Moment ideals of local Dirac mixtures

Alexandros Grosdos and Markus Wageringel

Osnabrück University

Graduate Student Meeting on Applied Algebra and Combinatorics, 18–20 February 2019

Dirac distributions and their moments

- Dirac distribution centered at $\xi \in \mathbb{K}$

$$\delta_\xi(x) = \begin{cases} +\infty & \text{if } x = \xi \\ 0 & \text{if } x \neq \xi \end{cases}, \qquad \int \varphi(x) \mathrm{d}\delta_\xi(x) = \varphi(\xi)$$

- Moments of a Dirac $m_i = \int x^i \mathrm{d}\delta_\xi(x) = \xi^i$

Toy problems

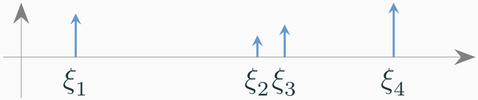
- What is the variety consisting of the closure of all points $\{[m_0 : m_1 : \dots : m_d] \mid m_i = \xi^i, \xi \in \mathbb{K}\}$?
- Assuming a sample coming from Diracs, how do we estimate the parameter ξ?

In [Améndola–Faugère–Sturmfels 2016] the corresponding questions were answered for Gaussians and their mixtures.

Let $\xi_j \in \mathbb{K}$ be points, $1 \leq j \leq r$.

Mixture of Dirac distributions

$\mu(x) := \sum_{j=1}^r \lambda_j \delta_{\xi_j}(x)$, with $0 \leq \lambda_j$ and $\lambda_1 + \cdots + \lambda_r = 1$.



Moments

$m_i = \int_{\mathbb{K}} x^i \mathrm{d}\mu(x) = \sum_{j=1}^r \lambda_j \xi_j^i$

Solved problems

- Defining equations for the moment variety.
- Recovery of parameters ξ_j, λ_j from moments m_i.

For a single Dirac we get $m_i = \xi^i$. These give the Veronese variety defined by the vanishing of all 2×2 minors of

$$H_{1,d-1} := (m_{i+j})_{\substack{0 \le i \le 1, \\ 0 \le j \le d-1}} = \left(\begin{smallmatrix} m_0 & m_1 & \dots & m_{d-1} \\ m_1 & m_2 & \dots & m_d \end{smallmatrix}\right).$$

Moment variety for mixtures of Diracs

- Parametric description:
  $\left\{[m_0 : m_1 : \dots : m_d] \mid m_i = \sum_{j=1}^r \lambda_j \xi_j^i, \lambda_j \in \mathbb{K}, \xi_j \in \mathbb{K}\right\}.$
- Implicit description (defining equations):
  $(r+1) \times (r+1)$-minors of the moment matrix

$$H_{r,d-r} := (m_{i+j})_{\substack{0 \le i \le r, \\ 0 \le j \le d-r}} = \begin{pmatrix} m_0 & m_1 & \dots & m_{d-r} \\ m_1 & m_2 & \dots & m_{d-r+1} \\ \vdots & \vdots & \ddots & \vdots \\ m_r & m_{r+1} & \dots & m_d \end{pmatrix}.$$

- Local mixture of a Dirac:

$$\mu_\xi := \delta_\xi - \alpha\delta'_\xi, \quad \alpha \in \mathbb{K},$$

so that $\int \varphi(x) \mathrm{d}\mu_\xi(x) = \varphi(\xi) + \alpha\varphi'(\xi)$.
Moments: $m_i = \xi^i + \alpha i \xi^{i-1}$.

- Mixture of local mixtures of Diracs:

$$\mu := \sum_{j=1}^{r} \lambda_j \mu_{\xi_j} = \sum_{j=1}^{r} \lambda_j \left(\delta_{\xi_j} + \alpha_j \delta'_{\xi_j}\right)$$

Problems

- Defining equations for the moment variety.
- Recovery of parameters $\xi_j, \alpha_j, \lambda_j$ from minimal number of moments.

Moment variety

The moment variety of a single local mixture is

$$\overline{\{[m_0 : m_1 : \ldots : m_d] \in \mathbb{P}^d \mid m_i = \xi^i + \alpha i \xi^{i-1} \text{ for } \xi, \alpha \in \mathbb{K}\}}.$$

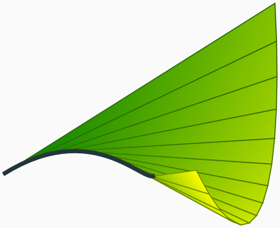
Theorem (Eisenbud 1992; —, Wageringel 2018)

For $d \geq 5$, the moment variety is defined by the relations

$$(j-i+3)m_i m_j - 2(j-i+2)m_{i+1}m_{j-1} + (j-i+1)m_{i+2}m_{j-2}$$

for all $2 \leq i \leq j \leq d-2$.

