Making Do with Less: An Introduction to Compressed Sensing

Tanya Leise

Amherst College

March 5, 2012

Ingredients

Given

- Seven American Eagle half-ounce gold coins, which weigh 16.966 grams each. But one may be counterfeit...

How can we find the bad coin with the fewest weighings?

Ingredients

Given

- Seven American Eagle half-ounce gold coins, which weigh 16.966 grams each. But one may be counterfeit...
- An electronic balance.

How can we find the bad coin with the fewest weighings?

Pooling Design Approach

- Label the coins with numbers 1 through 7.
- For the first weighing place coins 1, 3, 5, and 7 on the scale.
- For the second weighing use coins 2, 3, 6, and 7.
- For the third weighing use coins 4, 5, 6, and 7.

Combinatorial Group Test

$$\mathbf{\Phi} = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}$$

(Binary encoding)

- Label the coins with numbers 1 through 7.
- For the first weighing place coins 1, 3, 5, and 7 on the scale.
- For the second weighing use coins 2, 3, 6, and 7.
- For the third weighing use coins 4, 5, 6, and 7.

Combinatorial Group Test

$$\mathbf{\Phi} = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}$$

(Binary encoding)

- Pattern of weights that deviate from the expected value reveals the bad coin:
- If only the 1st weighing is "off," then coin 1 is the counterfeit.
- If both the 1st and 2nd weighings are "off," then coin 3 is the counterfeit.

General strategy

- If there are $k \geq 2$ bad coins (of unknown weights) the problem gets a lot harder.
- We need a general strategy: how should we choose the subsets?
- Randomly!

Random Subsets

- Suppose we now have 100 coins. Number them 1 to 100.
- Choose a random subset where each coin is included with probability $\frac{1}{2}$ (flip each coin–if heads, put it in the subset; if tails, leave it out).
- Weigh the subset and record.
- Repeat the previous two steps a total of n times, say, $n = 25$.

Claim: If there are only a few bad coins (say $k \leq 3$) and we weigh $n = 25$ subsets, we almost certainly have enough information to identify the bad ones.

Random Subsets

- Let X_i denote the true DEVIATION of the ith coin's weight from nominal (16.966 grams).
- We expect $X_i = 0$ for most i.
- Suppose the coins in the first random subset have indices $i = i_1, i_2, \ldots, i_m$ ($m \approx 50$), and this subset weighs a_1 grams. Then
$$X_{i_1} + X_{i_2} + \cdots + X_{i_m} = a_1 - 16.966m.$$
- A similar equation holds for each of the other random subsets that we choose.

Random Subsets

Let x_i be our estimate of the true value X_i.

We end up with $n = 25$ linear equations and $N = 100$ unknowns (the mass deviations X_i), of the form

$$\mathbf{\Phi x} = \mathbf{b}.$$

$\mathbf{\Phi}$ is the *sensing matrix*, a $0 - 1$ matrix with $\Phi_{ij} = 1$ if the jth coin was included in the ith subset, and

b_j is the mass of the jth subset minus 16.966 times the number of coins in the subset.

Random Subsets Example

For example, with $N = 10$ coins and $n = 3$ weighings we might have

$$\mathbf{\Phi} = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

if the first weighing involves coins $1, 4, 5, 7$, the second involves coins $3, 4, 5, 6, 8$, the third involves coins $2, 3, 4, 5, 9, 10$.

Random Subsets Example

In the coin problem with $N = 100$, suppose that the true values are $X_{13} = -0.3, X_{37} = 0.44$, and $X_{71} = -0.33$, and all other $X_i = 0$ (but we don't know this – the goal is to deduce this information).

Random Subsets Example

In the coin problem with $N = 100$, suppose that the true values are $X_{13} = -0.3, X_{37} = 0.44$, and $X_{71} = -0.33$, and all other $X_i = 0$ (but we don't know this – the goal is to deduce this information).

