An algebraic perspective on integer sparse recovery

Lenny Fukshansky
Claremont McKenna College
(joint work with Deanna Needell and Benny Sudakov)

Combinatorics Seminar
USC
October 31, 2018

Compressed sensing

From Wikipedia:

"Compressed sensing (also known as compressive sensing, compressive sampling, or sparse sampling) is a signal processing technique for efficiently acquiring and reconstructing a signal, by finding solutions to underdetermined linear systems."

Compressed sensing

From Wikipedia:

"Compressed sensing (also known as compressive sensing, compressive sampling, or sparse sampling) is a signal processing technique for efficiently acquiring and reconstructing a signal, by finding solutions to underdetermined linear systems."

The main goal of compressed sensing is sparse recovery – the robust reconstructon of a sparse signal from a small number of linear measurements.

Compressed sensing

From Wikipedia:

"Compressed sensing (also known as compressive sensing, compressive sampling, or sparse sampling) is a signal processing technique for efficiently acquiring and reconstructing a signal, by finding solutions to underdetermined linear systems."

The main goal of compressed sensing is sparse recovery – the robust reconstructon of a sparse signal from a small number of linear measurements.

Ideally, given a signal $\boldsymbol{x} \in \mathbb{R}^d$, the goal is to accurately reconstruct $\boldsymbol{x}$ from its noisy measurements

$$\boldsymbol{b} = A\boldsymbol{x} + \boldsymbol{e} \in \mathbb{R}^m.$$

Here, A is an underdetermined matrix $A \in \mathbb{R}^{m \times d}$ $(m \ll d)$, and $\boldsymbol{e} \in \mathbb{R}^m$ is a vector modeling noise in the system.

Sparsity

Since the system is highly underdetermined, the problem is ill-posed until one imposes additional constraints. We say that a vector $\boldsymbol{x} \in \mathbb{R}^d$ is s-**sparse** if it has at most s nonzero entries:

$$\|\boldsymbol{x}\|_0 := |\{i : x_i \neq 0\}| \leq s \ll d.$$

Sparsity

Since the system is highly underdetermined, the problem is ill-posed until one imposes additional constraints. We say that a vector $\boldsymbol{x} \in \mathbb{R}^d$ is s-**sparse** if it has at most s nonzero entries:

$$\|\boldsymbol{x}\|_0 := |\{i : x_i \neq 0\}| \leq s \ll d.$$

Any matrix A that is one-to-one on $2s$-sparse signals will allow reconstruction in the noiseless case ($\boldsymbol{e} = 0$). However, compressed sensing seeks the ability to reconstruct even in the presence of noise.

Sparsity

Since the system is highly underdetermined, the problem is ill-posed until one imposes additional constraints. We say that a vector $\boldsymbol{x} \in \mathbb{R}^d$ is s-**sparse** if it has at most s nonzero entries:

$$\|\boldsymbol{x}\|_0 := |\{i : x_i \neq 0\}| \leq s \ll d.$$

Any matrix A that is one-to-one on $2s$-sparse signals will allow reconstruction in the noiseless case ($\boldsymbol{e} = 0$). However, compressed sensing seeks the ability to reconstruct even in the presence of noise.

This problem has been extensively studied on real signals. We focus on the case when the signal $\boldsymbol{x}$ comes from the integer lattice $\mathbb{Z}^d$.

Basic setup

Let

$$\mathbb{Z}_s^d := \left\{ \boldsymbol{x} \in \mathbb{Z}^d : \|\boldsymbol{x}\|_0 \leq s \right\}.$$

To reconstruct $\boldsymbol{x} \in \mathbb{Z}_s^d$ from its image $A\boldsymbol{x}$, we need to be sure that there does not exist $\boldsymbol{x} \neq \boldsymbol{y} \in \mathbb{Z}_s^d$ such that $A\boldsymbol{x} = A\boldsymbol{y}$, i.e.

$$A(\boldsymbol{x} - \boldsymbol{y}) \neq \boldsymbol{0} \; \forall \; \boldsymbol{y} \in \mathbb{Z}_s^d.$$

Basic setup

Let

$$\mathbb{Z}_s^d := \left\{ \boldsymbol{x} \in \mathbb{Z}^d : \|\boldsymbol{x}\|_0 \leq s \right\}.$$

To reconstruct $\boldsymbol{x} \in \mathbb{Z}_s^d$ from its image $A\boldsymbol{x}$, we need to be sure that there does not exist $\boldsymbol{x} \neq \boldsymbol{y} \in \mathbb{Z}_s^d$ such that $A\boldsymbol{x} = A\boldsymbol{y}$, i.e.

$$A(\boldsymbol{x} - \boldsymbol{y}) \neq \boldsymbol{0} \ \forall \ \boldsymbol{y} \in \mathbb{Z}_s^d.$$

This is equivalent to requiring that

$$A\boldsymbol{z} \neq \boldsymbol{0} \ \forall \ \boldsymbol{z} \in \mathbb{Z}_{2s}^d,$$

which is to say that

no $2s$ columns of the $m \times d$ matrix A are $\mathbb{Q}$-linearly dependent,

where $2s \leq m < d$.

Basic setup

We start with a trivial observation.