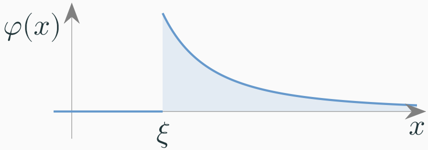
This is the tangent variety of the Veronese curve, i. e. the closure of the union of all lines tangent to the curve.



The Pareto distribution

Let $\xi, \alpha \in \mathbb{R}_{>0}$.

$$\varphi(x) := \frac{\alpha \xi^\alpha}{x^{\alpha+1}} \mathbb{1}_{\{x \geq \xi\}}, \qquad m_i = \begin{cases} \frac{\alpha}{\alpha - i} \xi^i, & i < \alpha, \\ \infty, & i \geq \alpha. \end{cases}$$



Theorem (—, Wageringel 2018)

The moment variety of the Pareto distribution is the closure of the image of the moment variety of local mixtures of Diracs under the map

$$m_i \longmapsto m_i^{-1}.$$

Let $r = 2$, i. e. $m_i = \lambda(\xi_1^i + \alpha_1 i \xi_1^{i-1}) + (1-\lambda)(\xi_2^i + \alpha_2 i \xi_2^{i-1})$.

Theorem (—, Wageringel 2018)

Let $R = \mathbb{K}[m_1, ..., m_5]$. Then there exist polynomials $g_0, g_1 \in R[x]$ of degree 4, such that

$$g_0(\xi_1 + \xi_2) = 0 = g_1(\xi_1 \xi_2).$$

Moment map: $(\xi_1, \xi_2, \alpha_1, \alpha_2, \lambda) \mapsto (m_1, m_2, ..., m_d)$

Algebraic identifiability: finitely many solutions. The moment map is finite-to-one for $d = 5$.

Rational identifiability: unique solution given rationally in the m_i. The moment map is one-to-one for $d = 6$.

Let $Z = \xi_1 + \xi_2, Y = \xi_1\xi_2$.

Strategy

1. Compute the 4 solutions of $g_0(Z) = 0$.
2. From Z and the moments m_i, uniquely determine Y via

$$(2Zk_2 - 2k_3)Y + 6Zk_2^2 - Z^2k_3 - 10k_2k_3 + 2Zk_4 - k_5 = 0$$

   (where each k_i is some polynomial in $m_1, m_2, ..., m_i$.)
3. Uniquely recover ξ_1, ξ_2 from the system $Z = \xi_1 + \xi_2, Y = \xi_1\xi_2$.
4. From ξ_1, ξ_2, uniquely recover $\lambda, \alpha_1, \alpha_2$ from the moment relations (linear problem).

Different approach: Prony's method (apolarity)

For r-mixtures of Diracs $m_i = \sum_{j=1}^{r} \lambda_j \xi_j^i$, we have:

$$\begin{aligned} H_{r-1,r} \colon \mathbb{K}[x]_{\leq r} &\longrightarrow \mathbb{K}[x]^*_{\leq r-1}, \\ p &\longmapsto \Big(q \mapsto \sum_{j=1}^{r} \lambda_j p(\xi_j) q(\xi_j) \Big). \end{aligned}$$

Recovery of parameters via Prony's method (Prony 1795)

1. Recover points ξ_j by solving $H_{r-1,r}(p) = 0$ for $p = \prod_{j=1}^{r} (x - \xi_j)$.
2. Recover parameters λ_j from linear system $m_i = \sum_{j=1}^{r} \lambda_j \xi_j^i$.

Local mixture setting: pairwise colliding nodes

$p = \prod_{j=1}^{r} \lim_{\xi_j' \to \xi_j} (x - \xi_j')(x - \xi_j) = \prod_{j=1}^{r} (x - \xi_j)^2$.

Recovery of parameters of an r-mixture from moments

An r-mixture of local mixtures of Diracs is a degenerate $2r$-mixture of Diracs for which each point has multiplicity 2 (cf. Mourrain 2017).

Strategy

Apply Prony's method to $H_{2r-1,2r} = (m_{i+j})_{\substack{0\le i\le 2r-1,\\ 0\le j\le 2r}}$ to get Prony polynomial $p^2 = \prod_{j=1}^{r}(x-\xi_j)^2 = \left(\sum_{i=0}^{r} p_i x^i\right)^2$.

