Super-resolution by means of Beurling minimal extrapolation

John J. Benedetto and Weilin Li

Norbert Wiener Center
Department of Mathematics
University of Maryland, College Park
http://www.norbertwiener.umd.edu

Acknowledgements
DTRA 1-13-1-0015, ARO W911NF-15-1-0112,
ARO W911NF-16-1-0008, ARO W911NF17-1-0014

Outline

What is super-resolution?

Broadly speaking, super-resolution is concerned with recovering fine details (high-frequency) from coarse information (low-frequency).

There are two main categories of super-resolution:

- Spectral extrapolation – Optical, radar, geophysics, astronomy, medical imaging, e.g., MRI, problems;
- Spatial interpolation – Geometrical or image-processing, e.g., in-painting, problems.

Remark We shall deal with spectral extrapolation, and we shall not deal with the important role of non-uniform sampling and multiple measurements, nor the critical setting of noisy environments.

Background and notation

Our super-resolution model is based on the theory of Candès and Fernandez-Granda [4], [5] for discrete measures, *and* our main idea was inspired by classical work of Beurling [2], [3].

- $\mathbb{T}^d$ is the d-dimensional torus group.
- $M(\mathbb{T}^d)$ is the space of complex Radon measures on the torus.
- $\|\cdot\|$ is the total variation norm.
- The Fourier transform of μ is the function $\widehat{\mu}\colon \mathbb{Z}^d \to \mathbb{C}$, defined as

$$\widehat{\mu}(m) = \int_{\mathbb{T}^d} e^{-2\pi i m x}\, d\mu(x).$$

- $\Lambda \subseteq \mathbb{Z}^d$ is a finite set.

Norbert Wiener Center
for Harmonic Analysis and Applications

The super-resolution problem

The *unknown information* is modeled as $\mu \in M(\mathbb{T}^d)$, not only discrete measures. There are two reasons for $\mu \in M(\mathbb{T}^d)$:

- Objects (images) are not necessarily supported by discrete sets;
- Fine features can be supported in measure 0 non-discrete sets.

The *given low-frequency information* is modeled as spectral data, $F(n)$, $n \in \Lambda$. To recover μ from F, we pose the *super-resolution problem*,

$$\inf \|\nu\| \quad \text{subject to} \quad \nu \in M(\mathbb{T}^d) \quad \text{and} \quad \widehat{\nu} = F \quad \text{on} \quad \Lambda. \tag{SR}$$

Remark Using weak-$*$ compactness arguments, we can show that this problem is well-posed (the inf can be replaced with a min), but not without significant theoretical and computational challenges.

Norbert Wiener Center
for Harmonic Analysis and Applications

Connection with compressed sensing

If the unknown measure μ is of the form,

$$\mu = \sum_{m=0}^{N-1} x_m \delta_{\frac{m}{N}} \in M(\mathbb{T}),$$

where $x \in \mathbb{C}^N$, $x = (x_0, \dots, x_{N-1})$, then

$$\widehat{\mu}(n) = \sum_{m=0}^{N-1} x_m \, e^{-2\pi i m n/N} = \mathcal{F}_N(x)(n),$$

the DFT of x. This shows that Problem (SR) is a generalization of the *basis pursuit algorithm* for under-sampled DFT data F:

For given $F(n)$, $n \in \Omega \subseteq \mathbb{Z}/N\mathbb{Z}$, solve

$$\min \|y\|_{\ell^1} \quad \text{subject to} \quad y \in \mathbb{C}^N \quad \text{and} \quad \mathcal{F}_N y = F \quad \text{on} \quad \Omega \subseteq \mathbb{Z}/N\mathbb{Z},$$

For this reason, super-resolution is a continuous theory of *compressed sensing*.

Outline

Norbert Wiener Center for Harmonic Analysis and Applications

A theorem of Candès and Fernandez-Granda, $d = 1$

The following theorem for $d = 1$ shows that one can reconstruct a discrete measure whose support satisfies a minimum separation condition.

Theorem, Candès and Fernandez-Granda [5]

Let $\Lambda_M = \{-M, -M+1, \ldots, M\}$ for some integer $M \geq 128$ and let $F = \widehat{\mu}$ on Λ_M, where $\mu \in M(\mathbb{T})$ is a discrete measure for which

$$\inf_{x,y \in \mathrm{supp}(\mu),\ x \neq y} |x - y| \geq \frac{2}{M}.$$

Then, μ is the unique solution to Problem (SR) given F on Λ_M.

A theorem of Candès and Fernandez-Granda, $d > 1$

The following theorem for $d > 1$ shows that one can reconstruct a discrete measure whose support satisfies a minimum separation condition.

