Mathematical Foundation for Compressed Sensing

Jan-Olov Strömberg

Royal Institute of Technology, Stockholm, Sweden

Lecture 12, March 6, 2012

An outline for today

- A short summary from last lectures:

An outline for today

- A short summary from last lectures:
  - A lower estimate for RIP.

An outline for today

- A short summary from last lectures:
  - A lower estimate for RIP.
  - RIP of Random matrices .

An outline for today

- A short summary from last lectures:
    - A lower estimate for RIP.
    - RIP of Random matrices .
    - A construction of a RIP matrix .

An outline for today

- A short summary from last lectures:
  - A lower estimate for RIP.
  - RIP of Random matrices .
  - A construction of a RIP matrix .
- Incoherent bases and Structured Random matrices.

An outline for today

- A short summary from last lectures:
  - A lower estimate for RIP.
  - RIP of Random matrices .
  - A construction of a RIP matrix .
- Incoherent bases and Structured Random matrices.
- RIP estimates for Structured Random matrices (no proof today).

An outline for today

- A short summary from last lectures:
  - A lower estimate for RIP.
  - RIP of Random matrices .
  - A construction of a RIP matrix .
- Incoherent bases and Structured Random matrices.
- RIP estimates for Structured Random matrices (no proof today).
- Non-uniform versus uniform recovery of sparse vectors.

Remains to do

Things that remains to be done:

- The proof of RIP estimate for Structured Random matrices.

Remains to do

Things that remains to be done:

- The proof of RIP estimate for Structured Random matrices.
- More about non-uniform recovery.

Remains to do

Things that remains to be done:

- The proof of RIP estimate for Structured Random matrices.
- More about non-uniform recovery.
- The recoverey of almost s-sparse vector with noise.

A short summary from last lecture:

A short summary from last lecture: Recall the definition property of a matrix:

A short summary from last lecture:

A short summary from last lecture: Recall the definition property of a matrix:

The $m \times N$ matrix $\mathbf{A}$ satisfies Restricted Isometry Property (RIP) with constant δ_s if

A short summary from last lecture:

A short summary from last lecture: Recall the definition property of a matrix:

The $m \times N$ matrix $\mathbf{A}$ satisfies Restricted Isometry Property (RIP) with constant δ_s if

$$\left| \|\mathbf{A}\mathbf{x}\|_2^2 - \|\mathbf{x}\|_2^2 \right| \leq \delta_s \|\mathbf{x}\|_2^2,$$

for all s- sparse vectors $\mathbf{x}$.

A short summary from last lecture:

A short summary from last lecture: Recall the definition property of a matrix:

The $m \times N$ matrix $\mathbf{A}$ satisfies Restricted Isometry Property (RIP) with constant δ_s if

$$||\mathbf{Ax}\|_2^2 - \|\mathbf{x}\|_2^2| \leq \delta_s \|\mathbf{x}\|_2^2,$$

for all s- sparse vectors $\mathbf{x}$.

Theorem 4.1: If $\delta_s < 1$ there is a constant $C > 0$ such that if any $m \times N$ matrix $\mathbf{A}$ har RIP with constant δ_s, then

$$m > Cs \log(Ne/s).$$

A short summary from last lecture:

A short summary from last lecture: Recall the definition property of a matrix:

The $m \times N$ matrix $\mathbf{A}$ satisfies Restricted Isometry Property (RIP) with constant δ_s if

$$|\|\mathbf{A}\mathbf{x}\|_2^2 - \|\mathbf{x}\|_2^2| \leq \delta_s \|\mathbf{x}\|_2^2,$$

for all s- sparse vectors $\mathbf{x}$.

Theorem 4.1: If $\delta_s < 1$ there is a constant $C > 0$ such that if any $m \times N$ matrix $\mathbf{A}$ har RIP with constant δ_s, then

$$m > Cs \log(Ne/s).$$

More precisely we get

$$m \geq \frac{s \log(2Ne/s) - \log 4}{2(\log(\sqrt{1-\delta_s} + 4\sqrt{1+\delta_s}) - \log(1-\delta_s))}.$$

RIP for Gaussian Matrices

For Gaussian random matrices we have

Theorem 4.5: There is a constant $C > 0$ such that for any $\epsilon > 0$, and $0 < \delta \leq 1$, the following holds: If

$$m > C\frac{s}{\delta^2}\log(N/(s\epsilon) + \text{ lower order terms},$$

then the Gaussian $m \times N$ matrix $\mathbf{A}$ (as above) satisfy RIP constant $\delta_s = \delta$ with a probability not less than $1 - \epsilon$.