To estimate the coin masses, we choose 25 random subsets, each of size about 50, and weigh each subset. We obtain a system

$$\mathbf{\Phi x} = \mathbf{b}.$$

Random Subsets Example

In the coin problem with $N = 100$, suppose that the true values are $X_{13} = -0.3, X_{37} = 0.44$, and $X_{71} = -0.33$, and all other $X_i = 0$ (but we don't know this – the goal is to deduce this information).

To estimate the coin masses, we choose 25 random subsets, each of size about 50, and weigh each subset. We obtain a system

$$\mathbf{\Phi x} = \mathbf{b}.$$

The matrix $\mathbf{\Phi}$ is 25 rows by 100 columns—there will be at least 75 free variables. Solving for the x_i in any meaningful way looks hopeless.

Minimum ℓ^2 Norm Solution

A traditional approach to "solving" a consistent under determined system of equations $\mathbf{Mx} = \mathbf{b}$ is to choose that (unique) vector $\mathbf{x}^*$ with the properties that

Minimum ℓ^2 Norm Solution

A traditional approach to “solving” a consistent under determined system of equations $\mathbf{M}\mathbf{x} = \mathbf{b}$ is to choose that (unique) vector $\mathbf{x}^*$ with the properties that

1. $\mathbf{M}\mathbf{x}^* = \mathbf{b}$ ($\mathbf{x}^*$ actually satisfies the equations);

Minimum ℓ^2 Norm Solution

A traditional approach to "solving" a consistent under determined system of equations $\mathbf{Mx} = \mathbf{b}$ is to choose that (unique) vector $\mathbf{x}^*$ with the properties that

1. $\mathbf{Mx}^* = \mathbf{b}$ ($\mathbf{x}^*$ actually satisfies the equations);
2. $\mathbf{x}^*$ has minimum ℓ^2 norm, that is, if $\mathbf{x}^{**} \neq \mathbf{x}^*$ satisfies $\mathbf{Mx}^{**} = \mathbf{b}$ then $\|\mathbf{x}^*\|_2 < \|\mathbf{x}^{**}\|_2$ where

$$\|\mathbf{x}\|_2 = \left(\sum_i |x_i|^2 \right)^{1/2}$$

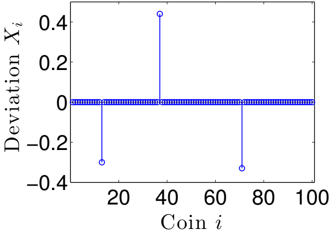
is the usual Euclidean norm.

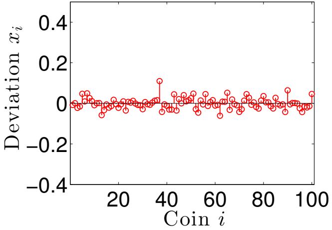
Minimum ℓ^2 Norm Solution

Computing the ℓ^2 minimum solution is a standard calculus/matrix algebra problem.

Minimum ℓ^2 Norm Solution

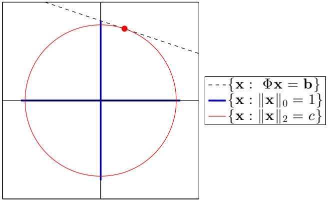
Computing the ℓ^2 minimum solution is a standard calculus/matrix algebra problem. In the present case we find





Minimum ℓ^2 Norm Solution

Geometric intuition:



The point of tangency will usually have many nonzero components, so this approach rarely leads to a sparse solution.

Minimum ℓ^0 Norm Solution

Since we suspect the solution $\mathbf{X}$ is sparse we really ought to seek a vector $\mathbf{x}$ such that

- $\mathbf{\Phi x} = \mathbf{b}$ and
- $\mathbf{x}$ has as few nonzero components as possible, that is, $\mathbf{x}$ minimizes the quantity

$$\|\mathbf{x}\|_0 = \#\{x_i : x_i \neq 0\}$$

Unfortunately this optimization problem is too difficult to solve in any practical way.

