

Super-Resolution of Point Sources Down to the Rayleigh Limit from Multiple Observations

December 1, 2017

1 Introduction

1.1 Single Observation Problem Statement

Suppose we observe a true signal whose representation in the physical domain is

$$x^*(t) = \sum_{j=1}^s a_j \delta_{t_j}(t),$$

for $t \in \mathbb{T}$ where $\mathbb{T} = \mathbb{R}/\mathbb{Z}$, which we will usually think of as the unit interval $[0, 1]$ with its endpoints identified, and $T = \{t_j\}_{j=1}^s \subset \mathbb{T}$ a discrete support set. In the Fourier domain, this signal takes the form

$$\widehat{x^*}(k) = \sum_{j=1}^s a_j \exp(-2\pi i k t_j)$$

for $k \in \mathbb{Z}$.

Now, we are interested in recovering this true signal x from an observation that has been corrupted due to poor resolution, which we treat as convolution of $x(t)$ with a point-spread function $\phi(t)$. We then observe the signal

$$(\phi * x^*)(t) = \sum_{j=1}^s a_j \phi(t - t_j).$$

In the simplest case, ϕ is a Dirichlet kernel with cutoff frequency f_c , in which case the Fourier transform of the above is simply the truncation of the Fourier transform of x^* . Calling the Fourier transform of the observed signal y , we have

$$y(k) = \left(\sum_{j=1}^s a_j \exp(-2\pi i k t_j) \right) \mathbb{1}_{\{|k| \leq f_c\}}.$$

We may then think of the data we observe as simply the $2f_c + 1$ values $y(-f_c), \dots, y(f_c)$. We write \mathcal{F}_n for the sensing operator mapping x^* to these n Fourier coefficients. A popular technique for solving this problem is *total variation minimization*, which attempts to recover x^* by solving the

convex problem

$$\operatorname{argmin}_x \|x\|_{\text{TV}} \text{ subject to } \mathcal{F}_n x = y.$$

1.2 Extension to Multiple Observations

We now consider a slight modification of the problem presented in the previous section, where we make *several* observations signals sharing the support $T = \{t_1, \dots, t_s\}$ of x^* , but having varying amplitudes a_j . To formalize this, our true signal is now

$$x_k^*(t) = \sum_{j=1}^s a_{j,k} \delta_{t_j} \text{ for } k \in \{1, \dots, m\}, \quad (1)$$

and our observations are $y_{\ell,k} = (\mathcal{F}_n x_k^*)_{\ell}$ for $\ell \in \{-f_c, \dots, f_c\}$ and $k \in \{1, \dots, m\}$. We think of the $a_{j,k}$ as organized into a matrix $\mathbf{A} \in \mathbb{C}^{s \times m}$ and the $y_{\ell,k}$ as organized into a matrix $\mathbf{Y} \in \mathbb{C}^{n \times m}$. The *group total variation minimization* is the natural extension of the previous convex problem to this setting, where we solve

$$\operatorname{argmin}_x \|x\|_{\text{gTV}} \text{ subject to } \begin{bmatrix} \mathcal{F}_n x_1 & \mathcal{F}_n x_2 & \dots & \mathcal{F}_n x_m \end{bmatrix} = \mathbf{Y}. \quad (2)$$

1.3 Dual Certificates

Definition 1 Let $\mu^0 \subset \mathbb{C}$ denote the complex unit circle. A sign pattern on a set is an assignment of points of μ^0 to each point of the set.

Definition 2 For a sign pattern $v \in (\mu^0)^T$, a low-pass trigonometric polynomial $q : \mathbb{T} \rightarrow \mathbb{C}$ given by

$$q(t) = \sum_{k=-f_c}^{f_c} c_k e^{2\pi i k t} \quad (3)$$

is a (single-observation) dual certificate for v if $q(t_j) = v_j$ for $t_j \in T$, and $|q(t)| < 1$ for $t \notin T$.

Proposition 1 (TV Norm Minimization Duality) If a dual certificate for the sign pattern $v_j = \frac{a_j}{|a_j|}$ exists, then x^* is the unique solution of TV minimization for the super-resolution problem.

