Overview of the Mathematics of Compressive Sensing

Simon Foucart

Reading Seminar on “Compressive Sensing, Extensions, and Applications”
Texas A&M University
15 October 2015
RIP for Random Matrices
Concentration Inequality

Let $A \in \mathbb{R}^{m \times N}$ be a random matrix with entries $a_{i,j} = g_{i,j} \sqrt{m}$ where the $g_{i,j}$ are independent $\mathcal{N}(0,1)$.

For a fixed $x \in \mathbb{R}^N$, note that $(Ax)_i = \sum_{j=1}^N a_{i,j} x_j$, hence $E((Ax)_i^2) = \sum x_j^2 V(a_{i,j}) = \|x\|_2^2 m$.

In fact, $\|Ax\|_2^2$ concentrates around its mean: for $t \in (0,1)$,

\[
P\left(\|Ax\|_2^2 - \|x\|_2^2 > t \|x\|_2^2\right) \leq 2 \exp\left(-ct^2 m\right).
\]
Concentration Inequality

Let $A \in \mathbb{R}^{m \times N}$ be a random matrix with entries

$$a_{i,j} = \frac{g_{i,j}}{\sqrt{m}}$$

where the $g_{i,j}$ are independent $\mathcal{N}(0, 1)$. 

In fact, $\|Ax\|_2^2$ concentrates around its mean: for $t \in (0, 1)$, 

$$(CI) \quad P\left(\|Ax\|_2^2 - \|x\|_2^2 > t\|x\|_2^2\right) \leq 2 \exp\left(-c t^2 m\right).$$
Concentration Inequality

Let $A \in \mathbb{R}^{m \times N}$ be a random matrix with entries $a_{i,j} = \frac{g_{i,j}}{\sqrt{m}}$ where the $g_{i,j}$ are independent $\mathcal{N}(0, 1)$.

For a fixed $x \in \mathbb{R}^N$, note that $(Ax)_i = \sum_{j=1}^N a_{i,j}x_j$, hence

$$E((Ax)_i^2) = \mathbb{V}(\sum_{j=1}^N a_{i,j}x_j) = \sum_{j=1}^N x_j^2 \mathbb{V}(a_{i,j}) = \frac{\|x\|_2^2}{m},$$
$$E(\|Ax\|_2^2) = \|x\|_2^2.$$
Concentration Inequality

- Let $A \in \mathbb{R}^{m \times N}$ be a random matrix with entries
  \[ a_{i,j} = \frac{g_{i,j}}{\sqrt{m}} \]
  where the $g_{i,j}$ are independent $\mathcal{N}(0, 1)$.

- For a fixed $x \in \mathbb{R}^N$, note that $(Ax)_i = \sum_{j=1}^{N} a_{i,j}x_j$, hence
  \[
  \mathbb{E}((Ax)_i^2) = \mathbb{V}(\sum a_{i,j}x_j) = \sum x_j^2 \mathbb{V}(a_{i,j}) = \frac{\|x\|_2^2}{m},
  \]
  \[
  \mathbb{E}(\|Ax\|_2^2) = \|x\|_2^2.
  \]

- In fact, $\|Ax\|_2^2$ concentrates around its mean: for $t \in (0, 1)$,
  \[
  (CI) \quad \mathbb{P}(\|Ax\|_2^2 - \|x\|_2^2 > t\|x\|_2^2) \leq 2 \exp(-ct^2m).
  \]
Covering Arguments

Suppose that the random matrix $A \in \mathbb{R}^{m \times N}$ satisfies (CI). Let $S \subseteq [N]$ with $\text{card}(S) = s$. Then

$$P(\|A^*S A_S - \text{Id}\|_2^2 > \delta) \leq 2 \exp(-c\delta^2 m)$$

provided $m \geq c'\delta^2 s$. The argument relies on the following fact:

A subset $U$ of the unit ball of $\mathbb{R}^k$ relative to a norm $\|\cdot\|$ has covering and separating numbers satisfying

$$N(U, \|\cdot\|, \rho) \leq S(U, \|\cdot\|, \rho) \leq (1 + 2\rho)^k.$$
Covering Arguments

Suppose that the random matrix $A \in \mathbb{R}^{m \times N}$ satisfies (CI).

Let $S \subseteq [N]$ with $\text{card}(S) = s$. Then $P(\|A^*SA - I\|_2 > \delta) \leq 2 \exp(-c\delta^2 m)$ provided $m \geq c'_\delta^2 s$. The argument relies on the following fact: A subset $U$ of the unit ball of $\mathbb{R}^k$ relative to a norm $\|\cdot\|$ has covering and separating numbers satisfying $N(U, \|\cdot\|, \rho) \leq S(U, \|\cdot\|, \rho) \leq (1 + 2\rho)^k$. 
Covering Arguments

Suppose that the random matrix $A \in \mathbb{R}^{m \times N}$ satisfies (CI). Let $S \subseteq [N]$ with $\text{card}(S) = s$. 

