Robust recovery for super-resolution methods via optimal transport

Johan Karlsson
Department of Mathematics
KTH Royal Institute of Technology

Lipeng Ning
Brigham & Women's Hospital
Harvard Medical School

MTNS
University of Minnesota, Minneapolis, July 2016

Let $\mathbf{x}_{\text{true}}$ be sparse.

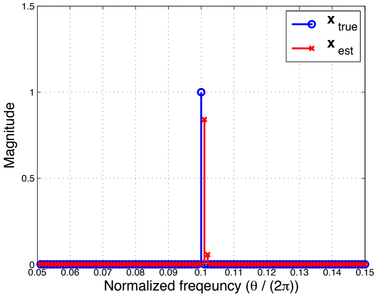
$$\begin{aligned}\text{Data:} \quad \mathbf{y} &= \mathbf{A}\mathbf{x}_{\text{true}} + \text{noise} \\ \text{Recovery:} \quad \mathbf{x}_{\text{est}} &= \arg\min \|\mathbf{x}\|_1 \\ &\quad \text{subject to } \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2 \leq \text{error}\end{aligned}$$

Distance between $\mathbf{x}_{\text{true}}$ and $\mathbf{x}_{\text{est}}$?

A incoherent: Robust recovery (ℓ_2) by Candès, Donoho, Tao, Tropp, etal.

A coherent or integral operator?

- Typical in spectral estimation, radar, etc.
- Robust recovery (ℓ_2) impossible.
- For $\mathbf{x}_{\text{true}}$ with sparse separated support:
  - Error bounds in Transportation distance
  - $\mathbf{x}_{\text{est}} \xrightarrow{w^*} \mathbf{x}_{\text{true}}$ in prob. as $\#\text{data} \to \infty$.



Outline

1. Sparse methods and spectral estimation
2. Reconstruction with a continuous dictionary
3. Motivating example and transportation distances
4. Worst case bounds
5. Convergence in probability

Model (discrete setting)

$$\mathbf{y} = \sum_{k=1}^{K} \mathbf{a}_k x_k + \mathbf{w} = \mathbf{A}\mathbf{x} + \mathbf{w}$$

Measurements:	$\mathbf{y} \in \mathbb{C}^N$	
Steering matrix:	$\mathbf{A} \in \mathbb{C}^{N \times K}$	typically $N \ll K$
Matrix columns (dictionary):	$\mathbf{a}_k \in \mathbb{C}^N$, for $k = 1, \ldots, K$	assume $\|\mathbf{a_k}\|_2 = 1$
Noise:	$\mathbf{w} \in \mathbb{C}^N$	assume $\|\mathbf{w}\|_2 \leq \epsilon$
Sought sparse vector:	$\mathbf{x} \in \mathbb{C}^K$	

Sparse recovery problem

Given measurement $\mathbf{y}$ and ϵ, find sparsest $\mathbf{x}$:

$$\min \#\{k \mid x_k \neq 0\} \quad \text{subject to} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2 \leq \epsilon$$

Combinatorial problem in general. Approach: Use a convex surrogate problem.

Sparse methods

- Find sparsest solution: a combinatorial problem.
- Use ℓ_1-norm as surrogate for the cardinality:

$$\arg\min_{\mathbf{x}\in\mathbb{C}^K} \|\mathbf{x}\|_1 \quad \text{subject to } \|\mathbf{y}-\mathbf{A}\mathbf{x}\|_2 \leq \epsilon. \tag{1}$$

Empirically: good recovery of the true $\mathbf{x}_{\text{true}}$.

