Basis pursuit II

Arianna Rast

LMU Munich

June 1, 2025

Summary

Review Exact Reconstruction

- For $A \in \mathbb{K}^{m\times N}$ and $s \in \mathbb{N}$ we have that

$$\left.\begin{array}{c} \forall x \in \mathbb{C}^N \text{ } s\text{-sparse}: \\ x \text{ is the unique minimizer} \\ \text{of } \left\{ \|z\|_1 \,\middle|\, Az = Ax \right\} \end{array}\right\} \iff A \text{ satisfies the NSP of order } s.$$

- Also,

$$\left.\begin{array}{c} \forall x \in \mathbb{C}^N \text{ } s\text{-sparse}: \\ x \text{ is the unique minimizer} \\ \text{of } \left\{ \|z\|_1 \,\middle|\, Az = Ax \right\} \end{array}\right\} \iff \left\{\begin{array}{c} \forall x \in \mathbb{C}^N \text{ } s\text{-sparse}: \\ \arg\min \left\{ \|z\|_1 \,\middle|\, Az = Ax \right\} \\ = \arg\min \left\{ \|z\|_0 \,\middle|\, Az = Ax \right\} \end{array}\right.$$

Stimmt die Rueckrichtung auch? Hence, the NSP of order s is a necessary and sufficient condition for the exact reconstruction of every s-sparse vector via the basis pursuit.

Review Stability

- The basis pursuit is stable under a sparsity defect in the vector x, if the measurement matrix satisfies the stable null space property (SNSP).
- For a matrix $A \in \mathbb{C}^{m\times N}$, we have

$$\left.\begin{array}{c}\forall x, z \in \mathbb{C}^N : \\ \|z - x\|_1 \leq \frac{1+\rho}{1-\rho}\left(\|z\|_1 - \|x\|_1 + 2\|x_{\overline{S}}\|_1\right)\end{array}\right\} \iff \begin{array}{c} A \text{ satisfies} \\ \mathrm{SNSP}(\rho, S)\end{array} .$$

- In particular, if A satisfies the $\mathrm{SNSP}(\rho, S)$, any minimizer $x^{\#}$ of $\left\{\|z\|_1 \,\middle|\, Ax = Az\right\}$ satisfies

$$\|x - x^{\#}\|_1 \leq \frac{2(1+\rho)}{1-\rho}\sigma_s(x)_1 .$$

Robustness

In this chapter we also want to handle noise in the measurement in addition to a sparsity deficit of the data.
What additional assumptions do we need to obtain similar results, if we consider the problem

$$\min \left\{ \|z\|_1 \,\middle|\, z \in \mathbb{C}^N, \;\; \|Az - y\| \leq \eta \right\} ? \tag{$P_{1,\eta}$}$$

It depends on the norm, in which we measure the error, i.e. the distance of TODO

At first, we consider the situation where we measure the noise in the ℓ^1-norm, i.e. $\|Az - y\| = \|Az - y\|_1$.

Definition 1.1: Robust null space property

A matrix $A \in \mathbb{C}^{m\times N}$ is said to satisfy the **robust null space property** with respect to $\|\cdot\|$ with the constants $\rho \in (0,1)$ and $\tau > 0$ relative to a set $S \subseteq [N]$ iff

$$\forall v \in \mathbb{C}^N : \quad \|v_S\|_1 \leq \rho \|v_{\overline{S}}\|_1 + \tau \|Av\|\,. \qquad (\mathrm{RNSP}(\|\cdot\|, \rho, \tau, S))$$

A satisfies the robust null space property of **order** s with respect to $\|\cdot\|$ with the constants $\rho \in (0,1)$ and $\tau > 0$ $\mathrm{RNSP}(\|\cdot\|, \rho, \tau, s)$ iff A satisfies $\mathrm{RNSP}(\|\cdot\|, \rho, \tau, S)$ for all sets $S \subseteq [N]$ with $|S| \leq s$.

Intuition for the null space property? Is it a common property or rather rare? Is it true, that it is relatively hard to verify?

- Maybe the content of this slide is just done on the board.
- Note that $\mathrm{RNSP}(\|\cdot\|, \rho, 0, S) \iff \mathrm{SNSP}(\rho, S)$ for all $\|\cdot\|$, ρ and S.
- Furthermore, $\mathrm{RNSP}(\|\cdot\|, \rho, \tau, S) \implies \mathrm{SNSP}(\rho, S)$ for all $\|\cdot\|$, ρ, τ and S.
- Hence, all statements are in particular statements on SNSP

The main result is the following theorem.

Theorem 1.2

A matrix $A \in \mathbb{C}^{m \times N}$ satisfies $\mathrm{RNSP}(\|\cdot\|, \rho, \tau, S)$ if and only if

$$\forall x, z \in \mathbb{C}^N :$$
$$\|z - x\|_1 \leq \frac{1+\rho}{1-\rho}(\|z\|_1 - \|x\|_1 + 2\|x_S\|) + \frac{2\tau}{1-\rho}\|A(x-z)\|.$$

This is a generalisation of the previously discussed theorem for the SNSP.