Therefore, to prove the effectiveness of TV minimization it suffices to show that a dual certificate exists under some conditions on T . The main result of [1] is thus equivalent to the following.

Theorem 1 (Proposition 2.3 of [1]) If $\Delta(T) \geq \frac{1.26}{f_c}$ and $f_c \geq 10^3$, then a dual certificate exists for any sign pattern on T .

The same ideas extend to the multiple observation case in a straightforward fashion.

Definition 3 Let $\mu^{m-1} \subset \mathbb{C}^m$ denote the m -dimensional complex unit sphere. A m -dimensional sign pattern on a set is an assignment of points of μ^{m-1} to each point of the set.

Definition 4 For a sign pattern $v \in (\mu^{m-1})^T$, a low-pass trigonometric polynomial $q : \mathbb{T} \rightarrow \mathbb{C}^m$ given by $q(t) = (q_1(t), \dots, q_m(t))$ and

$$q_{\ell}(t) = \sum_{k=-f_c}^{f_c} c_k \exp(2\pi i k t) \text{ for } \ell \in \{1, \dots, m\} \quad (4)$$

is an m -observation dual certificate for v if $q(t_j) = v_j$ for $t_j \in T$, and $\|q(t)\|_{\ell^2} < 1$ for $t \notin T$.

Proposition 2 (gTV Norm Minimization Duality) *If a dual certificate for the sign pattern $v_j = a_{j,\bullet}/\|a_{j,\bullet}\|_{\ell^2}$ exists, then x^* is the unique solution of gTV minimization for the super-resolution problem.*

However, while numerical experiments in [1] suggest that when the amplitudes \mathbf{A} are taken randomly then recovery is possible down to a lower critical minimum separation, the same uniform argument cannot apply—indeed, it is always possible that the observations $a_{j,\bullet}$ are all identical, in which case clearly it cannot be possible to recover x^* any more effectively than in the one observation case. Thus, to show that gTV norm minimization substantially improves on TV norm minimization, it is necessary to make some probabilistic assumptions about the data generating process.

2 Dual Certificate Construction

Recall that we are interested in building a low-pass trigonometric polynomial $q : \mathbb{T} \rightarrow \mathbb{C}^m$ interpolating the points (t_j, v_j) while remaining strictly inside the unit sphere in \mathbb{C}^m elsewhere. The idea of the construction of [1] when $m = 1$ is to build q via kernel interpolation, and adjust it so that $q'(t_j) = 0$ for each j , which one would hope might encourage the polynomial to remain in the unit sphere. The naive extension of this to the m -dimensional is to require $\nabla q(t_j) = 0$ for each t_j . However, numerical experiments show that the regime where this construction succeeds essentially does not depend on m , which is not surprising—after all, this is repeating the one-dimensional construction in each coordinate. We propose a different dual certificate, which does succeed numerically in certifying reconstruction down to the Rayleigh limit as $m \rightarrow \infty$.

2.1 Step 1: Relaxing Gradient Condition

The first observation is that the condition $\nabla q(t_j) = 0$ is overkill for encouraging q to remain in the unit sphere. Intuitively, all that is really necessary is that $\nabla q(t_j)$ not have a component pointing out of the sphere, i.e. in the direction of v_j . Thus, a reasonable condition is to require $\nabla q(t_j)$ and v_j , viewed as vectors, to be orthogonal. However, it is not possible to express this constraint in complex arithmetic, so we must rewrite the dual certificate in terms of its real and complex parts (each of which is still low-pass).

Thus, we write $q_\ell(t) = q_\ell^R(t) + iq_\ell^I(t)$ for real-valued low-pass trigonometric polynomials q_ℓ^R and q_ℓ^I . We combine these into a single real-valued $2m$ -dimensional trigonometric polynomial,

$$\tilde{q}(t) = (q_1^R(t), q_1^I(t), \dots, q_m^R(t), q_m^I(t)) \in \mathbb{R}^{2m}. \quad (5)$$

Likewise writing $v_{j,\ell} = v_{j,\ell}^R + iv_{j,\ell}^I$ and $\tilde{v}_j = (v_{j,1}^R, v_{j,1}^I, \dots, v_{j,m}^R, v_{j,m}^I)$, we can express the interpolation conditions that the dual certificate q must satisfy in terms of these new variables in essentially the same way: $\tilde{q}(t_j) = \tilde{v}_j$, and $\|\tilde{q}(t)\|_{\ell^2} < 1$ for all $t \notin T$. We can now formulate the conditions we will require of our interpolant.

