Nullspace Condition, Uncertainty Relation, and Recovery Guarantee for Signals with Low Density

Christoph Studer

C. Studer, “Recovery of Signals with Low Density,”
arXiv preprint: arxiv.org/abs/1507.02821, July 2015

Some well-known results

Let $\mathbf{A} \in \mathbb{C}^{M \times N_a}$ be a **dictionary** with coherence μ_a:

- $\|\mathbf{a}_i\|_2 = 1$ for $i = 1, \ldots, N_a$
- $\mu_a = \max_{i \neq j} |\mathbf{a}_i^H \mathbf{a}_j|$

Let $\mathbf{x} \in \mathbb{C}^{N_a}$ be an **n_x-sparse vector**

- $n_x = \|\mathbf{x}\|_0 =$ number of non-zero entries in $\mathbf{x}$

Theorem (Nullspace condition)

If $n_x < 1 + \frac{1}{\mu_a}$, then $\mathbf{A}\mathbf{x} \neq \mathbf{0}_{M \times 1}$

Corollary (Uniqueness condition)

Assume $\mathbf{y} = \mathbf{A}\mathbf{x}$. If $n_x < \frac{1}{2}(1 + \frac{1}{\mu_a})$, then $\mathbf{x}$ is unique

Some well-known results (cont'd)

How can we recover $\mathbf{x}$ from $\mathbf{y} = \mathbf{A}\mathbf{x}$?

- Combinatorial problem: minimize $\|\mathbf{x}\|_0$ subject to $\mathbf{y} = \mathbf{A}\mathbf{x}$
- Convex problem: minimize $\|\mathbf{x}\|_1$ subject to $\mathbf{y} = \mathbf{A}\mathbf{x}$
- Orthogonal matching pursuit (OMP)
- And many more...

Theorem (Recovery guarantee)

If $n_x < \frac{1}{2}(1 + \frac{1}{\mu_a})$, then $\mathbf{x}$ can be recovered from $\mathbf{y} = \mathbf{A}\mathbf{x}$ using ℓ_0-norm and ℓ_1-norm minimization, as well as OMP

Some well-known results (cont'd)

Let $\mathbf{B} \in \mathbb{C}^{M \times N_b}$ be another **dictionary** with coherence μ_b:

Let μ_m be the mutual coherence between **A** and **B**

- $\mu_m = \max_{i,j} |\mathbf{a}_i^H \mathbf{b}_j|$

Let $\mathbf{z} \in \mathbb{C}^{N_a}$ be another n_z**-sparse vector**

Theorem (Uncertainty relation)

Let $\mathbf{x}$ *and* $\mathbf{z}$ *be two vectors satisfying* $\mathbf{Ax} = \mathbf{Bz}$*. Then, the following holds* $$[1 - \mu_a(n_x - 1)]^+[1 - \mu_b(n_z - 1)]^+ \leq n_x n_z \mu_m^2$$

Corollary (UR for pairs of orthonormal bases)

Let **A** *and* **B** *be orthonormal bases. If* $\mathbf{Ax} = \mathbf{Bz}$*, then* $\frac{1}{\mu_m^2} \leq n_x n_z$

Do we really need sparsity?

- Signal sparsity is ubiquitous and central in the signal-recovery and compressive-sensing literature
- **Is sparsity is the key for such nullspace conditions, signal recovery problems, and uncertainty relations?**

Definition (δ-density)

For a non-zero signal $\mathbf{x} \in \mathbb{C}^{N_a}$, we define the δ-density as

$$\delta(\mathbf{x}) = \frac{\|\mathbf{x}\|_1}{\|\mathbf{x}\|_\infty}.$$

For an all-zero signal $\mathbf{x} = \mathbf{0}_{N_a \times 1}$, we define $\delta(\mathbf{x}) = 0$

Key properties of the δ-density

- For arbitrary signals $\mathbf{x} \in \mathbb{C}^{N_a}$, the δ-density satisfies

$$0 \leq \delta(\mathbf{x}) \leq \|\mathbf{x}\|_0 \leq N_a$$

- **Equality $\delta(\mathbf{x}) = \|\mathbf{x}\|_0$ holds if and only if the non-zero entries of x have constant modulus**

