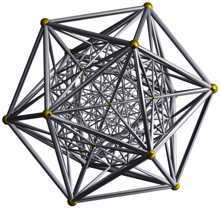
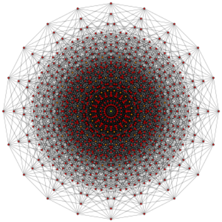
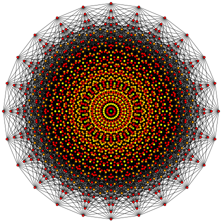
Computational Principles for High-dim Data Analysis

(Lecture Five)

Yi Ma

EECS Department, UC Berkeley

September 14, 2021

Convex Methods for Sparse Signal Recovery
(Towards Stronger Correctness Results)

1. Restricted Isometry Property (RIP)
2. Restricted Strong Convexity (RSC)
3. Success of ℓ^1 Minimization under RIP

"Algebra is but written geometry; geometry is but drawn algebra."
– Sophie Germain

From Incoherence to Isometry

Consider two columns $\boldsymbol{A}_{\mathsf{I}} = [\boldsymbol{a}_i \mid \boldsymbol{a}_j] \in \mathbb{R}^{m \times 2}$ of $\boldsymbol{A}$,

$$\boldsymbol{A}_{\mathsf{I}}^* \boldsymbol{A}_{\mathsf{I}} = \begin{bmatrix} 1 & \boldsymbol{a}_i^* \boldsymbol{a}_j \\ \boldsymbol{a}_j^* \boldsymbol{a}_i & 1 \end{bmatrix} \quad \mathbb{R}^{2 \times 2}. \tag{1}$$

If $|\boldsymbol{a}_i^* \boldsymbol{a}_j| \le \mu(\boldsymbol{A})$ is small, this matrix is well conditioned:

$$1 - \mu(\boldsymbol{A}) \;\le\; \sigma_{\min}(\boldsymbol{A}_{\mathsf{I}}^* \boldsymbol{A}_{\mathsf{I}}) \;\le\; \sigma_{\max}(\boldsymbol{A}_{\mathsf{I}}^* \boldsymbol{A}_{\mathsf{I}}) \;\le\; 1 + \mu(\boldsymbol{A}). \tag{2}$$

$\forall\, \mathsf{I}$ of size $\le k$,

$$1 - \underbrace{k\mu(\boldsymbol{A})}_{\delta} \;\le\; \sigma_{\min}(\boldsymbol{A}_{\mathsf{I}}^* \boldsymbol{A}_{\mathsf{I}}) \;\le\; \sigma_{\max}(\boldsymbol{A}_{\mathsf{I}}^* \boldsymbol{A}_{\mathsf{I}}) \;\le\; 1 + \underbrace{k\mu(\boldsymbol{A})}_{\delta}. \tag{3}$$

If $\delta = k\mu(\boldsymbol{A})$ is small, all $\sigma(\boldsymbol{A}_{\mathsf{I}}^* \boldsymbol{A}_{\mathsf{I}})$ are close to 1.

Restricted Isometry Property

Definition (Restricted Isometry Property)

The matrix $\boldsymbol{A}$ satisfies the *restricted isometry property (RIP)* of *order* k, with constant $\delta \in [0,1)$, if

$$\forall\, \boldsymbol{x} \;\; k\text{-sparse}, \quad (1-\delta)\,\|\boldsymbol{x}\|_2^2 \;\le\; \|\boldsymbol{A}\boldsymbol{x}\|_2^2 \;\le\; (1+\delta)\,\|\boldsymbol{x}\|_2^2\,. \tag{4}$$

The *order-k restricted isometry constant* $\delta_k(\boldsymbol{A})$ is the smallest number δ such that the above inequality holds.

