$\ell_1$ minimization

Bhaskar D Rao
University of California, San Diego
Email: brao@ucsd.edu
$\ell_p$ penalty function

The $\ell_p$ norm of a $x \in \mathbb{R}^n$ be given by

$$\|x\|_p = \left( \sum_{i=1}^{M} |x_i|^p \right)^{1/p}$$

$p = 2$ is the popular 2-norm. $p = 1$ is the popular 1-norm used for enforcing sparsity. $p \geq 1$ are norms and convex functions of $x$.

The counting measure is "zero-norm"

$$\|x\|_0 = |\text{support}(x)|$$

The 2-norm of a matrix $A$ (also known as the spectral norm) is given by

$$\|A\|_2 = \max_{\|x\|_2 = 1} \|Ax\|_2 = \sigma_{\text{MAX}} = \sqrt{\lambda_{\text{MAX}}(AA^T)}$$

where $\sigma_{\text{MAX}}$ refers to the maximum singular value of $A$. 
The $\ell_p$ norm of a $x \in \mathbb{R}^n$ be given by

$$\|x\|_p = \left( \sum_{i=1}^{M} |x_i|^p \right)^{\frac{1}{p}}$$

$p = 2$ is the popular 2-norm. $p = 1$ is the popular 1-norm used for enforcing sparsity. $p \geq 1$ are norms and convex functions of $x$. 
\( \ell_p \) penalty function

The \( \ell_p \) norm of a \( x \in \mathbb{R}^n \) be given by

\[
\|x\|_p = \left( \sum_{i=1}^{M} |x_i|^p \right)^{\frac{1}{p}}
\]

\( p = 2 \) is the popular 2-norm. \( p = 1 \) is the popular 1-norm used for enforcing sparsity. \( p \geq 1 \) are norms and convex functions of \( x \).

The counting measure is ”zero-norm”

\[
\|x\|_0 = |\text{support}(x)|
\]
The $l_p$ norm of a $x \in \mathbb{R}^n$ be given by

$$\|x\|_p = \left( \sum_{i=1}^{M} |x_i|^p \right)^{\frac{1}{p}}$$

$p = 2$ is the popular 2-norm. $p = 1$ is the popular 1-norm used for enforcing sparsity. $p \geq 1$ are norms and convex functions of $x$.

The counting measure is ”zero-norm”

$$\|x\|_0 = |\text{support}(x)|$$

The 2-norm of a matrix $A$ (also known as the spectral norm) is given by

$$\|A\|_2 = \max_{\|x\|_2=1} \|Ax\|_2 = \sigma_{\text{MAX}} = \sqrt{\lambda_{\text{MAX}}(AA^T)}$$

where $\sigma_{\text{MAX}}$ refers to the maximum singular value of $A$. 
\[ \ell_1 \text{ Regularization} \]

\[
\begin{align*}
\text{Noiseless Scenario} \\
\min_x & \quad Ax = y \\
\text{s.t.} & \quad \|x\|_1 = 1 
\end{align*}
\]

In the presence of noisy, there are two variants:

\[
\begin{align*}
\text{LASSO} & \quad \min_x & \quad \|y - Ax\|_2^2 + \lambda \|x\|_1, \quad \lambda > 0 \\
\text{Basis Pursuit} & \quad \min_x & \quad \|x\|_1 \\
\text{s.t.} & \quad \|y - Ax\|_2^2 \leq \beta 
\end{align*}
\]
$\ell_1$ Regularization

Noiseless Scenario

$$\min_{x: Ax = y} \|x\|_1$$

(1)
\( \ell_1 \) Regularization

**Noiseless Scenario**

\[
\min_{x : Ax = y} \| x \|_1 \tag{1}
\]

In the presence of noisy, there are two variants
$\ell_1$ Regularization

Noiseless Scenario

$$\min_{x : Ax = y} \| x \|_1$$  \hspace{1cm} (1)