- **Signals with decaying magnitude profile have $\delta(\mathbf{x}) < \|\mathbf{x}\|_0$**
- Example: $x_i = \alpha^{i-1}$, $i = 1, \ldots, N_a$ with $0 < \alpha < 1 - 1/N_a$

$$\delta(\mathbf{x}) \leq (1-\alpha)^{-1} < \|\mathbf{x}\|_0 = N_a$$

The δ-density $\delta(\mathbf{x})$ behaves similarly to the sparsity $\|\mathbf{x}\|_0$,
but captures properties of the signal's magnitude

Nullspace condition for signals with low density

Theorem (ℓ_∞-norm restricted isometry)

Let $\mathbf{A} \in \mathbb{C}^{M \times N_a}$ *be a dictionary with coherence* μ_a *and* $\mathbf{x} \in \mathbb{C}^{N_a}$ *a nonzero signal. Then, the following inequalities hold:*

$$1 - \mu_a(\delta(\mathbf{x}) - 1) \leq \frac{\left\|\mathbf{A}^H \mathbf{A}\mathbf{x}\right\|_\infty}{\|\mathbf{x}\|_\infty} \leq 1 + \mu_a(\delta(\mathbf{x}) - 1)$$

Corollary (δ-density-based nullspace condition)

If $\delta(\mathbf{x}) < 1 + \frac{1}{\mu_a}$, *then* $\mathbf{A}\mathbf{x} \neq \mathbf{0}_{M \times 1}$.

- Example: For the α-decaying signal $x_i = \alpha^{i-1}$, $i = 1, \ldots, N_a$, with $0 < \alpha < 1 - (1 + \frac{1}{\mu_a})^{-1}$, we have $\|\mathbf{x}\|_0 = N_a$ but $\mathbf{A}\mathbf{x} \neq \mathbf{0}_{M \times 1}$
- For signals with constant-modulus non-zero entries, we recover the well-known nullspace condition $\delta(\mathbf{x}) = \|\mathbf{x}\|_0 < 1 + \frac{1}{\mu_a}$

δ-density-based uncertainty relation

Theorem (Uncertainty relation)

Let $\mathbf{x}$ *and* $\mathbf{z}$ *be two vectors satisfying* $\mathbf{Ax} = \mathbf{Bz}$. *Then, the following holds* $[1 - \mu_a(\delta(\mathbf{x}) - 1)]^+[1 - \mu_b(\delta(\mathbf{z}) - 1)]^+ \leq \delta(\mathbf{x})\delta(\mathbf{z})\mu_m^2$

Corollary (UR for pairs of orthonormal bases)

Let $\mathbf{A}$ *and* $\mathbf{B}$ *be orthonormal bases. If* $\mathbf{Ax} = \mathbf{Bz}$, *then* $\frac{1}{\mu_m^2} \leq \delta(\mathbf{x})\delta(\mathbf{z})$

- Example: For the Fourier-identity pair we have $N \leq \delta(\mathbf{x})\delta(\mathbf{z})$
- Implies that signals $\mathbf{x}$ and $\mathbf{z}$ cannot have low density in Fourier and identity domain the same time
- For signals with constant-modulus non-zero entries, we recover the Donoho–Stark uncertainty relation $N \leq \|\mathbf{x}\|_0\|\mathbf{z}\|_0$

How about signal recovery?

Can we recover $\mathbf{x}$ from $\mathbf{y} = \mathbf{A}\mathbf{x}$ if signal $\mathbf{x}$ has low density?