Lemma 1

Let $A = (\alpha_1 \ \ldots \ \alpha_d)$ *be a* $1 \times d$ *matrix with real* $\mathbb{Q}$*-linearly independent entries. Then the equation* $A\boldsymbol{x} = 0$ *has no solutions in* $\mathbb{Z}^d$ *except for* $\boldsymbol{x} = \mathbf{0}$.

Basic setup

We start with a trivial observation.

Lemma 1

Let $A = (\alpha_1 \ \ldots \ \alpha_d)$ be a $1 \times d$ matrix with real $\mathbb{Q}$-linearly independent entries. Then the equation $A\boldsymbol{x} = 0$ has no solutions in $\mathbb{Z}^d$ except for $\boldsymbol{x} = \mathbf{0}$.

In practice, we want to be able to tolerate noise in the system and decode robustly. The noise $\boldsymbol{e}$ typically scales with the entries of A, so we ask for the following two properties:

i. the entries of A are uniformly bounded in absolute value

ii. $\|A\boldsymbol{z}\|$ is bounded away from zero for any nonzero vector $\boldsymbol{z}$ in our signal space.

Problem formulation

Hence we have the following optimization problem.

Problem 1

Construct an $m \times d$ matrix A with $m < d$ such that

$$|A| := \max\{|a_{ij}| : 1 \leq i \leq m, 1 \leq j \leq d\} \leq C_1,$$

and for every nonzero $\boldsymbol{x} \in \mathbb{Z}_s^d$, $s \leq m$,

$$\|A\boldsymbol{x}\| \geq C_2,$$

where $C_1, C_2 > 0$.

Existence of such matrices

Theorem 2 (F., Needell, Sudakov – 2017)

There exist $m \times d$ integer matrices A with $m < d$ and bounded $|A|$ such that for any nonzero $\boldsymbol{x} \in \mathbb{Z}_s^d$, $0 < s \leq m$,

$$\|A\boldsymbol{x}\| \geq 1. \tag{1}$$

In fact, for sufficiently large m, there exist such matrices with

$$|A| = 1 \text{ and } d = 1.2938\, m,$$

and there also exist such matrices with

$$|A| = k \text{ and } d = \Omega(\sqrt{k}\, m).$$

Existence of such matrices

Theorem 2 (F., Needell, Sudakov – 2017)

There exist $m \times d$ integer matrices A with $m < d$ and bounded $|A|$ such that for any nonzero $\boldsymbol{x} \in \mathbb{Z}_s^d$, $0 < s \leq m$,

$$\|A\boldsymbol{x}\| \geq 1. \tag{1}$$

In fact, for sufficiently large m, there exist such matrices with

$$|A| = 1 \text{ and } d = 1.2938\, m,$$

and there also exist such matrices with

$$|A| = k \text{ and } d = \Omega(\sqrt{k}\, m).$$

The bound (1) can be replaced by $\|A\boldsymbol{x}\| \geq \ell > 1$ multiplying A by ℓ at the expense of making $|A|$ larger by the constant factor of ℓ.

Proof of Theorem 2

First we need to prove the existence of an $m \times d$ matrix A with

$$|A| = 1 \text{ and } d = 1.2938\, m,$$

so that every $m \times m$ submatrix is nonsingular.

Proof of Theorem 2

First we need to prove the existence of an $m \times d$ matrix A with

$$|A| = 1 \text{ and } d = 1.2938\, m,$$

so that every $m \times m$ submatrix is nonsingular.

We use a powerful result of Bourgain, Vu and Wood (2010):

Let M_m be an $m \times m$ random matrix whose entries are 0 with probability $1/2$ and ± 1 with probability $1/4$ each. Then the probability that matrix M_m is singular is at most $(1/2 - o(1))^m$.

Proof of Theorem 2

First we need to prove the existence of an $m \times d$ matrix A with

$$|A| = 1 \text{ and } d = 1.2938\, m,$$

so that every $m \times m$ submatrix is nonsingular.

We use a powerful result of Bourgain, Vu and Wood (2010):

Let M_m be an $m \times m$ random matrix whose entries are 0 with probability $1/2$ and ± 1 with probability $1/4$ each. Then the probability that matrix M_m is singular is at most $(1/2 - o(1))^m$.

Form an $m \times d$ random matrix A by taking its entries to be 0 with probability $1/2$ and ± 1 with probability $1/4$ each. Then $|A| = 1$ and any m columns of A form a matrix distributed according to M_m.

Proof of Theorem 2

Therefore the probability that any $m \times m$ submatrix of A is singular is at most

$$(1/2 - o(1))^m.$$

Since the number of such submatrices is $\binom{d}{m}$ we have (by union bound) that the probability that A contains an $m \times m$ singular submatrix is at most

$$\binom{d}{m}(1/2 - o(1))^m.$$

Proof of Theorem 2

Therefore the probability that any $m \times m$ submatrix of A is singular is at most

$$(1/2 - o(1))^m.$$

Since the number of such submatrices is $\binom{d}{m}$ we have (by union bound) that the probability that A contains an $m \times m$ singular submatrix is at most

$$\binom{d}{m}(1/2 - o(1))^m.$$

Bounding $\binom{d}{m}$ we see that this probability is < 1 for $d \leq 1.2938\, m$, and hence an $m \times d$ matrix A with $|A| = 1$ and all $m \times m$ submatrices nonsingular exists.