RIP for Gaussian Matrices

For Gaussian random matrices we have

Theorem 4.5: There is a constant $C > 0$ such that for any $\epsilon > 0$, and $0 < \delta \leq 1$, the following holds: If

$$m > C\frac{s}{\delta^2}\log(N/(s\epsilon) + \text{ lower order terms},$$

then the Gaussian $m \times N$ matrix $\mathbf{A}$ (as above) satisfy RIP constant $\delta_s = \delta$ with a probability not less than $1 - \epsilon$.

RIP for Gaussian Matrices

For Gaussian random matrices we have
Theorem 4.5: There is a constant $C > 0$ such that
for any $\epsilon > 0$, and $0 < \delta \leq 1$, the following holds: If

$$m > C\frac{s}{\delta^2}\log(N/(s\epsilon) + \text{ lower order terms},$$

More precisely

$$m > \frac{2s\log(N/s) + 4s\log\log(N/s)}{\delta_s - \log(1+\delta_s)} + \text{ lower order terms }.$$

then the Gaussian $m \times N$ matrix $\mathbf{A}$ (as above) satisfy RIP constant $\delta_s = \delta$ with a probability not less than $1 - \epsilon$.

Construction of matrices satisfyin RIP

Construction of RIP matrices:

We constructed an $m \times N$ matrix $\mathbf{A}$ statistfying the RIP property with constant δ_s with

$$m \sim C \frac{s^2}{\delta_s^2} \log N.$$

Construction of matrices satisfyin RIP

Construction of RIP matrices:
We constructed an $m \times N$ matrix $\mathbf{A}$ statistfying the RIP property with constant δ_s with

$$m \sim C \frac{s^2}{\delta_s^2} \log N.$$

We constructed a matrix $\mathbf{A}$ such that for any index-subset S, $|S| = s$, the corresponding matrix $\mathbf{A}_S$ of columns vectors of $\mathbf{A}$ is almost orthonormal.

Construction of matrices satisfyin RIP

Construction of RIP matrices:

We constructed an $m \times N$ matrix $\mathbf{A}$ statistfying the RIP property with constant δ_s with

$$m \sim C \frac{s^2}{\delta_s^2} \log N.$$

We constructed a matrix $\mathbf{A}$ such that for any index-subset S, $|S| = s$, the corresponding matrix $\mathbf{A}_S$ of columns vectors of $\mathbf{A}$ is almost orthonormal.

More precisely:

$$\mathbf{A}_S = \mathbf{U} + \mathbf{V},$$

where $\mathbf{U}$ is orthonormal and the column in $\mathbf{V}$ has length less than $\delta_s/\sqrt{s}$.

Structured Random Matrices

Definition: Structured Random matrices:

Structured Random Matrices

Definition: Structured Random matrices:

- A structured $m \times N$ random matrix $\mathbf{A}$ is genererated by one Random vector variable $Z = (Z_1, Z_2, \ldots, Z_N)$.

Structured Random Matrices

Definition: Structured Random matrices:

- A structured $m \times N$ random matrix $\mathbf{A}$ is genererated by one Random vector variable $Z = (Z_1, Z_2, \dots, Z_N)$.
- The random components Z_j, $1 \leq j \leq N$
  - need not to be independent-

Structured Random Matrices

Definition: Structured Random matrices:

- A structured $m \times N$ random matrix $\mathbf{A}$ is genererated by one Random vector variable $Z = (Z_1, Z_2, \ldots, Z_N)$.
- The random components Z_j, $1 \leq j \leq N$
  - need not to be independent-
  however: $E(Z_i Z_j) = 0$ for $i \neq j$, and $E(Z_j^2) = 1$ for all j.

Structured Random Matrices

Definition: Structured Random matrices:

- A structured $m \times N$ random matrix $\mathbf{A}$ is genererated by one Random vector variable $Z = (Z_1, Z_2, \ldots, Z_N)$.
- The random components Z_j, $1 \leq j \leq N$
  - need not to be independent-
  however: $E(Z_i Z_j) = 0$ for $i \neq j$, and $E(Z_j^2) = 1$ for all j.
- All Z_j are uniformly bounded: there is a constant $K \geq 1$ such that $|Z_j| \leq K$ for all j.