Brute Force ℓ^0 Solution

An expensive strategy for finding a sparse solution to $\mathbf{\Phi x} = \mathbf{b}$ ($\mathbf{\Phi}$ dimensions $n \times N$):

- Look for a 1-sparse solution (try $\mathbf{x} = (0, \ldots, 0, x_i, 0, \ldots, 0)$ for each i, need N linear solves). If that doesn't work

Brute Force ℓ^0 Solution

An expensive strategy for finding a sparse solution to $\mathbf{\Phi x} = \mathbf{b}$ ($\mathbf{\Phi}$ dimensions $n \times N$):

- Look for a 1-sparse solution (try $\mathbf{x} = (0, \ldots, 0, x_i, 0, \ldots, 0)$ for each i, need N linear solves). If that doesn't work
- Try 2-sparse solutions $\mathbf{x} = (\mathbf{0}, x_{i_1}, \mathbf{0}, x_{i_2}, \mathbf{0})$, for each (i_1, i_2) pair, need $N(N-1)/2$ linear solves. If that doesn't work

Brute Force ℓ^0 Solution

An expensive strategy for finding a sparse solution to $\mathbf{\Phi x} = \mathbf{b}$ ($\mathbf{\Phi}$ dimensions $n \times N$):

- Look for a 1-sparse solution (try $\mathbf{x} = (0, \ldots, 0, x_i, 0, \ldots, 0)$ for each i, need N linear solves). If that doesn't work
- Try 2-sparse solutions $\mathbf{x} = (\mathbf{0}, x_{i_1}, \mathbf{0}, x_{i_2}, \mathbf{0})$, for each (i_1, i_2) pair, need $N(N-1)/2$ linear solves. If that doesn't work
- Try 3-sparse solutions, 4-sparse, k-sparse, etc.

Brute Force ℓ^0 Solution

An expensive strategy for finding a sparse solution to $\mathbf{\Phi x} = \mathbf{b}$ ($\mathbf{\Phi}$ dimensions $n \times N$):

- Look for a 1-sparse solution (try $\mathbf{x} = (0, \ldots, 0, x_i, 0, \ldots, 0)$ for each i, need N linear solves). If that doesn't work
- Try 2-sparse solutions $\mathbf{x} = (\mathbf{0}, x_{i_1}, \mathbf{0}, x_{i_2}, \mathbf{0})$, for each (i_1, i_2) pair, need $N(N-1)/2$ linear solves. If that doesn't work
- Try 3-sparse solutions, 4-sparse, k-sparse, etc.

If N and k are of any significant size, this is completely hopeless!

Minimum ℓ^1 Norm Solution

A "compromise": Seek an $\mathbf{x}$ that satisfies $\mathbf{\Phi x} = \mathbf{b}$ and minimizes the ℓ^1 norm

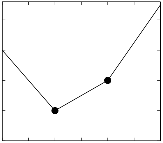
$$\|\mathbf{x}\|_1 = \sum_{i=1}^{N} |x_i|.$$

Minimum ℓ^1 Norm Solution

A "compromise": Seek an $\mathbf{x}$ that satisfies $\mathbf{\Phi x} = \mathbf{b}$ and minimizes the ℓ^1 norm

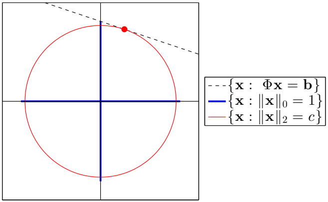
$$\|\mathbf{x}\|_1 = \sum_{i=1}^{N} |x_i|.$$

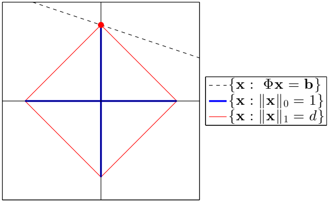
This might look hard since $|x|$ isn't differentiable, but it can be converted into a standard "easy" linear programming problem.