Theorem, Candès and Fernandez-Granda [5]

Given $S = \{s_j\}_{j=1}^{J} \subseteq \mathbb{T}^d$ and $\mu \in M(\mathbb{T}^d)$ for which $\mathrm{supp}\,(\mu) \subseteq S$. Let $\Lambda_M = \{-M, -M+1, \ldots, M\}^d$, let F be spectral data on Λ_M, and let $\widehat{\mu} = F$ on Λ_M. There exist $C_d,\ M_d > 0$ such that if $M \geq M_d$ and

$$\inf_{1 \leq j < k \leq J} \|s_j - s_k\|_{\ell^\infty(\mathbb{T}^d)} \geq \frac{C_d}{M},$$

then μ is the *unique solution* to Problem (SR).

Non-discrete measures?

Motivated by applications, we develop a super-resolution theory for non-discrete measures. This is carried out by connecting the Candès and Fernandez-Granda theory of super-resolution [5] with Beurling's theory of minimal extrapolation [2], [3]. To this end –

- Let ϵ be the minimum value attained by Problem (SR), i.e.,

$$\epsilon = \epsilon(\Lambda, F) = \inf\{\|\nu\| : \widehat{\nu} = F \text{ on } \Lambda\}.$$

- Let $\mathcal{E}$ be the set of all solutions to Problem (SR), i.e.,

$$\mathcal{E} = \mathcal{E}(\Lambda, F) = \{\nu \in M(\mathbb{T}^d) \colon \|\nu\| = \epsilon \text{ and } \widehat{\nu} = F \text{ on } \Lambda\}.$$

  If $\nu \in \mathcal{E}$, then we say ν is a *minimal extrapolation* from Λ.
- Our theory depends essentially on the set,

$$\Gamma = \Gamma(\Lambda, F) = \{m \in \Lambda \colon |F(m)| = \epsilon\}.$$

Theorem

Theorem [1]

Let $\Lambda \subseteq \mathbb{Z}^d$ be a finite set and let F be spectral data defined on Λ.

(a) Suppose $\Gamma = \emptyset$. Then, there exists a closed set S of d-dimensional Lebesgue measure zero such that each minimal extrapolation is a singular measure supported in S.

(b) Suppose $\#\Gamma \geq 2$. For each distinct pair $m, n \in \Gamma$, define $\alpha_{m,n} \in \mathbb{R}/\mathbb{Z}$ by $e^{2\pi i \alpha_{m,n}} = F(m)/F(n)$. Define the closed set,

$$S = \bigcap_{\substack{m,n\in\Gamma \\ m\neq n}} \{x \in \mathbb{T}^d : x \cdot (m-n) + \alpha_{m,n} \in \mathbb{Z}\},$$

which is an intersection of $\binom{\#\Gamma}{2}$ periodic hyperplanes. Then, each minimal extrapolation is a singular measure supported in S.

Illustration of the theorem

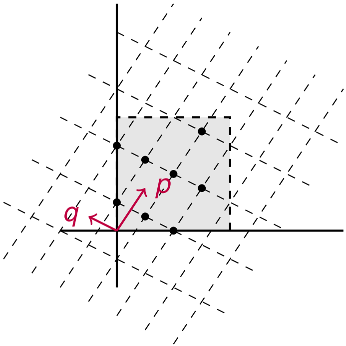

Figure: An illustration of the second case of the theorem. The hyperplanes in the theorem are represented by the dashed lines. The vectors $p = (1/4, 3/8)$ and $q = (-1/4, 1/8)$ are normal to the hyperplanes and their lengths determine the separation of the hyperplanes.

The role of uniqueness

Why uniqueness is important:

- If $\mu \in \mathcal{E}(\Lambda, F)$ is unique, then any numerical solution to Problem (SR) approximates μ.
- Without uniqueness, even if $\mu \in \mathcal{E}(\Lambda, F)$, it is possible that a numerical solution to Problem (SR) does not approximate μ.

Outline

1. Introduction
2. Theorems
3. Examples
4. Conclusion

Uniqueness and super-resolution reconstruction

Let $\Lambda = \{-1, 0, 1\}$. Define F in the following ways.

- $F(0) = 0,\ F(\pm 1) = 2$. Define $\mu = \delta_0 - \delta_{1/2} \in M(\mathbb{T})$. $\Gamma = \{-1, 1\}$.
- $F(0) = 0, F(\pm 1) = 1 \pm i$. Define $\mu = \delta_0 - \delta_{1/4} \in M(\mathbb{T})$. $\Gamma = \{-1, 1\}$.
- $F(-1) = 0,\ F(0) = 1 + e^{\pi i/3},\ F(1) = 1 + e^{-\pi i/3}$. Define $\mu = \delta_0 + e^{\pi i/3}\delta_{1/3} \in M(\mathbb{T})$. $\Gamma = \{0, 1\}$.

In each case μ can be proved to be the unique minimal extrapolation, and so super-resolution reconstruction of μ from the values of F on Λ is possible.