Structured Random Matrices

Definition: Structured Random matrices:

- A structured $m \times N$ random matrix $\mathbf{A}$ is genererated by one Random vector variable $Z = (Z_1, Z_2, \ldots, Z_N)$.
- The random components Z_j, $1 \leq j \leq N$
  - need not to be independent-
  however: $E(Z_i Z_j) = 0$ for $i \neq j$, and $E(Z_j^2) = 1$ for all j.
- All Z_j are uniformly bounded: there is a constant $K \geq 1$ such that $|Z_j| \leq K$ for all j.
- Each row $\mathbf{A}_i$ is an independent copy of the random variable

Structured Random Matrices

Definition: Structured Random matrices:

- A structured $m \times N$ random matrix $\mathbf{A}$ is genererated by one Random vector variable $Z = (Z_1, Z_2, \ldots, Z_N)$.
- The random components Z_j, $1 \leq j \leq N$
  - need not to be independent-
  however: $E(Z_i Z_j) = 0$ for $i \neq j$, and $E(Z_j^2) = 1$ for all j.
- All Z_j are uniformly bounded: there is a constant $K \geq 1$ such that $|Z_j| \leq K$ for all j.
- Each row $\mathbf{A}_i$ is an independent copy of the random variable
- Often the $m \times N$ matrix $\mathbf{A}$ is normalised:

$$\mathbf{A} = \frac{1}{\sqrt{m}} (\mathbf{A}_j)_{j=1}^{N}.$$

Structured Random Matrices

Definition: Structured Random matrices:

- A structured $m \times N$ random matrix $\mathbf{A}$ is genererated by one Random vector variable $Z = (Z_1, Z_2, \ldots, Z_N)$.
- The random components Z_j, $1 \leq j \leq N$
  - need not to be independent-
  however: $E(Z_i Z_j) = 0$ for $i \neq j$, and $E(Z_j^2) = 1$ for all j.
- All Z_j are uniformly bounded: there is a constant $K \geq 1$ such that $|Z_j| \leq K$ for all j.
- Each row $\mathbf{A}_i$ is an independent copy of the random variable
- Often the $m \times N$ matrix $\mathbf{A}$ is normalised:

$$\mathbf{A} = \frac{1}{\sqrt{m}} (\mathbf{A}_j)_{j=1}^{N}.$$

- Note that $\mathbf{A}_i$ is the columns of $\mathbf{A}$ then

$$E(\langle \mathbf{A}_i, \mathbf{A}_j \rangle) = 1 \text{ for } i \neq j \text{ and } E(\langle \mathbf{A}_i, \mathbf{A}_i \rangle) = 1 \text{ all } j.$$

Structured Random Matrices

Random Matrix from an Orthonormal $N \times N$ Matrix:

Structured Random Matrices

Random Matrix from an Orthonormal $N \times N$ Matrix:

- Let $\mathbf{B}$ be an orthonormal $N \times N$ matrix

Structured Random Matrices

Random Matrix from an Orthonormal $N \times N$ Matrix:

- Let $\mathbf{B}$ be an orthonormal $N \times N$ matrix
- We assume there is a constant K such that elements $\mathbf{B}_{ij}$ of $\mathbf{B}$ satisfies $|B_{ij}| \leq K\sqrt{N}$.

Structured Random Matrices

Random Matrix from an Orthonormal $N \times N$ Matrix:

- Let $\mathbf{B}$ be an orthonormal $N \times N$ matrix
- We assume there is a constant K such that elements $\mathbf{B}_{ij}$ of $\mathbf{B}$ satisfies $|B_{ij}| \leq K\sqrt{N}$.
- Let Z be the any row of $\mathbf{B}$ multiplied by $\sqrt{N}$, each row chosen by equal probability $\frac{1}{N}$.

Example: The discrete Fourier matrix

Structured Random Matrices

Random Matrix from an Orthonormal basis

Structured Random Matrices

Random Matrix from an Orthonormal basis

- Let $\{\varphi_j(t)\}_j$ be an orthononomal basis on an interval I.

Structured Random Matrices

Random Matrix from an Orthonormal basis

- Let $\{\varphi_j(t)\}_j$ be an orthononomal basis on an interval I.
- Assume that there is a constant $K \geq 1$ such that $|\varphi(t)| \leq K$ for all $t \in I$ and all j.

Structured Random Matrices

Random Matrix from an Orthonormal basis

- Let $\{\varphi_j(t)\}_j$ be an orthononomal basis on an interval I.
- Assume that there is a constant $K \geq 1$ such that $|\varphi(t)| \leq K$ for all $t \in I$ and all j.
- Let T be a random variable with its values uniformly distributed on the interval I.

Structured Random Matrices

Random Matrix from an Orthonormal basis

- Let $\{\varphi_j(t)\}_j$ be an orthononomal basis on an interval I.
- Assume that there is a constant $K \geq 1$ such that $|\varphi(t)| \leq K$ for all $t \in I$ and all j.
- Let T be a random variable with its values uniformly distributed on the interval I.
- Let $Z = (Z_1, Z_2, \ldots, Z_N)$, where $Z_j = \varphi(T)$.

Example: A wavelet basis and ..

Example: A wavelet basis and ..
the noiselet basis.