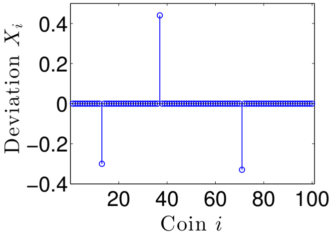
Minimum ℓ^1 Norm Solution

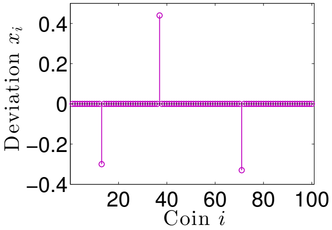
Geometric intuition:





Minimum ℓ^1 Norm Solution

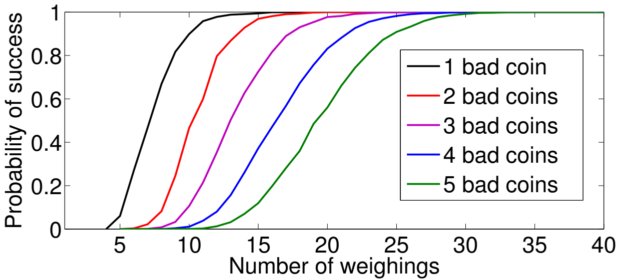




The ℓ^1 minimization recovers **X** *exactly!*

Probability of ℓ^1 Success

The more weighings we do (for any fixed number of defective coins) the better chance of success ℓ^1 minimization has.



Main Points of This Talk

- If $\mathbf{x}^*$ is a sparse (mostly zero) vector, $\mathbf{\Phi}$ a random matrix, and $\mathbf{\Phi x}^* = \mathbf{b}$ then $\mathbf{x}^*$ is almost certainly the only sparse solution to $\mathbf{\Phi x} = \mathbf{b}$, and
- With high probability, minimizing $\|\mathbf{x}\|_1$ subject to $\mathbf{\Phi x} = \mathbf{b}$ will recover $\mathbf{x} = \mathbf{x}^*$ *exactly*.

Bit of History

- The idea of recovering sparse signals using ℓ^1 minimization appeared in the mid 1980s, in signal processing and inverse problems, and were extended by Donoho et al. around 2002.
- A more comprehensive framework was developed around 2006 with landmark papers by Candes, Romberg, Donoho, Tao.
- Compressed sensing is now a very hot area in mathematics, statistics, computer science, and engineering.

Some Terminology

A vector $\mathbf{x} \in \mathbb{R}^N$ is *k-sparse* if it has at most k non-zero entries:

1-sparse:

$$\mathbf{x} = (\mathbf{0}, x_{i_1}, \mathbf{0})$$

2-sparse:

$$\mathbf{x} = (\mathbf{0}, x_{i_1}, \mathbf{0}, x_{i_2}, \mathbf{0})$$

3-sparse:

$$\mathbf{x} = (\mathbf{0}, x_{i_1}, \mathbf{0}, x_{i_2}, \mathbf{0}, x_{i_3}, \mathbf{0})$$

k-sparse:

$$\mathbf{x} = (\mathbf{0}, x_{i_1}, \mathbf{0}, x_{i_2}, \mathbf{0}, x_{i_3}, \ldots, x_{i_k}, \mathbf{0})$$

Uniqueness of Sparse Solutions

Let $\mathbf{\Phi}$ be an $n \times N$ matrix and suppose $\mathbf{\Phi x} = \mathbf{b}$ has a k-sparse solution $\mathbf{X}$. We want conditions under which $\mathbf{X}$ is unique ($\mathbf{x} = \mathbf{X}$ is the ONLY k-sparse solution).

Uniqueness of Sparse Solutions

Let $\mathbf{\Phi}$ be an $n \times N$ matrix and suppose $\mathbf{\Phi x} = \mathbf{b}$ has a k-sparse solution $\mathbf{X}$. We want conditions under which $\mathbf{X}$ is unique ($\mathbf{x} = \mathbf{X}$ is the ONLY k-sparse solution).

