Optimización Semidefinida
Clase 15 - El problema de los momentos

Segundo Cuatrimestre 2021

Facultad de Ciencias Exactas y Naturales, UBA

Motivación: optimización global de polinomios

Dado $f \in \mathbb{R}[\boldsymbol{x}]$, consideramos el problema de optimización

$$\begin{aligned} f^\star = \quad & \sup \quad f(\boldsymbol{x}) \\ \text{sujeto a:} \quad & \boldsymbol{x} \in X \end{aligned}$$

y el *problema generalizado de momentos* asociado

$$\begin{aligned} \rho_{\mathsf{mom}} = \quad & \sup_{\mu \in \mathcal{M}(X)} \int_X f \, d\mu \\ \text{sujeto a:} \quad & \int_X d\mu = 1. \end{aligned}$$

Ambos problemas son equivalentes. Es decir, $f^\star = \rho_{\mathsf{mom}}$.

El problema de los momentos - Introducción

Sea X un subconjunto cerrado de $\mathbb{R}^n$ y $\mathcal{B} = \mathcal{B}(X)$ la σ-algebra generada por los conjuntos abiertos de X.

Consideramos una medida de Borel en X

$$\mu : \mathcal{B}(X) \to [0, +\infty]$$

tal que $\mu(K) < +\infty$ para cualquier subconjunto compacto K de X.

Para una medida μ en $\mathcal{B}(\mathbb{R}^n)$, definimos el soporte de μ, $\operatorname{supp}(\mu)$, como el complemento en $\mathbb{R}^n$ del mayor conjunto abierto $O \subset \mathbb{R}^n$ tal que $\mu(O) = 0$.

Medidas de Borel - Ejemplos

Ejemplos:

1. Dada una función continua no-negativa g en X, definimos para $M \in \mathcal{B}$
$$\mu(M) = \int_M g \, dx,$$
la integral de Lebesgue de g en M.
2. Dado un subconjunto finito E de X, y una función $g : E \to \mathbb{R}_{\geq 0}$, definimos para $M \in \mathcal{B}$
$$\mu(M) = \sum_{\boldsymbol{x} \in M \cap E} g(\boldsymbol{x}).$$
3. (Medida de Dirac) Como caso particular, dado un punto $\boldsymbol{x} \in X$ definimos la medida
$$\delta_{\boldsymbol{x}}(M) = \begin{cases} 1 & \text{si } \boldsymbol{x} \in M, \\ 0 & \text{si no.} \end{cases}$$

Integral de Lebesgue

A partir de una medida μ podemos definir la integral de Lebesgue con respecto a μ, $f \to \int f d\mu$.

En los ejemplos anteriores, tenemos

1. Si $\mu(M) = \int_M g dx$,
$$\int_X f d\mu = \int_X f(x)g(x)dx.$$
2. Si $\mu(M) = \sum_{x \in M \cap E} g(x)$,
$$\int_X f d\mu = \sum_{x \in M \cap E} f(x)g(x).$$

El problema generalizado de los momentos

Para $\mathcal{M}(X)$ el espacio de medidas de Borel en X positivas, definimos el problema generalizado de momentos (GMP)

$$
\begin{aligned}
\rho_{\text{mom}} = \quad & \sup_{\mu \in \mathcal{M}(X)} \int_X f \, d\mu \\
\text{sujeto a: } & \int_X h_j d\mu \leq \gamma_j, \quad j \in \Gamma_1 \\
& \int_X h_j d\mu = \gamma_j, \quad j \in \Gamma_2
\end{aligned}
$$

Aplicación: optimización global de polinomios

Dado $f \in \mathbb{R}[\boldsymbol{x}]$, consideramos el problema de optimización

$$\begin{aligned} f^\star = \quad & \sup \quad f(\boldsymbol{x}) \\ & \text{sujeto a:} \quad \boldsymbol{x} \in X \end{aligned}$$

y el problema generalizado de momentos asociado

$$\begin{aligned} \rho_{\mathsf{mom}} = \quad & \sup_{\mu \in \mathcal{M}(X)} \int_X f \, d\mu \\ & \text{sujeto a:} \int_X d\mu = 1. \end{aligned}$$

Teorema

Ambos problemas son equivalentes. Es decir, $f^\star = \rho_{\mathsf{mom}}$.

