Math 591 – Real Algebraic Geometry and Convex Optimization

Lecture 9: The dual perspective: moments
Cynthia Vinzant, Spring 2019

The moment problem. A linear function $L : \mathbb{R}[x_1, \ldots, x_n] \to \mathbb{R}$ is specified by its values at the basis of monomials:

$$L(x_1^{\alpha} \cdots x_n^{\alpha_n}) \;=\; y_\alpha \;\in\; \mathbb{R}.$$

Then $L(\sum_\alpha c_\alpha x_1^{\alpha} \cdots x_n^{\alpha_n}) = \sum_\alpha c_\alpha y_\alpha$. A classical question in functional analysis is: when does L come from integration against a measure? That is, for what values of $(y_\alpha)_{\alpha \in \mathbb{N}^n}$ does L have the form

$$L(f) \;=\; \int f \; d\mu$$

for some positive Borel measure μ on $\mathbb{R}^n$. We might also ask when we can take μ to be supported on some semialgebraic set $S = \{x \in \mathbb{R}^n : g_1(x) \geq 0, \ldots, g_s(x) \geq 0\}$. That is, a measure for which $\mu(A) = \mu(A \cap S)$ for all measurable subsets $A \subseteq \mathbb{R}^n$.

Example. One important example of a linear function is evaluation at a point $p \in \mathbb{R}^n$

$$\mathrm{ev}_p(f) \;=\; f(p),$$

which we can think of as integration against a measure supported at a single point.

We could also integrate against a measure that is supported on all of $\mathbb{R}$.

Example. For $n = 1$, we could take L to be a integration against a measure supported on all of $\mathbb{R}$. For example,

$$L(f) \;=\; \int_{-\infty}^{\infty} f(x) e^{-x^2} dx.$$

One necessary condition for L to have a representing measure μ supported on S, is that $L(f)$ is nonnegative for all polynomials $f \in \mathcal{P}(S)$ that are nonnegative on S.

Theorem (Haviland, 1930's)**.** *$L : \mathbb{R}[x_1, \ldots, x_n] \to \mathbb{R}$ has the form $L(f) \;=\; \int f \; d\mu$ for some measure μ if and only if $L(f) \geq 0$ for all $f \in \mathcal{P}(S)$.*

This is a characterization of the *dual cone* of $\mathcal{P}(S)$. Namely,

$$\mathcal{P}(S)^* \;=\; \left\{ L \in \mathbb{R}[x_1, \ldots, x_n]^* : L(f) = \int f \; d\mu \text{ for some measure } \mu \right\}.$$

The univariate case (n=1). Consider a linear function $L : \mathbb{R}[x] \to \mathbb{R}$ with

$$y_0 = L(1), \;\; y_1 = L(x), \;\; y_2 = L(x^2), \;\; y_3 = L(x^3), \ldots$$

Recall that $\mathcal{P}_1$ is the set of polynomials $f \in \mathbb{R}[x]$ that are nonnegative on $\mathbb{R}$. We saw in a previous class, that every nonnegative polynomial in $\mathbb{R}[x]$ is a sum of squares, i.e. $\mathcal{P}_1 = \mathrm{SOS}_1$.

Haviland's Theorem says that

$$L(f) = \int f\, d\mu \text{ for some measure } \mu \text{ on } \mathbb{R} \;\Leftrightarrow\; L(f) \geq 0 \text{ for all } f \in \mathcal{P}_1 = \mathrm{SOS}_1$$

$$\Leftrightarrow\; L\left(\sum_{i=1}^{k} h_i^2\right) \geq 0 \text{ for all } h_i \in \mathbb{R}[x]$$

$$\Leftrightarrow\; L\left(h^2\right) \geq 0 \text{ for all } h \in \mathbb{R}[x]$$

$$\Leftrightarrow\; \text{the quadratic form } h \mapsto L\left(h^2\right) \geq 0 \text{ is PSD on } \mathbb{R}[x].$$

We can make this more concrete by restricting the degree. That is, considering linear functions $L : \mathbb{R}[x]_{\leq 2d} \to \mathbb{R}$. Suppose $L(x^k) = y_k \in \mathbb{R}$ for $k = 0, \ldots, 2d$.