Theorem (Candés, Wakin, 2008)

Assume that for $\delta_{2s} < \sqrt{2}-1$, the inequality

$$(1-\delta_{2s})\|\mathbf{x}\|_2^2 \leq \|\mathbf{A}\mathbf{x}\|_2^2 \leq (1+\delta_{2s})\|\mathbf{x}\|_2^2 \tag{RIP}$$

holds for all $2s-$sparse x. Let $\mathbf{y} = \mathbf{A}\mathbf{x}_{\text{true}} + \mathbf{w}$ where $\mathbf{x}_{\text{true}}$ is s-sparse and $\|\mathbf{w}\|_2 \leq \epsilon$, then the minimizer $\mathbf{x}_{\text{est}}$ of (1) satisfy

$$\|\mathbf{x}_{\text{true}} - \mathbf{x}_{\text{est}}\|_2 \leq C(\delta_{2s})\epsilon.$$

□

- Useful bound in several case: e.g., random matrix **A**.
- Coherency of **A**: $\delta_2 = \max_{k,\ell} |\mathbf{a}_k^*\mathbf{a}_\ell|$
- **A** highly coherent $\Rightarrow$ robustness results not applicable.
- What can we say in this case?

Example: Spectral estimation

Discrete-time signal $y_n \in \mathbb{C}$, $n \in \mathbb{Z}$. Sinusoids in noise:

$$y_n = \sum_{\ell=1}^{L} x_\ell e^{in\lambda_\ell} + w_n, \quad \text{for } n = 0, 1 \ldots, N-1$$

$\lambda_\ell \in [-\pi, \pi]$ frequency, $x_\ell \in \mathbb{C}$ magnitude and phase, $w_n \in \mathbb{C}$ error/noise.

Discretize frequency grid $\implies$ linear system

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{w} \tag{2}$$

columns of $\mathbf{A} \in \mathbb{C}^{N \times K}$ overcomplete Fourier basis:

$$\mathbf{a}(\theta_k) = \frac{1}{\sqrt{N}} \left(1, e^{i\theta_k}, \ldots, e^{i(N-1)\theta_k}\right)^T$$

for $\Omega = \{\theta_1, \theta_2, \ldots, \theta_K\}$. Here $\mathbf{x} \in \mathbb{C}^K$, $\mathbf{w} \in \mathbb{C}^N$, $N < K$.

(9) is underdetermined. ℓ_1 regularization popular to single out a unique solution.
(Chen, Donoho, 1998)

Example: Spectral estimation

Example:

- One sinusoid $\mathbf{y} := \mathbf{a}(0.1) + \mathbf{w}$
- $\|\mathbf{w}\|_2 = \epsilon = 10\%$
- Results in relative ℓ_2-error of 130%.

$\mathbf{A}$ is highly coherent: robust recovery in the sense of ℓ_2-norm is not possible.

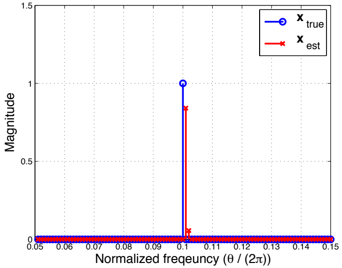


Figure: Here SNR:= $\|\mathbf{Ax}\|_2 / \|\mathbf{w}\|_2 = 10$, $N = 100$, and $K = 1000$.

Note: In high resolution spectral estimation,
typically $\delta_2 \geq 0.95 > \sqrt{2} - 1 \approx 0.41 \Rightarrow$ (RIP) condition not applicable.

Related work on Spectral estimation problem

- Candès and Fernandez-Grada (2014, Comm. on Pure and Appl. Math.): Assuiming support of $\mathbf{x}_{\text{true}}$ is $4K/N$ separated and $\|\mathbf{A}^*\mathbf{w}\|_1 \leq \epsilon$, then

$$\|\mathbf{x}_{\text{true}} - \mathbf{x}_{\text{est}}\|_1 \leq C\frac{K}{N^2}\epsilon.$$

  (+) Transparant condition on separation.
  (+) Exact recovery in the noiseless case.
  (-) Error bound in noisy case grows quickly as K/N increase.

- K., Ning (2014, IEEE CDC):
  (+) Error bounds on magnitude in terms of transportation distance
  (+) Convergence in probability of magnitude in weak topology
  (-) Only magnitude result
  (-) Less transparent separation condition.

- Tang, Bhaskar, Recht (2015, IEEE TIT) use atomic norm minimization:
  (+) Convergence in probability in the weak topology.
  (-) No explicit bounds for finite N.