Demostración

Consideramos el caso $f^\star < +\infty$.

Como $f(\boldsymbol{x}) \leq f^\star$ para todo $\boldsymbol{x} \in X$,

$$\int_X f d\mu \leq \int_X f^\star d\mu = f^\star$$

y por lo tanto $\rho_{\text{mom}} \leq f^\star$.

Recíprocamente, para $\boldsymbol{x} \in X$ tomamos la medida de Dirac

$$\delta_{\boldsymbol{x}} \in \mathcal{M}(X),$$

que es una solución factible del problema, con valor $f(\boldsymbol{x})$, y por lo tanto $\rho_{\text{mom}} \geq f^\star$.

Momentos - Motivación

Dado un polinomio $f \in \mathbb{R}[x_1, \ldots, x_n]$,

$$f = f_{\boldsymbol{\alpha}_1} \boldsymbol{x}^{\boldsymbol{\alpha}_1} + f_{\boldsymbol{\alpha}_2} \boldsymbol{x}^{\boldsymbol{\alpha}_2} + \cdots + f_{\boldsymbol{\alpha}_s} \boldsymbol{x}^{\boldsymbol{\alpha}_s},$$

con $\boldsymbol{\alpha}_i \in \mathbb{N}_0^n$ podemos calcular su integral

$$\int_X f d\mu = f_{\boldsymbol{\alpha}_1} \int_X \boldsymbol{x}^{\boldsymbol{\alpha}_1} d\mu + f_{\boldsymbol{\alpha}_2} \int_X \boldsymbol{x}^{\boldsymbol{\alpha}_2} d\mu + \cdots + f_{\boldsymbol{\alpha}_s} \int_X \boldsymbol{x}^{\boldsymbol{\alpha}_s} d\mu,$$

conociendo las integrales de los monomios $\int_X \boldsymbol{x}^{\boldsymbol{\alpha}}$ para todo $\boldsymbol{\alpha} \in \mathbb{N}_0^n$.

Esto motiva a considerar el problema de los posibles valores que pueden tomar dichas integrales para entender las posibles medidas en $\mathcal{M}(X)$.

El problema de los momentos

Dado un conjunto semi-algebraico básico

$$S = \{\boldsymbol{x} \in \mathbb{R}^n \mid g_1(\boldsymbol{x}) \geq 0, \ldots, g_m(\boldsymbol{x}) \geq 0\} \subset \mathbb{R}^n$$

definido por polinomios $g_1, \ldots, g_m \in \mathbb{R}[\boldsymbol{x}]$, consideramos los problemas:

El problema de momentos completo. Dada una sucesión infinita $\boldsymbol{y} = (y_{\boldsymbol{\alpha}}) \subset \mathbb{R}$ de números reales, con $\boldsymbol{\alpha} \in \mathbb{N}_0^n$, ¿existe una medida μ con soporte en S tal que

$$y_{\boldsymbol{\alpha}} = \int_S \boldsymbol{x}^{\boldsymbol{\alpha}} d\mu \quad \forall \boldsymbol{\alpha} \in \mathbb{N}_0^n?$$

El problema de momentos truncado. Dado un conjunto finito $\Delta \subset \mathbb{N}_0^n$ y una sucesión finita $\boldsymbol{y} = (y_{\boldsymbol{\alpha}})_{\boldsymbol{\alpha} \in \Delta} \subset \mathbb{R}$ de números reales, ¿existe una medida μ con soporte en S tal que

$$y_{\boldsymbol{\alpha}} = \int_S \boldsymbol{x}^{\boldsymbol{\alpha}} d\mu \quad \forall \boldsymbol{\alpha} \in \Delta?$$