Structured matrix form an incoherent pair of bases

Let $\{\mathbf{e}_j\}$ and $\{\mathbf{f}_k\}$ be a pair of orthonormal bases on $\mathbb{R}^N$.

Structured matrix form an incoherent pair of bases

Let $\{\mathbf{e}_j\}$ and $\{\mathbf{f}_k\}$ be a pair of orthonormal bases on $\mathbb{R}^N$.

Definition:

Structured matrix form an incoherent pair of bases

Let $\{\mathbf{e}_j\}$ and $\{\mathbf{f}_k\}$ be a pair of orthonormal bases on $\mathbb{R}^N$.

Definition:

A pair of orthonormal bases is incoherent if all inner products $< \mathbf{e}_j, \mathbf{f}_k >$ are small.

Structured matrix form an incoherent pair of bases

Let $\{\mathbf{e}_j\}$ and $\{\mathbf{f}_k\}$ be a pair of orthonormal bases on $\mathbb{R}^N$.

Definition:

A pair of orthonormal bases is incoherent if all inner products $< \mathbf{e}_j, \mathbf{f}_k >$ are small.

Since

$$\sum_j < \mathbf{e}_j, \mathbf{f}_k >^2 = \sum_k < \mathbf{e}_j, \mathbf{f}_k >^2 = 1,$$

an upper bound for these innerproducts has to be at least $1/\sqrt{N}$.

Structured matrix form an incoherent pair of bases

Let $\{\mathbf{e}_j\}$ and $\{\mathbf{f}_k\}$ be a pair of orthonormal bases on $\mathbb{R}^N$.

Definition:
A pair of orthonormal bases is incoherent if all inner products $< \mathbf{e}_j, \mathbf{f}_k >$ are small.

Since

$$\sum_j < \mathbf{e}_j, \mathbf{f}_k >^2 = \sum_k < \mathbf{e}_j, \mathbf{f}_k >^2 = 1,$$

an upper bound for these innerproducts has to be at least $1/\sqrt{N}$.

Definition:

Structured matrix form an incoherent pair of bases

Let $\{\mathbf{e}_j\}$ and $\{\mathbf{f}_k\}$ be a pair of orthonormal bases on $\mathbb{R}^N$.

Definition:

A pair of orthonormal bases is incoherent if all inner products $< \mathbf{e}_j, \mathbf{f}_k >$ are small.

Since

$$\sum_j < \mathbf{e}_j, \mathbf{f}_k >^2 = \sum_k < \mathbf{e}_j, \mathbf{f}_k >^2 = 1,$$

an upper bound for these innerproducts has to be at least $1/\sqrt{N}$.

Definition:

The pair of orthonormal bases is incoherent with constant K if all the inner products above are bounded by $K/\sqrt{N}$.

Random matrix from incoherent bases

How to get the Structured random Matrix from the pair of incoherent bases:

Random matrix from incoherent bases

How to get the Structured random Matrix from the pair of incoherent bases:

- Let $\mathbf{E}^{-1}$ be the $N \times N$-matrix with rows $\mathbf{e}_j$.

Random matrix from incoherent bases

How to get the Structured random Matrix from the pair of incoherent bases:

- Let $\mathbf{E}^{-1}$ be the $N \times N$-matrix with rows $\mathbf{e}_j$.
- Let $\mathbf{F}$ be the matrix with columns $\mathbf{f}_k$.

Random matrix from incoherent bases

How to get the Structured random Matrix from the pair of incoherent bases:

- Let $\mathbf{E}^{-1}$ be the $N \times N$-matrix with rows $\mathbf{e}_j$.
- Let $\mathbf{F}$ be the matrix with columns $\mathbf{f}_k$.
- Then $\mathbf{B} = \mathbf{E}^{-1}\mathbf{F}$ with entries $< \mathbf{e}_j, \mathbf{f}_k >$ is an orthonormal $N \times N$-matrix as above with constant K.

Random matrix from incoherent bases

How to get the Structured random Matrix from the pair of incoherent bases:

- Let $\mathbf{E}^{-1}$ be the $N \times N$-matrix with rows $\mathbf{e}_j$.
- Let $\mathbf{F}$ be the matrix with columns $\mathbf{f}_k$.
- Then $\mathbf{B} = \mathbf{E}^{-1}\mathbf{F}$ with entries $< \mathbf{e}_j, \mathbf{f}_k >$ is an orthonormal $N \times N$-matrix as above with constant K.
- the $N \times N$ matrix $\mathbf{B}$ we get stuctured random matrix by selectiing each row by probaility $\frac{1}{N}$.