Observation: If there IS another k-sparse solution $\tilde{\mathbf{X}}$ then $\mathbf{w} = \mathbf{X} - \tilde{\mathbf{X}}$ satisfies $\mathbf{\Phi w} = \mathbf{0}$, so $\mathbf{w} \neq \mathbf{0}$ is in the null space $\mathcal{N}(\mathbf{\Phi})$ of $\mathbf{\Phi}$.

Sparse Vector Arithmetic

Fact: If $\mathbf{v}$ and $\mathbf{w}$ are k-sparse, then $\mathbf{v} \pm \mathbf{w}$ is $2k$-sparse, e.g., given 3-sparse vectors

$$\mathbf{v} = (0, 0, 0, v_4, 0, v_6, 0, 0, v_8, 0, 0)$$

and

$$\mathbf{w} = (w_1, 0, w_3, 0, 0, 0, 0, 0, 0, w_{10})$$

Sparse Vector Arithmetic

Fact: If $\mathbf{v}$ and $\mathbf{w}$ are k-sparse, then $\mathbf{v} \pm \mathbf{w}$ is $2k$-sparse, e.g., given 3-sparse vectors

$$\mathbf{v} = (0, 0, 0, v_4, 0, v_6, 0, 0, v_8, 0, 0)$$

and

$$\mathbf{w} = (w_1, 0, w_3, 0, 0, 0, 0, 0, 0, w_{10})$$

the sum

$$\mathbf{v} + \mathbf{w} = (w_1, 0, w_3, v_4, 0, v_6, 0, v_8, 0, w_{10})$$

is 6-sparse (maybe sparser).

The Nullspace of $\mathbf{\Phi}$

So if $\mathbf{X}$ and $\tilde{\mathbf{X}}$ are distinct k-sparse solutions to $\mathbf{\Phi x} = \mathbf{b}$, the vector $\mathbf{w} = \mathbf{X} - \tilde{\mathbf{X}} \neq \mathbf{0}$ is in the nullspace $\mathcal{N}(\mathbf{\Phi})$ and $\mathbf{w}$ is $2k$-sparse.

The Nullspace of $\mathbf{\Phi}$

So if $\mathbf{X}$ and $\tilde{\mathbf{X}}$ are distinct k-sparse solutions to $\mathbf{\Phi x} = \mathbf{b}$, the vector $\mathbf{w} = \mathbf{X} - \tilde{\mathbf{X}} \neq \mathbf{0}$ is in the nullspace $\mathcal{N}(\mathbf{\Phi})$ and $\mathbf{w}$ is $2k$-sparse.

Thus if k-sparse solutions to $\mathbf{\Phi x} = \mathbf{b}$ are not unique, $\mathcal{N}(\mathbf{\Phi})$ must contain non-zero $2k$-sparse vectors. So...

The Nullspace of $\mathbf{\Phi}$

So if $\mathbf{X}$ and $\tilde{\mathbf{X}}$ are distinct k-sparse solutions to $\mathbf{\Phi x} = \mathbf{b}$, the vector $\mathbf{w} = \mathbf{X} - \tilde{\mathbf{X}} \neq \mathbf{0}$ is in the nullspace $\mathcal{N}(\mathbf{\Phi})$ and $\mathbf{w}$ is $2k$-sparse.

Thus if k-sparse solutions to $\mathbf{\Phi x} = \mathbf{b}$ are not unique, $\mathcal{N}(\mathbf{\Phi})$ must contain non-zero $2k$-sparse vectors. So...

If we construct $\mathbf{\Phi}$ so $\mathcal{N}(\mathbf{\Phi})$ does NOT contain $2k$-sparse vectors, we can rest assured any k-sparse solution we find is the "right one."

Nullspace Example

Let

$$\mathbf{\Phi} = \begin{bmatrix} 1 & 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 0 & 1/\sqrt{2} & 1 & 1/\sqrt{2} \end{bmatrix}.$$

Nullspace Example

Let

$$\mathbf{\Phi} = \begin{bmatrix} 1 & 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 0 & 1/\sqrt{2} & 1 & 1/\sqrt{2} \end{bmatrix}.$$

There are no nonzero 2-sparse vectors in $\mathcal{N}(\mathbf{\Phi})$. For example, if $\mathbf{x} = (x_1, x_2, 0, 0)$ then

$$\mathbf{\Phi x} = x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

implies $x_1 = x_2 = 0$. The same goes for other possible 2-sparse vectors.