Covering Arguments

Suppose that the random matrix $A \in \mathbb{R}^{m \times N}$ satisfies (CI). Let $S \subseteq [N]$ with $\text{card}(S) = s$. Then

$$\mathbb{P}(\|A_S^*A_S - \text{Id}\|_{2\to2} > \delta) \leq 2 \exp(-c\delta^2 m)$$
Covering Arguments

Suppose that the random matrix $A \in \mathbb{R}^{m \times N}$ satisfies (CI). Let $S \subseteq [N]$ with $\text{card}(S) = s$. Then

$$
\mathbb{P}(\|A^* S A_S - \text{Id}\|_{2\to2} > \delta) \leq 2 \exp(-c \delta^2 m)
$$

provided

$$
m \geq \frac{c'}{\delta^2} s.
$$
Covering Arguments

Suppose that the random matrix $A \in \mathbb{R}^{m \times N}$ satisfies (CI). Let $S \subseteq [N]$ with $\text{card}(S) = s$. Then

$$\mathbb{P}(\|A^*_S A_S - \text{Id}\|_{2\to2} > \delta) \leq 2 \exp(-c\delta^2 m)$$

provided

$$m \geq \frac{c'}{\delta^2}s.$$

The argument relies on the following fact:

A subset $U$ of the unit ball of $\mathbb{R}^k$ relative to a norm $\| \cdot \|$ has covering and separating numbers satisfying

$$\mathcal{N}(U, \| \cdot \|, \rho) \leq \mathcal{S}(U, \| \cdot \|, \rho) \leq \left(1 + \frac{2}{\rho}\right)^k.$$
Restricted Isometry Property

Suppose that the random matrix $A \in \mathbb{R}^{m \times N}$ satisfies (CI). Then

$$P(\delta_s > \delta) \leq 2 \exp\left(-c\delta^2m\right)$$

provided $m \geq c'\delta^2 \ln\left(eN/s\right)$. The arguments are also valid for subgaussian matrices (e.g. Bernoulli matrices), since these satisfy (CI), too. For Gaussian matrices, more powerful techniques can provide an explicit value for $c'$. 
Suppose that the random matrix $A \in \mathbb{R}^{m \times N}$ satisfies (CI). Then

$$\mathbb{P}(\delta_s > \delta) \leq 2 \exp(-c\delta^2 m)$$

provided

$$m \geq \frac{c'}{\delta^2} s \ln(eN/s).$$
Suppose that the random matrix $A \in \mathbb{R}^{m \times N}$ satisfies (CI). Then

$$\mathbb{P}(\delta_s > \delta) \leq 2 \exp(-c\delta^2 m)$$

provided

$$m \geq \frac{c'}{\delta^2} s \ln(eN/s).$$

The arguments are also valid for subgaussian matrices (e.g. Bernoulli matrices), since these satisfy (CI), too.
Suppose that the random matrix \( A \in \mathbb{R}^{m \times N} \) satisfies (CI). Then
\[
\mathbb{P}(\delta_s > \delta) \leq 2 \exp(-c\delta^2 m)
\]
provided
\[
m \geq \frac{c'}{\delta^2} s \ln(eN/s).
\]

The arguments are also valid for subgaussian matrices (e.g. Bernoulli matrices), since these satisfy (CI), too.

For Gaussian matrices, more powerful techniques can provide an explicit value for \( c' \).
Summary

The RI conditions for \( s \)-sparse recovery are of the type \( \delta \kappa_s < \delta^* \). They guarantee stable and robust reconstructions in the form, say, \( (1) \)

\[
\| x - \Delta(Ax + e) \|_2 \leq C \sqrt{s} \sigma_s(x) + D \| e \|_2
\]

for all \( x \) and all \( e \).

Random matrices fulfill the RI conditions with high probability as soon as \( (2) \)

\[
m \geq c s \ln(\frac{N}{s})
\]

Next, we will see that this number of measurement is optimal, in the sense that estimates of type (1) require (2) to hold.

We will also examine the gain in replacing for all \( x \) in (1) by for a fixed \( x \).
Summary

The RI conditions for $s$-sparse recovery are of the type

$$\delta_{\kappa_s} < \delta_*. $$
Summary

The RI conditions for $s$-sparse recovery are of the type

$$\delta_{\kappa s} < \delta_\ast.$$ 

They guarantee stable and robust reconstructions in the form, say,

$$\|x - \Delta(Ax + e)\|_2 \leq \frac{C}{\sqrt{s}} \sigma_s(x)_1 + D\|e\|_2 \quad \text{for all } x \text{ and all } e.$$
Summary

The RI conditions for $s$-sparse recovery are of the type

$$\delta_{\kappa s} < \delta_*. $$

They guarantee stable and robust reconstructions in the form, say,

$$\|x - \Delta(Ax + e)\|_2 \leq \frac{C}{\sqrt{s}} \sigma_s(x)_1 + D\|e\|_2 \quad \text{for all } x \text{ and all } e.$$

