Approximating atomic measures from trigonometric moments

Paul Catala*. Ongoing work with M. Hockmann*, S. Kunis* and M. Wageringel*
*University of Osnabrück, Institute for Mathematics.

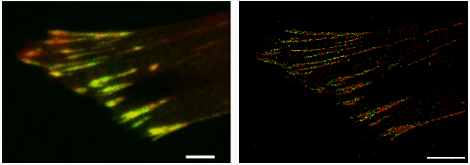
ICCHA, 15.09.21

Motivations

Sparse super-resolution. Estimate $\mu = \sum_{j=1}^{s} \lambda_j \delta_{x_j}$ (with $\lambda_j \in \mathbb{C}$, $x_j \in \mathbb{T}^d$) from

$$c_k = \int e^{-2\imath\pi\langle k, x\rangle} \mathrm{d}\mu(x)$$

for k such that $\|k\|_\infty \leqslant n$.


[1]

- ubiquitous problem in image processing (low-pass filtering)
- relevant model for compressed sensing, mixture estimation, ...
- sharp recovery algorithms / computational tractability

[1] images from the cell image library (http://cellimagelibrary.org/)

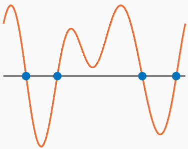
Decoding the Geometry of the Support

Support identification.

- Prony[1] / Subspace methods[2,3] (SVD)
  support = common roots *of a few polynomials*
- Dual certificates[4] (∞-dimensional optimization)
  support $\subset$ saturation points of one polynomial

Approximation.

- Christoffel functions[5,6] (SVD)
  support $\subset$ *level sets of a rational function*



1 *R. de Prony, 1795*, Sur les Lois de la Dilatabilité des Fluides Elastiques ...

2 *Schmidt, 1986*, Multiple emitter location and signal parameter estimation

3 *Roy and Kailath, 1989*, ESPRIT- Estimation of Signal Parameters vie Rotational Invariance Techniques

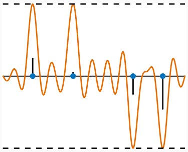
Decoding the Geometry of the Support

Support identification.

- Prony / Subspace methods (SVD)
  support = common roots of a few polynomials
- Dual certificates[1] (∞-dimensional optimization)
  support $\subset$ saturation *points of one polynomial*

Approximation.

- Christoffel functions[2,3] (SVD)
  support $\subset$ *level sets of a rational function*



[1] *Candès and Fernandez-Granda, 2014,* Towards a Mathematical Theory of Super-Resolution

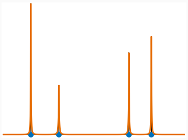
Decoding the Geometry of the Support

Support identification.

- Prony / Subspace methods (SVD)
  support = common roots of a few polynomials
- Dual certificates (∞-dimensional optimization)
  support $\subset$ saturation points of one polynomial

Approximation.

- Christoffel functions[1,2](SVD)
  support $\subset$ level sets *of a rational function*



1 *Lasserre and Pauwels, 2016,* Sorting Out Typicality with the Inverse Moment Matrix SOS polynomial

2 *Pauwels et al., 2020,* Data Analysis from Empirical Moments and the Christoffel Function

Decoding the Geometry of the Support

Support identification.

- Prony / Subspace methods (SVD)

  support = common roots of a few polynomials

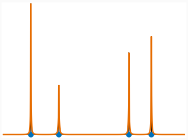
- Dual certificates (∞-dimensional optimization)

  support $\subset$ saturation points of one polynomial

Approximation.

- Christoffel functions[1,2](SVD)

  support $\subset$ level sets *of a rational function*



In this work.

$\rightarrow$ computationally efficient polynomial approximation of μ, with rate

$\rightarrow$ polynomial approximation of unweighted counterpart $\mu^1 = \sum \delta_{x_j}$

[1] *Lasserre and Pauwels, 2016*, Sorting Out Typicality with the Inverse Moment Matrix SOS polynomial

[2] *Pauwels et al., 2020*, Data Analysis from Empirical Moments and the Christoffel Function

Preliminaries

Topology of Radon Measures

$d \in \mathbb{N} \setminus \{0\}$, $\mathbb{T} \stackrel{\text{def.}}{=} \mathbb{R}/\mathbb{Z}$ Torus