We consider orthogonal matching pursuit (OMP):

- Initialize residual $\mathbf{r}^{(0)} = \mathbf{y}$ and empty support set $\mathcal{S}^{(0)} = \varnothing$, and repeat the following two steps for $t = 1, \ldots, t_{\max}$ iterations:
  1. Select a column of the dictionary $\mathbf{A}$ according to
  $$\hat{k}^{(t)} = \arg\max_{i \in \mathcal{R}^{(t-1)}} |\mathbf{a}_i^H \mathbf{r}^{(t-1)}|$$
  The set $\mathcal{R}^{(t-1)} = \{1, \ldots, N_a\} \backslash \mathcal{S}^{(t-1)}$ contains all remaining indices that are not (yet) in the support set $\mathcal{S}^{(t-1)}$
  2. Add $\hat{k}^{(t)}$ to set $\mathcal{S}^{(t)} = \mathcal{S}^{(t-1)} \cup \hat{k}^{(t)}$ and compute new residual
  $$\mathbf{r}^{(t)} = (\mathbf{I}_M - \mathbf{A}_{\mathcal{S}^{(t)}} \mathbf{A}^{\dagger}_{\mathcal{S}^{(t)}})\mathbf{y}$$

δ-density-based recovery guarantee

Theorem (Recovery guarantee)

Let the maximum number of iterations be $t_{max} < 1 + \frac{1}{\mu_a}$. Assume that in every iteration $t = 1, \ldots, t_{max}$ the entries in $\mathbf{x}$ satisfy

$$\delta(\mathbf{x}_{\mathcal{R}^{(t-1)}}) < \frac{1}{2}\Big(1 + \frac{1}{\mu_a} - (t-1)\Big).$$

Then, OMP will always select an atom associated to the largest coefficient from $\mathbf{x}_{\mathcal{R}^{(t)}}$ in iteration t. Furthermore, the set $\mathcal{S}^{(t)}$ contains the indices associated to the $t = |\mathcal{S}^{(t)}|$ largest entries in $\mathbf{x}$.

- Condition to succeed in first iteration is $\delta(\mathbf{x}) < \frac{1}{2}(1 + \frac{1}{\mu_a})$
- Condition to succeed in second iteration is

  $$\delta(\mathbf{x}_{\mathcal{R}^{(1)}}) < \frac{1}{2}(1 + \frac{1}{\mu_a} - 1),$$

  whose RHS is more restrictive. However, $\delta(\mathbf{x}_{\mathcal{R}^{(1)}})$ may be smaller; depends on magnitude decay profile of $\mathbf{x}$

Example: Recovery of α-decaying signal

- For signal $x_i = \alpha^{i-1}$, $i = 1, \ldots, N_a$, we have $\delta(\mathbf{x}) \leq (1-\alpha)^{-1}$
- Removing t largest entries still satisfies $\delta(\mathbf{x}_{\mathcal{R}^{(t)}}) \leq (1-\alpha)^{-1}$
- Decay condition becomes $t < 2 + \frac{1}{\mu_a} - 2(1-\alpha)^{-1}$

- For very fast coefficient decay, i.e., $\alpha \to 0$, we can perform up to $t_{\max} < \frac{1}{\mu_a}$ OMP iterations
- **OMP is able to identify the largest $t_{\max}$ coefficients, even for signals having up to N_a non-zero coefficients**

The condition $t_{\max} < \frac{1}{\mu_a}$ is roughly $2\times$ less restrictive than the sparsity-based recovery condition $\|\mathbf{x}\|_0 < \frac{1}{2}(1 + \frac{1}{\mu_a})$

Extensions for signals with small block-density

- We can also generalize our results to block-sparse signals
- Signal model $\mathbf{y} = \sum_{b=1}^{B} \mathbf{A}_b \mathbf{x}_b$ with B blocks

Definition (Block density)

For a non-zero signal $\mathbf{x} \in \mathbb{C}^{N_a}$, we define the block density as

$$\beta(\mathbf{x}) = \frac{\sum_{b=1}^{B} \|\mathbf{x}_b\|_2}{\max_{b=1,\ldots,B} \|\mathbf{x}_b\|_2}.$$

For an all-zero signal $\mathbf{x} = \mathbf{0}_{N_a \times 1}$, we define $\beta(\mathbf{x}) = 0$

- Define the block coherence of $\mathbf{A}$ as $\mu^A = \max_{b \neq \ell} \frac{\sigma_{\max}(\mathrm{A}_b^H \mathrm{A}_\ell)}{\sigma_{\min}^2(\mathrm{A}_b)}$