Sparse recovery with a continuous dictionary

In many cases the model is given by the integral operator $\mathbf{A} : \mathfrak{M}(\Omega) \to \mathbb{C}^N$:

$$\mathbf{A}\mathbf{x} := \int_{\omega \in \Omega} \mathbf{a}(\omega) d\mathbf{x}(\omega).$$

$\mathbf{x} \in \mathfrak{M}(\Omega)$: set of bounded complex measures on $\Omega \subset \mathbb{R}^d$.
Inverse problem: find a sparse measure $\mathbf{x}$ with consistent with the linear system

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{w}.$$

Continuous dictionary $\mathcal{A} = \{\mathbf{a}(\omega), \omega \in \Omega\} \subset \mathbb{C}^N$.

Sparse recovery with a continuous dictionary

In many cases the model is given by the integral operator $\mathbf{A} : \mathfrak{M}(\Omega) \to \mathbb{C}^N$:

$$\mathbf{Ax} := \int_{\omega \in \Omega} \mathbf{a}(\omega) d\mathbf{x}(\omega).$$

$\mathbf{x} \in \mathfrak{M}(\Omega)$: set of bounded complex measures on $\Omega \subset \mathbb{R}^d$.
Inverse problem: find a sparse measure $\mathbf{x}$ with consistent with the linear system

$$\mathbf{y} = \mathbf{Ax} + \mathbf{w}.$$

Continuous dictionary $\mathcal{A} = \{\mathbf{a}(\omega), \omega \in \Omega\} \subset \mathbb{C}^N$.

- Frequency estimation, spectral estimation, direction of arrival:

$$[\mathbf{a}(\omega)]_n = \frac{1}{\sqrt{N}} e^{i(n-1)\omega}, \quad n = 1 \dots N, \quad \omega \in [0, 2\pi).$$

- d-dimensional frequency estimation, e.g., synthetic aperture radar (SAR):

$$[\mathbf{a}(\omega)]_n = \frac{1}{\sqrt{N}} e^{ik_n^T \omega}, \; k_n \in \mathbb{Z}^d, n = 1 \dots N, \quad \omega \in [0, 2\pi)^d.$$

- Delay estimation, e.g., radar and sonar:

$$[\mathbf{a}(\omega)]_n = s(\omega - t_n), \quad n = 1 \dots N, \quad \omega \in [\omega_{\min}, \omega_{\max}],$$

and where $s(t)$ is the probing waveform and $t_1, \dots, t_N$ is a set of measuring points.

Sparse recovery with a continuous dictionary

- Measurements:

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{w} = \int_{\omega \in \Omega} \mathbf{a}(\omega) d\mathbf{x}(\omega) + \mathbf{w},$$

  where $\mathbf{y}, \mathbf{w} \in \mathbb{R}^N$, and $\|\mathbf{w}\|_2 \leq \epsilon$.

- Reconstruction:

$$\arg \min_{\mathbf{x} \in \mathfrak{M}(\Omega)} \|\mathbf{x}\|_1 \quad \text{subject to } \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2 \leq \epsilon.$$

In this setting $\|\mathbf{x}\|_1 = \int_{\omega \in \Omega} d|\mathbf{x}|(\omega)$ is the total variation of the measure.

What can be said about robustness of this problem?

Motivating Example

To illustrate the principle, assume all measures are nonnegative,[1] and assume $d\mathbf{x}_{\text{true}} = \delta_\lambda(\omega)d\omega$ is a measure with support in only one point λ.

The measurement is given by

$$\mathbf{y} = \mathbf{A}\mathbf{x}_{\text{true}} + \mathbf{w} = \mathbf{a}(\lambda) + \mathbf{w}, \quad \text{where } \|\mathbf{w}\|_2 \le \epsilon.$$

$\mathbf{x}_{\text{est}}$ recovered solution from

$$\mathbf{x}_{\text{est}} = \arg \min_{\mathbf{x} \in \mathfrak{M}_+(\Omega)} \|\mathbf{x}\|_1 \text{ subject to } \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2 \le \epsilon.$$

[1] Problem can be lifted to non-negative measures $\tilde{\mathbf{x}}(\tau,\omega) = |x(\omega)|\delta(\tau - \arg(x(\omega)))$ on $\Omega \times [-\pi,\pi]$, where $\arg(x(\omega))$ is the phase in the polar representation $dx(\omega) = e^{i\arg(x(\omega))}d|x(\omega)|$.