Example of Gaussian Matrices: If $\boldsymbol{A}_{\mathsf{I}}$ is a large $m \times k$ $(k < m)$ matrix with entries independent $\mathcal{N}(0, 1/m)$,

$$\sigma_{min}(\boldsymbol{A}_{\mathsf{I}}^*\boldsymbol{A}_{\mathsf{I}}) \approx (\sqrt{1} - \sqrt{k/m})^2 \ge 1 - 2\sqrt{k/m},$$

$$\sigma_{max}(\boldsymbol{A}_{\mathsf{I}}^*\boldsymbol{A}_{\mathsf{I}}) \approx (\sqrt{1} + \sqrt{k/m})^2 \le 1 + 3\sqrt{k/m}.$$

Restricted Isometry Property

Upper and lower bounds[1] for RIP constants of random Gaussian matrix $\boldsymbol{A} \in \mathbb{R}^{m \times n}$. Consider proportional growth of the size of (k, m, n): $\rho = \frac{k}{m}, \gamma = \frac{m}{n}$.

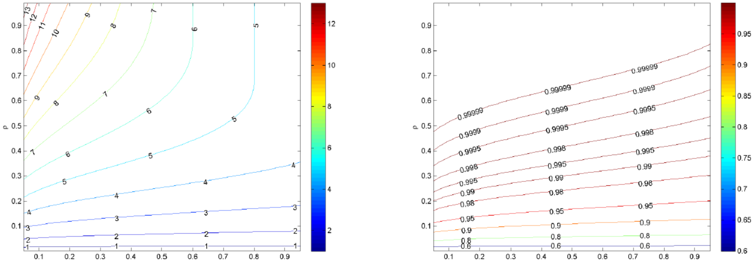


Figure: Left: $\delta_u = \lambda_{\max}(\rho, \gamma) - 1$; Right: $\delta_l = 1 - \lambda_{\min}(\rho, \gamma)$.

[1] *Improved Bounds on Restricted Isometry Constants for Gaussian Matrices,* B. Bah, J. Tanner, SIAM Journal on Matrix Analysis and Applications, 2010.

Restricted Isometry Property: Uniqueness

Theorem (ℓ^0 Recovery under RIP)

Suppose that $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}_o$*, with* $k = \|\boldsymbol{x}_o\|_0$*. If* $\delta_{2k}(\boldsymbol{A}) < 1$*, then* $\boldsymbol{x}_o$ *is the unique optimal solution to*

$$\min \|\boldsymbol{x}\|_0 \quad \text{s.t.} \quad \boldsymbol{A}\boldsymbol{x} = \boldsymbol{y}. \tag{5}$$

Proof.

Suppose on the contrary that there exists $\boldsymbol{x}' \neq \boldsymbol{x}_o$ with $\|\boldsymbol{x}'\|_0 \leq k$. Then $\boldsymbol{x}_o - \boldsymbol{x}' \in \operatorname{null}(\boldsymbol{A})$, and $\|\boldsymbol{x}_o - \boldsymbol{x}'\|_0 \leq 2k$. This implies that $\delta_{2k}(\boldsymbol{A}) \geq 1$, contradicting our assumption. □

Restricted Isometry Property: Correctness

Theorem (ℓ^1 Recovery under RIP)

Suppose that $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}_o$, with $k = \|\boldsymbol{x}_o\|_0$. If $\delta_{2k}(\boldsymbol{A}) < \sqrt{2} - 1$, then $\boldsymbol{x}_o$ is the unique optimal solution to

$$\min \|\boldsymbol{x}\|_1 \quad \text{s.t.} \quad \boldsymbol{A}\boldsymbol{x} = \boldsymbol{y}. \tag{6}$$

Some later developments:

- $\delta_{2k} < \sqrt{2} - 1 \approx 0.414$, Candes and Tao, 2006.
- $\delta_{2k} < 0.4531$, Fouchart and Lai, 2009.
- $\delta_{2k} < 0.472$, Cai, Wang, and Xu, 2009.
- $\delta_k < 0.307$, Cai, Wang, and Xu, 2010.[2]

---

[2] *New Bounds for Restricted Isometry Constants*, T. Cai, L. Wang, and G. Xu, IEEE Transactions on Information Theory, 56, 2010.