For what values of $(y_0, \ldots, y_{2d}) \in \mathbb{R}^{2d+1}$ does L belong to $(\mathcal{P}_{1,\leq 2d})^*$? Since nonnegative polynomials are sums of squares, this happens if and only if L is nonnegative on squares h^2 for $h \in \mathbb{R}[x]_{\leq d}$. Note that any $h \in \mathbb{R}[x]_{\leq d}$ can be written as $v^T m_d$ where $m_d = \begin{pmatrix} 1 & x & x^2 & \ldots & x^d \end{pmatrix}^T$ and $v \in \mathbb{R}^{d+1}$.

$$L(f) \geq 0 \text{ for all } f \in \mathcal{P}_{1,\leq 2d} \;\Leftrightarrow\; L(h^2) \geq 0 \text{ for all } h \in \mathbb{R}[x]_{\leq d}$$

$$\Leftrightarrow\; L((v^T m_d)^2) = L(v^T m_d m_d^T v) \geq 0 \text{ for all } v \in \mathbb{R}^{d+1}$$

$$\Leftrightarrow\; v^T L(m_d m_d^T) v \geq 0 \text{ for all } v \in \mathbb{R}^{d+1}$$

$$\Leftrightarrow\; L(m_d m_d^T) \in \mathrm{PSD}_{d+1}$$

Here, by $L(m_d m_d^T)$, we mean the $(d+1) \times (d+1)$ real matrix obtained by applying L to the entries of the polynomial matrix $m_d m_d^T$. Note that $m_d m_d^T$ is the matrix with (i,j)th entry x^{i+j-2}, so the (i,j)th entry of $L(m_d m_d^T)$ is y_{i+j-2}. Then $L \in (\mathcal{P}_{1,\leq 2d})^*$ if and only if the following matrix is positive semidefinite:

$$L(m_d m_d^T) \;=\; L\begin{pmatrix} 1 & x & x^2 & \ldots & x^d \\ x & x^2 & & \ddots & x^{d+1} \\ x^2 & & & & \\ \vdots & \ddots & & & x^{2d-1} \\ x^d & & & x^{2d-1} & x^{2d} \end{pmatrix} = \begin{pmatrix} y_0 & y_1 & y_2 & \cdots & y_d \\ y_1 & y_2 & & \ddots & y_{d+1} \\ y_2 & & & & \\ \vdots & \ddots & & & y_{2d-1} \\ y_d & & & y_{2d-1} & y_{2d} \end{pmatrix}.$$

Example. ($d = 1$) A linear function $L : \mathbb{R}[x]_{\leq 2} \to \mathbb{R}$ is nonnegative on nonnegative polynomials if and only if $L((a+bx)^2) \geq 0$ for all $(a,b)\mathbb{R}^2$. If $L(1) = y_0$, $L(x) = y_1$, and $L(x^2) = y_2$, then

$$L((a+bx)^2) = L\left(\begin{pmatrix} a & b \end{pmatrix}\begin{pmatrix} 1 & x \\ x & x^2 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix}\right) = \begin{pmatrix} a & b \end{pmatrix}\begin{pmatrix} L(1) & L(x) \\ L(x) & L(x^2) \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix}$$

So $L((a+bx)^2)$ is nonnegative for all $(a,b) \in \mathbb{R}^2$ if and only if the matrix

$$\begin{pmatrix} L(1) & L(x) \\ L(x) & L(x^2) \end{pmatrix} = \begin{pmatrix} y_0 & y_1 \\ y_1 & y_2 \end{pmatrix}$$

is positive semidefinite. We previously found that the dual cone of the set of nonnegative quadratic polynomials, $\mathcal{P}_{1,\leq 2}$, is

$$\overline{\mathrm{conicalHull}\{\mathrm{ev}_p \in \mathbb{R}[x]^*_{\leq 2} : p \in \mathbb{R}\}} \;\cong\; \overline{\mathrm{conicalHull}\{(1, p, p^2) : p \in \mathbb{R}\}}.$$

This equals the set of (y_0, y_1, y_2) for which $\begin{pmatrix} y_0 & y_1 \\ y_1 & y_2 \end{pmatrix}$ is positive semidefinite.