Nullspace Example

Let

$$\mathbf{\Phi} = \begin{bmatrix} 1 & 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 0 & 1/\sqrt{2} & 1 & 1/\sqrt{2} \end{bmatrix}.$$

There are no nonzero 2-sparse vectors in $\mathcal{N}(\mathbf{\Phi})$. For example, if $\mathbf{x} = (x_1, x_2, 0, 0)$ then

$$\mathbf{\Phi x} = x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

implies $x_1 = x_2 = 0$. The same goes for other possible 2-sparse vectors.

Any 1-sparse solution to $\mathbf{\Phi x} = \mathbf{b}$ is thus unique.

The Restricted Isometry Property

Instead of requiring that $\mathbf{\Phi x} \neq \mathbf{0}$ for all $2k$-sparse vectors $\mathbf{x}$ (an algebraic condition), require that:

The Restricted Isometry Property

Instead of requiring that $\mathbf{\Phi x} \neq \mathbf{0}$ for all $2k$-sparse vectors $\mathbf{x}$ (an algebraic condition), require that:

For some $c_1 > 0$ and all $2k$-sparse vectors we have

$$c_1\|\mathbf{x}\|_2 \leq \|\mathbf{\Phi x}\|_2.$$

The Restricted Isometry Property

Instead of requiring that $\mathbf{\Phi x} \neq \mathbf{0}$ for all $2k$-sparse vectors $\mathbf{x}$ (an algebraic condition), require that:

For some $c_1 > 0$ and all $2k$-sparse vectors we have

$$c_1 \|\mathbf{x}\|_2 \leq \|\mathbf{\Phi x}\|_2.$$

This inequality is more analytic in nature.

The Restricted Isometry Property

Instead of requiring that $\mathbf{\Phi x} \neq \mathbf{0}$ for all $2k$-sparse vectors $\mathbf{x}$ (an algebraic condition), require that:

For some $c_1 > 0$ and all $2k$-sparse vectors we have

$$c_1 \|\mathbf{x}\|_2 \leq \|\mathbf{\Phi x}\|_2.$$

This inequality is more analytic in nature.

Equivalently, we can state this as

$$c_1 \leq \|\mathbf{\Phi u}\|_2.$$

for all $2k$-sparse unit vectors.

The Restricted Isometry Property II

The inequality $\|\mathbf{\Phi u}\|_2 \leq c_2$ for some c_2 is automatic (because $\mathbf{u} \to \|\mathbf{\Phi u}\|_2$ is continuous and the unit ball is compact).

The Restricted Isometry Property II

The inequality $\|\mathbf{\Phi u}\|_2 \leq c_2$ for some c_2 is automatic (because $\mathbf{u} \to \|\mathbf{\Phi u}\|_2$ is continuous and the unit ball is compact).

We'll throw it in anyway and require $\mathbf{\Phi}$ to have the property that

$$c_1 \leq \|\mathbf{\Phi u}\|_2 \leq c_2.$$

for all $2k$-sparse unit vectors. This is one version of the *restricted isometry property* (RIP).

The Restricted Isometry Property II

The inequality $\|\mathbf{\Phi u}\|_2 \leq c_2$ for some c_2 is automatic (because $\mathbf{u} \to \|\mathbf{\Phi u}\|_2$ is continuous and the unit ball is compact).

We'll throw it in anyway and require $\mathbf{\Phi}$ to have the property that

$$c_1 \leq \|\mathbf{\Phi u}\|_2 \leq c_2.$$

for all $2k$-sparse unit vectors. This is one version of the *restricted isometry property* (RIP).