Norbert Wiener Center for Harmonic Analysis and Applications

Cantor measures and $\#\Gamma = 1$

$C_q = \bigcap_{k=0}^{\infty} C_{q,k}$, integer $q \geq 3$, is the *middle* $1/q$*-Cantor set*, where

$$C_{q,0} = [0,1] \text{ and } C_{q,k+1} = \frac{C_{q,k}}{q} \cup (1-q) + \frac{C_{q,k}}{q},$$

and let σ_q be the continuous singular Cantor-Lebesgue measure with

$$\widehat{\sigma_q}(m) = (-1)^m \prod_{k=1}^{\infty} \cos(\pi m q^{-k}(1-q)),$$

$$\forall n \in \mathbb{Z} \setminus \{0\}, \quad \widehat{\sigma_q}(q^n) \neq 0 \quad \text{takes the same constant value.}$$

Let $\Lambda \subseteq \mathbb{Z}$ be finite, assume $0 \in \Lambda$, and suppose F defined on Λ satisfies $F(0) = 1$, noting $\widehat{\sigma_q}(0) = \|\sigma_q\| = 1$. If $\sigma_q \in \mathcal{E}(\Lambda, F)$, then $\#\Gamma = 1$, and our present theory does not determine if σ_q is the unique minimal extrapolation.

Non-uniqueness: $\mu = \delta_0 + \delta_{1/2} \in M(\mathbb{T})$ and $\#\Gamma = 1$

- Given $\Lambda = \{-1, 0, 1\}$ and $F(0) = 2$, $F(\pm 1) = 0$. If $\mu = \delta_0 + \delta_{1/2} \in M(\mathbb{T})$, then $\widehat{\mu} = F$ on Λ.
- μ is a minimal extrapolation, $\epsilon = 2$, and $\Gamma = \{0\}$.
- There are uncountably many discrete minimal extrapolations. In fact, $x \in \mathbb{T}$ and any integer $N \geq 2$ define the discrete measure

$$\nu_{N,x} = \frac{2}{N} \sum_{n=0}^{N-1} \delta_{x+\frac{n}{N}},$$

and each $\nu_{N,x}$ is a minimal extrapolation.

$\mu = \delta_0 + \delta_{1/2} \in M(\mathbb{T})$ and $\#\Gamma = 1$, continued

- There are also uncountably many positive absolutely continuous minimal extrapolations. In fact, for any integer $N \geq 2$ and constant $0 < c \leq (2N+2)/(3N+1)$, extend F on Λ to the sequence $\{(a_{N,c})_n\}_{n\in\mathbb{Z}}$, where

$$(a_{N,c})_n = \begin{cases} 2 & \text{if } n = 0, \\ c\left(1 - \frac{|n|}{N+1}\right) & \text{if } 2 \leq |n| \leq N, \\ 0 & \text{otherwise.} \end{cases}$$

The non-negative real-valued function

$$f_{N,c}(x) = 2 + \sum_{n=-N}^{-2} (a_{N,c})_n \, e^{2\pi inx} + \sum_{n=2}^{N} (a_{N,c})_n \, e^{2\pi inx}$$

is a positive absolutely continuous minimal extrapolation.

Optimality in higher dimensions

In higher dimensions, geometry plays an important role.

- Let $\Lambda = \{-1, 0, 1\}^2 \setminus \{(1, -1), (-1, 1)\}$ and let $\mu = \delta_{(0,0)} + \delta_{(1/2,1/2)} \in M(\mathbb{T}^2)$.
- Then, μ is a minimal extrapolation, $\epsilon = 2$, and $\Gamma = \{(0, 0), (1, 1), (-1, -1)\}$.
- We can construct other discrete minimal extrapolations. For any $x \in \mathbb{R}$ and any integer $N \geq 2$, define the measure

$$\nu_{N,x} = \frac{2}{N} \sum_{n=0}^{N-1} \delta_{\left(x+\frac{n}{N}, 1-x-\frac{n}{N}\right)}.$$

Then, each $\nu_{N,x}$ is a minimal extrapolation.

Optimality in higher dimensions, continued

- We can also construct a continuous singular minimal extrapolation. According to the theorem, each minimal extrapolation is supported in the set,

  $$S = \{x \in \mathbb{T}^2 \colon x_1 + x_2 = 1\}.$$

  Let $\sigma = \sqrt{2}\sigma_S$, where σ_S is the *surface measure* of the Borel set S. We readily verify that σ is indeed a minimal extrapolation.

This example shows that the second statement of our theorem is optimal.