Motivating Example

To illustrate the principle, assume all measures are nonnegative,[1] and assume $d\mathbf{x}_{\text{true}} = \delta_\lambda(\omega)d\omega$ is a measure with support in only one point λ.

The measurement is given by

$$\mathbf{y} = \mathbf{A}\mathbf{x}_{\text{true}} + \mathbf{w} = \mathbf{a}(\lambda) + \mathbf{w}, \quad \text{where } \|\mathbf{w}\|_2 \leq \epsilon.$$

$\mathbf{x}_{\text{est}}$ recovered solution from

$$\mathbf{x}_{\text{est}} = \arg \min_{\mathbf{x} \in \mathfrak{M}_+(\Omega)} \|\mathbf{x}\|_1 \text{ subject to } \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2 \leq \epsilon.$$

Note that

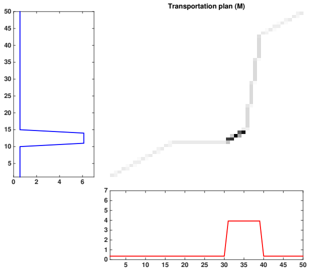
$$\begin{aligned}
\epsilon &\geq \|\mathbf{A}\mathbf{x}_{\text{est}} - \mathbf{y}\|_2 \geq |\mathbf{a}(\lambda)^*(\mathbf{A}\mathbf{x}_{\text{est}} - \mathbf{y})| \\
&= \left| \underbrace{\mathbf{a}(\lambda)^*\mathbf{a}(\lambda)}_{=1=\|\mathbf{x}_{\text{true}}\|_1} - \int_{\omega\in\Omega} \mathbf{a}(\lambda)^*\mathbf{a}(\omega)d\mathbf{x}_{\text{est}}(\omega) + \underbrace{\mathbf{a}(\lambda)^*\mathbf{w}}_{|\cdot|\leq\epsilon} \right| \\
&\geq \|\mathbf{x}_{\text{true}}\|_1 - \int_{\omega\in\Omega} \mathfrak{Re}(\mathbf{a}(\lambda)^*\mathbf{a}(\omega))d\mathbf{x}_{\text{est}}(\omega) - \epsilon \\
&= \int_{\omega\in\Omega} \underbrace{(1 - \mathfrak{Re}(\mathbf{a}(\lambda)^*\mathbf{a}(\omega)))d\mathbf{x}_{\text{est}}(\omega)}_{\text{"mass" transport}} + \underbrace{\|\mathbf{x}_{\text{true}}\|_1 - \|\mathbf{x}_{\text{est}}\|_1}_{\text{"mass" deviation}} - \epsilon.
\end{aligned}$$

[1] Problem can be lifted to non-negative measures $\tilde{\mathbf{x}}(\tau, \omega) = |x(\omega)|\delta(\tau - \arg(x(\omega)))$ on $\Omega \times [-\pi, \pi]$, where $\arg(x(\omega))$ is the phase in the polar representation $dx(\omega) = e^{i \arg(x(\omega))} d|x(\omega)|$.

Transportation distances

Transportation cost (Monge 1781, Kantorovich 1942):

$$
\begin{aligned}
T(\mathbf{x}_0, \mathbf{x}_1) := \min_{M \in \mathfrak{M}_+(\Omega \times \Omega)} \quad & \int_{\theta,\omega\in\Omega} c(\theta,\omega) dM(\theta,\omega) \\
\text{subject to} \quad & \int_{\omega\in\Omega} dM(\theta,\omega) = d\mathbf{x}_0(\theta) \\
& \int_{\theta\in\Omega} dM(\theta,\omega) = d\mathbf{x}_1(\omega) \\
& M(\theta,\omega) \geq 0, \quad \theta,\omega \in \Omega
\end{aligned}
$$