- I1. $\tilde{q}(t_j) = \tilde{v}_j$ for $j \in \{1, \dots, m\}$.
- I2. $\langle \nabla \tilde{q}(t_j), \tilde{v}_j \rangle = 0$ for $j \in \{1, \dots, m\}$.

2.2 Step 2: Kernel Interpolation

We now proceed just as in [1]: fix a real-valued low-pass kernel K , and define

$$q_\ell^R(t) = \sum_{j=1}^s \alpha_{j,\ell}^R K(t - t_j) + \sum_{j=1}^s \beta_{j,\ell}^R K'(t - t_j), \quad (6)$$

$$q_\ell^I(t) = \sum_{j=1}^s \alpha_{j,\ell}^I K(t - t_j) + \sum_{j=1}^s \beta_{j,\ell}^I K'(t - t_j), \quad (7)$$

for coefficients $\alpha_{j,\ell}^R, \alpha_{j,\ell}^I, \beta_{j,\ell}^R, \beta_{j,\ell}^I \in \mathbb{R}$ to be determined. Conditions I1 and I2 from the previous part may now be expressed as a system of linear constraints on these coefficients, which we introduce some notation to write down concisely.

Definition 5 For a function $K : \mathbb{T} \rightarrow \mathbb{R}$ and an ordered point set $T = \{t_1, \dots, t_s\}$, define $D(K, T) \in \mathbb{R}^{s \times s}$ to be the matrix whose entry at position k, ℓ is $K(t_k - t_\ell)$.

We then define our *constraint matrix* rather baroquely to be

$$D_{\text{ker}} := \begin{bmatrix} D(K, T) & D(K', T) \end{bmatrix}, \quad (8)$$

$$D^{(1)} := I_{2s} \otimes D_{\text{ker}}, \quad (9)$$

$$D'_{\text{ker}} := \begin{bmatrix} D(K', T) & D(K'', T) \end{bmatrix}, \quad (10)$$

$$D'_{\text{coeff}} := \begin{bmatrix} v_{\bullet,1}^R & v_{\bullet,1}^R & v_{\bullet,1}^I & v_{\bullet,1}^I & \cdots & v_{\bullet,m}^R & v_{\bullet,m}^R & v_{\bullet,m}^I & v_{\bullet,m}^I \end{bmatrix} \mathbf{1}^\top, \quad (11)$$

$$D^{(2)} := D'_{\text{coeff}} \odot (D'_{\text{ker}} \otimes \mathbf{1}), \quad (12)$$

$$D := \begin{bmatrix} D^{(1)} \\ D^{(2)} \end{bmatrix}. \quad (13)$$

If we now define the coefficient vector

$$\mathbf{c} := \begin{bmatrix} \alpha_{\bullet,1}^R & \beta_{\bullet,1}^R & \alpha_{\bullet,1}^I & \beta_{\bullet,1}^I & \cdots & \alpha_{\bullet,m}^R & \beta_{\bullet,m}^R & \alpha_{\bullet,m}^I & \beta_{\bullet,m}^I \end{bmatrix}^\top, \quad (14)$$

then I1 and I2 are encoded in the linear system

$$D\mathbf{c} = \begin{bmatrix} v_{\bullet,1}^R \\ v_{\bullet,1}^I \\ \vdots \\ v_{\bullet,m}^R \\ v_{\bullet,m}^I \\ \mathbf{0} \end{bmatrix}. \quad (15)$$

[TODO: reordering things can maybe clean this up?]

2.3 Step 3: Choosing Solution of Underconstrained System

References

- [1] Carlos Fernandez-Granda. Super-resolution of point sources via convex programming. *Information and Inference: A Journal of the IMA*, 5(3):251–303, 2016.