In the presence of noisy, there are two variants

LASSO

$$\min_x \| y - Ax \|_2^2 + \lambda \| x \|_1, \quad \lambda > 0.$$ \hspace{1cm} (2)
\( \ell_1 \) Regularization

Noiseless Scenario

\[
\min_{x \mid Ax = y} \|x\|_1
\]  \hspace{1cm} (1)

In the presence of noisy, there are two variants

LASSO

\[
\min_x \|y - Ax\|_2^2 + \lambda \|x\|_1, \quad \lambda > 0.
\]  \hspace{1cm} (2)

Basis Pursuit (BP)

\[
\min_x \|x\|_1 \quad \text{subject to} \quad \|y - Ax\|_2^2 \leq \beta.
\]  \hspace{1cm} (3)
Notation

Index Set \[ M \] = \{ 1, 2, \ldots, M \}

Let \( S \) be an index set such that \(|S| = \|x\|_0 = k\), \( S \subset \{ 1, 2, \ldots, M \} \),

Let \( x_{S} \in \mathbb{R}^n \) be the vector that coincides with \( x \) on \( S \) and is 0 elsewhere:

\[ x_{i}^{S} = \begin{cases} x_{i} \quad \text{if } i \in S \\ 0 \quad \text{otherwise} \end{cases} \]

Let \( S \) be a set such that \(|S| = \|x\|_0 = k\), \( S \subset \{ 1, 2, \ldots, M \} \) \( (4) \),

\( A_{S} \) is the matrix consisting of columns indexed by \( S \):

\[ A_{S} = \begin{bmatrix} a_{j} \end{bmatrix} : j \in S \]
Notation

Index Set $[M]$

\[ [M] = \{1, 2, \cdots, M\} \]
Index Set $[M]$

$[M] = \{1, 2, \cdots, M\}$

Let $S$ be an index set such that

$|S| = \|x\|_0 = k$, $S^c = [M] \setminus S$
Notation

Index Set $[M]$

$[M] = \{1, 2, \cdots, M\}$

Let $S$ be an index set such that

$|S| = \|x\|_0 = k, \quad S^c = [M] \setminus S$

Let $x_S \in \mathbb{R}^n$ be the vector that coincides with $x$ on $S$ and is 0 elsewhere:

$x_S^i = \begin{cases} 
    x_i & : \quad i \in S \\
    0 & : \quad \text{else}
\end{cases}$
Index Set $[M]$

$$[M] = \{1, 2, \ldots, M\}$$

Let $S$ be an index set such that

$$|S| = \|x\|_0 = k, \quad S^c = [M] \setminus S$$

Let $x_S \in \mathbb{R}^n$ be the vector that coincides with $x$ on $S$ and is 0 elsewhere:

$$x^i_S = \begin{cases} x_i &: i \in S \\ 0 &: \text{else} \end{cases}$$

Let $S$ be a set such that

$$|S| = \|x\|_0 = k, \quad S^c = [M] \setminus S \quad (4)$$
Notation

Index Set \([M]\)

\([M] = \{1, 2, \cdots, M\}\)

Let \(S\) be an index set such that

\[|S| = \|x\|_0 = k, \ S^c = [M] \setminus S\]

Let \(x_S \in \mathbb{R}^n\) be the vector that coincides with \(x\) on \(S\) and is 0 elsewhere:

\[x^i_S = \begin{cases} x_i & : i \in S \\ 0 & : \text{else} \end{cases}\]

Let \(S\) be a set such that

\[|S| = \|x\|_0 = k, \ S^c = [M] \setminus S\] (4)

\(A_S\) is the matrix consisting of columns indexed by \(S\):

\[A_S = [a_j : j \in S]\]
Null Space Property (NSP)

Definition
A matrix $A$ satisfies the NSP of order $k$ if for all subsets $S \subseteq [M]$ with $|S| \leq k$, we have

$$\|v_S\|_1 < \|v_S^C\|_1,$$

where $v \in N(A)$.