Relaxed transportation cost allowing for different masses (K., Georgiou, Takyar, 2009):

$$
\tilde{T}(\mathbf{x}_0, \mathbf{x}_1) \quad := \min_{\|\boldsymbol{\rho}_0\|_1 = \|\boldsymbol{\rho}_1\|_1} \left(T(\boldsymbol{\rho}_0, \boldsymbol{\rho}_1) + \sum_{j=0}^{1} \|\mathbf{x}_j - \boldsymbol{\rho}_j\|_1 \right).
$$

$$2\epsilon \geq \int_{\omega\in\Omega} \underbrace{(1 - \Re e(\mathbf{a}(\lambda)^* \mathbf{a}(\omega))) d\mathbf{x}_{\text{est}}(\omega)}_{\text{“mass” transport}} + \underbrace{\|\mathbf{x}_{\text{true}}\|_1 - \|\mathbf{x}_{\text{est}}\|_1}_{\text{“mass” deviation}}.$$

- Term 1: Transportation cost of transporting unit mass from ω to λ

$$c(\lambda, \theta) = 1 - \Re e(\mathbf{a}(\lambda)^* \mathbf{a}(\omega)).$$

- Term 2: Deviation in the total mass. By optimality $\|\mathbf{x}_{\text{true}}\|_1 - \|\mathbf{x}_{\text{est}}\|_1 \geq 0$.

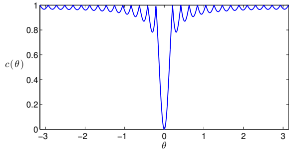
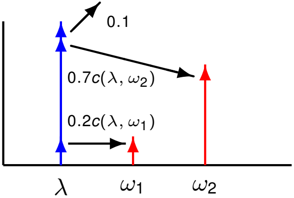


Figure: Spectral estimation: $c(0, \omega)$ for $N = 30$.

Example:

$\mathbf{x}_{\text{true}} = \mathbf{e}(\lambda)$ and $\mathbf{x}_{\text{est}} = 0.2\mathbf{e}(\omega_1) + 0.7\mathbf{e}(\omega_2)$

$$\Rightarrow 2\epsilon \geq 0.2c(\lambda, \omega_1) + 0.7c(\lambda, \omega_2) + 0.1.$$



Example: Spectral estimation

Example:

- One sinusoid $\mathbf{y} := \mathbf{a}(0.1) + \mathbf{w}$
- $\|\mathbf{w}\|_2 = \epsilon = 10\%$
- Results in relative ℓ_2-error of 130%.

A is highly coherent: robust recovery in the sense of ℓ_2-norm is not possible.

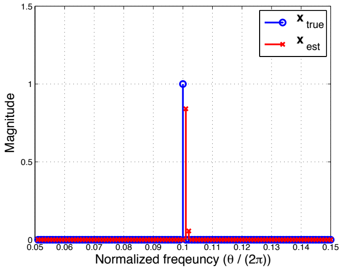


Figure: Here SNR:$= \|\mathbf{Ax}\|_2 / \|\mathbf{w}\|_2 = 10$, $N = 100$, and $K = 1000$.

Worst case bounds

General case: $$d\mathbf{x}_{\text{true}} = \sum_{\lambda \in \Lambda} x_\lambda \delta_\lambda(\omega) d\omega, \quad \text{supp}(\mathbf{x}_{\text{true}}) = \Lambda \subset \Omega$$

Definition

Let **A** be a dictionary with index set Ω and let $\Lambda \subset \Omega$. Define

$$\mu_\Lambda \quad := \quad \max_{\theta \in \Omega} \left(\sum_{\lambda \in \Lambda} |\mathbf{a}(\lambda)^* \mathbf{a}(\theta)| - \max_{\lambda \in \Lambda} |\mathbf{a}(\lambda)^* \mathbf{a}(\theta)| \right),$$

which we denote by the *cumulative intercoherence*.

μ_Λ: quantifies sparsity and separateness of Λ .