If U is the input signal we use the basis $\{\mathbf{e}_j\}$ and random measurement

$$\mathbf{y}_j = \langle U, \mathbf{e}_j \rangle,$$

If U is the input signal we use the basis $\{\mathbf{e}_j\}$ and random measurement

$$\mathbf{y}_j = \langle U, \mathbf{e}_j \rangle,$$

and U is sparsely represented int the basis $\{\mathbf{f}_k\}$ i.e

$$U = \sum_k \mathbf{x}_k \mathbf{f}_k,$$

for $\mathbf{x} = (\mathbf{x}_k)$ sparse.

If U is the input signal we use the basis $\{\mathbf{e}_j\}$ and random measurement

$$\mathbf{y}_j = \langle U, \mathbf{e}_j \rangle,$$

and U is sparsely represented int the basis $\{\mathbf{f}_k\}$ i.e

$$U = \sum_k \mathbf{x}_k \mathbf{f}_k,$$

for $\mathbf{x} = (\mathbf{x}_k)$ sparse. This gives the equation

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

RIP for Structured Random Matrices

For structured random matrices $\mathbf{A}$ with uniform bound K the best known result (ref. Holger Rauhut, 2011)

Rauhut gives the proof of

RIP for Structured Random Matrices

For structured random matrices **A** with uniform bound K the best known result (ref. Holger Rauhut, 2011)

Rauhut gives the proof of
Theorem 6.1
There is a constant $C > 0$ and $D < 0$ such that
for any $\epsilon > 0$, and $0 < \delta \leq 1$, the following holds:
If

RIP for Structured Random Matrices

For structured random matrices **A** with uniform bound K the best known result (ref. Holger Rauhut, 2011)

Rauhut gives the proof of
Theorem 6.1
There is a constant $C > 0$ and $D < 0$ such that
for any $\epsilon > 0$, and $0 < \delta \leq 1$, the following holds:
If

$$m > D\frac{K^2 s}{\delta^2}\ln^2(100s)\ln(4N)\ln(10m)\ln(\frac{7}{\epsilon}),$$

then

RIP for Structured Random Matrices

For structured random matrices **A** with uniform bound K the best known result (ref. Holger Rauhut, 2011)

Rauhut gives the proof of
Theorem 6.1
There is a constant $C > 0$ and $D < 0$ such that
for any $\epsilon > 0$, and $0 < \delta \leq 1$, the following holds:
If

$$m > D\frac{K^2 s}{\delta^2}\ln^2(100s)\ln(4N)\ln(10m)\ln(\frac{7}{\epsilon}),$$

then the $m \times N$ structured random matrix **A** (as above) satisfy RIP constant $\delta_s = \delta$ with

RIP for Structured Random Matrices

For structured random matrices **A** with uniform bound K the best known result (ref. Holger Rauhut, 2011)

Rauhut gives the proof of
Theorem 6.1
There is a constant $C > 0$ and $D < 0$ such that
for any $\epsilon > 0$, and $0 < \delta \leq 1$, the following holds:
If

$$m > D\frac{K^2 s}{\delta^2} \ln^2(100s) \ln(4N) \ln(10m) \ln(\frac{7}{\epsilon}),$$

then the $m \times N$ structured random matrix **A** (as above) satisfy RIP constant $\delta_s = \delta$ with a probability not less than $1 - \epsilon$.

RIP for Structured Random Matrices

For structured random matrices **A** with uniform bound K the best known result (ref. Holger Rauhut, 2011)

Rauhut gives the proof of
Theorem 6.1
There is a constant $C > 0$ and $D < 0$ such that
for any $\epsilon > 0$, and $0 < \delta \leq 1$, the following holds:
If

$$m > D\frac{K^2 s}{\delta^2} \ln^2(100s) \ln(4N) \ln(10m) \ln(\frac{7}{\epsilon}),$$

then the $m \times N$ structured random matrix **A** (as above) satisfy RIP constant $\delta_s = \delta$ with a probability not less than $1 - \epsilon$.

The constant D is very large, it satisfies $D < 163931.48$.