Theorem
Let $A \in \mathbb{R}^{N \times M}$. Then, every $k$-sparse vector $x$ is the unique solution of the minimization problem

$$\min_{x: Ax = y} \|x\|_1 (5)$$

if and only if $A$ satisfies the NSP of order $k$. 
Definition
A matrix $A$ satisfies the NSP of order $k$ if for all subsets $S \subset [M]$ with $|S| \leq k$, we have

$$\|v_S\|_1 < \|v_{Sc}\|_1, \ \forall \ v \in N(A)$$
Null Space Property (NSP)

**Definition**
A matrix $A$ satisfies the NSP of order $k$ if for all subsets $S \subset [M]$ with $|S| \leq k$, we have

$$
\|v_S\|_1 < \|v_{Sc}\|_1, \quad v \in \mathcal{N}(A)
$$

**Theorem**
Let $A \in \mathbb{R}^{N \times M}$. Then, every $k$ sparse vector $x$ is the unique solution of the minimization problem

$$
\min_{x: Ax = y} \|x\|_1 \quad (5)
$$

if and only if $A$ satisfies the NSP of order $k$. 
Proof.

We will start by proving the forward direction: if minimization of (5) generates the sparsest possible solution, then $A$ satisfies the NSP. We start by picking a $k$ sparse $x$ and generate $y = Ax$. Then, we solve (5) to find $x^*$, which, by assumption, is equal to $x$. We now seek to show that $\|v_S\|_1 < \|v_{Sc}\|$. We start with some $v \in N(A)$, which can always be decomposed as $v = v_S + v_{Sc}$. Since we are free to choose any $k$ sparse $x$, we choose $x = v_S$. Now, since $Av = 0$, we have that $Av_S = -Av_{Sc}$. We know that

$$\arg \min_{x: Ax = Av_S} \|x\|_1$$

is solved by $x^* = v_S$. But we also know that $-v_{Sc}$ is a candidate solution to (6) and it was not chosen as the minimum $\ell_1$ solution, so this must mean that $\|v_S\|_1 < \|v_{Sc}\|$. \qed
We will now show that if $A$ satisfies the NSP, then the minimizer of (5) is the sparsest solution. Given any $x$ that is $k$ sparse over $S$ and $z \neq x$ such that $Az = Ax$, we can form:

$$v = z - x \in N(A)$$

Now, we have

$$\|x\|_1 = \|x - z_S + z_S\|_1$$
$$\leq \|x - z_S\|_1 + \|z_S\|_1$$
$$= \|v_S\|_1 + \|z_S\|_1$$
$$\leq \|v_{Sc}\|_1 + \|z_S\|_1 \text{ by NSP}$$
$$= \|z_{Sc}\|_1 + \|z_S\|_1$$
$$= \|z\|_1$$

Therefore, $x$ has to be the unique minimizer of (5).
Restricted Isometry Property (RIP)

The NSP is a good theoretical tool, but it is nearly impossible to verify. Therefore, we turn our attention to a different property, the RIP. The RIP is also nearly impossible to verify, but it has been shown that, with arbitrarily high probability, matrices can be constructed which satisfy the RIP.

Definition

A restricted isometry constant \( \delta_k \) of a matrix \( A \in \mathbb{R}^{n \times m} \) is defined as the smallest \( \delta_k \) such that

\[
(1 - \delta_k) \|x\|_2^2 \leq \|Ax\|_2^2 \leq (1 + \delta_k) \|x\|_2^2
\]

for all \( x \) that are \( k \)-sparse or less.
Restricted Isometry Property (RIP)

The NSP is a good theoretical tool, but it is nearly impossible to verify. Therefore, we turn our attention to a different property, the RIP. The RIP is also nearly impossible to verify, but it has been shown that, with arbitrarily high probability, matrices can be constructed which satisfy the RIP.