Restricted Isometry Property: Universality

Computing RIP constant $\delta_k(\boldsymbol{A})$ is in general NP-hard; and in fact, certifying a matrix is (k, δ)-RIP is also hard when $k \gg \sqrt{m}$.[3] **However:**

Theorem (RIP of Gaussian Matrices)

There exists a numerical constant $C > 0$ such that if $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ is a random matrix with entries independent $\mathcal{N}\left(0, \frac{1}{m}\right)$ random variables, with high probability, $\delta_k(\boldsymbol{A}) < \delta$, provided

$$m \geq Ck \log(n/k)/\delta^2. \tag{7}$$

Compare with Incoherence: With incoherence, we need $m \geq \Omega(k^2)$. Here this result allows (k, m, n) to scale proportionally:

$$m \geq \Omega(k).$$

[3] The Average-Case Time Complexity of Certifying the Restricted Isometry Property, Y. Ding, D. Kunisky, A. Wein, and A. Bandeira, https://arxiv.org/pdf/2005.11270.pdf.

ℓ^1 Recovery under RIP

How to prove ℓ^1 minimization succeeds under RIP?

Null Space Property

Given $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}_o$, try to recover $\boldsymbol{x}_o$ from

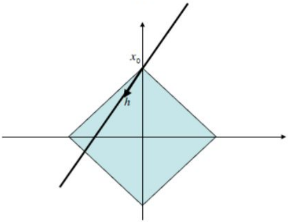
$$\min \|\boldsymbol{x}\|_1 \quad \text{subject to} \quad \boldsymbol{A}\boldsymbol{x} = \boldsymbol{y}. \tag{8}$$

Let $\boldsymbol{x}_{\ell^1}$ is the optimal solution. If $\boldsymbol{h} = \boldsymbol{x}_{\ell^1} - \boldsymbol{x}_o \neq \mathbf{0}$. Since $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}_o = \boldsymbol{A}\boldsymbol{x}_{\ell^1}$, we also have $\boldsymbol{A}\boldsymbol{h} = \mathbf{0}$. We must have

$$\begin{aligned} 0 &\geq \|\boldsymbol{x}_{\ell^1}\|_1 - \|\boldsymbol{x}_o\|_1 = \|\boldsymbol{x}_o + \boldsymbol{h}\|_1 - \|\boldsymbol{x}_o\|_1 \\ &\geq \|\boldsymbol{x}_o\|_1 - \|\boldsymbol{h}_{\mathsf{I}}\|_1 + \|\boldsymbol{h}_{\mathsf{I}^c}\|_1 - \|\boldsymbol{x}_o\|_1 \\ &= -\|\boldsymbol{h}_{\mathsf{I}}\|_1 + \|\boldsymbol{h}_{\mathsf{I}^c}\|_1. \end{aligned}$$

That is, we have

$$\|\boldsymbol{h}_{\mathsf{I}^c}\|_1 \leq \|\boldsymbol{h}_{\mathsf{I}}\|_1.$$



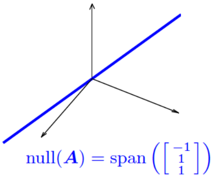
Null Space Property

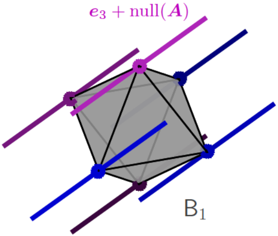
Definition (Null Space Property)

The matrix $\boldsymbol{A}$ satisfies the *null space property* of order k if for every $\boldsymbol{h} \in \mathrm{null}(\boldsymbol{A}) \setminus \{\boldsymbol{0}\}$ and every I of size at most k,

$$\|\boldsymbol{h}_{\mathsf{I}}\|_1 < \|\boldsymbol{h}_{\mathsf{I}^c}\|_1 . \tag{9}$$

Example: $\boldsymbol{A} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & -1 \end{bmatrix} \in \mathbb{R}^{2\times 3}$:





Null Space Property

Lemma (Success from Null Space Property)

Suppose that $\boldsymbol{A}$ satisfies the null space property of order k. Then for any $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}_o$, with $\|\boldsymbol{x}_o\|_0 \le k$, $\boldsymbol{x}_o$ is the unique optimal solution to the ℓ^1 problem

$$\min \|\boldsymbol{x}\|_1 \quad \text{s.t.} \quad \boldsymbol{A}\boldsymbol{x} = \boldsymbol{y}. \tag{10}$$

Proof.