Proof of Theorem 2

Next, we want to prove existence of an $m \times d$ matrix A with

$$|A| = k \text{ and } d = \Omega(\sqrt{k}\, m),$$

so that every $m \times m$ submatrix is nonsingular.

Proof of Theorem 2

Next, we want to prove existence of an $m \times d$ matrix A with

$$|A| = k \text{ and } d = \Omega(\sqrt{k}\, m),$$

so that every $m \times m$ submatrix is nonsingular.

We use another result of Bourgain, Vu and Wood (2010) from the same paper:

If N_m is an $m \times m$ random matrix whose entries come from the set $\{-k, \cdots, k\}$ with equal probability, then the probability that N_m is singular is at most $(1/\sqrt{2k} - o(1))^m$.

Proof of Theorem 2

Next, we want to prove existence of an $m \times d$ matrix A with

$$|A| = k \text{ and } d = \Omega(\sqrt{k}\, m),$$

so that every $m \times m$ submatrix is nonsingular.

We use another result of Bourgain, Vu and Wood (2010) from the same paper:

If N_m is an $m \times m$ random matrix whose entries come from the set $\{-k, \cdots, k\}$ with equal probability, then the probability that N_m is singular is at most $(1/\sqrt{2k} - o(1))^m$.

Consider an $m \times d$ random matrix A whose entries come from the set $\{-k, \cdots, k\}$ with equal probability. Then $|A| \leq k$ and the probability that any $m \times m$ submatrix of A is singular is at most

$$(1/\sqrt{2k} - o(1))^m.$$

Proof of Theorem 2

Since the number of such submatrices is $\binom{d}{m}$ we have that the probability that A contains an $m \times m$ singular submatrix is at most

$$\binom{d}{m}\left(\frac{1}{\sqrt{2k}} - o(1)\right)^m.$$

Proof of Theorem 2

Since the number of such submatrices is $\binom{d}{m}$ we have that the probability that A contains an $m \times m$ singular submatrix is at most

$$\binom{d}{m}\left(\frac{1}{\sqrt{2k}} - o(1)\right)^m.$$

Again, estimating the binomial coefficient and choosing d to be a sufficiently small multiple of $\sqrt{k}m$, this probability can be made smaller than 1.

Proof of Theorem 2

Since the number of such submatrices is $\binom{d}{m}$ we have that the probability that A contains an $m \times m$ singular submatrix is at most

$$\binom{d}{m}\left(\frac{1}{\sqrt{2k}} - o(1)\right)^m.$$

Again, estimating the binomial coefficient and choosing d to be a sufficiently small multiple of $\sqrt{k}m$, this probability can be made smaller than 1.

Thus with positive probability A does not have singular $m \times m$ submatrices, implying again that for any $\boldsymbol{x} \in \mathbb{Z}_s^d$, $0 < s \leq m$, $\|A\boldsymbol{x}\| \geq 1$.

How much bigger than m can d be?

Theorem 3 (F., Needell, Sudakov – 2017)

For any integers $m \geq 3$, $k \geq 1$ and $m \times d$ integer matrix A with $|A| = k$ satisfying (1) for all $s \leq m$, we must have

$$d \leq (2k^2 + 2)(m - 1) + 1.$$

How much bigger than m can d be?

Theorem 3 (F., Needell, Sudakov – 2017)

For any integers $m \geq 3$, $k \geq 1$ and $m \times d$ integer matrix A with $|A| = k$ satisfying (1) for all $s \leq m$, we must have

$$d \leq (2k^2 + 2)(m - 1) + 1.$$

Theorem 4 (Konyagin – 2018)

If $m \geq 2(\log k + 1)$, then

$$d < 144k(\log k + 1)m.$$

How much bigger than m can d be?

Theorem 3 (F., Needell, Sudakov – 2017)

For any integers $m \geq 3$, $k \geq 1$ and $m \times d$ integer matrix A with $|A| = k$ satisfying (1) for all $s \leq m$, we must have

$$d \leq (2k^2 + 2)(m - 1) + 1.$$

Theorem 4 (Konyagin – 2018)

If $m \geq 2(\log k + 1)$, then

$$d < 144k(\log k + 1)m.$$

Theorem 5 (Sudakov – 2018)

For sufficiently large m,

$$d = O(k\sqrt{\log k}\ m).$$

Proof of Theorem 3

Let

$$C_m(k) = \{\boldsymbol{x} \in \mathbb{Z}^m : |\boldsymbol{x}| \leq k\},$$

then $|C_m(k)| = (2k+1)^m$. Let $\ell = (2k^2+2)(m-1)+1$ and let $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_\ell$ be any ℓ vectors from $C_m(k)$.

Proof of Theorem 3

Let

$$C_m(k) = \{\boldsymbol{x} \in \mathbb{Z}^m : |\boldsymbol{x}| \leq k\},$$

then $|C_m(k)| = (2k+1)^m$. Let $\ell = (2k^2+2)(m-1)+1$ and let $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_\ell$ be any ℓ vectors from $C_m(k)$.

