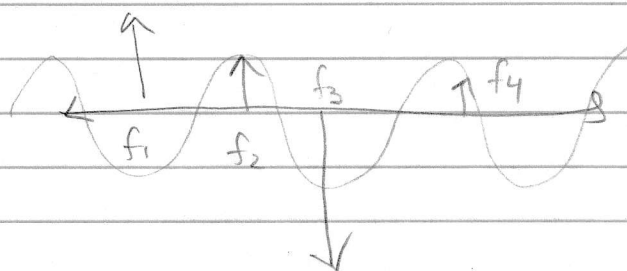


Super resolution

Many optical devices are inherently lowpass

Main Question: Can we recover fine-grained structure from coarse measurements?

[Donoho] introduced the setup



measurement
frequency ω

$$|\omega| \leq m$$

↑
cutoff frequency

$$x(t) = \sum_{j=1}^k u_j \delta_{f_j}(t)$$

delta function at f_j

$$P_\omega = \int_0^1 e^{i2\pi\omega t} x(t) dt$$

$$= \sum_{j=1}^k u_j e^{i2\pi f_j \omega}$$

Can we recover the parameters (coefficients u_j and frequencies f_j) from low frequency measurements?

Proposition [Prony] when there is no noise, \exists polynomial time algorithm for exact recovery from any $2k+1$ distinct measurements

This theorem seems forgotten in the literature, e.g.

Compressed Sensing: can recover a k -sparse vector $x \in \mathbb{R}^n$ from $O(k \log \frac{n}{k})$ noisy linear measurements

"We can beat Shannon-Nyquist by assuming sparsity"

Applications in MRI, etc

But notice Prony's method:

(1) gets $O(k)$ measurements

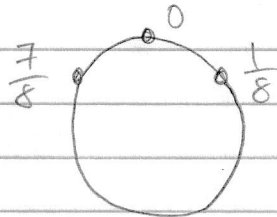
(2) doesn't require points on a grid

But is it stable? More on this later

definition: The wrap around distance d_w is

$$d_w(x, y) = \min(|x - y|, |1 + x - y|)$$

e.g.



$$d_w(1/8, 7/8) = 1/4$$

History, minus Prony

Let $\Delta = \min_{j \neq j'} d_w(f_j, f_{j'})$

[Donoho] Asymptotic ^{stability} bounds for $m = 1/\Delta$ on grid

[Candes, Fernandez-Granda] Convex program for $m \geq 2/\Delta$, no noise

[Fernandez-Granda] Same, but with noise

[Liao, Fannjiang] (independent) Algorithm for $m = \frac{1+C(\Delta)}{\Delta}$, with noise

^[Mokha]
Theorem. There is a polynomial time algorithm that works with noise, for $m \geq 1/\Delta^{1+\epsilon}$ and for any $m \leq \frac{(1-\epsilon)}{\Delta}$ it is impossible (i.e. noise must be exponentially small)

Matrix Pencil Method (No noise)

definition: Let $\alpha_j = e^{i2\pi f_j}$ and let

$$(V \Rightarrow) V_m^k = \begin{bmatrix} 1 & 1 & 1 \\ \alpha_1 & \alpha_2 & \alpha_k \\ \vdots & \vdots & \vdots \\ \alpha_1^{m-1} & \alpha_2^{m-1} & \alpha_k^{m-1} \end{bmatrix}$$

Now consider the matrix

$$A = V D_u V^H$$

↑
diagonal matrix of u_j 's

and B

Claim: The entries of A_n correspond to measurements at integer frequencies $-m+1 \leq \omega \leq m-1$

Proof: By direct computation

$$A_{jj'} = [\alpha_1^{j-1}, \alpha_2^{j-1}, \dots, \alpha_k^{j-1}] \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_k \end{bmatrix} \begin{bmatrix} \alpha_1^{-j'+1} \\ \alpha_2^{-j'+1} \\ \vdots \\ \alpha_k^{-j'+1} \end{bmatrix}$$

$$= P_{j-j'}$$

Similarly define

$$B = V D_u D_\alpha V^H$$

↑
diagonal matrix of α_j 's

Fact: If α_j 's are distinct and $m \geq k$, then V has full column rank

Lemma: Under same conditions, and u_j 's are nonzero then $\lambda = 1/\alpha_j$ are unique solns to

$$Ax = \lambda Bx$$

Proof: By the fact V^H has a right inverse R s.t.

$$V^H R = I_k$$

then $r_j = j^{\text{th}}$ column of R satisfies

$$Ar_j = VD_u V^H r_j = VD_u e_j$$

$$= u_j V e_j = u_j V_j$$

↑
jth column of V

$$Br_j = VD_u D_\alpha V^H r_j = u_j \alpha_j V_j$$

This is a generalized eigenvalue problem,
similar uniqueness. \square

Thus our algorithm is

Construct A and B

For $Ax = \lambda Bx$, find all values of $\lambda = \frac{1}{\alpha_j}$

Solve $V \begin{bmatrix} u_1 \\ \vdots \\ u_k \end{bmatrix} = \begin{bmatrix} p_1 \\ \vdots \\ p_k \end{bmatrix}$ to find coefficients

Noise stability

When is the ~~MFA~~ robust to noise?
generalized eigenvalue problem

i.e. can we get accuracy guarantees that are
polynomial in magnitude of noise?