Ejercicio

En $\mathbb{R}^2$, para un conjunto S tal que $(3,-1) \in S$, calcular la sucesión

$$y_{\boldsymbol{\alpha}} = \int_S \boldsymbol{x}^{\boldsymbol{\alpha}} d\mu$$

para la medida

$$\mu(X) = \delta_{(3,-1)}(X).$$

Utilizar esta sucesión para calcular $p(3,-1)$ para

$$p(x,y) = x^2y^3 - 2x + 1.$$

La integral como funcional lineal

Recordemos que para $f(\boldsymbol{x}) = \sum_{\boldsymbol{\alpha}\in\mathbb{N}^n} f_{\boldsymbol{\alpha}} \boldsymbol{x}^{\boldsymbol{\alpha}}$,

$$\int_X f d\mu = \sum_{\boldsymbol{\alpha}\in\mathbb{N}^n} f_{\boldsymbol{\alpha}} \int_X \boldsymbol{x}^{\boldsymbol{\alpha}} d\mu.$$

Motivados por esta fórmula, para una secuencia infinita $\boldsymbol{y} = (y_{\boldsymbol{\alpha}}) \subset \mathbb{R}$ definimos la funcional lineal $L_{\boldsymbol{y}} : \mathbb{R}[\boldsymbol{x}] \to \mathbb{R}$ por

$$f(\boldsymbol{x}) = \sum_{\boldsymbol{\alpha}\in\mathbb{N}^n} f_{\boldsymbol{\alpha}} \boldsymbol{x}^{\boldsymbol{\alpha}} \mapsto L_{\boldsymbol{y}}(f) = \sum_{\boldsymbol{\alpha}\in\mathbb{N}^n} f_{\boldsymbol{\alpha}} y_{\boldsymbol{\alpha}}.$$

El teorema de Riesz-Haviland

El siguiente teorema (que vemos sin demostración) será clave para resolver algunos casos del problema de momentos.

Teorema (Riesz-Haviland)

Sea $\boldsymbol{y} = (y_{\boldsymbol{\alpha}})_{\boldsymbol{\alpha}\in\mathbb{N}_0^n} \subset \mathbb{R}$ *y sea* $K \subset \mathbb{R}^n$ *un conjunto cerrado. Existe una medida de Borel finita* μ *en* K *tal que*

$$\int_K \boldsymbol{x}^{\boldsymbol{\alpha}} d\mu = y_{\boldsymbol{\alpha}}, \quad \forall \boldsymbol{\alpha} \in \mathbb{N}_0^n,$$

si y solo si $L_{\boldsymbol{y}}(f) \geq 0$ *para todos los polinomios* $f \in \mathbb{R}[\boldsymbol{x}]$ *no-negativos en* K.

Observación: La implicación $\Rightarrow$ es inmediata.

Matrices de Hankel

Dada una sucesión $\boldsymbol{y} = (y_i)_{i\in\mathbb{N}_0} \subset \mathbb{R}$, definimos las matrices de Hankel $H_n(\boldsymbol{y}) \in \mathbb{R}^{(n+1)\times(n+1)}$ por

$$H_n(\boldsymbol{y})(i,j) := y_{i+j-2},$$

para todo $i, j \in \mathbb{N}$, $1 \leq i, j \leq n+1$.

Ejemplo

Para $\boldsymbol{y} = (1, 2, 3, 4, 5, 6, 7, 8, 0, 0, 0, \dots)$,

$$H_3 = \begin{pmatrix} y_0 & y_1 & y_2 & y_3 \\ y_1 & y_2 & y_3 & y_4 \\ y_2 & y_3 & y_4 & y_5 \\ y_3 & y_4 & y_5 & y_6 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \\ 4 & 5 & 6 & 7 \end{pmatrix}.$$

El problema de los momentos de Hamburger

Teorema

Sea $\boldsymbol{y} = (y_j)_{j\in\mathbb{N}_0} \subset \mathbb{R}$. *Existe una medida de Borel* μ *que representa a* $\boldsymbol{y}$ *en* $\mathbb{R}$ *si y solo si la forma cuadrática*

$$\boldsymbol{q} \mapsto s_n(\boldsymbol{q}) := \sum_{i,j=0}^{n} y_{i+j} q_i q_j \tag{1}$$

es positiva semidefinida para todo $n \in \mathbb{N}$.
Equivalentemente, $H_n(\boldsymbol{y}) \succeq 0$ *para todo* $n \in \mathbb{N}$.