We can also do this over semialgebraic sets. Let's first do this in an example. Last class, we saw that a polynomial $f \in \mathbb{R}[x]_{\leq 2d}$ is nonnegative on the interval $[-1, 1]$ if and only if $f = \sigma_0 + \sigma_1(1 - x^2)$ where $\sigma_0 \in \mathrm{SOS}_{1,\leq 2d}$ and $\sigma_1 \in \mathrm{SOS}_{1,\leq 2d-2}$. Then

$$
\begin{aligned}
L(f) \geq 0 \text{ for all } f \in \mathcal{P}([-1,1]) \\
&\Leftrightarrow\ L(\sigma_0 + \sigma_1(1-x^2)) \geq 0 \text{ for all } \sigma_0 \in \mathrm{SOS}_{1,\leq 2d}, \sigma_1 \in \mathrm{SOS}_{1,\leq 2d-2} \\
&\Leftrightarrow\ L(h^2) \geq 0 \ \text{ for all } \ h \in \mathbb{R}[x]_{\leq d}, \text{ and} \\
&\quad\ L(\tilde{h}^2(1-x^2)) \geq 0 \text{ for all } \ \tilde{h} \in \mathbb{R}[x]_{\leq d-1}
\end{aligned}
$$

As before, we can translate this to a PSD condition on matrices whose entries are linear in the values $L(x^k) = y_k$. Namely, L is nonnegative on $\mathcal{P}([-1,1])$, the two matrices

$$
L(m_d m_d^T) = \begin{pmatrix} y_0 & y_1 & y_2 & \cdots & y_d \\ y_1 & y_2 & & \ddots & y_{d+1} \\ y_2 & & & & \\ \vdots & \ddots & & & y_{2d-1} \\ y_d & & & y_{2d-1} & y_{2d} \end{pmatrix}
$$

and

$$
L((1-x)^2 \cdot m_{d-1} m_{d-1}^T) = \begin{pmatrix} y_0 - y_2 & y_1 - y_3 & y_2 & \cdots & y_{d-1} - y_{d+1} \\ y_1 - y_3 & y_2 - y_4 & & \ddots & y_d - y_{d+2} \\ y_2 - y_4 & & & & \\ \vdots & \ddots & & & y_{2d-3} - y_{2d-1} \\ y_{d-1} - y_{d+1} & & & y_{2d-3} - y_{2d-1} & y_{2d-2} - y_{2d} \end{pmatrix}
$$

are positive semidefinite.

Example. For $d = 1$, consider the convex cone $K = \{(a, b, c) : a + bx + cx^2 \geq 0 \text{ on } [-1, 1]\}$. By the arguments above, the dual cone is

$$
K^* \ = \ \left\{ (y_0, y_1, y_2) \in \mathbb{R}^3 \ : \ \begin{pmatrix} y_0 & y_1 \\ y_1 & y_2 \end{pmatrix} \succeq 0 \ \text{ and } y_0 - y_2 \geq 0 \right\}.
$$

Connections with point evaluations and duality. For a given semialgebraic set, $S = \{p \in \mathbb{R}^n : g_1(p) \geq 0, \ldots, g_s(p) \geq 0\}$, consider the convex cone

$$
\mathcal{P}(S)_{\leq d} \ = \{f \in \mathbb{R}[x_1, \ldots, x_n]_{\leq d} : f \geq 0 \text{ on } S\}
$$

of polynomials of degree $\leq d$ that are nonnegative on S. Then, by definition, $\mathcal{P}(S)_{\leq d}$ is the dual cone of

$$
\mathcal{M}(S)_{\leq d} \ = \ \text{conic hull}(\{\mathrm{ev}_p : p \in S\}) \subset \mathbb{R}[x_1 \ldots, x_n]^*_{\leq d}
$$

where $\mathrm{ev}_p(f) = f(p)$. Indeed, we can check that

$$
\begin{aligned}
f \in \mathcal{P}(S)_{\leq d} \ &\Leftrightarrow\ \mathrm{ev}_p(f) \geq 0 \text{ for all } p \in S \\
&\Leftrightarrow\ \sum_{i=1}^k \lambda_i \mathrm{ev}_{p_i}(f) \geq 0 \text{ for all } p_1, \ldots, p_k \in S \text{ and } \lambda_1, \ldots, \lambda_k \in \mathbb{R}_{\geq 0} \\
&\Leftrightarrow\ L(f) \geq 0 \text{ for all } L \in \text{conic hull}(\{\mathrm{ev}_p : p \in S\}).
\end{aligned}
$$

Since $(K^*)^* = \overline{K}$ for any convex cone in a finite dimensional vector space, we find that the dual cone of $\mathcal{P}(S)_{\leq d}$ is

$$\mathcal{P}(S)^*_{\leq d} = (\mathcal{M}(S)^*_{\leq d})^* = \overline{\text{conic hull}(\{\text{ev}_p : p \in S\})}.$$

On the other hand, consider the set of polynomials

$$\begin{aligned} P(g_1,\ldots,g_s)_{\leq d} &= \{\sigma_0 + \sigma_1 g_1 + \ldots + \sigma_s g_s : \sigma_i \in \text{SOS}_n \text{ and } \deg(\sigma_i g_i) \leq d\} \\ &= \text{SOS}_{n,\leq d} + \text{SOS}_{n,\leq d_1} g_1 + \ldots + \text{SOS}_{n,\leq d_s} g_s \end{aligned}$$

where $d_i = d - \deg(g_i)$ for each $i = 1,\ldots,s$.