If there are m vectors with the first or second coordinate equal to 0, then they all lie in the same $(m-1)$-dimensional subspace.

Proof of Theorem 3

Let

$$C_m(k) = \{\boldsymbol{x} \in \mathbb{Z}^m : |\boldsymbol{x}| \leq k\},$$

then $|C_m(k)| = (2k+1)^m$. Let $\ell = (2k^2+2)(m-1)+1$ and let $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_\ell$ be any ℓ vectors from $C_m(k)$.

If there are m vectors with the first or second coordinate equal to 0, then they all lie in the same $(m-1)$-dimensional subspace.

If not, there are $2k^2 m$ vectors with the first two coordinates nonzero. Multiplying some of these vectors by -1, if necessary (does not change linear independence properties) we can assume that all of them have positive first coordinate.

Proof of Theorem 3

Let

$$C_m(k) = \{\boldsymbol{x} \in \mathbb{Z}^m : |\boldsymbol{x}| \leq k\},$$

then $|C_m(k)| = (2k+1)^m$. Let $\ell = (2k^2+2)(m-1)+1$ and let $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_\ell$ be any ℓ vectors from $C_m(k)$.

If there are m vectors with the first or second coordinate equal to 0, then they all lie in the same $(m-1)$-dimensional subspace.

If not, there are $2k^2 m$ vectors with the first two coordinates nonzero. Multiplying some of these vectors by -1, if necessary (does not change linear independence properties) we can assume that all of them have positive first coordinate.

Hence there are a total of $k \times 2k = 2k^2$ choices for the first two coordinates, so there must exist a subset of m of these vectors that have these first two coordinates the same, let these be $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m$.

Proof of Theorem 3

Then there exists a vector $\boldsymbol{y} = (a, b, 0, \ldots, 0)^\top \in C_m(k)$ such that the vectors

$$\boldsymbol{z}_1 = \boldsymbol{x}_1 - \boldsymbol{y}, \ \ldots, \ \boldsymbol{z}_m = \boldsymbol{x}_m - \boldsymbol{y}$$

all have the first two coordinates equal to 0. This means that these vectors lie in an $(m-2)$-dimensional subspace

$$V = \{\boldsymbol{z} \in \mathbb{R}^m : z_1 = z_2 = 0\}$$

of $\mathbb{R}^m$. Then let $V' = \operatorname{span}_{\mathbb{R}}\{V, \boldsymbol{y}\}$, so $\dim_{\mathbb{R}} V' = m - 1$. On the other hand, $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m \in V'$, and hence the $m \times m$ matrix with rows $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m$ must have determinant equal to 0.

Proof of Theorem 3

Then there exists a vector $\boldsymbol{y} = (a, b, 0, \ldots, 0)^\top \in C_m(k)$ such that the vectors

$$\boldsymbol{z}_1 = \boldsymbol{x}_1 - \boldsymbol{y}, \ \ldots, \ \boldsymbol{z}_m = \boldsymbol{x}_m - \boldsymbol{y}$$

all have the first two coordinates equal to 0. This means that these vectors lie in an $(m-2)$-dimensional subspace

$$V = \{\boldsymbol{z} \in \mathbb{R}^m : z_1 = z_2 = 0\}$$

of $\mathbb{R}^m$. Then let $V' = \operatorname{span}_{\mathbb{R}}\{V, \boldsymbol{y}\}$, so $\dim_{\mathbb{R}} V' = m - 1$. On the other hand, $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m \in V'$, and hence the $m \times m$ matrix with rows $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m$ must have determinant equal to 0.

Therefore if an $m \times d$ matrix A with column vectors in $C_m(k)$ has all nonsingular $m \times m$ submatrices, then

$$d \leq (2k^2 + 2)(m - 1) + 1.$$

Question and example

Question 1

What is the optimal upper bound on d in terms of k and m? In particular, is it true that $d = O(km)$?

Question and example

Question 1

What is the optimal upper bound on d in terms of k and m? In particular, is it true that $d = O(km)$?

Here is a low-dimensional example.

Example 1

Let $m = 3$, $d = 6$, $k = 1$, and define a 3×6 matrix

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & -1 & -1 \\ 1 & 0 & 1 & -1 & 0 & -1 \end{pmatrix}.$$

This matrix has $|A| = 1$ and any three of its columns are linearly independent. Then for $s \leq 3$ and any $\boldsymbol{x} \in \mathbb{Z}_s^6$, $\|A\boldsymbol{x}\| \geq 1$.

Algebraic matrices

While our Theorem 2 gives the image vector $A\boldsymbol{x}$ bounded away from 0, many of its coordinates may still be 0. It is natural to ask if a matrix A can be constructed so that $A\boldsymbol{x}$ bounded away from 0 **and** its coordinates are nonzero?

Algebraic matrices

While our Theorem 2 gives the image vector $A\boldsymbol{x}$ bounded away from 0, many of its coordinates may still be 0. It is natural to ask if a matrix A can be constructed so that $A\boldsymbol{x}$ bounded away from 0 **and** its coordinates are nonzero?