Theorem (Nullspace condition)

If $\beta(\mathbf{x}) < 1 + \frac{1}{\mu^A}$, then $\sum_{b=1}^{B} \mathbf{A}_b \mathbf{x}_b \neq \mathbf{0}_{M \times 1}$

Extensions for signal with small block-density (cont'd)

- Define the block coherence of $\mathbf{B}$ as $\mu^B = \max_{b \neq \ell} \frac{\sigma_{\max}(\mathsf{B}_b^H \mathsf{B}_\ell)}{\sigma_{\min}^2(\mathsf{B}_b)}$
- Define the following mutual block coherences:

$$\mu_m^A = \max_{b,\ell} \frac{\sigma_{\max}(\mathbf{A}_b^H \mathbf{B}_\ell)}{\sigma_{\min}(\mathbf{A}_b)} \quad \text{and} \quad \mu_m^B = \max_{b,\ell} \frac{\sigma_{\max}(\mathbf{A}_b^H \mathbf{B}_\ell)}{\sigma_{\min}(\mathbf{B}_b)}$$

Theorem (Uncertainty relation)

Let $\sum_{b=1}^{B} \mathbf{A}_b \mathbf{x}_b = \sum_{b'=1}^{B'} \mathbf{B}_{b'} \mathbf{z}_{b'}$. *Then, the following holds*

$$[1 - \mu_a(\beta(\mathbf{x}) - 1)]^+ [1 - \mu_b(\beta(\mathbf{z}) - 1)]^+ \leq \beta(\mathbf{x})\beta(\mathbf{z})\mu_m^A \mu_m^B$$

- Generalizes our results to un-normalized matrices $\mathbf{A}$ and $\mathbf{B}$
- **We can derive a block OMP recovery guarantee**
- We can also analyze the **case of bounded measurement noise**

Extensions to other density measures

There are other ways of measuring the signal density!

Definition (γ-density or "effective sparsity")

For a non-zero signal $\mathbf{x} \in \mathbb{C}^{N_a}$*, we define the* γ*-density as*

$$\gamma(\mathbf{x}) = \frac{\|\mathbf{x}\|_1^2}{\|\mathbf{x}\|_2^2}.$$

For an all-zero signal $\mathbf{x} = \mathbf{0}_{N_a \times 1}$*, we define* $\gamma(\mathbf{x}) = 0$

- **We can derive equivalent results by replacing δ by γ**
- However, the γ-density is more restrictive, i.e., we have

$$\delta(\mathbf{x}) \leq \gamma(\mathbf{x}) \leq \|\mathbf{x}\|_0$$

Conclusions

- Sparsity is ubiquitous in the literature on signal recovery, compressive sensing, and uncertainty relations
- In practice, signals are not necessarily sparse
- The decay profile of the signals of interest does matter

- With a proper definition of signal density, we obtain nullspace conditions, uncertainty relations, and recovery guarantees
- The delta density $\delta(\mathbf{x}) = \|\mathbf{x}\|_1 / \|\mathbf{x}\|_\infty$ enables us to capture crucial magnitude information, beyond the number of nonzeros

- Non-sparse signals with sufficiently fast decaying coefficients cannot be in the nullspace of a matrix
- OMP-based recovery allows up to $2\times$ more non-zero entries if the magnitudes decay quickly

Two open problems

Problem 1:

- There is a density that is less-restrictive than the δ-density, i.e.,
$$\sigma(\mathbf{x}) = \frac{\|\mathbf{x}\|_2^2}{\|\mathbf{x}\|_\infty^2} \leq \delta(\mathbf{x})$$
- Can we derive nullspace conditions, uncertainty relations, and recovery guarantees with this σ-density?

Problem 2:

- What can we say about the recovery performance of algorithms that solve the following problem:
$$\text{minimize } \delta(\mathbf{x}) \text{ subject to } \mathbf{y} = \mathbf{A}\mathbf{x} \quad ?$$
- And how can we solve such problems efficiently?

Visit csl.cornell.edu/∼studer for more information

Thanks to R. Baraniuk, H. Bőlcskei, and T. Goldstein