Recall that for two convex cones C, K, the dual cone of the sum is the intersection of the dual cones, $(C+K)^* = C^* \cap K^*$. Therefore, the dual cone of $P(g_1,\ldots,g_s)_{\leq d}$ is the intersection of the dual cones of $\text{SOS}_{n,\leq d}$ and $\text{SOS}_{n,\leq d_i} g_i$. This gives that the dual cone of $P(g_1,\ldots,g_s)_{\leq d}$ is

$$\begin{aligned} P(g_1,\ldots,g_s)^*_{\leq d} &= \big\{L \in \mathbb{R}[x_1,\ldots,x_n]_{\leq d} : L(h^2) \geq 0 \text{ for } h \in \mathbb{R}[x_1,\ldots,x_n]_{\leq d}, \text{ and} \\ &\qquad L(h^2 g_i) \geq 0 \text{ for } h \in \mathbb{R}[x_1,\ldots,x_n]_{\leq d_i/2}\big\} \\ &= \big\{L \in \mathbb{R}[x_1,\ldots,x_n]_{\leq d} : L(m_{d/2} m_{d/2}^T) \succeq 0 \text{ and } L(g_i m_{d_i/2} m_{d_i/2}^T) \succeq 0 \text{ for all } i\big\} \end{aligned}$$

where m_d is the vector of monomials in $x_1,\ldots,x_n$ of degree $\leq d$. This lets us write down $P(g_1,\ldots,g_s)^*_{\leq d}$ as a spectrahedron!

Note that since $P(g_1,\ldots,g_s)_{\leq d} \subseteq \mathcal{P}(S)_{\leq d}$, we immediately have $\mathcal{P}(S)^*_{\leq d} \subseteq P(g_1,\ldots,g_s)^*_{\leq d}$. If it so happens that $P(g_1,\ldots,g_s)_{\leq d}$ equals $\mathcal{P}(S)_{\leq d}$, then we have equality in their duals as well. In particular, this lets use write (the closure) of the conic hull of point evaluations as a spectrahedron.

Example. ($n = 1$, $d = 3$) For $S = [0,1] \subset \mathbb{R}$, we saw that

$$\mathcal{P}(S)_{\leq 3} = \{x\sigma_0 + (1-x)\sigma_1 : \sigma_0, \sigma_1 \in \text{SOS}_{1,\leq 2}\}.$$

This gives equalities in the dual cones as well.

$$\begin{aligned} &\text{conic hull}\{(1,t,t^2,t^3) : t \in [0,1]\} \\ &\quad = \mathcal{P}(S)^*_{\leq 3} \\ &\quad = (x \cdot \text{SOS}_{1,\leq 2} + (1-x) \cdot \text{SOS}_{1,\leq 2})^* \\ &\quad = \{L \in \mathbb{R}[x]^*_{\leq 3} : L(xh^2) \geq 0 \text{ and } L((1-x)h^2) \geq 0 \text{ for all } h \in \mathbb{R}[x]_{\leq 1}\} \\ &\quad = \left\{L \in \mathbb{R}[x]^*_{\leq 3} : \begin{pmatrix} L(x) & L(x^2) \\ L(x^2) & L(x^3) \end{pmatrix} \succeq 0 \text{ and } \begin{pmatrix} L(1)-L(x) & L(x)-L(x^2) \\ L(x)-L(x^2) & L(x^2)-L(x^3) \end{pmatrix} \succeq 0\right\} \end{aligned}$$

Restricting to the plane $L(1) = 1$, this gives that

$$\text{conv}(\{(t,t^2,t^3) : t \in [0,1]\}) = \left\{(y_1,y_2,y_3) \in \mathbb{R}^3 : \begin{pmatrix} y_1 & y_2 \\ y_2 & y_3 \end{pmatrix} \succeq 0 \text{ and } \begin{pmatrix} 1-y_1 & y_1-y_2 \\ y_1-y_2 & y_2-y_3 \end{pmatrix} \succeq 0\right\}.$$