore, any element of cone(M_{2d}) is a valid moment vector.*

From this perspective, the dual of SOS univariate polynomials is the set of valid moment vectors (up to closure). □

Primal and dual formulation of polynomial optimization.

We have seen that

$$\min_{x \in \mathbb{R}} p(x) \qquad = \qquad \max \gamma \quad \text{subject to} \ \ p - \gamma \in P_{2d}. \tag{6}$$

where $p(x)$ is a polynomial of degree $2d$. Note that (6) is a particular conic program over P_{2d}. We can derive its dual. Let $y \in P_{2d}^*$ be the dual variable for the constraint $p - \gamma \in P_{2d}$. Then we have

$$\langle p - \gamma, y \rangle \geq 0 \qquad \Rightarrow \qquad \gamma y_0 \leq \langle p, y \rangle.$$

Since we are interested in deriving an upper bound on γ, we let $y_0 = 1$. Then, the dual problem becomes

$$\begin{aligned} \min_{y} \quad & \langle p, y \rangle \\ \text{subject to} \quad & y \in P_{2d}^*, \ \ y_0 = 1. \end{aligned} \tag{7}$$

Again, this is a conic program over the dual cone P_{2d}^*. We know that P_{2d}^* corresponds to moments of nonnegative measures (up to closure). The constraint $y_0 = 1$ means that we restrict ourselves to probability measures. Thus, the dual problem (7) is equivalent to

$$\begin{aligned} \min_{\mu} \quad & \int_{\mathbb{R}} p(x) d\mu \\ \text{subject to} \quad & \mu \text{ is a probability measure on } \mathbb{R}. \end{aligned} \tag{8}$$

It is interesting to compare (8) with (6). It is not difficult to see that these two problems have the same cost value. Indeed, let $p^* = \min_{x \in \mathbb{R}} p(x)$ and x^* is be a minimizer of $p(x)$. Then, for any probability measure μ, we have

$$\int_{\mathbb{R}} p(x) d\mu \geq \int_{\mathbb{R}} p^* d\mu = p^*.$$

On the other hand, if we choose $\mu = \delta_{x^*}$, we have

$$\int_{\mathbb{R}} p(x) d\mu = p^*.$$

Even though $p(x)$ can be a nonconvex polynomial, (8) is a linear optimization problem in μ. However, (8) is an infinite-dimensional problem since the underlying space is the space of probability measures on $\mathbb{R}$. Finally, note that the objective function in (8) only depends on the moments up to degree $2d$ of the measure μ; indeed, it is equivalent to

$$\begin{aligned} \min_{y} \quad & \sum_{k=0}^{2d} y_k p_k \\ \text{subject to} \quad & y_k = \int_{\mathbb{R}} x^k d\mu, k = 0, \dots, 2d \\ & \mu \text{ is a probability measure on } \mathbb{R}, \end{aligned} \tag{9}$$

which is the same as (7).

1.2 SDP representation of P_{2d}^*

The cone P_{2d} has an SDP representation in (2). Here, we derive an SDP representation for P_{2d}^*. By definition, given $y \in P_{2d}^*$, we have

$$\langle y, p \rangle \geq 0, \forall p \in P_{2d}. \tag{10}$$

Since P_{2d} is the set of nonnegative univariate polynomials, each element in P_{2d} can be written as an SOS polynomial of degree d, i.e.,

$$p(x) = q(x)^2 = \sum_{k=0}^{2d} \left(\sum_{i+j=k} q_i q_j \right) x^k.$$

Therefore, we have

$$\begin{aligned}
\langle y, p \rangle &= \langle y, q^2 \rangle \\
&= \sum_{0 \leq i,j \leq d} q_i q_j y_{i+j} \\
&= q^\mathsf{T} H(y) q \geq 0, \qquad \forall q \in \mathbb{R}^{d+1}
\end{aligned}$$

where $H(y)$ is the Hankel matrix associated with y:

$$H(y) = \begin{bmatrix} y_0 & y_1 & \cdots & y_d \\ y_1 & y_2 & \cdots & y_{d+1} \\ \vdots & \vdots & \ddots & \vdots \\ y_d & y_{d+1} & \cdots & y_{2d} \end{bmatrix}. \tag{11}$$

The reasoning above indicates that $y \in P_{2d}^*$ if and only if $H(y) \succeq 0$.

Theorem 16.1. *$P_{2d}^* = \{(y_0, y_1, \ldots, y_{2d}) \in \mathbb{R}^{2d+1} \mid H(y) \succeq 0\}$ where $H(y)$ is the Hankel matrix associated with y, defined in* (11).