Corollary 6 (F., Needell, Sudakov – 2017)

Let B be the $d \times m$-transpose of a matrix satisfying (1) as guaranteed by Theorem 2. Let θ be an algebraic integer of degree m, and let $\theta = \theta_1, \theta_2, \ldots, \theta_m$ be its algebraic conjugates. For each $1 \leq i \leq m$, let $\boldsymbol{\theta}_i = (1 \ \theta_i \ \ldots \theta_i^{m-1})^\top$, compute the $d \times m$ matrix

$$B \begin{pmatrix} \boldsymbol{\theta}_1 & \ldots & \boldsymbol{\theta}_m \end{pmatrix},$$

and let A be its transpose. Then $|A| = O(|B|m)$, for any $\boldsymbol{x} \in \mathbb{Z}_s^d$, $0 < s \leq m$, $\|A\boldsymbol{x}\| \geq \sqrt{m}$ and the vector $A\boldsymbol{x}$ has **all nonzero coordinates**.

Proof of Corollary 6

We use the AM-GM inequality:

$$\begin{aligned}
\frac{1}{m}\|A\boldsymbol{x}\|^2 &= \frac{1}{m}\sum_{i=1}^{m}|(B\boldsymbol{\theta}_i)\boldsymbol{x}|^2 \\
&\geq \left(\prod_{i=1}^{m}|(B\boldsymbol{\theta}_i)\boldsymbol{x}|^2\right)^{1/m} \\
&= |\mathbb{N}_K((B\boldsymbol{\theta}_1)\boldsymbol{x})|^{2/m},
\end{aligned}$$

where $\mathbb{N}_K$ is the field norm, which is $\geq$ since $(B\boldsymbol{\theta}_1)\boldsymbol{x}$ is an algebraic integer.

Proof of Corollary 6

We use the AM-GM inequality:

$$\begin{aligned}\frac{1}{m}\|A\boldsymbol{x}\|^2 &= \frac{1}{m}\sum_{i=1}^{m}|(B\boldsymbol{\theta}_i)\boldsymbol{x}|^2 \\ &\geq \left(\prod_{i=1}^{m}|(B\boldsymbol{\theta}_i)\boldsymbol{x}|^2\right)^{1/m} \\ &= |\mathbb{N}_K((B\boldsymbol{\theta}_1)\boldsymbol{x})|^{2/m},\end{aligned}$$

where $\mathbb{N}_K$ is the field norm, which is $\geq$ since $(B\boldsymbol{\theta}_1)\boldsymbol{x}$ is an algebraic integer.

Since $(B\boldsymbol{\theta}_1)\boldsymbol{x} \neq 0$, its algebraic conjugates, which are the rest of the coordinates of the vector $A\boldsymbol{x}$ must all be nonzero.

Algebraic matrix example

Let $m = 3$, $d = 6$, and take $K = \mathbb{Q}(\theta)$, where $\theta = 2^{1/3}$, then

$$\boldsymbol{\theta} = \begin{pmatrix} 1 \\ \theta \\ \theta^2 \end{pmatrix}.$$

Algebraic matrix example

Let $m = 3$, $d = 6$, and take $K = \mathbb{Q}(\theta)$, where $\theta = 2^{1/3}$, then

$$\boldsymbol{\theta} = \begin{pmatrix} 1 \\ \theta \\ \theta^2 \end{pmatrix}.$$

Let $k = 1$ and take B to be the transpose of the matrix from Example 1, i.e.

$$B = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & -1 \\ 1 & -1 & 0 \\ 1 & -1 & -1 \end{pmatrix}.$$

Algebraic matrix example

The number field K has three embeddings into $\mathbb{C}$, given by

$$\theta \mapsto \theta, \ \theta \mapsto \xi\theta, \ \theta \mapsto \xi^2\theta,$$

where $\xi = e^{\frac{2\pi i}{3}}$ is a third root of unity, i.e. θ is mapped to roots of its minimal polynomial by injective field homomorphisms that fix $\mathbb{Q}$.

Algebraic matrix example

The number field K has three embeddings into $\mathbb{C}$, given by

$$\theta \mapsto \theta,\ \theta \mapsto \xi\theta,\ \theta \mapsto \xi^2\theta,$$

where $\xi = e^{\frac{2\pi i}{3}}$ is a third root of unity, i.e. θ is mapped to roots of its minimal polynomial by injective field homomorphisms that fix $\mathbb{Q}$.

Hence we get the following 3×6 matrix:

$$A = \begin{pmatrix} 1+\theta+\theta^2 & 1+\theta & 1+\theta^2 & 1-\theta^2 & 1-\theta & 1-\theta-\theta^2 \\ 1+\xi\theta+\xi^2\theta^2 & 1+\xi\theta & 1+\xi^2\theta^2 & 1-\xi^2\theta^2 & 1-\xi\theta & 1-\xi\theta-\xi^2\theta^2 \\ 1+\xi^2\theta+\xi\theta^2 & 1+\xi^2\theta & 1+\xi\theta^2 & 1-\xi\theta^2 & 1-\xi^2\theta & 1-\xi^2\theta-\xi\theta^2 \end{pmatrix}$$

with $|A| \leq 3\sqrt[3]{2}$ and $\|A\boldsymbol{x}\| \geq \sqrt{3}$ for every $\boldsymbol{x} \in \mathbb{Z}_s^6$, $s \leq 3$.