Let $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}_o$, with $\|\boldsymbol{x}_o\|_0 \le k$, and let $\mathsf{I} = \mathrm{supp}(\boldsymbol{x}_o)$. Let $\hat{\boldsymbol{x}}_{\ell^1}$ be the optimal solution, so $\boldsymbol{h} = \hat{\boldsymbol{x}}_{\ell^1} - \boldsymbol{x}_o \in \mathrm{null}(\boldsymbol{A})$. If $\boldsymbol{h} \neq \boldsymbol{0}$, then

$$\|\hat{\boldsymbol{x}}_{\ell^1}\|_1 = \|\boldsymbol{x}_o + \boldsymbol{h}\|_1 \ge \|\boldsymbol{x}_o\|_1 - \|\boldsymbol{h}_{\mathsf{I}}\|_1 + \|\boldsymbol{h}_{\mathsf{I}^c}\|_1 > \|\boldsymbol{x}_o\|_1,$$

contradicting the optimality of $\hat{\boldsymbol{x}}_{\ell^1}$. □

No direction h in $\mathrm{null}(\boldsymbol{A})$ could further reduce $\|\boldsymbol{x}_o\|_1$.

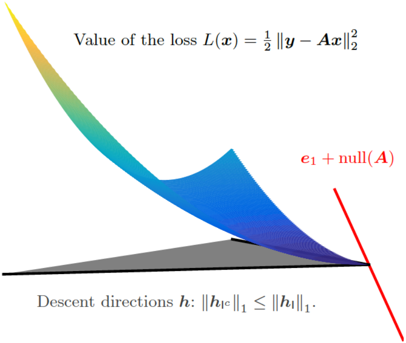
Restricted Strong Convexity Condition

The null space property is equivalent to:

$$\|\boldsymbol{A}\boldsymbol{h}\|_2^2 > 0 \quad \forall \boldsymbol{h} \;\; \|\boldsymbol{h}_{\mathsf{I}^c}\|_1 \leq \|\boldsymbol{h}_{\mathsf{I}}\|_1. \tag{11}$$

$\|\boldsymbol{A}\boldsymbol{h}\|_2^2$ must attain its minimum $\mu > 0$ on a compact set. The above is equivalent to:

$$\|\boldsymbol{A}\boldsymbol{h}\|_2^2 \geq \mu\|\boldsymbol{h}\|_2^2, \quad \forall \boldsymbol{h} \;\; \|\boldsymbol{h}_{\mathsf{I}^c}\|_1 \leq \|\boldsymbol{h}_{\mathsf{I}}\|_1. \tag{12}$$



Restricted Strong Convexity Condition

Definition (Restricted Strong Convexity)

The matrix $\boldsymbol{A}$ satisfies the *restricted strong convexity* (RSC) condition of order k, with parameters $\mu > 0$, $\alpha \geq 1$, if for every I of size at most k and for all nonzero $\boldsymbol{h}$ satisfying $\|\boldsymbol{h}_{\mathsf{I}^c}\|_1 \leq \alpha \|\boldsymbol{h}_{\mathsf{I}}\|_1$,

$$\|\boldsymbol{A}\boldsymbol{h}\|_2^2 \geq \mu \|\boldsymbol{h}\|_2^2. \tag{13}$$

Lemma (Success from RSC Condition)

Suppose that $\boldsymbol{A}$ satisfies the restricted strong convexity condition of order k with constant $\alpha \geq 1$, for some $\mu > 0$. Then for any $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}_o$, with $\|\boldsymbol{x}_o\|_0 \leq k$, $\boldsymbol{x}_o$ is the unique optimal solution to the ℓ^1 problem

$$\min \|\boldsymbol{x}\|_1 \quad \text{subject to} \quad \boldsymbol{A}\boldsymbol{x} = \boldsymbol{y}. \tag{14}$$

RIP Preserves Incoherence

Lemma (RIP Preserves Incoherence of Images of Sparse Vectors)

If $\boldsymbol{x}$, $\boldsymbol{z}$ *are vectors with disjoint support, and* $|\mathrm{supp}(\boldsymbol{x})| + |\mathrm{supp}(\boldsymbol{z})| \le k$, *then*

$$|\langle \boldsymbol{A}\boldsymbol{x}, \boldsymbol{A}\boldsymbol{z}\rangle| \;\le\; \delta_k(\boldsymbol{A}) \left\|\boldsymbol{x}\right\|_2 \left\|\boldsymbol{z}\right\|_2 . \tag{15}$$

Proof.