Demostración

$(\Rightarrow)$ Si $y_n = \int_{\mathbb{R}} z^n d\mu(z)$, entonces

$$
\begin{aligned}
s_n(\boldsymbol{q}) &= \sum_{i,j=0}^{n} y_{i+j} q_i q_j \\
&= \sum_{i,j=0}^{n} q_i q_j \int_{\mathbb{R}} z^{i+j} d\mu(z) \\
&= \int_{\mathbb{R}} \left(\sum_{i,j=0}^{n} q_i q_j z^{i+j} \right) d\mu(z) = \int_{\mathbb{R}} \sum_{i,j=0}^{n} (q_i z^i)(q_j z^j) d\mu(z) \\
&= \int_{\mathbb{R}} (\sum_{i=0}^{n} q_i z^i)^2 d\mu(z) \geq 0.
\end{aligned}
$$

Demostración

Recíprocamente, si $H_n(\boldsymbol{y}) \succeq 0$ para todo $n \in \mathbb{N}$ y

$$f = \sum_{k=0}^{2d} f_k x^k \in \mathbb{R}[x]$$

un polinomio no-negativo en $\mathbb{R}$, veamos que $\sum_{k=0}^{2d} f_k y_k = L_{\boldsymbol{y}}(f) \geq 0$.

Recordemos que f se puede escribir como una suma de cuadrados $f = \sum_{j=1}^{r} g_j^2$. Luego,

$$L_{\boldsymbol{y}}(f) = L_{\boldsymbol{y}}(\sum_{j=1}^{r} g_j^2) = \sum_{j=1}^{r} L_{\boldsymbol{y}}(g_j^2).$$

Veamos que $L_{\boldsymbol{y}}(g^2) \geq 0$ para cualquier polinomio $g \in \mathbb{R}[x]$.

Demostración

Para todo $\boldsymbol{q} \in \mathbb{R}^{n+1}$ tenemos $\boldsymbol{q}^t H_n(\boldsymbol{y})\boldsymbol{q} \geq 0$.

Ahora

$$\begin{aligned} L_{\boldsymbol{y}}(g^2) &= L_{\boldsymbol{y}}\left((\sum_{j=0}^{d} g_j x^j)^2 \right) \\ &= L_{\boldsymbol{y}}\left(\sum_{i,j} g_i g_j x^{i+j} \right) \\ &= \sum_{0 \leq i,j \leq d} g_i g_j y_{i+j} = \boldsymbol{g}^t H_d(\boldsymbol{y})\boldsymbol{g} \geq 0. \end{aligned}$$

Como $f \geq 0$ es arbitrario, obtenemos por el Teorema 2 que $y_i = \int_{\mathbb{R}} x^i d\mu$ para alguna medida μ en $\mathbb{R}$.

Ejemplo

Repasamos la construcción paso a paso en los siguientes ejemplos:

1. Hallar $\max_{x\in\mathbb{R}} 1 - x^2$.
2. Hallar $\min_{x\in\mathbb{R}} x^2 - 2x + 5$.

En el primer caso, el problema es equivalente a

$$\rho_{\mathsf{mom}} = \sup_{\mu\in\mathcal{M}(\mathbb{R})} \int_{\mathbb{R}} 1 - x^2 d\mu$$

$$\text{sujeto a: } \int_{\mathbb{R}} d\mu = 1.$$

Ejemplo

Por linealidad,

$$\int_{\mathbb{R}} 1 - x^2 d\mu = \int_{\mathbb{R}} 1 d\mu - \int_{\mathbb{R}} x^2 d\mu$$

y definiendo $y_i = \int_{\mathbb{R}} x^i d\mu$,

$$\int_{\mathbb{R}} 1 - x^2 d\mu = y_0 - y_2.$$

La restricción $\int_{\mathbb{R}} d\mu = 1$ nos da $y_0 = 1$.

Por lo tanto, nuestro problema se convierte en

$$\rho_{\text{mom}} = \sup_{\boldsymbol{y}=(y_0,y_1,y_2,\dots)} y_0 - y_2$$

sujeto a: $y_0 = 1$,

existe una medida μ que representa a $\boldsymbol{y}$ en $\mathbb{R}$.