**Definition**

A restricted isometry constant \( \delta_k \) of a matrix \( A \in \mathbb{R}^{n \times m} \) is defined as the smallest \( \delta_k \) such that

\[
(1 - \delta_k) \| x \|_2^2 \leq \| Ax \|_2^2 \leq (1 + \delta_k) \| x \|_2^2
\]

for all \( x \) that are \( k \)-sparse or less.
The NSP is a good theoretical tool, but it is nearly impossible to verify. Therefore, we turn our attention to a different property, the RIP. The RIP is also nearly impossible to verify, but it has been shown that, with arbitrarily high probability, matrices can be constructed which satisfy the RIP.

**Definition**

A restricted isometry constant $\delta_k$ of a matrix $A \in \mathbb{R}^{n \times m}$ is defined as the smallest $\delta_k$ such that $(1 - \delta_k)\|x\|_2^2 \leq \|Ax\|_2^2 \leq (1 + \delta_k)\|x\|_2^2$ for all $x$ that are $k$ sparse or less.
Properties of the RIP

1. $\delta_1 \leq \delta_2 \leq \delta_3 \cdots$

2. $\delta_k = \max_{S \subset [m], |S| = k} \| A_T S A_S - I \|_2$

3. If $u, v \in \mathbb{R}^m$ have disjoint support, and $\| u \|_0 \leq k$, $\| v \|_0 \leq s$, then $|\langle A u, A v \rangle| \leq \delta_k + s \| u \|_2 \| v \|_2$
1. $\delta_1 \leq \delta_2 \leq \delta_3 \cdots$
Properties of the RIP

1. $\delta_1 \leq \delta_2 \leq \delta_3 \cdots$

2. 

$$\delta_k = \max_{S \subseteq [m], |S|=k} \|A_S^T A_S - I\|_2$$
Properties of the RIP

1. $\delta_1 \leq \delta_2 \leq \delta_3 \cdots$

2. 

$$\delta_k = \max_{S \subset [m], |S| = k} \|A_S^T A_S - I\|_2$$

3. If $u, v \in \mathbb{R}^m$ have disjoint support, and $\|u\|_0 \leq k$, $\|v\|_0 \leq s$, then

$$|\langle Au, Av \rangle| \leq \delta_{k+s} \|u\|_2 \|v\|_2$$
Theorem

Suppose the restricted isometry constant (RIC) of $A$ satisfies $\delta_{2k} < \frac{1}{3}$, then $A$ satisfies the NSP of order $k$.

In particular, every $k$-sparse vector is recovered by $l_1$ minimization.
Compressible Signals

Compressible $x$, i.e. $x$ which are not sparse but approximately sparse.
Compressible Signals

Compressible $x$, i.e. $x$ which are not sparse but approximately sparse.

**Definition**

$$\sigma_k(x)_p = \min_{z: \|z\|_0 = k} \|z - x\|_p, \ p \geq 1$$
Compressible Signals

Compressible $x$, i.e. $x$ which are not sparse but approximately sparse.

**Definition**

$$\sigma_k(x)_p = \min_{z: \|z\|_0 = k} \|z - x\|_p, \ p \geq 1$$

We consider $x^*$ given by

$$x^* = \arg \min_{z:Az=y} \|z\|_1 \quad (7)$$

and ask how close $x$ and $x^*$ are.
Compressible Signals

Compressible $x$, i.e. $x$ which are not sparse but approximately sparse.

**Definition**

$$\sigma_k(x)_p = \min_{z: \|z\|_0 = k} \|z - x\|_p, \ p \geq 1$$

We consider $x^*$ given by

$$x^* = \arg \min_{z: Az = y} \|z\|_1$$

(7)

and ask how close $x$ and $x^*$ are.

Modified NSP of order $k$:

$$\|v_S\|_1 \leq \gamma \|v_{Sc}\|_1$$

where parameter $\gamma \in (0, 1)$. 
Compressible Signals

Compressible $x$, i.e. $x$ which are not sparse but approximately sparse.