Rauhut also state a better result:

Theorem 6.2

There is a constant $C > 0$ and $D < 0$ such that for any $\epsilon > 0$, and $0 < \delta \leq 1$, the following holds:

Rauhut also state a better result:

Theorem 6.2

There is a constant $C > 0$ and $D < 0$ such that for any $\epsilon > 0$, and $0 < \delta \leq 1$, the following holds: If

$$\frac{m}{\ln(10m)} > C\frac{K^2 s}{\delta^2}\ln^2(100s)\ln(N),$$

Rauhut also state a better result:

Theorem 6.2

There is a constant $C > 0$ and $D < 0$ such that for any $\epsilon > 0$, and $0 < \delta \leq 1$, the following holds: If

$$\frac{m}{\ln(10m)} > C\frac{K^2 s}{\delta^2}\ln^2(100s)\ln(N),$$

$$\frac{m}{\ln(10m)} > D\frac{K^2 s}{\delta^2}\ln(\frac{1}{\epsilon}).$$

then

Rauhut also state a better result:

Theorem 6.2

There is a constant $C > 0$ and $D < 0$ such that for any $\epsilon > 0$, and $0 < \delta \leq 1$, the following holds: If

$$\frac{m}{\ln(10m)} > C\frac{K^2 s}{\delta^2}\ln^2(100s)\ln(N),$$
$$\frac{m}{\ln(10m)} > D\frac{K^2 s}{\delta^2}\ln(\frac{1}{\epsilon}).$$

then the $m \times N$ structured random matrix **A** (as above) satisfy RIP constant $\delta_s = \delta$ with

Rauhut also state a better result:

Theorem 6.2

There is a constant $C > 0$ and $D < 0$ such that for any $\epsilon > 0$, and $0 < \delta \leq 1$, the following holds: If

$$\frac{m}{\ln(10m)} > C\frac{K^2 s}{\delta^2}\ln^2(100s)\ln(N),$$

$$\frac{m}{\ln(10m)} > D\frac{K^2 s}{\delta^2}\ln(\frac{1}{\epsilon}).$$

then the $m \times N$ structured random matrix **A** (as above) satisfy RIP constant $\delta_s = \delta$ with a probability not less than $1 - \epsilon$.

Rauhut also state a better result:

Theorem 6.2
There is a constant $C > 0$ and $D < 0$ such that
for any $\epsilon > 0$, and $0 < \delta \le 1$, the following holds: If

$$\frac{m}{\ln(10m)} > C\frac{K^2 s}{\delta^2}\ln^2(100s)\ln(N),$$
$$\frac{m}{\ln(10m)} > D\frac{K^2 s}{\delta^2}\ln(\frac{1}{\epsilon}).$$

then the $m \times N$ structured random matrix **A** (as above) satisfy
RIP constant $\delta_s = \delta$ with a probability not less than $1 - \epsilon$.

The constants C and D are very large:

Rauhut also state a better result:

Theorem 6.2

There is a constant $C > 0$ and $D < 0$ such that for any $\epsilon > 0$, and $0 < \delta \leq 1$, the following holds: If

$$\frac{m}{\ln(10m)} > C\frac{K^2 s}{\delta^2}\ln^2(100s)\ln(N),$$

$$\frac{m}{\ln(10m)} > D\frac{K^2 s}{\delta^2}\ln(\frac{1}{\epsilon}).$$

then the $m \times N$ structured random matrix $\mathbf{A}$ (as above) satisfy RIP constant $\delta_s = \delta$ with a probability not less than $1 - \epsilon$.

The constants C and D are very large:
The constants satisfy $C < 17190$ and $D < 456$.

Uniform versus non-uniform recovery with random matrices

Let $\mathbf{A}$ be $m \times N$ random matrix. Let sparsness level s and let $\epsilon > 0$ be given

Uniform versus non-uniform recovery with random matrices

Let **A** be $m \times N$ random matrix. Let sparsness level s and let $\epsilon > 0$ be given

- Uniform recovery: With large probability $(1 - \epsilon)$ a realisation of **A** have the property that all s sparse vectors **x** can uniqely be recorvered from the equation $\mathbf{Ax} = \mathbf{y}$ by l_1 opitmization:

Uniform versus non-uniform recovery with random matrices

Let **A** be $m \times N$ random matrix. Let sparsness level s and let $\epsilon > 0$ be given

- Uniform recovery: With large probability $(1 - \epsilon)$ a realisation of **A** have the property that all s sparse vectors **x** can uniqely be recorvered from the equation $\mathbf{Ax} = \mathbf{y}$ by l_1 opitmization:

$$\min_{\bar{\mathbf{x}} \in R^N} \|\bar{\mathbf{x}}\|_1 \text{ subject to } \mathbf{A}\bar{\mathbf{x}} = \mathbf{y}.$$

Uniform versus non-uniform recovery with random matrices

Let $\mathbf{A}$ be $m \times N$ random matrix. Let sparsness level s and let $\epsilon > 0$ be given

- Uniform recovery: With large probability $(1-\epsilon)$ a realisation of $\mathbf{A}$ have the property that all s sparse vectors $\mathbf{x}$ can uniqely be recorvered from the equation $\mathbf{A}\mathbf{x} = \mathbf{y}$ by l_1 opitmization:

$$\min_{\bar{\mathbf{x}} \in R^N} \|\bar{\mathbf{x}}\|_1 \text{ subject to } \mathbf{A}\bar{\mathbf{x}} = \mathbf{y}.$$