Proposition (Spectral estimation)

$$\mu_\Lambda \leq \frac{(|\Lambda| - 1)\pi}{N\Delta(\Lambda)}, \qquad \textit{where}$$

$$\Delta(\Lambda) = \min_{\lambda_0, \lambda_1 \in \Lambda, \lambda_0 \neq \lambda_1} |\lambda_0 - \lambda_1|.$$

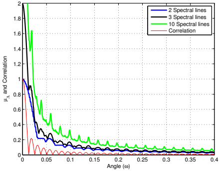


Figure: Intercoherence/$\Delta\Lambda$: $|\Lambda| = L$ where $L \in \{2, 3, 10\}$.

Worst case bounds

Let $\mathbf{y} = \mathbf{A}\mathbf{x}_{\text{true}} + \mathbf{w}$ and take $\mathbf{x}_{\text{est}}$ to be the minimizer of

$$\arg \min_{\mathbf{x} \in \mathfrak{M}(\Omega)} \|\mathbf{x}\|_1 \quad \text{subject to } \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2 \leq \epsilon,$$

We will derive the bound based on the following properties

$$\begin{aligned} 3a) &\quad \mathbf{x}_{\text{true}} \text{ has support } \Lambda, \\ 3b) &\quad \|\mathbf{x}_{\text{est}}\|_1 \leq \|\mathbf{x}_{\text{true}}\|_1, \\ 3c) &\quad \|\mathbf{A}\mathbf{x}_{\text{est}} - \mathbf{A}\mathbf{x}_{\text{true}}\|_2 \leq 2\epsilon. \end{aligned} \tag{3}$$

Theorem

Let $\mathbf{x}_{\text{true}} \in \mathbb{C}^{K \times 1}$ *be a vector with support* Λ*, and let* $\|\mathbf{w}\|_2 \leq \epsilon$*. Then the optimal solution* $\mathbf{x}_{\text{est}}$ *satisfies*

$$\tilde{T}(\tilde{\mathbf{x}}_{\text{est}}, \tilde{\mathbf{x}}_{\text{true}}) \leq 6 \left(\epsilon \sqrt{|\Lambda|(1 + \mu_\Lambda)} + \mu_\Lambda \|\mathbf{x}_{\text{true}}\|_1 \right),$$

where $\tilde{\mathbf{x}}(\tau, \omega) = |x(\omega)|\delta(\tau - \arg(x(\omega)))$ *on* $\Omega \times [-\pi, \pi]$*, where* $\arg(x(\omega))$ *is the phase in the polar representation* $dx(\omega) = e^{i \arg(x(\omega))} d|x(\omega)|$.

Error bounds with given confidence level

If noise $\geq$ signal, worst case bounds useless.

If $\mathbf{w}$ nearly orthogonal to all dictionary elements of $\mathbf{A}$, i.e., $\kappa \ll 1$ and

$$|\mathbf{a}(\omega)^* \mathbf{w}| \leq \kappa \|\mathbf{w}\|_2 \quad \text{for all} \quad \omega \in \Omega. \tag{4}$$

In addition assume $\|\mathbf{w}\|_2 = \epsilon$ and

$$\|\mathbf{A}\mathbf{x}_{\text{est}} - \mathbf{A}\mathbf{x}_{\text{true}} - \mathbf{w}\|_2 \leq \epsilon.$$

Then it follows that

$$\begin{aligned} \|\mathbf{A}\mathbf{x}_{\text{est}} - \mathbf{A}\mathbf{x}_{\text{true}}\|_2^2 &\leq 2|\mathbf{w}^* \mathbf{A}(\mathbf{x}_{\text{est}} - \mathbf{x}_{\text{true}})| \\ &\leq 2\kappa\epsilon(\|\mathbf{x}_{\text{est}}\|_1 + \|\mathbf{x}_{\text{true}}\|_1) \\ &\leq 4\kappa\epsilon\|\mathbf{x}_{\text{true}}\|_1. \end{aligned}$$