Random matrices fulfill the RI conditions with high probability as soon as

$$m \geq c s \ln(N/s).$$
Summary

The RI conditions for $s$-sparse recovery are of the type

$$\delta_{\kappa s} < \delta_\star.$$  

They guarantee stable and robust reconstructions in the form, say,

$$\|x - \Delta(Ax + e)\|_2 \leq \frac{C}{\sqrt{s}} \sigma_s(x)_1 + D\|e\|_2 \quad \text{for all } x \text{ and all } e.$$

Random matrices fulfill the RI conditions with high probability as soon as

$$m \geq c s \ln(N/s).$$

Next, we will see that this number of measurement is optimal, in the sense that estimates of type (1) require (2) to hold.
Summary

The RI conditions for $s$-sparse recovery are of the type

$$\delta_{\kappa s} < \delta_*.$$  

They guarantee stable and robust reconstructions in the form, say,

$$\|x - \Delta(Ax + e)\|_2 \leq \frac{C}{\sqrt{s}}\sigma_s(x)_1 + D\|e\|_2 \quad \text{for all } x \text{ and all } e.$$  

Random matrices fulfill the RI conditions with high probability as soon as

$$m \geq c s \ln(N/s).$$  

Next, we will see that this number of measurement is optimal, in the sense that estimates of type (1) require (2) to hold. We will also examine the gain in replacing \textit{for all } $x$ in (1) by \textit{for a fixed } $x$. 
Other types of random matrices

Under proper normalization of the matrices, partial Fourier matrices with \( m \geq c \delta^{-2} s \ln(3)(N) \) rows selected at random satisfy the RIP with high probability.

Pregaussian (e.g. Laplace) random matrices with \( m \geq c \delta s \ln(eN/s) \) rows satisfy with overwhelming probability

\[ (1 - \delta)z \leq \|Az\|_1 \leq (1 + \delta)z \]

for all \( s \)-sparse \( z \in \mathbb{R}^N \), where the slanted norm is comparable to the \( \ell_2 \)-norm.

Adjacency matrices of lossless expanders (which exist with nonzero probability) satisfy

\[ (1 - \theta)\|z\|_1 \leq \|Az\|_1 \leq \|z\|_1 \]

for all \( s \)-sparse \( z \in \mathbb{R}^N \).
Other types of random matrices

Under proper normalization of the matrices,
Other types of random matrices

Under proper normalization of the matrices,

- Partial Fourier matrices with $m \geq c\delta^{-2}s \ln^3(N)$ rows selected at random satisfy the RIP with high probability.

- Pregaussian (e.g. Laplace) random matrices with $m \geq c\delta^{-2}s \ln(eN/s)$ rows satisfy with overwhelming probability

$$(1-\delta) \|z\|_1 \leq \|Az\|_1 \leq (1+\delta) \|z\|_1$$

for all $s$-sparse $z \in \mathbb{R}^N$, where the slanted norm is comparable to the $\ell_2$-norm.

- Adjacency matrices of lossless expanders (which exist with nonzero probability) satisfy

$$(1-\theta) \|z\|_1 \leq \|Az\|_1 \leq \|z\|_1$$

for all $s$-sparse $z \in \mathbb{R}^N$. 

Other types of random matrices

Under proper normalization of the matrices,

- Partial Fourier matrices with $m \geq c\delta^{-2} s \ln^3(N)$ rows selected at random satisfy the RIP with high probability.

- Pregaussian (e.g. Laplace) random matrices with $m \geq c\delta s \ln(eN/s)$ rows satisfy with overwhelming probability

\[(1-\delta) \|z\|_1 \leq \|Az\|_1 \leq (1+\delta) \|z\|_1 \text{ for all } s\text{-sparse } z \in \mathbb{R}^N,\]

where the slanted norm is comparable to the $\ell_2$-norm.

- Adjacency matrices of lossless expanders (which exist with nonzero probability) satisfy

\[(1-\theta) \|z\|_1 \leq \|Az\|_1 \leq \|z\|_1 \text{ for all } s\text{-sparse } z \in \mathbb{R}^N,\]
Other types of random matrices

Under proper normalization of the matrices,

- Partial Fourier matrices with $m \geq c\delta^{-2}s \ln^3(N)$ rows selected at random satisfy the RIP with high probability.

- Pregaussian (e.g. Laplace) random matrices with $m \geq c\delta s \ln(eN/s)$ rows satisfy with overwhelming probability

$$\left(1-\delta\right)\|z\| \leq \|Az\| \leq \left(1+\delta\right)\|z\|$$

for all $s$-sparse $z \in \mathbb{R}^N$, where the slanted norm is comparable to the $\ell_2$-norm.

- Adjacency matrices of lossless expanders (which exist with nonzero probability) satisfy

$$\left(1-\theta\right)\|z\| \leq \|Az\| \leq \|z\|$$

for all $s$-sparse $z \in \mathbb{R}^N$. 