WLOG, $\|\boldsymbol{x}\|_2 = \|\boldsymbol{z}\|_2 = 1$. Notice that $\|\boldsymbol{p}+\boldsymbol{q}\|_2^2 - \|\boldsymbol{p}-\boldsymbol{q}\|_2^2 = 4\langle \boldsymbol{p}, \boldsymbol{q}\rangle$. Hence,

$$|\langle \boldsymbol{A}\boldsymbol{x}, \boldsymbol{A}\boldsymbol{z}\rangle| \;\le\; \frac{1}{4}\left|\|\boldsymbol{A}\boldsymbol{x}+\boldsymbol{A}\boldsymbol{z}\|_2^2 - \|\boldsymbol{A}\boldsymbol{x}-\boldsymbol{A}\boldsymbol{z}\|_2^2\right| \tag{16}$$

$$\le\; \frac{1}{4}\left|(1+\delta_k)\,\|\boldsymbol{x}+\boldsymbol{z}\|_2^2 - (1-\delta_k)\,\|\boldsymbol{x}-\boldsymbol{z}\|_2^2\right|. \tag{17}$$

Since $\boldsymbol{x}$ and $\boldsymbol{z}$ have disjoint support, $\|\boldsymbol{x}+\boldsymbol{z}\|_2^2 = \|\boldsymbol{x}-\boldsymbol{z}\|_2^2 = 2$. □

Bounds between Norms

Lemma (Bounds between Norms of Sparse Vectors)

For any vector z *with* $\|z\|_0 \leq k$, $\|z\|_1 \leq \sqrt{k}\,\|z\|_2$ *and* $\|z\|_2 \leq \sqrt{k}\,\|z\|_\infty$.

Proof.

First inequality: since x^2 is a convex function, we have:

$$\left(\frac{a_1 + a_2 + \cdots + a_k}{k}\right)^2 \leq \frac{a_1^2 + a_2^2 + \cdots + a_k^2}{k}. \tag{18}$$

Second inequality:

$$\frac{a_1^2 + a_2^2 + \cdots + a_k^2}{k} \leq \max_j \{a_j^2\}. \tag{19}$$

□

RIP Implies RSC

Theorem (RIP Implies RSC)

If a matrix $\boldsymbol{A}$ satisfies RIP with $\delta_{2k}(\boldsymbol{A}) < \frac{1}{1+\alpha\sqrt{2}}$, then $\boldsymbol{A}$ satisfies the RSC condition of order k with constant α.

Proof (A Sketch): We want to show:

$$\forall \boldsymbol{h} \in \mathbb{R}^n : \|\boldsymbol{h}_{\mathsf{I}^c}\|_1 \leq \alpha \cdot \|\boldsymbol{h}_{\mathsf{I}}\|_1,\ |\mathsf{I}| = k \implies \|\boldsymbol{A}\boldsymbol{h}\|_2 \geq \mu\|\boldsymbol{h}\|_2. \tag{20}$$

Partition the indices of the entries in $\boldsymbol{h}_{\mathsf{I}^c}$ based on their magnitudes:

J_1 indexes the k largest (in magnitude) elements of $\boldsymbol{h}_{\mathsf{I}^c}$,

J_2 indexes the k largest (in magnitude) elements of $\boldsymbol{h}_{(\mathsf{I}\cup\mathsf{J}_1)^c}$,

J_3 indexes the k largest (in magnitude) elements of $\boldsymbol{h}_{(\mathsf{I}\cup\mathsf{J}_1\cup\mathsf{J}_2)^c}$,

$$\vdots$$

Then $\forall\ i \geq 1,\ \|\boldsymbol{h}_{\mathsf{J}_i}\|_1 \ \geq\ k \cdot \left\|\boldsymbol{h}_{\mathsf{J}_{i+1}}\right\|_\infty$.