Upper bound on $\|A\boldsymbol{x}\|$

While these results show the existence of matrices A such that $\|A\boldsymbol{x}\|$ is bounded away from $\mathbf{0}$ on sparse vectors, it is also clear that for any $m \times d$ matrix A there exist sparse vectors with $\|A\boldsymbol{x}\|$ not too large: for instance, if $\boldsymbol{x} \in \mathbb{Z}^d$ is a standard basis vector, then

$$\|A\boldsymbol{x}\| \leq \sqrt{m}\ |A|. \tag{2}$$

Upper bound on $\|A\boldsymbol{x}\|$

While these results show the existence of matrices A such that $\|A\boldsymbol{x}\|$ is bounded away from $\mathbf{0}$ on sparse vectors, it is also clear that for any $m \times d$ matrix A there exist sparse vectors with $\|A\boldsymbol{x}\|$ not too large: for instance, if $\boldsymbol{x} \in \mathbb{Z}^d$ is a standard basis vector, then

$$\|A\boldsymbol{x}\| \leq \sqrt{m}\ |A|. \tag{2}$$

We prove a determinantal upper bound on $\|A\boldsymbol{x}\|$ in the spirit of Minkowski's Geometry of Numbers, which is often better.

Theorem 7 (F., Needell, Sudakov – 2017)

Let A be an $m \times d$ real matrix of rank $m \leq d$, and let A' be the $d \times m$ real matrix so that AA' is the $m \times m$ identity matrix. There exists a nonzero point $\boldsymbol{x} \in \mathbb{Z}_m^d$ such that

$$\|A\boldsymbol{x}\| \leq \sqrt{m}\ \left|\det\left((A')^\top A'\right)\right|^{-1/2m}. \tag{3}$$

Upper bound examples

Example 2

Let $d = 5$, $m = 3$, and let

$$A = \begin{pmatrix} 15 & 15 & 4 & 13 & 15 \\ 2 & -1 & -15 & 2 & -13 \\ -13 & 2 & 1 & -15 & 4 \end{pmatrix},$$

then

$$A' = \begin{pmatrix} 3392/3905 & 23/355 & 3021/3905 \\ -1949/2130 & 3/710 & -1697/2130 \\ -6409/9372 & -19/284 & -5647/9372 \\ -6407/9372 & -17/284 & -6353/9372 \\ 13869/15620 & 1/1420 & 12047/15620 \end{pmatrix},$$

and so the bound of (3) is 8.375..., which is better than 25.980..., the bound given by (2).

Upper bound examples

Example 3

Let $d = 6$, $m = 3$, and let

$$A = \begin{pmatrix} 50000 & 20 & 40 & 3 & -50000 & 30 \\ -1 & -50000 & 20 & 40 & 4 & -50000 \\ -50000 & -1 & -50000 & -50000 & 20 & 40 \end{pmatrix},$$

then the bound of (3) is 7651.170... and (2) is 86602.540...

Upper bound examples

Example 3

Let $d = 6$, $m = 3$, and let

$$A = \begin{pmatrix} 50000 & 20 & 40 & 3 & -50000 & 30 \\ -1 & -50000 & 20 & 40 & 4 & -50000 \\ -50000 & -1 & -50000 & -50000 & 20 & 40 \end{pmatrix},$$

then the bound of (3) is 7651.170... and (2) is 86602.540...
Let $d = 8$, $m = 4$, and let

$$A = \begin{pmatrix} 6 & 13 & 13 & 11 & 6 & 12 & 11 & 10 \\ 7 & 12 & 6 & 13 & 7 & 11 & 11 & 9 \\ 8 & 11 & 12 & 9 & 12 & 12 & 12 & 11 \\ 13 & 10 & 7 & 8 & 13 & 13 & 13 & 13 \end{pmatrix},$$

then the bound of (3) is 2.412... and (2) is 26.

Sparse Geometry of Numbers

Let us now sketch the strategy of proof of Theorem 7. For this, we develop sparse analogues of some classical theorems in Minkowski's Geometry of Numbers. We start with a result of J. D. Vaaler (1979).

Sparse Geometry of Numbers

Let us now sketch the strategy of proof of Theorem 7. For this, we develop sparse analogues of some classical theorems in Minkowski's Geometry of Numbers. We start with a result of J. D. Vaaler (1979).

Lemma 8 (Vaaler's Cube Slicing Inequality)

Let $\mathcal{C}^d(1)$ be a cube of sidelength 1 centered at the origin in $\mathbb{R}^d$, i.e.

$$\mathcal{C}^d(1) = \{\boldsymbol{x} \in \mathbb{R}^d : |x_i| \leq 1/2 \ \forall \ 1 \leq i \leq d\}.$$

Let V be an m-dimensional subspace of $\mathbb{R}^d$, $m \leq d$. Then the m-dimensional volume of the section $\mathcal{C}_d(1) \cap V$ is

$$\mathrm{Vol}_m(\mathcal{C}_d(1) \cap V) \geq 1.$$

Sparse Geometry of Numbers

We can use Vaaler's lemma to prove a sparse version of Minkowski's Convex Body Theorem for parallelepipeds.