**Definition**

$$\sigma_k(x)_p = \min_{z: \|z\|_0 = k} \|z - x\|_p, \ p \geq 1$$

We consider $x^*$ given by

$$x^* = \arg \min_{z: Az = y} \|z\|_1 \tag{7}$$

and ask how close $x$ and $x^*$ are.

Modified NSP of order $k$:

$$\|v_S\|_1 \leq \gamma \|v_{Sc}\|_1$$

where parameter $\gamma \in (0, 1)$.

The main result is as follows:

$$\|x - x^*\|_1 \leq \frac{2(1 + \gamma)}{1 - \gamma} \sigma_k(x)_1$$
Noisy Case with $x$-sparse

Theorem

1 Assume $y = Ax_0 + e$ where $x_0$ is $k$-sparse and the noise vector is bounded: $\|e\|_2 \leq \epsilon$. Suppose $x^*$ is the solution to the following $\ell_1$ minimization problem

$$\min_{x : \|Ax - y\| \leq \epsilon} \|x\|_1$$

and $A$ has RIP constants $\delta_{3k} + 3\delta_{4k} < 2$. Then

$$\|x_0 - x^*\|_2 \leq c_k \epsilon.$$ 

For $\delta_{4k} = \frac{1}{5}$ then $c_k \approx 8.82$.

---

1By Candes and Tao
A general result with a slightly different definition of NSP which is given below. Here Λ is an index set, same as S before.

**Definition**

Matrix $A$ satisfies the NSP of order $k$ if there exists a constant $c > 0$ such that

$$\|h_\Lambda\|_2 \leq c \|h_\Lambda c\|_1 \sqrt{k}, \forall h \in \mathbb{N}(A), |\Lambda| = k,$$

Previously $\|h_\Lambda\|_1 \leq \gamma \|h_\Lambda c\|_1$. We know that $\|h_\Lambda\|_1 \leq \sqrt{k} \|h_\Lambda\|_2$. So if matrix $A$ satisfies NSP with new definition, it satisfies the NSP with previous definition.
A general result with a slightly different definition of NSP which is given below. Here Λ is a index set, same as $S$ before.
A general result with a slightly different definition of NSP which is given below. Here $\Lambda$ is an index set, same as $S$ before.

**Definition**

Matrix $A$ satisfies the NSP of order $k$ if there exists a constant $c > 0$ such that

$$
\| h_\Lambda \|_2 \leq c \frac{\| h_\Lambda^c \|_1}{\sqrt{k}}, \quad \forall h \in \mathcal{N}(A), \quad |\Lambda| = k
$$
A general result with a slightly different definition of NSP which is given below. Here $\Lambda$ is a index set, same as $S$ before.

**Definition**
Matrix $A$ satisfies the NSP of order $k$ if there exists a constant $c > 0$ such that

$$||h_\Lambda||_2 \leq c \frac{||h_{\Lambda^c}||_1}{\sqrt{k}}, \quad \forall h \in \mathcal{N}(A), \quad |\Lambda| = k$$

Previously $||h_\Lambda||_1 \leq \gamma ||h_{\Lambda^c}||_1$. We know that $||h_\Lambda||_1 \leq \sqrt{k} ||h_\Lambda||_2$. So if matrix $A$ satisfies NSP with new definition, it satisfies the NSP with previous definition.
Theorem

Suppose matrix $A$ satisfies the RIP of order $2k$ with $\delta_{2k} < \sqrt{2} - 1$. 
Main Result

**Theorem**
Suppose matrix $A$ satisfies the RIP of order $2k$ with $\delta_{2k} < \sqrt{2} - 1$. Let $x, \hat{x} \in \mathbb{R}^m$ be given and $h = \hat{x} - x$. 