$\mu \in \mathcal{M}(\mathbb{T}^d)$

$\varphi \in \mathscr{C}(\mathbb{T}^d)$

Topology of Radon Measures

$d \in \mathbb{N} \setminus \{0\}$, $\mathbb{T} \overset{\text{def.}}{=} \mathbb{R}/\mathbb{Z}$ Torus

$\mu \in \mathcal{M}(\mathbb{T}^d)$

weak-* topology

$\varphi \in \mathscr{C}(\mathbb{T}^d)$

weak topology

Topology of Radon Measures

$d \in \mathbb{N} \setminus \{0\}$, $\mathbb{T} \overset{\text{def.}}{=} \mathbb{R}/\mathbb{Z}$ Torus

$\mu \in \mathcal{M}(\mathbb{T}^d)$

weak-* topology

$\varphi \in \mathscr{C}(\mathbb{T}^d)$

weak topology

$$\langle \varphi, \mu \rangle = \int_{\mathbb{T}^d} \varphi(x) \mathrm{d}\mu(x)$$

Def. (Weak-* convergence of measures)

$$\mu_n \rightharpoonup \mu \iff \langle \varphi, \mu_n - \mu \rangle \to 0$$

Moment Matrix

$\mu \in \mathcal{M}(\mathbb{T}^d)$, $k \in \mathbb{Z}^d$,
k-th trigonometric moment:

$$c_k(\mu) \overset{\text{def.}}{=} \int_{\mathbb{T}^d} e^{-2\imath\pi\langle k,\, x\rangle} \mathrm{d}\mu(x)$$

Def. $n \in \mathbb{N}$, $N \overset{\text{def.}}{=} (n+1)^d$, truncated moment matrix

$$T_n \overset{\text{def.}}{=} (c_{k-l})_{k,l \in \mathbb{N}_n^d} \in \mathbb{C}^{N \times N}$$

Vandermonde decomposition. When $\mu = \sum_{j=1}^s \lambda_j \delta_{x_j}$, then

$$T_n = \sum_{j=1}^{s} \lambda_j a_j a_j^* = A_n \Lambda A_n^*, \quad \text{where} \quad a_j \overset{\text{def.}}{=} \left[e^{-2\imath\pi\langle k,\, x_j\rangle}\right]_{k \in \mathbb{N}_n^d}$$

Moment Matrix

$\mu \in \mathcal{M}(\mathbb{T}^d)$, $k \in \mathbb{Z}^d$,
k-th trigonometric moment:

$$c_k(\mu) \stackrel{\text{def.}}{=} \int_{\mathbb{T}^d} e^{-2\imath\pi\langle k, x\rangle} \mathrm{d}\mu(x)$$

Def. $n \in \mathbb{N}$, $N \stackrel{\text{def.}}{=} (n+1)^d$, truncated moment matrix

$$T_n \stackrel{\text{def.}}{=} (c_{k-l})_{k,l \in \mathbb{N}_n^d} \in \mathbb{C}^{N \times N}$$

Vandermonde decomposition. When $\mu = \sum_{j=1}^s \lambda_j \delta_{x_j}$, then

$$T_n = \sum_{j=1}^s \lambda_j a_j a_j^* = A_n \Lambda A_n^*, \quad \text{where} \quad a_j \stackrel{\text{def.}}{=} \left[e^{-2\imath\pi\langle k, x_j\rangle}\right]_{k \in \mathbb{N}_n^d}$$

Singular value decomposition. $T_n = \sum_{j=1}^s \sigma_j u_j v_j^* = U\Sigma V^*$, where $U, V \in \mathbb{C}^{N\times s}$ and $\Sigma \in \mathbb{C}^{s \times s}$

Polynomial estimates

Weighted approximation

Let

$$p_n(x) \stackrel{\text{def.}}{=} \frac{1}{N} a(x)^* T_n a(x) = \frac{1}{N} \sum_{i=1}^{s} \sigma_j u_j(x) \overline{v_j(x)}$$

where $a(x) = (e^{-2\imath\pi\langle k, x\rangle})_{k \in \mathbb{N}_n^d}$

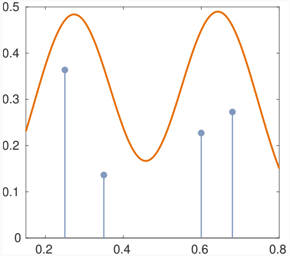
Lemma. (Weighted sum of kernels) One has

$$p_n(x) = \sum_j \lambda_j F_n(x - x_j),$$

where F_n is the Fejér kernel, *i.e.* $F_n(y) \stackrel{\text{def.}}{=} \frac{1}{N} \prod_{i=1}^{d} \frac{\sin((n+1)\pi y_i)}{\sin(\pi y_i)}$.