We'll show a sharp phase transition in the
condition number of the complex Vandermonde

Theorem $\|V_m^k u\|^2 = (m-1 \pm \frac{1}{\Delta}) \|u\|^2$

Theorem: If $m = \frac{(1-\epsilon)}{\Delta}$, $\exists \alpha_j$'s and u_j 's s.t.

$$\|V_m^k u\|^2 \leq e^{-\epsilon k} \|u\|^2$$

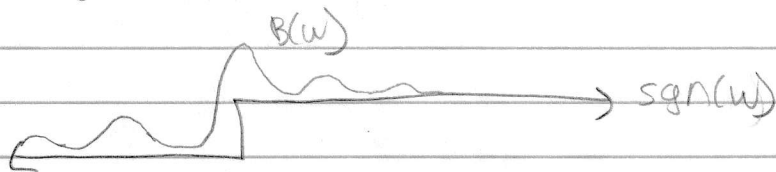
Thus the noise needs to be exponentially small to superresolve

Extremal Functions

Theorem: There is a function $B(\omega)$ satisfying

(1) $\text{sgn}(\omega) \leq B(\omega)$

i.e. it majorizes



(2) $\hat{B}(x) = \text{fourier transform of } B(\omega)$

is supported on $[-1, 1]$

i.e. it is smooth

(3) $\int_{-\infty}^{\infty} B(\omega) - \text{sgn}(\omega) d\omega = 1$

Explicitly we have

$$B(\omega) = \left(\frac{\sin(\pi\omega)}{\pi} \right)^2 \left(\sum_{j=0}^{\infty} (\omega-j)^{-2} - \sum_{j=1}^{\infty} (\omega+j)^{-2} + \frac{2}{\omega} \right)$$

This comes from reconstruction formula for functions whose Fourier transform is supported on $[-1, 1]$

Similarly there is a minorant $b(\omega)$

$$(1) \ b(\omega) \leq \text{sgn}(\omega)$$

$$(2) \ \hat{b}(\omega) \text{ supported in } [-1, 1]$$

$$(3) \ \int_{-\infty}^{\infty} \text{sgn}(\omega) - b(\omega) \, d\omega = 1$$

and these are sharp

Relatedly we can sandwich the indicator of an interval $E \subseteq \mathbb{R}$

Corollary There are functions $C_E^u(\omega)$ and $C_E^l(\omega)$ with $E = [0, m-1]$ that satisfy

$$(1) \ C_E^l(\omega) \leq \mathbb{1}_E(\omega) \leq C_E^u(\omega)$$

$$(2) \ \hat{C}_E^l(x) \text{ and } \hat{C}_E^u(x) \text{ are supported on } [-\Delta, \Delta]$$

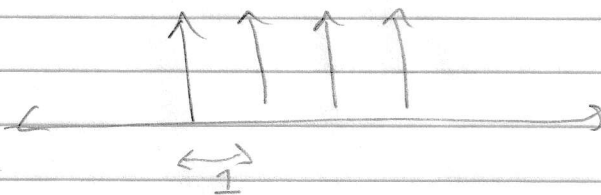
$$(3) \ \int_{-\infty}^{\infty} C_E^u(\omega) - \mathbb{1}_E(\omega) \, d\omega = \int_{-\infty}^{\infty} \mathbb{1}_E(\omega) - C_E^l(\omega) \, d\omega = \frac{1}{\Delta}$$

Now let's prove the condition number bound

Proof: We have

$$\|V_m^k u\|^2 = \sum_{\omega=0}^{m-1} |P_\omega|^2 = (*)$$

Now let $h(\omega) =$ Dirac comb function, i.e.



i.e. $h(\omega) = \sum_{t=-\infty}^{\infty} \delta_t(\omega)$

then $(*) = \int_{-\infty}^{\infty} h(\omega) I_E(\omega) |P_\omega|^2 d\omega$

$\leq \int_{-\infty}^{\infty} h(\omega) C_E^u(\omega) |P_\omega|^2 d\omega$

(1)