Note that in the above lemma, $x$ is not necessarily the optimal solution of the $\ell_1$ norm problem.
Main Result

Theorem

Suppose matrix $A$ satisfies the RIP of order $2k$ with $\delta_{2k} < \sqrt{2} - 1$. Let $x, \hat{x} \in \mathbb{R}^m$ be given and $h = \hat{x} - x$. Define $\Lambda_0$ and $\Lambda_1$ as following:

\[ \Lambda_0: \text{index set corresponding to the } k \text{ largest entries of } x \text{ in magnitude}. \]

\[ \Lambda_1: \text{index set corresponding to the } k \text{ largest entries of } x \text{ in } \Lambda^c_0 \text{ in magnitude}. \]

Let $\Lambda = \Lambda_0 \cup \Lambda_1$. If $\|\hat{x}\|_1 \leq \|x\|_1$, then

\[ \|h\|_2 \leq 2(1 - (1 - \sqrt{2})) \delta_{2k} \sigma_k(x) \sqrt{k} + 2\|Ah\|_2 \cdot \frac{|h\|_2}{\|h\|_2}. \]

Note that in the above lemma, $x$ is not necessarily the optimal solution of the $\ell_1$ norm problem.
Main Result

Theorem

Suppose matrix $A$ satisfies the RIP of order $2k$ with $\delta_{2k} < \sqrt{2} - 1$. Let $x, \hat{x} \in \mathbb{R}^m$ be given and $h = \hat{x} - x$. Define $\Lambda_0$ and $\Lambda_1$ as following:

$\Lambda_0$: index set corresponding to the $k$ largest entries of $x$ in magnitude.

$\Lambda_1$: index set corresponding to the $k$ largest entries of $x$ in magnitude.

Let $\Lambda = \Lambda_0 \cup \Lambda_1$. If $\|\hat{x}\|_1 \leq \|x\|_1$, then $\|h\|_2 \leq c_0 \sigma_k(x) \sqrt{k} + c_1 |A^h|$. 

Note that in the above lemma, $x$ is not necessarily the optimal solution of the $\ell_1$ norm problem.
Main Result

**Theorem**
Suppose matrix $A$ satisfies the RIP of order $2k$ with $\delta_{2k} < \sqrt{2} - 1$. Let $x, \hat{x} \in \mathbb{R}^m$ be given and $h = \hat{x} - x$. Define $\Lambda_0$ and $\Lambda_1$ as following:

- $\Lambda_0$: index set corresponding to the $k$ largest entries of $x$ in magnitude.
- $\Lambda_1$: index set corresponding to the $k$ largest entries of $x$ in $\Lambda_0^c$ in magnitude.

Note that in the above lemma, $x$ is not necessarily the optimal solution of the $\ell_1$ norm problem.
**Main Result**

**Theorem**

Suppose matrix $A$ satisfies the RIP of order $2k$ with $\delta_{2k} < \sqrt{2} - 1$. Let $x, \hat{x} \in \mathbb{R}^m$ be given and $h = \hat{x} - x$. Define $\Lambda_0$ and $\Lambda_1$ as following:

- $\Lambda_0$: index set corresponding to the $k$ largest entries of $x$ in magnitude.
- $\Lambda_1$: index set corresponding to the $k$ largest entries of $x$ in $\Lambda_0^c$ in magnitude.

Let $\Lambda = \Lambda_0 \cup \Lambda_1$. If $\|\hat{x}\|_1 \leq \|x\|_1$, then

$$
\|h\|_2 \leq c_0 \frac{\sigma_k(x)_1}{\sqrt{k}} + c_1 \left| \langle Ah_\Lambda, Ah \rangle \right| \|h_\Lambda\|_2
$$

where

$$
c_0 = \frac{2(1 - (1 - \sqrt{2}))\delta_{2k}}{1 - (1 + \sqrt{2})\delta_{2k}}, \quad c_1 = \frac{2}{1 - (1 + \sqrt{2})\delta_{2k}}
$$

(8)

Note that in the above lemma, $x$ is not necessarily the optimal solution of the $\ell_1$ norm problem.
Special case 1: Noiseless case