Rem. Since $p_n = F_n * \mu$, it can be evaluated very efficiently using Fast Fourier Transforms.

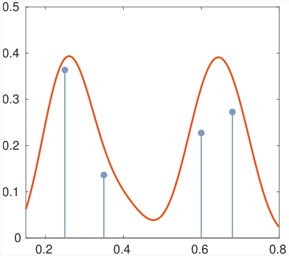
Weak-* convergence



$n = 3$

Thm. (Weak-* convergence) Let μ_n be the measure with density p_n. Then $\mu_n \rightharpoonup \mu$.

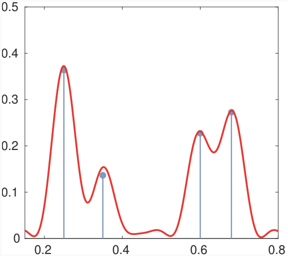
Weak-* convergence



$n = 6$

Thm. (Weak-* convergence) Let μ_n be the measure with density p_n. Then $\mu_n \rightharpoonup \mu$.

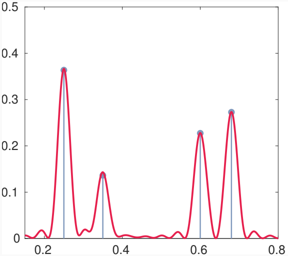
Weak-* convergence



$n = 12$

Thm. (Weak-* convergence) Let μ_n be the measure with density p_n. Then $\mu_n \rightharpoonup \mu$.

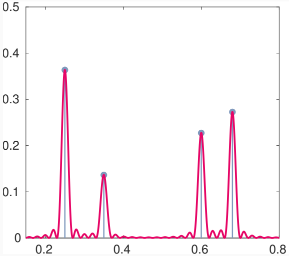
Weak-* convergence



$n = 24$

Thm. (Weak-* convergence) Let μ_n be the measure with density p_n. Then $\mu_n \rightharpoonup \mu$.

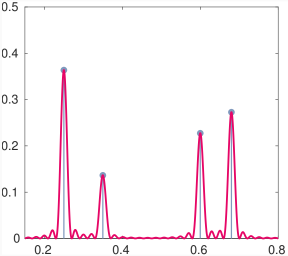
Weak-* convergence



$n = 48$

Thm. (Weak-* convergence) Let μ_n be the measure with density p_n. Then $\mu_n \rightharpoonup \mu$.

Weak-* convergence



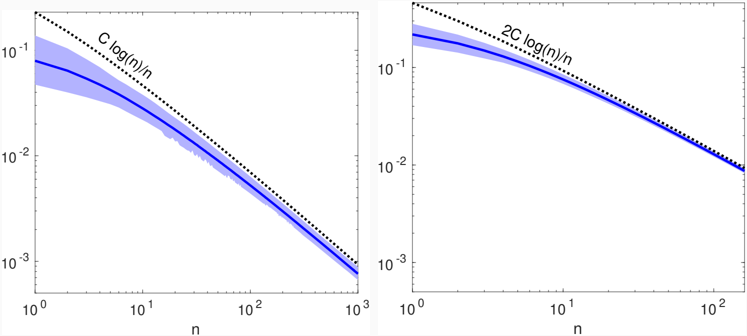
$n = 48$

Thm. (Weak-* convergence) Let μ_n be the measure with density p_n. Then $\mu_n \rightharpoonup \mu$.

Thm. (Wasserstein-1 convergence) We have

$$\mathcal{W}_1(\mu_n, \mu) = O\left(\frac{d \log(n)}{n}\right)$$

Numerics



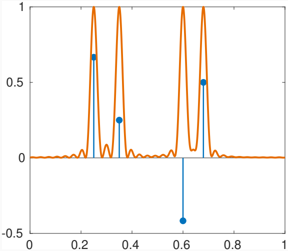
Wasserstein-1 distance between μ_n and μ, averaged over 100 randomly generated measures with sparsity $2 \leqslant s \leqslant 20$, in 1D (left) and 2D (right).