This requires $4r$ moments (linear).

Recovery of parameters of an r-mixture from moments

An r-mixture of local mixtures of Diracs is a degenerate $2r$-mixture of Diracs for which each point has multiplicity 2 (cf. Mourrain 2017).

Strategy

Apply Prony's method to $H_{2r-1,2r} = (m_{i+j})_{\substack{0\leq i\leq 2r-1,\\ 0\leq j\leq 2r}}$ to get Prony polynomial $p^2 = \prod_{j=1}^{r}(x-\xi_j)^2 = \left(\sum_{i=0}^{r} p_i x^i\right)^2$.

This requires $4r$ moments (linear).

Refinement: Note that $p^2 \in \ker H_{r-1,2r}$, i. e.,

$$\begin{array}{ccccc} \mathbb{C}[x]_{\leq r} & \longrightarrow & \mathbb{C}[x]_{\leq 2r} & \xrightarrow{H_{r-1,2r}} & \mathbb{C}[x]_{\leq r-1}, \\ p & \longmapsto & p^2 & \longmapsto & 0, \end{array}$$

requiring $3r$ moments (non-linear), i. e.,

$$H_{r-1,2r} \cdot (p_0^2, 2p_0p_1, 2p_0p_2 + p_1^2, \dots, p_r^2)^\top = 0.$$

Question: Can this system of quadratic equations be solved efficiently?

Recovery of parameters of an r-mixture from moments (refined)

Algorithm

Input: Number of components $r \in \mathbb{N}$, moments $m_0, m_1, \dots$

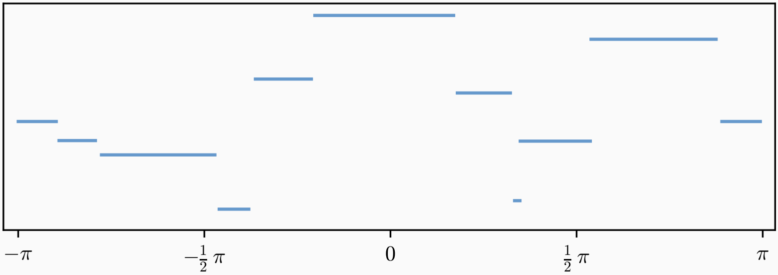
1. Let $s := r$.
2. Solve $\{p \in \mathbb{C}[x]_{\leq r} \mid H_{s,2r}(p^2) = 0\}$.
   If solution is not unique, increment s and repeat.
3. Compute roots $\xi_1, \dots, \xi_r$ of $p = \prod_{j=1}^{r}(x - \xi_j)$.
4. Compute weights $\lambda_j, \alpha_j, 1 \leq j \leq r$.

Output: Parameters satisfying $m_i = \sum_{j=1}^{r} \lambda_j(\xi_j^i + \alpha_j i \xi_j^{i-1})$.

Best case: Moments up to m_{3r} needed (if $s = r$).

Remark: The moment variety has dimension $\min(3r-1, d)$, so algebraic identifiability holds if $d \geq 3r - 1$.

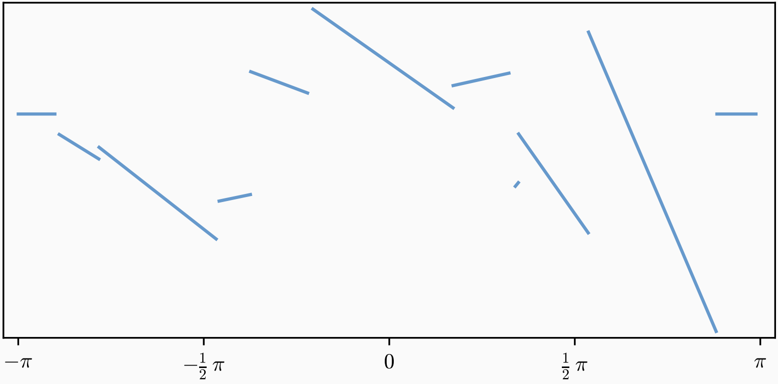
Application: Piecewise-constant functions



- A piecewise-constant function with jumps at $t_j \in [-\pi, \pi[$ corresponds to $\sum_{j=1}^{r} \lambda_j \delta_{\xi_j}, \xi_j := \mathrm{e}^{\mathrm{i}t_j}$.
- Reconstruction from Fourier samples.