Now it is clear that the dual problem (7) is an SDP of the form

$$\begin{aligned}
\min_{y} \quad & \langle p, y \rangle \\
\text{subject to} \quad & \begin{bmatrix} y_0 & y_1 & \cdots & y_d \\ y_1 & y_2 & \cdots & y_{d+1} \\ \vdots & \vdots & \ddots & \vdots \\ y_d & y_{d+1} & \cdots & y_{2d} \end{bmatrix} \succeq 0, \quad y_0 = 1.
\end{aligned} \tag{12}$$

We have seen previously the original polynomial optimization problem (6) is also an SDP of the form

$$\begin{aligned}
\max_{\gamma} \quad & \gamma \\
\text{subject to} \quad & p_0 - \gamma = Q_{00}, \\
& p_k = \sum_{i+j=k} Q_{ij}, k = 1, \ldots, 2d, \\
& Q \succeq 0.
\end{aligned} \tag{13}$$

One can verify that (13) and (12) are also dual to each other.

Remark 16.4. *Theorem 16.1 characterizes the set P_{2d}^* that is the closure of all valid moment vectors on the real line. Therefore, if $H(y)$ is strictly positive definite, (i.e., y is in the interior of P_{2d}^*), then y is*

valid moment vector, and we can find a representing nonnegative measure. However, if $H(y)$ is only positive semidefinite, (i.e., y is on the boundary of P_{2d}^*), then y may not be a valid moment vector, and there may exist no representing nonnegative measure for y. An explicit counterexample is given in [1, Remark 3.147].
□

Example 16.1 ([1, Remark 3.147])**.** *As mentioned above, $H(y) \succeq 0$ characterizes the closure of the set of valid moment vectors, but not necessarily the whole set. Consider $y = (1, 0, 0, 0, 1)$, and its Hankel matrix is*

$$H(y) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \succeq 0.$$

Although this matrix is positive semidefinite, there is no nonneagtive measure corresponding to those moments (note that it requires $y_2 = \int x^2 d\mu = 0$). However, the parameterized atomic measure by

$$\mu_\epsilon = \frac{\epsilon^4}{2}\delta_{\frac{1}{\epsilon}} + (1 - \epsilon^4)\delta_0 + \frac{\epsilon^4}{2}\delta_{-\frac{1}{\epsilon}},$$

has first five moments as

$$(1, 0, \epsilon^2, 0, 1),$$

and thus as $\epsilon \to 0$ they converge to $y = (1, 0, 0, 0, 1)$. □

2 Nonnegative measures on intervals

Recall from Lecture 13. We have characterized the nonnegative univariate polynomial on intervals $[-1, 1]$. For convenience, we state the theorem below.

Theorem 16.2. *Consider a univariate polynomial $p(x)$.*

- *If $p(x)$ is of even degree $2d$, then $p(x) \geq 0, \forall x \in [-1, 1]$ if and only if there exists SOS polynomial $s_1(x)$ of degree $2d$ and $s_2(x)$ of degree $2d - 2$ such that*

$$p(x) = s_1(x) + (1 - x^2)s_2(x). \tag{14}$$

- *If $p(x)$ is of odd degree $2d + 1$, then $p(x) \geq 0, \forall x \in [-1, 1]$ if and only if there exists SOS polynomial $s_1(x)$ of degree $2d$ and $s_2(x)$ of degree $2d$ such that*

$$p(x) = (1 - x)s_1(x) + (1 + x)s_2(x). \tag{15}$$

We can also characterize their duals. Let $P_{2d}[-1, 1]$ be the cone of polynomials of degree $2d$ nonnegative on $[-1, 1]$. Then, we can represent this set as

$$P_{2d}[-1, 1] = \{(p_0, \ldots, p_{2d}) \in \mathbb{R}^{2d+1} \mid \exists s_1 \in P_{2d}, s_2 \in P_{2d-2}, \text{ s.t. } p(x) = s_1(x) + (1 - x^2)s_2(x)\}.$$

The dual $P_{2d}^*[-1, 1]$ has a moment interpretation of nonnegative measures on interval $[-1, 1]$.