Proposition 9 (F., Needell, Sudakov – 2017)

Let $m \leq d$ be positive integers. Let $A \in \mathrm{GL}_d(\mathbb{R})$, and let $P_A = A\mathcal{C}^d(1)$. Assume that for some $I \subset [d]$ with $|I| = m$,

$$\sqrt{|\det(A_I^\top A_I)|} \geq 2^m. \tag{4}$$

Then P_A contains a nonzero point of $\mathbb{Z}_m^d$.

Sparse Geometry of Numbers

This implies a sparse version of Minkowski's Linear Forms Theorem.

Theorem 10 (F., Needell, Sudakov – 2017)

Let $m \leq d$ be positive integers and $B \in \mathrm{GL}_d(\mathbb{R})$. For each $1 \leq i \leq d$, let $L_i(X_1, \ldots, X_d) = \sum_{j=1}^{d} b_{ij} X_j$ be the linear form with entries of the i-th row of B for its coefficients. Let $c_1, \ldots, c_d$ be positive real numbers such that for some $I = \{1 \leq j_1 < \cdots < j_m \leq d\} \subset [d]$,

$$c_{j_1} \cdots c_{j_m} \geq \left| \det \left((B^{-1})_I^{\top} (B^{-1})_I \right) \right|^{-1/2}. \tag{5}$$

Then there exists a nonzero point $\boldsymbol{x} \in \mathbb{Z}_m^d$ such that

$$|L_{j_i}(\boldsymbol{x})| \leq c_{j_i} \ \forall \ 1 \leq i \leq m. \tag{6}$$

Sparse Geometry of Numbers

Taking $I = \{1, \ldots, m\}$ and

$$c_i = \left|\det\left((B^{-1})_I^\top (B^{-1})_I\right)\right|^{-1/2m}$$

for each $1 \leq i \leq m$ in Theorem 10 implies –

Corollary 11 (F., Needell, Sudakov – 2017)

Let A be an $m \times d$ real matrix of rank $m \leq d$. Let $B \in \mathrm{GL}_d(\mathbb{R})$ be a matrix whose first m rows are the rows of A. Let $I = \{1, \ldots, m\}$. Then there exists a nonzero point $\boldsymbol{x} \in \mathbb{Z}_m^d$ such that

$$\|A\boldsymbol{x}\| \leq \sqrt{m}\left|\det\left((B^{-1})_I^\top (B^{-1})_I\right)\right|^{-1/2m}.$$

Sparse Geometry of Numbers

Taking $I = \{1, \ldots, m\}$ and

$$c_i = \left|\det\left((B^{-1})_I^\top (B^{-1})_I\right)\right|^{-1/2m}$$

for each $1 \leq i \leq m$ in Theorem 10 implies –

Corollary 11 (F., Needell, Sudakov – 2017)

Let A be an $m \times d$ real matrix of rank $m \leq d$. Let $B \in \mathrm{GL}_d(\mathbb{R})$ be a matrix whose first m rows are the rows of A. Let $I = \{1, \ldots, m\}$. Then there exists a nonzero point $\boldsymbol{x} \in \mathbb{Z}_m^d$ such that

$$\|A\boldsymbol{x}\| \leq \sqrt{m}\left|\det\left((B^{-1})_I^\top (B^{-1})_I\right)\right|^{-1/2m}.$$

Theorem 7 follows immediately.

Closest Vector Problem

Let us also say a few words about the reconstruction algorithm for our matrix construction in the situation when $s = m$.

Closest Vector Problem

Let us also say a few words about the reconstruction algorithm for our matrix construction in the situation when $s = m$.

First recall the Closest Vector Problem (CVP) in $\mathbb{R}^m$:

Input: A matrix $C \in \mathrm{GL}_n(\mathbb{R})$ and a point $\boldsymbol{y} \in \mathbb{R}^m$.
Output: A point $\boldsymbol{x}$ in the lattice $\Lambda := C\mathbb{Z}^n$ such that

$$\|\boldsymbol{x} - \boldsymbol{y}\| = \min\{\|\boldsymbol{z} - \boldsymbol{y}\| : \boldsymbol{z} \in \Lambda\}.$$

Closest Vector Problem

Let us also say a few words about the reconstruction algorithm for our matrix construction in the situation when $s = m$.

First recall the Closest Vector Problem (CVP) in $\mathbb{R}^m$:

Input: A matrix $C \in \mathrm{GL}_n(\mathbb{R})$ and a point $\boldsymbol{y} \in \mathbb{R}^m$.
Output: A point $\boldsymbol{x}$ in the lattice $\Lambda := C\mathbb{Z}^n$ such that

$$\|\boldsymbol{x} - \boldsymbol{y}\| = \min\{\|\boldsymbol{z} - \boldsymbol{y}\| : \boldsymbol{z} \in \Lambda\}.$$

It is known that CVP in $\mathbb{R}^m$ can be solved by a deterministic $O(2^{2m})$ time and $O(2^m)$ space algorithm, or by a randomized $2^{m+o(m)}$-time and space algorithm.