- Non-uniform recovery: For all s sparse vectors $\mathbf{x}$ there is a large probability $(1-\epsilon)$ that at reaslisation of $\mathbf{A}$ has the propery that $\mathbf{x}$ can be recovered from the equation $\mathbf{A}\mathbf{x} = \mathbf{y}$ by l_1 optimization:

Uniform versus non-uniform recovery with random matrices

Let $\mathbf{A}$ be $m \times N$ random matrix. Let sparsness level s and let $\epsilon > 0$ be given

- Uniform recovery: With large probability $(1 - \epsilon)$ a realisation of $\mathbf{A}$ have the property that all s sparse vectors $\mathbf{x}$ can uniqely be recorvered from the equation $\mathbf{Ax} = \mathbf{y}$ by l_1 opitmization:

$$\min_{\bar{\mathbf{x}} \in R^N} \|\bar{\mathbf{x}}\|_1 \text{ subject to } \mathbf{A}\bar{\mathbf{x}} = \mathbf{y}.$$

- Non-uniform recovery: For all s sparse vectors $\mathbf{x}$ there is a large probability $(1 - \epsilon)$ that at reaslisation of $\mathbf{A}$ has the propery that $\mathbf{x}$ can be recovered from the equation $\mathbf{Ax} = \mathbf{y}$ by l_1 optimization:

$$\min_{\bar{\mathbf{x}} \in R^N} \|\bar{\mathbf{x}}\|_1 \text{ subject to } \mathbf{A}\bar{\mathbf{x}} = \mathbf{y}.$$

Uniform recovery for Structured R. Matrices

Collorary Theorem 6.2

There is a constant $C_1 > 0$ and $D_1 < 0$ such that for any $\epsilon > 0$, the following holds:

Uniform recovery for Structured R. Matrices

Collorary Theorem 6.2

There is a constant $C_1 > 0$ and $D_1 < 0$ such that for any $\epsilon > 0$, the following holds: If

$$\frac{m}{\ln(10m)} > C_1 K^2 s \ln^2(100s) \ln(N),$$

Uniform recovery for Structured R. Matrices

Collorary Theorem 6.2

There is a constant $C_1 > 0$ and $D_1 < 0$ such that for any $\epsilon > 0$, the following holds: If

$$\frac{m}{\ln(10m)} > C_1 K^2 s \ln^2(100s) \ln(N),$$

$$\frac{m}{\ln(10m)} > D_1 K^2 s \ln(\frac{1}{\epsilon}),$$

then

Uniform recovery for Structured R. Matrices

Collorary Theorem 6.2

There is a constant $C_1 > 0$ and $D_1 < 0$ such that for any $\epsilon > 0$, the following holds: If

$$\frac{m}{\ln(10m)} > C_1 K^2 s \ln^2(100s) \ln(N),$$

$$\frac{m}{\ln(10m)} > D_1 K^2 s \ln(\frac{1}{\epsilon}),$$

then the $m \times N$ structured random matrix **A** (as above) will have the uniform recovery property with

Uniform recovery for Structured R. Matrices

Collorary Theorem 6.2

There is a constant $C_1 > 0$ and $D_1 < 0$ such that for any $\epsilon > 0$, the following holds: If

$$\frac{m}{\ln(10m)} > C_1 K^2 s \ln^2(100s) \ln(N),$$

$$\frac{m}{\ln(10m)} > D_1 K^2 s \ln(\frac{1}{\epsilon}),$$

then the $m \times N$ structured random matrix **A** (as above) will have the uniform recovery property with a probability not less than $1 - \epsilon$.

Non-uniform recovery for Structured R. Matrices

Refer to Cande's and Yaniv Plan

Theorem 6.3

There is a constant $C_0 > 0$ such that
for any $\beta > 0$, $K \geq 1$, the following holds:

Non-uniform recovery for Structured R. Matrices

Refer to Cande's and Yaniv Plan

Theorem 6.3

There is a constant $C_0 > 0$ such that
for any $\beta > 0$, $K \geq 1$, the following holds: If

$$m > C_0(1 + \log(\epsilon))K^2 s \ln(N),$$

then

Non-uniform recovery for Structured R. Matrices

Refer to Cande's and Yaniv Plan

Theorem 6.3

There is a constant $C_0 > 0$ such that
for any $\beta > 0$, $K \geq 1$, the following holds: If

$$m > C_0(1 + \log(\epsilon))K^2 s \ln(N),$$

then the structured $m \times N$ random matrix $\mathbf{A}$
(as above - with non-coherence constant K)
will have the non-uniform recovery property with

Non-uniform recovery for Structured R. Matrices

Refer to Cande's and Yaniv Plan

Theorem 6.3

There is a constant $C_0 > 0$ such that
for any $\beta > 0$, $K \geq 1$, the following holds: If

$$m > C_0(1 + \log(\epsilon))K^2 s \ln(N),$$

then the structured $m \times N$ random matrix $\mathbf{A}$
(as above - with non-coherence constant K)
will have the non-uniform recovery property with
a probability not less than $1 - \frac{5}{n} - \epsilon$.