Ejemplo

A la vez, este último problema es equivalente a

$$\rho_{\mathsf{mom}} = \sup_{\boldsymbol{y}=(y_0,y_1,y_2,\dots)} y_0 - y_2$$

$$\text{sujeto a: } y_0 = 1,$$

$$\begin{pmatrix} y_0 & y_1 & y_2 & \cdots & y_n \\ y_1 & y_2 & y_3 & & \\ y_2 & y_3 & y_4 & & \vdots \\ \vdots & & & & \\ y_n & & \cdots & & y_{2n} \end{pmatrix} \succeq 0 \quad \text{para todo } n \in \mathbb{N}$$

Ejemplo

En particular

$$\begin{pmatrix} 1 & y_1 \\ y_1 & y_2 \end{pmatrix} \succeq 0$$

y esto nos da las restricciones $y_2 \geq 0$ y $y_2 \geq y_1^2$.

El máximo de $1 - y_2$ en esta región se da para $y_2 = 0$. Por lo tanto, obtenemos

$$\max_{x \in \mathbb{R}} 1 - x^2 \leq 1.$$

Tomando la sucesión $\boldsymbol{y} = (1, 0, 0, 0, \dots)$, vemos que

$$\max_{x \in \mathbb{R}} 1 - x^2 = 1.$$

Ejemplo

En el caso de $\min_{x\in\mathbb{R}} x^2 - 2x + 5$, obtenemos en forma análoga el problema equivalente

$$\rho_{\text{mom}} = \inf_{\boldsymbol{y}=(y_0,y_1,y_2,\dots)} 5y_0 - 2y_1 + y_2$$

$$\text{sujeto a: } y_0 = 1,$$

$$\begin{pmatrix} y_0 & y_1 & y_2 & \cdots & y_n \\ y_1 & y_2 & y_3 & & \\ y_2 & y_3 & y_4 & & \vdots \\ \vdots & & & & \\ y_n & & \cdots & & y_{2n} \end{pmatrix} \succeq 0 \quad \text{para todo } n \in \mathbb{N}$$

Vemos que al modificar el polinomio, solo cambia la función objetivo, pero no la región de factibilidad.

Ejemplo

En este caso, el mínimo de $5 - 2y_1 + y_2$ en la región $y_2 \geq 0$ y $y_2 \geq y_1^2$ se alcanza para

$$y_1 = y_2 = 1$$

y por lo tanto

$$\min_{x \in \mathbb{R}} x^2 - 2x + 5 \geq 5 - 2 + 1 = 4.$$

Tomando la sucesión $\boldsymbol{y} = (1, 1, 1, 1, \dots)$, vemos que

$$\min_{x \in \mathbb{R}} x^2 - 2x + 5 = 4.$$

Encontrar condiciones para las cuales es posible extender una sucesión finita $(y_0, y_1, \dots, y_n)$ a una sucesión representada por una medida μ se conoce como el problema truncado de los momentos.

El problema de los momentos de Stieltjes

Si en lugar de tomar $X = \mathbb{R}$ tomamos $X = \mathbb{R}_{\geq 0}$, el problema correspondiente se conoce como el problema de los momentos de Stieltjes.

Dada una sucesión $y = (y_i) \subset \mathbb{R}$, definimos la matricez de Hankel $B_n(y) \in \mathbb{R}^{(n+1)\times(n+1)}$ por $B_n(y)(i,j) := y_{i+j-1}$, para todo $i, j \in \mathbb{N}$, $1 \leq i, j \leq n+1$.

Teorema

Sea $\boldsymbol{y} = (y_j)_{j\in\mathbb{N}_0} \subset \mathbb{R}$. Existe una medida de Borel μ que representa a $\boldsymbol{y}$ en $\mathbb{R}_{\geq 0}$ si y solo si $\boldsymbol{H}_n(\boldsymbol{y}) \succeq 0$ y $\boldsymbol{B}_n(\boldsymbol{y}) \succeq 0$ para todo $n \in \mathbb{N}$.

Demostración: Ejercicio. Utilizar que $p \in \mathbb{R}[x]$ es no-negativo en $[0, +\infty)$ si y solo si $p = f_0 + x f_1$ para f_0, f_1 polinomios sumas de cuadrados.