Reconstruction algorithm

Let A be an $m \times d$ matrix with no m column vectors linearly dependent, so that for all $\boldsymbol{x} \in \mathbb{Z}_m^d$,

$$\|A\boldsymbol{x}\| \geq \alpha$$

for some real $\alpha > 0$. Let $[d] = \{1, \ldots, d\}$ and define

$$\mathcal{J} = \{I \subset [d] : |I| = m\},$$

then $|\mathcal{J}| = \binom{d}{m}$. For each $I \in \mathcal{J}$, let A_I be the $m \times m$ submatrix of A indexed by the elements of I and let $\Lambda_I = A_I \mathbb{Z}^m$ be the corresponding lattice of rank m in $\mathbb{R}^m$. Suppose now that $\boldsymbol{x} \in \mathbb{Z}_m^d$, then $A\boldsymbol{x}$ is a vector in some Λ_I.

Reconstruction algorithm

Let A be an $m \times d$ matrix with no m column vectors linearly dependent, so that for all $\boldsymbol{x} \in \mathbb{Z}_m^d$,

$$\|A\boldsymbol{x}\| \geq \alpha$$

for some real $\alpha > 0$. Let $[d] = \{1, \ldots, d\}$ and define

$$\mathcal{J} = \{I \subset [d] : |I| = m\},$$

then $|\mathcal{J}| = \binom{d}{m}$. For each $I \in \mathcal{J}$, let A_I be the $m \times m$ submatrix of A indexed by the elements of I and let $\Lambda_I = A_I \mathbb{Z}^m$ be the corresponding lattice of rank m in $\mathbb{R}^m$. Suppose now that $\boldsymbol{x} \in \mathbb{Z}_m^d$, then $A\boldsymbol{x}$ is a vector in some Λ_I. Let us write

$$\mathcal{J} = \{I_1, \ldots, I_t\},$$

where $t = \binom{d}{m}$. Given a CVP oracle, we can propose the following reconstruction algorithm for our problem.

Reconstruction algorithm

1. *Input:* A vector $\boldsymbol{y} = A\boldsymbol{x} + \boldsymbol{e}$ (here $\boldsymbol{x} \in \mathbb{Z}_m^d$ and error $\boldsymbol{e} \in \mathbb{R}^m$ with $\|\boldsymbol{e}\| < \alpha/2$).
2. *CVP:* Make t calls to the CVP oracle in $\mathbb{R}^m$ with the input Λ_{l_j} and $\boldsymbol{y}$ for each $1 \leq j \leq t$; let

$$\boldsymbol{z}_1 \in \Lambda_{l_1}, \ldots, \boldsymbol{z}_t \in \Lambda_{l_t}$$

   be the vectors returned.
3. *Comparison:* Out of $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_t$, pick $\boldsymbol{z}_i$ such that

$$\|\boldsymbol{z}_i - \boldsymbol{y}\| < \alpha/2.$$

   By our construction, there can be only one such vector.
4. *Matrix inverse:* Compute $(A_{l_i})^{-1}$.
5. *Reconstruction:* Take $\boldsymbol{x} = (A_{l_i})^{-1}\boldsymbol{z}_i$.

Reconstruction algorithm

On the other hand, suppose we had an oracle for some reconstruction algorithm with the error bound α. Given a point $\boldsymbol{y} \in \mathbb{R}^m$, make a call to this oracle, returning a vector $\boldsymbol{x} \in \mathbb{Z}_m^d$. Compute $\boldsymbol{z} = A\boldsymbol{x}$, then $\boldsymbol{z}$ is in one of the lattices $\Lambda_{l_1}, \ldots, \Lambda_{l_t}$, and, assuming that $\|\boldsymbol{z} - \boldsymbol{y}\| < \alpha/2$, we have

$$\|\boldsymbol{z} - \boldsymbol{y}\| = \min \left\{ \|\boldsymbol{u} - \boldsymbol{y}\| : \boldsymbol{u} \in \bigcup_{j=1}^{t} \Lambda_{l_j} \right\}.$$

Hence $\boldsymbol{z}$ is a CVP solution for $\boldsymbol{y}$ in $\bigcup_{j=1}^{t} \Lambda_{l_j}$. In other words, the problem of reconstructing the sparse signal from the image under such a matrix A in $\mathbb{R}^m$ has essentially the same computational complexity as CVP in $\mathbb{R}^m$.

A final question

Classical compressed sensing methods offer far more efficient complexity, but also require the sparsity level s to be much less than m. Our theoretical framework allows any $s \leq m$, which is a much taller order.

A final question

Classical compressed sensing methods offer far more efficient complexity, but also require the sparsity level s to be much less than m. Our theoretical framework allows any $s \leq m$, which is a much taller order.

This consideration suggests a natural question for future investigation.

Question 2

Can our construction be improved to yield a faster reconstruction algorithm under the assumption that $s \ll m$?

Reference

L. Fukshansky, D. Needell, B. Sudakov, *An algebraic perspective on integer sparse recovery*, Applied Mathematics and Computation, vol. 340 (2019), pg. 31–42

Preprint is available at:

`http://math.cmc.edu/lenny/research.html`

Reference

L. Fukshansky, D. Needell, B. Sudakov, *An algebraic perspective on integer sparse recovery*, Applied Mathematics and Computation, vol. 340 (2019), pg. 31–42

Preprint is available at:

`http://math.cmc.edu/lenny/research.html`

Thank you!