Theorem

Let $\mathbf{x}_{\text{true}} \in \mathbb{C}^{K\times 1}$ be a vector with support Λ, and let $\|\mathbf{w}\|_2 = \epsilon$ which satisfies (4). Then the optimal solution $\mathbf{x}_{\text{est}}$ satisfies

$$\tilde{T}(\tilde{\mathbf{x}}_{\text{est}}, \tilde{\mathbf{x}}_{\text{true}}) \leq 6\left(\sqrt{2\kappa\epsilon\|\mathbf{x}_{\text{true}}\|_1 |\Lambda|(1+\mu_\Lambda)} + \mu_\Lambda \|\mathbf{x}_{\text{true}}\|_1\right).$$

Error bounds with given confidence level (spectral estimation)

- When is the near orthogonality assumption justified?
- For spectral estimation:

Proposition (K., Ning, 2014)

Let $\kappa \in (0,1)$ be given and let $\mathbf{w} \in \mathbb{C}^N$ be a random vector that is complex Gaussian with zero mean and unit variance. Then

$$\mathrm{Prob}\left(\max_{0\le\theta\le 2\pi} |\mathbf{a}(\theta)^*\mathbf{w}| \ge \kappa\|\mathbf{w}\|_2\right) \le (1-\delta)^{N-3/2}\left(1+\kappa\frac{e^2}{\sqrt{6}}N^{3/2}\right).$$

$\Rightarrow \max_{0\le\theta\le 2\pi} \frac{|\mathbf{a}(\theta)^*\mathbf{w}|}{\|\mathbf{w}\|_2} \to 0$ as $N \to \infty$ in probability.

Error bounds with given confidence level (spectral estimation)

Let $N \in \mathbb{N}$ and let the signal in noise

$$y_n = \sum_{\lambda \in \Lambda}^{L} e^{in\lambda} x(\lambda) + w_n, \quad \text{for } n = 0, \dots, N-1$$

where $w_n \in CN(0, \sigma^2)$ is white Gaussian noise.

Let

$$\mathbf{x}_{\text{est}}^N = \arg \min_{\mathbf{x}_N \in \mathfrak{M}(\Omega)} \|\mathbf{x}_N\|_1 \quad \text{subject to } \|\mathbf{y}_N - \mathbf{A}_N \mathbf{x}_N\|_2 \leq \epsilon_N = \|\mathbf{w}_N\|,$$

Then

$$\tilde{T}_N(\tilde{\mathbf{x}}_{\text{est}}^N, \tilde{\mathbf{x}}_{\text{true}}) \to 0 \quad \text{in probability, as } N \to \infty.$$

And hence

$$\mathbf{x}_{\text{est}}^N \xrightarrow{\text{w}^*} \mathbf{x}_{\text{true}} \quad \text{in probability, as } N \to \infty.$$

Conclusions and further directions

Conclusions

- Standard ℓ_2 distance is not appropriate for quantifying uncertainty in high-resolution spectral estimation based on ℓ_1 regularization.
- Framework for sparse recovery for structured dictionaries with error bounds based on transportation distance.
- Error bounds: worst-case (general) and bounds that hold with a guaranteed probability (spectral estimation).
- The latter bound go to 0 (in probability) as number of data points go to ∞.

Further work

- Extend confidence bounds to general dictionaries.
- Extend framework to certain non-sparse cases, i.e. support regions.
- Consider problems with, e.g., sparse gradient.
- Towards a quantitative framework for comparing spectral estimation methods.

Conclusions and further directions

Conclusions

- Standard ℓ_2 distance is not appropriate for quantifying uncertainty in high-resolution spectral estimation based on ℓ_1 regularization.
- Framework for sparse recovery for structured dictionaries with error bounds based on transportation distance.
- Error bounds: worst-case (general) and bounds that hold with a guaranteed probability (spectral estimation).
- The latter bound go to 0 (in probability) as number of data points go to ∞.

Further work

- Extend confidence bounds to general dictionaries.
- Extend framework to certain non-sparse cases, i.e. support regions.
- Consider problems with, e.g., sparse gradient.
- Towards a quantitative framework for comparing spectral estimation methods.

Thank you for your attention!