Unweighted approximation

$$p_n^1(x) = a(x)^* UU^* a(x) = \sum_{i=1}^{r} |u_i(x)|^2$$

Rem. UU^* is **not** the moment matrix of $\mu^1 \overset{\text{def.}}{=} \sum_j \delta_{x_j}$

Prop. (Support identification) For n large enough, $0 \leqslant p_n^1 \leqslant N$ and $p_n^1(x) = N$ if and only if $x \in \operatorname{Supp} \mu$.



$N^{-1} p_n^1,\ n = 30$

Weak-* convergence

Prop. (Weak-* convergence) Let $\mu_n^1(x)$ be the measure with density p_n^1. Then $\mu_n^1 \rightharpoonup \mu^1$.

Sketch of proof. Consider the QR decomposition $A_n = QR$ with $Q \in \mathbb{C}^{N \times s}$ unitary

Weak-* convergence

Prop. (Weak-* convergence) Let $\mu_n^1(x)$ be the measure with density p_n^1. Then $\mu_n^1 \rightharpoonup \mu^1$.

Sketch of proof. Consider the QR decomposition $A_n = QR$ with $Q \in \mathbb{C}^{N\times s}$ unitary

$$a = \sum |a_j(x)|^2 = \sum F_n(x - x_j)$$

Weak-* convergence

Prop. (Weak-* convergence) Let $\mu_n^1(x)$ be the measure with density p_n^1. Then $\mu_n^1 \rightharpoonup \mu^1$.

Sketch of proof. Consider the QR decomposition $A_n = QR$ with $Q \in \mathbb{C}^{N \times s}$ unitary

$$a = \sum |a_j(x)|^2 = \sum F_n(x - x_j)$$

$$q = \sum |q_j(x)|^2 \text{ satisfies } q = p_n^1$$

Weak-* convergence

Prop. (Weak-* convergence) Let $\mu_n^1(x)$ be the measure with density p_n^1. Then $\mu_n^1 \rightharpoonup \mu^1$.

Sketch of proof. Consider the QR decomposition $A_n = QR$ with $Q \in \mathbb{C}^{N\times s}$ unitary

$a = \sum |a_j(x)|^2 = \sum F_n(x - x_j)$

$q = \sum |q_j(x)|^2$ satisfies $q = p_n^1$

Lemma. one has $\|a - q\|_1 = O(1/n)$

Weak-* convergence

Prop. (Weak-* convergence) Let $\mu_n^1(x)$ be the measure with density p_n^1. Then $\mu_n^1 \rightharpoonup \mu^1$.

Sketch of proof. Consider the QR decomposition $A_n = QR$ with $Q \in \mathbb{C}^{N\times s}$ unitary

$a = \sum |a_j(x)|^2 = \sum F_n(x - x_j)$

$q = \sum |q_j(x)|^2$ satisfies $q = p_n^1$

Lemma. one has $\|a - q\|_1 = O(1/n)$

Then, for φ Lipschitz,

$$\left| \int \varphi(x) p_n^1(x) \mathrm{d}x - \sum \varphi(x_j) \right| \leqslant \|a - q\|_1 \|\varphi\|_\infty + C \int F_n(x)|x| \mathrm{d}x$$

Conclusion

Conclusion

Summary.

- New tools to estimate discrete measures from moments
- Computationally efficient polynomial approximations
- Convergence with respect to Wasserstein-1 distances

Outlook.

- Extend results to non-discrete measures
- Derive convergence rate for Christoffel in the discrete case

Thank you for your attention!

References

Candès, E. and Fernandez-Granda, C. (2014). Towards a mathematical theory of super-resolution. *Comm. Pure Appl. Math.*, 67(6):906–956.

Lasserre, J. and Pauwels, E. (2016). Sorting out typicality with the inverse moment matrix sos polynomial. In *NIPS*.

Pauwels, E., Putinar, M., and Lasserre, J. (2020). Data analysis from empirical moments and the Christoffel function. *F. Comp. Math.*

R. de Prony, G. (1795). Essai expérimental et analytique: Sur les lois de la dilatabilité des fluides élastiques et sur celles de la force expansive de la vapeur de l'eau et de la vapeur de l'alkool, à différentes températures. *Journal de l'École Polytechnique Floréal et Plairial*, 1(cahier 22):24–76.

Roy, R. and Kailath, T. (1989). ESPRIT-estimation of signal parameters via rotational invariance techniques. *IEEE Trans. Acoustics Speech Signal Process.*, 37(7):984–995.

Schmidt, R. (1986). Multiple emitter location and signal parameter estimation. *IEEE Trans. Antennas Propagation*, 34(3):276–280.