RIP Implies RSC

Proof (Continued):

Step 1: Since $\boldsymbol{h}_{\mathsf{I}}$ and $\boldsymbol{h}_{\mathsf{J}_1}$ likely have the largest entries, first try to show:

$$\|\boldsymbol{A}\boldsymbol{h}\|_2 \geq C \left\|\boldsymbol{h}_{\mathsf{I}\cup\mathsf{J}_1}\right\|_2, \quad \text{for some} \quad C > 0. \tag{21}$$

From $\boldsymbol{A}\boldsymbol{h}_{\mathsf{I}} + \boldsymbol{A}\boldsymbol{h}_{\mathsf{J}_1} = \boldsymbol{A}\boldsymbol{h} - \boldsymbol{A}\boldsymbol{h}_{\mathsf{J}_2} - \boldsymbol{A}\boldsymbol{h}_{\mathsf{J}_3} - \cdots$, we have:

$$\begin{aligned}
(1-\delta_{2k})\|\boldsymbol{h}_{\mathsf{I}\cup\mathsf{J}_1}\|_2^2 \;&\leq\; \|\boldsymbol{A}\boldsymbol{h}_{\mathsf{I}\cup\mathsf{J}_1}\|_2^2 \\
&= \langle \boldsymbol{A}\boldsymbol{h}_{\mathsf{I}} + \boldsymbol{A}\boldsymbol{h}_{\mathsf{J}_1}, -\boldsymbol{A}\boldsymbol{h}_{\mathsf{J}_2} - \boldsymbol{A}\boldsymbol{h}_{\mathsf{J}_3} - \cdots \rangle + \langle \boldsymbol{A}\boldsymbol{h}_{\mathsf{I}} + \boldsymbol{A}\boldsymbol{h}_{\mathsf{J}_1}, \boldsymbol{A}\boldsymbol{h} \rangle \\
&\leq \sum_{j=2}^{\infty} \left(\left|\langle \boldsymbol{A}\boldsymbol{h}_{\mathsf{I}}, \boldsymbol{A}\boldsymbol{h}_{\mathsf{J}_j}\rangle\right| + \left|\langle \boldsymbol{A}\boldsymbol{h}_{\mathsf{J}_1}, \boldsymbol{A}\boldsymbol{h}_{\mathsf{J}_j}\rangle\right| \right) + \|\boldsymbol{A}\boldsymbol{h}_{\mathsf{I}\cup\mathsf{J}_1}\|_2 \|\boldsymbol{A}\boldsymbol{h}\|_2 \\
&\leq \delta_{2k}(\|\boldsymbol{h}_{\mathsf{I}}\|_2 + \|\boldsymbol{h}_{\mathsf{J}_1}\|_2) \sum_{j=2}^{\infty} \left\|\boldsymbol{h}_{\mathsf{J}_j}\right\|_2 + (1+\delta_{2k})^{1/2} \|\boldsymbol{h}_{\mathsf{I}\cup\mathsf{J}_1}\|_2 \|\boldsymbol{A}\boldsymbol{h}\|_2 \\
&\leq \delta_{2k}\sqrt{2} \left\|\boldsymbol{h}_{\mathsf{I}\cup\mathsf{J}_1}\right\|_2 \left\|\boldsymbol{h}_{\mathsf{I}^c}\right\|_1 / \sqrt{k} + (1+\delta_{2k})^{1/2} \|\boldsymbol{h}_{\mathsf{I}\cup\mathsf{J}_1}\|_2 \|\boldsymbol{A}\boldsymbol{h}\|_2.
\end{aligned} \tag{22}$$

RIP Implies RSC

Proof (Continued):