Before proving this theorem, note the following corollary.

Corollary 1.3

Assume that $A \in \mathbb{C}^{m\times N}$ satisfies $\mathrm{RNSP}(\|\cdot\|, \rho, \tau, s)$ with $0 < \rho < 1$ and $\tau > 0$ and let $x \in \mathbb{C}^N$. Then, if

$$\mathcal{L}_{x,\eta} := \partial B^{\|\cdot\|_1}_{\min\{\|z_1\| \mid \|Ax-Az\|\leq\eta\}}(0) \cap A^{-1}(Ax)$$

is the solution set of the problem $(P_{1,\eta})$ with $y = Ax$, then

$$\sup_{x^\# \in \mathcal{L}_x} \|x - x^\#\|_1 \leq \frac{2(1+\rho)}{1-\rho}\sigma_s(x)_1 + \frac{4\tau}{1-\rho}\eta\,,$$

i.e. the solution set $\mathcal{L}_x$ is contained in a ball of radius $\frac{2(1+\rho)}{1-\rho}\sigma_s(x)_1 + \frac{4\tau}{1-\rho}\eta$ around x in the ℓ^1-norm.

Macht es Sinn, dazu ein Bild zu malen?

- The content of this slide might be done only on the board.
- Explanation why the corollary follows from the theorem.
- Proof of the theorem (maybe shifted to the second part of the presentation, since similar to last time).

- Since $\| \cdot \|_p \leq \| \cdot \|_q$ for $p \leq q$, it is harder to bound the ℓ^q-error from above than the ℓ^p error for $p \leq q$.
- Of course, the norms are equivalent, but the constant for the other direction depends on the dimension, which we assume is large.
- But if we assume an adapted, stronger version of the RNSP, we get a useful bound for the error.

Definition 1.4: ℓ^q-robust null space property

Let $q \geq 1$. A matrix $A \in \mathbb{C}^{m\times N}$ satisfies the ℓ^q**-robust null space property** of **order** $s \in \mathbb{N}$ (with respect to $\|\cdot\|$) with the constants ρ
$in(0,1)$ and $\tau \geq 0$ iff

$$TODO$$

Warum tau echt groesser 0?

So far, our problem was to reconstruct x from the information what Ax is and that x is sparse (or knowing the support of x). We saw that the convex relaxation of the corresponding minimization problem reconstructs the minimizer iff the null space property holds for A. What, if we have additional a priori information on x? Maybe then there is a way to solve it in an acceptable computational complexity? But this is not discussed here, we still consider the convex relaxation, but now assume conditions on A and x. What does finer mean?

- As we will see, there is a difference between the real and complex case.
- Some other remarks.

Theorem 2.1

Let $A \in \mathbb{C}^{m\times N}$ and $x \in \mathbb{C}^N$ with support $S \subseteq [N]$ be given. Then, the following conditions are equivalent:

(i) $\forall v \in \ker(A) \setminus \{0\} : \quad |\langle v, \operatorname{sgn}(x)\rangle| < \|v_{\overline{S}}\|_1$.

(ii) A_S is injective and

$$\exists h \in \mathbb{C}^m : \left\{ \begin{array}{cl} (A^*h)_j = \operatorname{sgn}(x_j), & j \in S \quad \text{i.e. } A_S^*h = \operatorname{sgn}(x_S) \\ (A^*h)_j < 1, & j \in \overline{S} \end{array} \right. .$$

If one (and hence both) of these conditions hold, then x is the unique minimizer of the problem

$$\min\left\{\|z\|_1 \,\middle|\, Az = Ax\right\} .$$

Remarks

- A satisfies the NSP relative $S \iff \forall x \in \mathbb{C}^N$ with $\operatorname{supp}(x) \subseteq S$: (a) holds.
- If (a) holds for $x \in \mathbb{C}^N$ with $\operatorname{supp}(x) \subseteq S$, (a) also holds for all

$$\{x' \in \mathbb{C}^N \,|\, \operatorname{supp}(x') \subseteq S \text{ and } \operatorname{sgn}(x)_{\operatorname{supp}(x')} = \operatorname{sgn}(x')_{\operatorname{supp}(x')}\}\,.$$

- There are stable and robust versions of this theorem, which yield slightly weaker error bounds than of the previous theorems.