Recovery with appoximative sparsness and noise

Refer to Cande's et al.

Assume that the vector **x** to be recoverd is not s sparse but can be well approximated by s-sparse functions for some s.

Recovery with appoximative sparsness and noise

Refer to Cande's et al.

Assume that the vector $\mathbf{x}$ to be recoverd is not s sparse but can be well approximated by s-sparse functions for some s. Assume also we have some noise $\sigma\mathbf{z}$, i.e.

Recovery with appoximative sparsness and noise

Refer to Cande's et al.
Assume that the vector $\mathbf{x}$ to be recoverd is not s sparse but can be well approximated by s-sparse functions for some s. Assume also we have some noise $\sigma \mathbf{z}$, i.e.

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \sigma \mathbf{z},$$

where $\mathbf{z}$ is Gaussian $\mathbf{z}$ $N(0,1)$.
(However assuming $\|\mathbf{A}\mathbf{z}\|_\infty \leq \lambda_N$ for some $\lambda_N \geq 0$.)

Recovery with appoximative sparsness and noise

Refer to Cande's et al.
Assume that the vector $\mathbf{x}$ to be recoverd is not s sparse but can be well approximated by s-sparse functions for some s. Assume also we have some noise $\sigma \mathbf{z}$, i.e.

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \sigma \mathbf{z},$$

where $\mathbf{z}$ is Gaussian $\mathbf{z}$ $N(0,1)$.
(However assuming $\|\mathbf{A}\mathbf{z}\|_\infty \leq \lambda_N$ for some $\lambda_N \geq 0$.)
We consider the l_1 regularized least-square problem

$$\min_{\bar{\mathbf{x}} \in R^N} \|\mathbf{A}\bar{\mathbf{x}} - \mathbf{y}\|_2 + \lambda \|\bar{\mathbf{x}}\|_1, \qquad (*)$$

Theorem 6.4: Let $\mathbf{x}$ be an arbitary vector in $\mathbb{R}^n$. Then with probability at least $1 - \frac{6}{n} - 6\epsilon$ the solution to $(*)$ with $\lambda = 10\sqrt{\log N}$ obeys

Theorem 6.4: Let $\mathbf{x}$ be an arbitary vector in $\mathbb{R}^n$. Then with probability at least $1 - \frac{6}{n} - 6\epsilon$ the solution to $(*)$ with $\lambda = 10\sqrt{\log N}$ obeys

$$
\begin{aligned}
\|\bar{\mathbf{x}} - \mathbf{x}\|_2 &\leq \\
\leq \quad \min_{1 \leq s \leq \bar{s}} &\quad C\sqrt{(1 + \log(\frac{1}{\epsilon})s \log^5 N} \left[\frac{\|\mathbf{x} - \mathbf{x}_s\|_1}{\sqrt{s}} + \sigma \sqrt{\frac{s \log N}{m}} \right],
\end{aligned}
$$

Theorem 6.4: Let $\mathbf{x}$ be an arbitary vector in $\mathbb{R}^n$. Then with probability at least $1 - \frac{6}{n} - 6\epsilon$ the solution to $(*)$ with $\lambda = 10\sqrt{\log N}$ obeys

$$\begin{aligned} \|\bar{\mathbf{x}} - \mathbf{x}\|_2 &\leq \\ \leq \quad \min_{1 \leq s \leq \bar{s}} \quad & C\sqrt{(1 + \log(\frac{1}{\epsilon})s \log^5 N} \left[\frac{\|\mathbf{x} - \mathbf{x}_s\|_1}{\sqrt{s}} + \sigma\sqrt{\frac{s \log N}{m}} \right], \end{aligned}$$

or with error in l_1 norm:

$$\begin{aligned} \|\bar{\mathbf{x}} - \mathbf{x}\|_1 &\leq \\ \leq \quad \min_{1 \leq s \leq \bar{s}} \quad & C\sqrt{(1 + \log(\frac{1}{\epsilon})s \log^5 N} \left[\|\mathbf{x} - \mathbf{x}_s\|_1 + s\sigma\sqrt{\frac{\log N}{m}} \right]. \end{aligned}$$