Piecewise-linear function with jumps at $t_j \in [-\pi, \pi[$ corresponds to

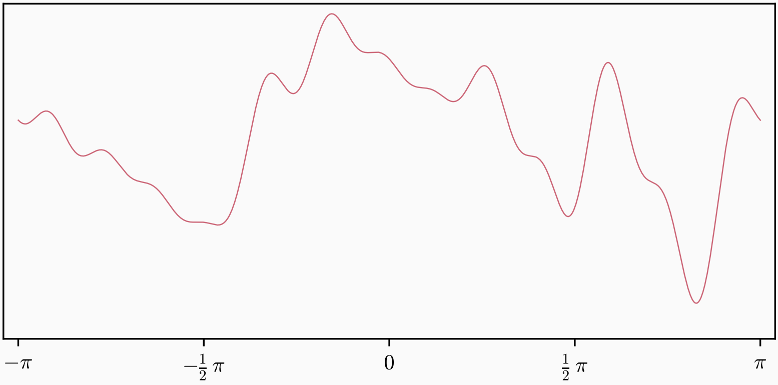
$$\sum_{j=1}^{r} \lambda_j \delta_{\xi_j} + \lambda_j' \delta_{\xi_j}', \quad \xi_j := \mathrm{e}^{\mathrm{i}t_j}.$$



Example: $r = 10$ jumping points, $3r + 1 = 31$ samples.

Piecewise-linear function with jumps at $t_j \in [-\pi, \pi[$ corresponds to

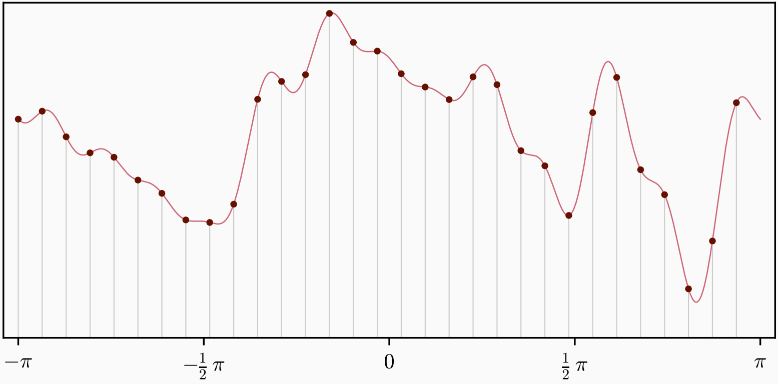
$$\sum_{j=1}^{r} \lambda_j \delta_{\xi_j} + \lambda_j' \delta_{\xi_j}', \quad \xi_j := \mathrm{e}^{\mathrm{i}t_j}.$$



Example: $r = 10$ jumping points, $3r + 1 = 31$ samples.

Piecewise-linear function with jumps at $t_j \in [-\pi, \pi[$ corresponds to

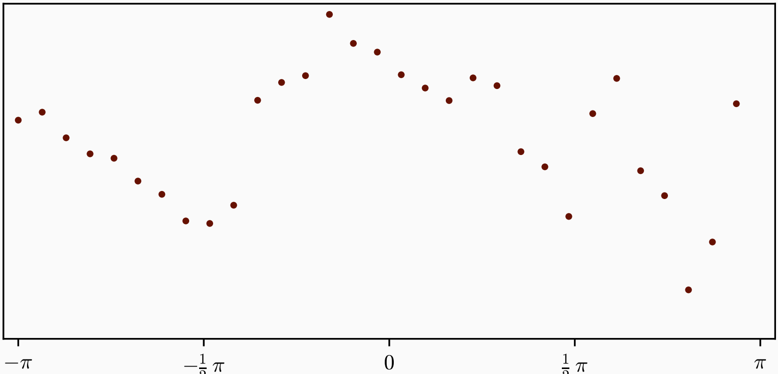
$$\sum_{j=1}^{r} \lambda_j \delta_{\xi_j} + \lambda_j' \delta_{\xi_j}', \quad \xi_j := \mathrm{e}^{\mathrm{i}t_j}.$$



Example: $r = 10$ jumping points, $3r + 1 = 31$ samples.