Now the Fourier transform of a comb is a comb so

$$h(\omega) = \sum_{t=-\infty}^{\infty} e^{i2\pi t \omega}$$

(2)

we get that

Now since $|P_w|^2 = P_w^* P_w$ and

$$P_w = \sum_{j=1}^k u_j e^{i2\pi f_j w}$$

we get that the RHS is

$$\sum_{j=1}^k \sum_{j'=1}^k u_j u_{j'}^* \int_{-b}^{\infty} h(w) C_E^u(w) e^{i2\pi(f_j - f_{j'})w} dw$$

$$= \sum_{t=-\infty}^{\infty} \sum_{j=1}^k \sum_{j'=1}^k u_j u_{j'}^* \int_{-b}^{\infty} e^{i2\pi t w} C_E^u(w) e^{i2\pi(f_j - f_{j'})w} dw$$

$$\hat{C}_E^u(f_j - f_{j'} + t)$$

But the f_j 's are Δ separated and \hat{C}_E^u is supported on $[-\Delta, \Delta]$ that the cross terms are zero!

$$= \sum_{j=1}^k |u_j|^2 \hat{C}_E^u(0)$$

ℓ_1 norm of $C_E^u = |E| + \frac{1}{\Delta}$

$$= \sum_{j=1}^k |u_j|^2 (|E| + \frac{1}{\Delta}) \quad \square$$

The same proof works for minorant. \square

Thus we have

extremal functions \Rightarrow sharp bounds for condition # \Rightarrow stability of MPM

\Rightarrow sharp bounds for super-resolution

Orbit Recovery

Setup: Unknown signal $x \in \mathbb{R}^d$ and group G with group action

$$P: G \rightarrow \mathbb{R}^{d \times d}$$

We get measurements of the form

$$y = P(g)x + \mathcal{N}$$

\uparrow \uparrow
random group element from Haar noise

random group element from Haar

Goal: recover some \hat{x} close to the orbit

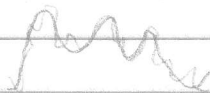
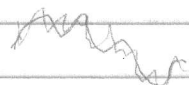
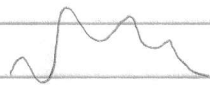
$$\{P(g)x \mid g \in G\}$$

why only up to orbit?

Examples

Multireference Alignment

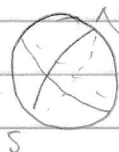
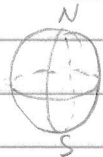
true signal



group

\mathbb{Z}_n or $SO(2)$

Cryo-electron Tomography



$SO(3)$

+ noise

Cryo-electron Microscopy: Cryo-ET + projection

[Perry, Weed, Bandeira, Rigollet, Singer]: Algorithms for discrete MRA with optimal sample complexity

Claim: we already know an algorithm

discrete MRA \longleftrightarrow mixture of Gaussians
where centers are cyclic shifts

Main Question: Can we exploit, rather than circumvent, group structure?

e.g. consider

$$T = \sum_{g \in G} (P(g)x)^{\otimes 3}$$

The rank of T is at most $|G|$. But what if the group is continuous / nonabelian?

Invariant Theory

definition: The ring of invariant polynomials consists of all polys q s.t.

$$q(P(g)x) = q(x) \quad \forall g \in G$$

Fact: For discrete MRA, the invariant ring is generated by

$$\prod_{j=1}^s \hat{X}_{l_j} \quad \text{s.t. } l_1 + l_2 + \dots + l_s = 0$$

↑
Fourier transform

When $s=3$, this is called bispectrum

Theorem: If G is compact and acts continuously on X , the invariant ring determines X up to orbit every

just on generators

$$\text{i.e. } \hat{x} \in \{P(g)x \mid g \in G\} \iff \nexists \text{ invariant } q$$
$$q(\hat{x}) = q(x)$$

This is the method of moments with group symmetries

How many moments suffice? \iff At what degree do invariant polynomials generate the full invariant ring?

Theorem [Takara] Provided that $\hat{x}_i \neq 0 \forall i$, the bispectrum determines X up to cyclic shift

This follows from Jennrich

Let d^* = degree at which you generate invariant ring generically

Theorem [Blum-smith, Bandeira, Kileel, Perry, Weed, Wein] For orbit recovery over compact G , the number of samples you need for generic list recovery is $\Theta(\sigma^{-2d^*})$

Caveats: (1) No efficient algorithms

(2) No bounds on d^*

(3) Bounds hold in the limit as $\sigma \rightarrow \infty$, and hidden constants can depend on d, G

Mean challenge: "orbit tensor decomposition"

$$\text{Given } T = \int_{g \in G} (p(g)x)^{\otimes 3} dg$$

can we recover x up to orbit?

Theorem [Liu, Moitra] There is a quasipolynomial time algorithm for cryo-electron tomography in the smoothed analysis model

Open Questions

① Solve orbit tensor decomposition

Can we make the method of invariants algorithmic?

② inverse problems in the sciences.

④ beyond worst-case analysis

e.g. reinforcement learning

③ linear dynamical systems