Theorem
Suppose matrix $A$ satisfies RIP of order $2k$ with $\delta_{2k} < \sqrt{2} - 1$. Let $y = Ax$ and $\hat{x}$ be the 1-norm solution to

$$\min_{z} \|z\|_1$$

$$s.t. Az = b$$

Therefore, $\|\hat{x}\|_1 \leq \|x\|_1$ and $h = \hat{x} - x \in \mathcal{N}(A)$ and we have

$$\|h\|_2 \leq \frac{c_0 \sigma_k(x)_1}{\sqrt{k}}$$

where $c_0$ is defined in (8).
Special Case 2: Sparse $x$

**Theorem**
Suppose matrix $A$ satisfies RIP of order $2k$ with $\delta_{2k} < \sqrt{2} - 1$. Let $y = Ax + e$ with bounded noise $\|e\| \leq \epsilon$ and $\hat{x}$ be the solution to

$$
\min_z \|z\|_1 \\
\text{s.t.} \quad \|Az - b\|_2 \leq \epsilon
$$

Then we have

$$
\|h\|_2 = \|x - \hat{x}\|_2 \leq c_0 \frac{\sigma_k(x)_1}{\sqrt{k}} + c_2 \epsilon
$$

where $c_2 = 2c_1 \sqrt{1 + \delta_{2k}}$ and $c_0$, $c_1$ are defined in (8).
Theorem

Let $A$ be a $N \times M$ Gaussian matrix with $N < M$. For $\eta, \epsilon \in (0, 1)$, assume that

$$N \geq 2\eta^{-2} \left( k \ln(\epsilon M/k) + \ln(2\epsilon^{-1}) \right).$$

Then with probability at least $(1 - \epsilon)$ the restricted isometry constant $\delta_k$ of $\frac{1}{\sqrt{N}} A$ satisfies

$$\delta_k \leq 2 \left( 1 + \frac{1}{\sqrt{2 \ln(eM/k)}} \right) \eta + \left( 1 + \frac{1}{\sqrt{2 \ln(eM/k)}} \right)^2 \eta^2.$$
RIP of Random Matrices

Theorem

Let $A$ be a $N \times M$ Gaussian matrix with $N < M$. For $\eta, \epsilon \in (0, 1)$, assume that

$$N \geq 2\eta^{-2} \left( k \ln(\epsilon M/k) + \ln(2\epsilon^{-1}) \right).$$

Then with probability at least $(1 - \epsilon)$ the restricted isometry constant $\delta_k$ of $\frac{1}{\sqrt{N}}A$ satisfies

$$\delta_k \leq 2 \left( 1 + \frac{1}{\sqrt{2\ln(eM/k)}} \right) \eta + \left( 1 + \frac{1}{\sqrt{2\ln(eM/k)}} \right)^2 \eta^2$$

Simpler variants

$$\delta_k \leq C\eta,$$ where $C \approx 6.3284$
RIP of Random Matrices

**Theorem**

Let $\mathbf{A}$ be a $N \times M$ Gaussian matrix with $N < M$. For $\eta, \epsilon \in (0, 1)$, assume that

$$N \geq 2\eta^{-2} \left( k \ln(\epsilon M/k) + \ln(2\epsilon^{-1}) \right).$$

Then with probability at least $(1 - \epsilon)$ the restricted isometry constant $\delta_k$ of $\frac{1}{\sqrt{N}} \mathbf{A}$ satisfies

$$\delta_k \leq 2 \left( 1 + \frac{1}{\sqrt{2 \ln(eM/k)}} \right) \eta + \left( 1 + \frac{1}{\sqrt{2 \ln(eM/k)}} \right)^2 \eta^2$$

Simpler variants

$$\delta_k \leq C\eta, \quad \text{where} \quad C \approx 6.3284$$

Alternately

$$N \geq \tilde{C}\delta^{-2} \left( k \ln(\epsilon M/k) + \ln(2\epsilon^{-1}) \right) \quad \text{with} \quad \tilde{C} \approx 80.098 \quad \text{implies} \quad \delta_k \leq \delta$$