Definition 16.2 (Truncated moment problem over an interval)**.** *Given a vector $y = (y_0, y_1, y_2, \ldots, y_{2d}) \in \mathbb{R}^{2d+1}$, does there exist a nonnegative measure μ on $[-1, 1]$ such that*

$$y_k = \int_{-1}^{1} x^k d\mu, \quad k = 0, \ldots, 2d\,?$$

The answer is fully characterized by $P_{2d}^*[-1, 1]$. If there exists a nonnegative measure μ on $[-1, 1]$ with moments $y = (y_0, y_1, y_2, \ldots, y_{2d}) \in \mathbb{R}^{2d+1}$, then we have

$$0 \leq \int_{-1}^{1} p(x) d\mu, \quad \forall p \in P_{2d}[-1, 1].$$

Sine $p(x) = q_1^2(x) + (1-x^2)q_2^2(x)$, we have

$$
\begin{aligned}
0 &\le \int_{-1}^{1} q_1^2(x) + (1-x^2)q_2^2(x) d\mu \\
&= q_1^\mathsf{T} \begin{bmatrix} y_0 & y_1 & \cdots & y_d \\ y_1 & y_2 & \cdots & y_{d+1} \\ \vdots & \vdots & \ddots & \vdots \\ y_d & y_{d+1} & \cdots & y_{2d} \end{bmatrix} q_1 + q_2^\mathsf{T} \left(\begin{bmatrix} y_0 & y_1 & \cdots & y_{d-1} \\ y_1 & y_2 & \cdots & y_d \\ \vdots & \vdots & \ddots & \vdots \\ y_{d-1} & y_d & \cdots & y_{2d-2} \end{bmatrix} - \begin{bmatrix} y_2 & y_3 & \cdots & y_{d+1} \\ y_3 & y_4 & \cdots & y_{d+2} \\ \vdots & \vdots & \ddots & \vdots \\ y_{d+1} & y_{d+2} & \cdots & y_{2d} \end{bmatrix} \right) q_2,
\end{aligned}
$$

for any $\forall q_1 \in \mathbb{R}^{d+1}, q_2 \in \mathbb{R}^d$. We have the following theorem.

Theorem 16.3. *$y \in P_{2d}^*[-1,1]$ if and only if*

$$
\begin{bmatrix} y_0 & y_1 & \cdots & y_d \\ y_1 & y_2 & \cdots & y_{d+1} \\ \vdots & \vdots & \ddots & \vdots \\ y_d & y_{d+1} & \cdots & y_{2d} \end{bmatrix} \succeq 0, \tag{16a}
$$

$$
\begin{bmatrix} y_0 & y_1 & \cdots & y_{d-1} \\ y_1 & y_2 & \cdots & y_d \\ \vdots & \vdots & \ddots & \vdots \\ y_{d-1} & y_d & \cdots & y_{2d-2} \end{bmatrix} - \begin{bmatrix} y_2 & y_3 & \cdots & y_{d+1} \\ y_3 & y_4 & \cdots & y_{d+2} \\ \vdots & \vdots & \ddots & \vdots \\ y_{d+1} & y_{d+2} & \cdots & y_{2d} \end{bmatrix} \succeq 0. \tag{16b}
$$

It is clear that (16a) corresponds to $q_1(x)^2$ and (16b) corresponds to $(1-x^2)q_2^2(x)$. Also, note that (16a) is the same as the Hankel matrix $H(y)$ associated with y in (11). This is also easy to understand, since any nonnegative measure on interval $[-1,1]$ is also a nonnegative measure on the real line, i.e.,

$$P_{2d}^*[-1,1] \subset P_{2d}^*.$$

This is also reflected in the simple fact that $P_{2d} \subset P_{2d}[-1,1]$.

Remark 16.5 (Constructing a measure)**.** *Theorem 16.1 and Theorem 16.3 present semidfinite characterizations of P_{2d}^* and $P_{2d}^*[-1,1]$, respectively. But, how can we find an atomic measure associated to a sequence of moments y? For the univariate case, we have good algorithms. For the multivariate case, a well-known sufficient condition is so-called flat extension theory [2]; see [1, Chapter 3.5] for more discussions.* □

Notes

The preparation of this lecture was based on [3, Lectures 11 & 12]. Further reading for this lecture can refer [1, Chapter 3] and [4, Chapter 3].

References

[1] Grigoriy Blekherman, Pablo A Parrilo, and Rekha R Thomas. *Semidefinite optimization and convex algebraic geometry.* SIAM, 2012.

[2] Raúl E Curto and Lawrence A Fialkow. *Solution of the truncated complex moment problem for flat data*, volume 568. American Mathematical Soc., 1996.

[3] Hamza Fawzi. Topics in Convex Optimisation, Michaelmas 2018.

[4] Jean Bernard Lasserre. *Moments, positive polynomials and their applications*, volume 1. World Scientific, 2009.