From the restricted cone condition, we have

$$\|\boldsymbol{h}_{\mathsf{I}^c}\|_1 \leq \alpha \|\boldsymbol{h}_{\mathsf{I}}\|_1 \leq \alpha\sqrt{k}\|\boldsymbol{h}_{\mathsf{I}}\|_2 \leq \alpha\sqrt{k}\|\boldsymbol{h}_{\mathsf{I}\cup\mathsf{J}_1}\|_2. \tag{23}$$

This gives:

$$\|\boldsymbol{A}\boldsymbol{h}\|_2 \geq \frac{1-\delta_{2k}(1+\alpha\sqrt{2})}{(1+\delta_{2k})^{1/2}}\|\boldsymbol{h}_{\mathsf{I}\cup\mathsf{J}_1}\|_2. \tag{24}$$

Step 2: Try to show:

$$\|\boldsymbol{h}_{\mathsf{I}\cup\mathsf{J}_1}\|_2^2 \geq C'\|\boldsymbol{h}\|_2^2, \quad \text{for some} \quad C' > 0. \tag{25}$$

Since the i-th element of $\boldsymbol{h}_{(\mathsf{I}\cup\mathsf{J}_1)^c}$ is no larger than the mean of the first i elements of $\boldsymbol{h}_{\mathsf{I}^c}$, we have

$$|\boldsymbol{h}_{(\mathsf{I}\cup\mathsf{J}_1)^c}|_{(i)} \leq \|\boldsymbol{h}_{\mathsf{I}^c}\|_1/i. \tag{26}$$

RIP Implies RSC

Proof (Continued):

Combining with the restriction (20), we have

$$\begin{aligned} \|\boldsymbol{h}_{(\mathsf{I}\cup\mathsf{J}_1)^c}\|_2^2 &\leq \|\boldsymbol{h}_{\mathsf{I}^c}\|_1^2 \sum_{i=k+1}^{\infty} \frac{1}{i^2} \leq \frac{\|\boldsymbol{h}_{\mathsf{I}^c}\|_1^2}{k} \\ &\leq \frac{\alpha^2\|\boldsymbol{h}_{\mathsf{I}}\|_1^2}{k} \leq \alpha^2\|\boldsymbol{h}_{\mathsf{I}}\|_2^2 \leq \alpha^2\|\boldsymbol{h}_{\mathsf{I}\cup\mathsf{J}_1}\|_2^2. \end{aligned}$$

So we have

$$\|\boldsymbol{h}\|_2^2 \leq (1+\alpha^2)\|\boldsymbol{h}_{\mathsf{I}\cup\mathsf{J}_1}\|_2^2. \tag{27}$$

Finally: Combine the results:

$$\|\boldsymbol{A}\boldsymbol{h}\|_2 \geq \frac{1-\delta_{2k}(1+\alpha\sqrt{2})}{(1+\delta_{2k})^{1/2}\sqrt{1+\alpha^2}} \|\boldsymbol{h}\|_2 \,. \tag{28}$$

RIP for ℓ^1 Minimization

To prove the RIP Theorem: We need $\delta_{2k} < \frac{1}{1+\alpha\sqrt{2}}$ to ensure RSC hence the NSP.
For ℓ^1 minimization

$$\min \|\boldsymbol{x}\|_1 \quad \text{s.t.} \quad \boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}. \tag{29}$$

to succeed, we need

$$\forall \boldsymbol{h} \in \text{null}(\boldsymbol{A}) : \ \|\boldsymbol{h}_{\mathsf{I}^c}\|_1 \leq \|\boldsymbol{h}_{\mathsf{I}}\|_1,$$

hence $\alpha = 1$ and the associated RIP constant should be:

$$\delta_{2k} < \frac{1}{1+\sqrt{2}} = \sqrt{2} - 1.$$

Conclusions

Conditions when ℓ^1 minimization find the correct k-sparse solution:

$$\min \|\boldsymbol{x}\|_1 \quad \text{subject to} \quad \boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}.$$

- **Mutual Coherence:**

$$m = O(k^2).$$

- **Restricted Isometry:**

$$m = O\big(k \log(n/k)\big).$$

Next: what $m \times n$ matrix A has a small RIP constant δ_k?

Assignments

- Reading: Section 3.3 of Chapter 3.
- Programming Homework #1.