Norbert Wiener Center for Harmonic Analysis and Applications

Final remarks

- Our theory shows that Γ provides significant information about the minimal extrapolations. In particular, when $\#\Gamma \neq 1$, they are always singular measures, but when $\#\Gamma = 1$, they could be absolutely continuous.
- We have not discussed how to solve Problem (SR) computationally. Candès and Fernandez-Granda provided an algorithm that potentially fails, but this occurs only if $\Gamma \neq \emptyset$. Hence, our theorem is capable of computing analytical solutions even when it is impossible to compute a numerical approximation.
- The theorem opens up the possibility of the super-resolution of continuous singular measures. Since we are concerned with Fourier samples, medical imaging is a natural application of this theory.

Final remarks, continued

- Our theorem does not require additional assumptions on $\mu \in M(\mathbb{T}^d)$ or on the finite subset $\Lambda \subseteq \mathbb{Z}^d$. Since the theorem also describes the support set of the minimal extrapolations of μ from Λ, it is useful for determining whether a given μ can be recovered by solving the super-resolution problem.
- The second statement of the theorem provides sufficient conditions for when the minimal extrapolations are supported in a lattice. As we have seen, such measures correspond to vectors solving the discrete compressed sensing problem. Thus, our theorem is a continuous-discrete correspondence result.
- Our results are closely related to Beurling's work on minimal extrapolation. He dealt with $\mathbb{R}^1$ instead of $\mathbb{T}^d$, so our theorem is an adaptation to the torus and a generalization to higher dimensions. There are non-trivial technical differences between working with $\mathbb{R}^1$ and $\mathbb{T}^d$.


Thats all folks!

Pre-dual of the super-resolution problem

The strategy is to analyze an appropriate dual formulation. As such, we define the *pre-dual* problem:

$$\max \left| \sum_{m\in\Lambda} a_m F(m) \right| \quad \text{subject to} \quad \forall x \in \mathbb{T}^d, \quad \left| \sum_{m\in\Lambda} a_m e^{2\pi i m\cdot x} \right| \le 1. \tag{SR$'$}$$

If $\{a_m\}_{m\in\Lambda}$ solves Problem (SR′), then for all $\nu \in \mathcal{E}(\Lambda, F)$,

$$\operatorname{supp}(\nu) \subseteq \left\{ x \in \mathbb{T}^d : \left| \sum_{m\in\Lambda} a_m e^{2\pi i m\cdot x} \right| = 1 \right\}.$$

Remark Problem (SR′) can be recast as a semi-definite program. It is unknown whether Problem (SR) can be.

Norbert Wiener Center for Harmonic Analysis and Applications

Admissibility range

A numerical approximation of ϵ can be obtained by solving Problem (SR′), but its exact value is typically unknown. On the other hand, if we are given finite $\Lambda \subseteq \mathbb{Z}^d$, spectral data F on Λ, and $\mu \in \mathbb{T}^d$, then

$$\sup_{m \in \Lambda} |F(m)| \leq \epsilon(\Lambda, F) \leq \|\mu\|.$$

- If the lower bound is attained, then $\Gamma \neq \emptyset$. Our theory is particularly strong for large $\#\Gamma$.
- The upper bound $\epsilon = \|\mu\|$ is a necessary condition for uniqueness of the super-resolution of μ from F.

On minimum separation

In view of the CFG theorem, it natural to ask whether separation is necessary in order to recover a discrete measure. We show that if two Dirac masses are too close, super-resolution is impossible.

- Let $\Lambda \subseteq \mathbb{Z}^d$ be a finite set and let $\mu_y = \delta_0 - \delta_y$ for some non-zero $y \in \mathbb{T}^d$.
- Let ν_y be the absolutely continuous measure,

$$\nu_y(x) = \sum_{m \in \Lambda} \widehat{\mu_y}(m) e^{2\pi i m \cdot x}.$$

By construction, $\widehat{\nu_y} = \widehat{\mu_y}$ on Λ. As $y \to 0$,

$$\|\nu_y\| = \int_{\mathbb{T}^d} \left| \sum_{m \in \Lambda} \widehat{\mu_y}(m) e^{2\pi i m \cdot x} \right| dx \to 0.$$

For $|y|$ sufficiently small, we see that $\mu_y \notin \mathcal{E}$.

Outline

Bibliography

John J. Benedetto and Weilin Li, *Super-resolution by means of Beurling minimal extrapolation*, submitted 2016.

Arne Beurling, *Balayage of Fourier-Stieltjes transforms*, The Collected Works of Arne Beurling **2** (1989), 341–350.

______, *Interpolation for an interval in* $\mathbb{R}^1$, The Collected Works of Arne Beurling **2** (1989), 351–365.

Emmanuel J. Candès and Carlos Fernandez-Granda, *Super-resolution from noisy data*, The Journal of Fourier Analysis and Applications **19** (2013), no. 6, 1229–1254.

______, *Towards a mathematical theory of super-resolution*, Communications on Pure and Applied Mathematics **LXVII** (2014), 906–956.