If RIP of order $2k$ holds, any k-sparse solution to $\mathbf{\Phi x} = \mathbf{b}$ is unique, and so (in principle) can be found.

The Restricted Isometry Property III

Observation: Multiplying $\mathbf{\Phi}$ by a constant doesn't change solvability of $\mathbf{\Phi x} = \mathbf{b}$, so

The Restricted Isometry Property III

Observation: Multiplying $\mathbf{\Phi}$ by a constant doesn't change solvability of $\mathbf{\Phi x} = \mathbf{b}$, so multiply

$$c_1 \leq \|\mathbf{\Phi u}\|_2 \leq c_2.$$

through by $2/(c_1 + c_2)$, and redefine $\mathbf{\Phi} = 2\mathbf{\Phi}/(c_1 + c_2)$.

The Restricted Isometry Property III

Observation: Multiplying $\mathbf{\Phi}$ by a constant doesn't change solvability of $\mathbf{\Phi x} = \mathbf{b}$, so multiply

$$c_1 \leq \|\mathbf{\Phi u}\|_2 \leq c_2.$$

through by $2/(c_1 + c_2)$, and redefine $\mathbf{\Phi} = 2\mathbf{\Phi}/(c_1 + c_2)$.

RIP becomes

$$(1 - \delta) \leq \|\mathbf{\Phi u}\|_2 \leq (1 + \delta).$$

with $\delta = (c_2 - c_1)/(c_2 + c_1)$. Note $0 \leq \delta < 1$. The closer δ is to zero, the better (to be explained...)

When Does RIP Hold?

- Verifying that RIP of any order holds for a specific $\mathbf{\Phi}$ is difficult, but...

When Does RIP Hold?

- Verifying that RIP of any order holds for a specific $\mathbf{\Phi}$ is difficult, but...
- Showing $(1-\delta) \leq \|\mathbf{\Phi u}\|_2 \leq (1+\delta)$ holds with high probability for many types of randomly generated $\mathbf{\Phi}$ is fairly easy.

When Does RIP Hold?

- Verifying that RIP of any order holds for a specific $\mathbf{\Phi}$ is difficult, but...
- Showing $(1-\delta) \leq \|\mathbf{\Phi u}\|_2 \leq (1+\delta)$ holds with high probability for many types of randomly generated $\mathbf{\Phi}$ is fairly easy.
- If $\mathbf{\Phi}$ is $n \times N$ with entries $\phi_{ij} = N(0, 1/n)$ then RIP of order k holds with probability $1-\epsilon$ if

$$n \geq C(\delta, \epsilon) k \ln(N/k).$$

Tools in the Proof

The proof is elementary and involves only

- Sum of normals is normal, mean of sum is sum of means, ditto variances.
- Elementary probability, e.g., the "union" bound $P(E_1 \cup E_2) \leq P(E_1) + P(E_2)$.
- Simple analysis, e.g., estimates for elementary integrals.

Summary

- If $\mathbf{\Phi}$ is an $n \times N$ random matrix ($n << N$), any k-sparse solution to

$$\mathbf{\Phi x} = \mathbf{b}$$

  is almost certainly unique, if k, n, and N stand in the right relation.
- The RIP condition also guarantees that ℓ^1 minimization will find the desired sparse solution.

Some References

- K. Bryan and T. Leise, *Making do with less: an introduction to compressed sensing*, SIAM Review, 55 (2013), 547-566.
- R. Baranuik, M. Davenport, R. DeVore, and M. Wakin, *A simple proof of the restricted isometry property for random matrices*, Constructive Approximation 28 (2008), 253-263.
- E. Candès and J. Romberg, ℓ_1-*MAGIC*, http://www.acm.caltech.edu/l1magic/ (accessed 5/21/11).
- E. Candès and T. Tao, *Decoding by linear programming*, IEEE Trans. Inf. Theory 51 (2005), 4203-4215.
- E. J. Candès and M.B. Wakin, *An introduction to compressive sampling*, IEEE Signal Processing Magazine 25 (2008), 21-30.