- If A_S is injective, then $A_S^* A_S$ is invertible.
- Furthermore, the Moore-Penrose pseudo-inverse of A_S is then a left-inverse of A_S and given by

$$A_S^\dagger = (A_S^* A_S)^{-1} A_S^* \,.$$

- Therefore the condition $A_S^* h = \operatorname{sgn}(x_S)$ is satisfied by $h := (A_S^\dagger)^* \operatorname{sgn}(x_S)$.
- Hence, if this choice of h also satisfies the rest of (b), then the Theorem holds:

Corollary 2.2

Let $A = (a_1, \ldots, a_N) \in \mathbb{C}^{m \times N}$ and $x \in \mathbb{C}^N$ with support $S \subseteq [N]$ be given. If A_S is injective and if

$$\forall l \in \overline{S}: \ \left|\left\langle A_S^{\dagger} a_l, \operatorname{sgn}(x_S)\right\rangle\right| = \left|\left\langle a_l, (A_S^{\dagger})^* \operatorname{sgn}(x_S)\right\rangle\right| \\ = \Big(A^*(A_S^{\dagger})^* \operatorname{sgn}(x_S)\Big)_l < 1\,,$$

then x is the unique minimizer of the problem

$$\min\left\{\|z\|_1 \,\middle|\, Az = Ax\right\}\,.$$

The converse direction

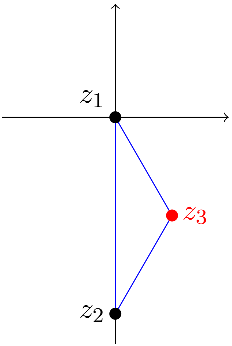
The converse of the Theorem 4.26 is not true in general, e.g. take

$$A = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \end{pmatrix}, \quad x = \begin{pmatrix} \mathrm{e}^{-\pi \mathrm{i}/3} \\ \mathrm{e}^{\pi \mathrm{i}/3} \\ 0 \end{pmatrix}.$$

The solution space of $Az = Ax$ is then

$$\left\{ \begin{pmatrix} 0 \\ \sqrt{3}\mathrm{i} \\ -\mathrm{e}^{-\mathrm{i}\pi/3} \end{pmatrix} + z \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \,\middle|\, z \in \mathbb{C} \right\},$$

and hence, the minimization problem consists in finding the point $z \in \mathbb{C}$ such that the sum of the distances to the points $z_1 = 0$, $z_2 = -\sqrt{3}\mathrm{i}$, $z_3 = \mathrm{e}^{-\mathrm{i}\pi/3}$ is minimal.



In the picture one can see that the sum of the distances is minimal exactly for $z = z_3$ (cf. first Fermat point, the angle at z_3 is $120°$). Hence, $x = (0, \sqrt{3}\mathrm{i}, \mathrm{e}^{-\mathrm{i}\pi/3})^T + z_3(1,1,1)^T$ is the unique minimizer of the problem.

However, the condition (a) does not hold: For $v = (z, z, z) \in \ker(A) \setminus \{0\}$, we have that

$$|\left\langle \mathrm{sgn}(x), v_{\{1,2\}} \right\rangle| = |\mathrm{e}^{\pi \mathrm{i}/3} z + \mathrm{e}^{-\pi \mathrm{i}/3} z| = |z| = \|v_{\{3\}}\|_1 .$$

In the real case, however, the converse does actually hold:

Theorem 2.3

Let $A \in \mathbb{R}^{m\times N}$ and $x \in \mathbb{R}^N$ with support $S \subseteq [N]$ be given. Then, the following conditions are equivalent:

(i) $\forall v \in \ker(A) \setminus \{0\} : \quad |\langle v, \operatorname{sgn}(x)\rangle| < \|v_{\overline{S}}\|_1$.

(ii) A_S is injective and

$$\exists h \in \mathbb{R}^m : \begin{cases} (A^*h)_j = \operatorname{sgn}(x_j), & j \in S \\ (A^*h)_j < 1, & j \in \overline{S} \end{cases} .$$

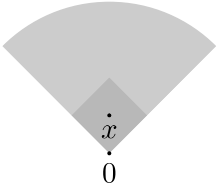
(iii) x is the unique minimizer of

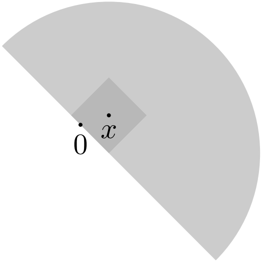
$$\min \left\{ \|z\|_1 \,\middle|\, z \in \mathbb{R}^N,\ Az = Ax \right\} .$$

Definition 2.4

Let $x \in \mathbb{R}^N$, then the **tangent cone** to the ℓ^1-ball at x is defined as

$$T(x) = \operatorname{cone}\left\{ z - x \,|\, z \in \mathbb{R}^N,\ \|z\|_1 \leq \|x\|_1 = \operatorname{cone}(\overline{B^{\|\cdot\|_1}_{\|x\|_1}(x)}) \right\}.$$





Theorem 2.5

Let $A \in \mathbb{R}^{m \times N}$ and $x \in \mathbb{R}^N$. Then

$$\left.\begin{array}{c} x \text{ is the unique minimizer} \\ \text{of } \left\{ \|z\|_1 \,\middle|\, Az = Ax \right\} \end{array}\right\} \iff \ker(A) \cap T(x) = \{0\} \, .$$