Piecewise-linear function with jumps at $t_j \in [-\pi, \pi[$ corresponds to

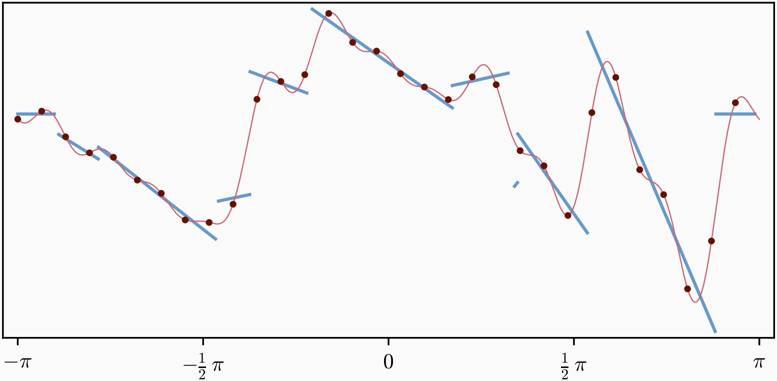
$$\sum_{j=1}^{r} \lambda_j \delta_{\xi_j} + \lambda_j' \delta_{\xi_j}', \qquad \xi_j := \mathrm{e}^{\mathrm{i} t_j}.$$



Example: $r = 10$ jumping points, $3r + 1 = 31$ samples.

Piecewise-linear function with jumps at $t_j \in [-\pi, \pi[$ corresponds to

$$\sum_{j=1}^{r} \lambda_j \delta_{\xi_j} + \lambda'_j \delta'_{\xi_j}, \quad \xi_j := \mathrm{e}^{\mathrm{i}t_j}.$$



Example: $r = 10$ jumping points, $3r + 1 = 31$ samples.

2^{ND}-ORDER LOCAL MIXTURES OF DIRACS

Local mixture of a Dirac:

$$\mu_\xi := \delta_\xi - \alpha\delta'_\xi + \beta\delta''_\xi, \quad \alpha, \beta \in \mathbb{K},$$

so that $\int \varphi(x)\mathrm{d}\mu_\xi(x) = \varphi(\xi) + \alpha\varphi'(\xi) + \beta\varphi''(\xi)$.

Moments: $m_i = \xi^i + \alpha i \xi^{i-1} + \beta i(i-1)\xi^{i-2}$.

Conjecture

For $d \geq 12$ the 2^{nd}-order moment ideal is generated by

$$c_0 m_{i+3} m_j + c_1 m_{i+2} m_{j+1} + c_2 m_{i+1} m_{j+2} + c_3 m_i m_{j+3}$$

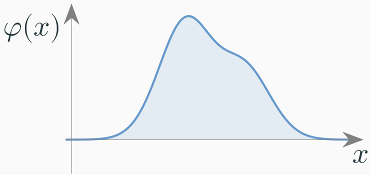
for $i \geq 0$, $j \geq 0$ and $i \geq j - 3$, where

$$\begin{aligned} c_0 &= (j-i+1)(j-i+2) & c_1 &= -3(j-i-1)(j-i+2) \\ c_2 &= 3(j-i+1)(j-i-2) & c_3 &= -(j-i-1)(j-i-2). \end{aligned}$$

Application in Gaussian local mixtures

[Marriott 2002] considers local mixture models, for example the local Gaussian distribution has p.d.f.

$$\varphi(x) := \varphi_{\xi,\sigma}(x) + \alpha \frac{\partial}{\partial \xi} \varphi_{\xi,\sigma}(x) + \beta \frac{\partial^2}{\partial \xi^2} \varphi_{\xi,\sigma}(x).$$



These can be expressed as a convolution of $\mathcal{N}_{0,\sigma}$ and $\mu_\xi = \delta_\xi - \alpha\delta'_\xi + \beta\delta''_\xi$.

Thank you!

References

Améndola, C., J.-C. Faugère, and B. Sturmfels (2016). “Moment varieties of Gaussian mixtures”. In: *J. Algebr. Stat.* 7.1, pp. 14–28.

Eisenbud, D. (1992). “Green’s conjecture: an orientation for algebraists”. In: *Free resolutions in commutative algebra and algebraic geometry (Sundance, UT, 1990)*. Vol. 2. Res. Notes Math. Jones and Bartlett, Boston, MA, pp. 51–78.

Mourrain, B. (Oct. 2017). “Polynomial–Exponential Decomposition From Moments”. In: *Foundations of Computational Mathematics*.

Marriott, P. (Mar. 2002). “On the local geometry of mixture models”. In: *Biometrika* 89, pp. 77–93.

Grosdos Koutsoumpelias, A. and M. Wageringel (Sept. 2018). “Moment ideals of local Dirac mixtures”. In: *ArXiv e-prints*.