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

Concentration Inequality

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

$$a_{i,j} = \frac{g_{i,j}}{\sqrt{m}} \qquad \text{where the } g_{i,j} \text{ are independent } \mathcal{N}(0,1).$$

Concentration Inequality

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

$$a_{i,j} = \frac{g_{i,j}}{\sqrt{m}} \qquad \text{where the } g_{i,j} \text{ are independent } \mathcal{N}(0,1).$$

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

$$\mathbb{E}\big((A\mathbf{x})_i^2\big) = \mathbb{V}\big(\sum a_{i,j} x_j\big) = \sum x_j^2 \mathbb{V}(a_{i,j}) = \frac{\|\mathbf{x}\|_2^2}{m},$$
$$\mathbb{E}\big(\|A\mathbf{x}\|_2^2\big) = \|\mathbf{x}\|_2^2.$$

Concentration Inequality

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

$$a_{i,j} = \frac{g_{i,j}}{\sqrt{m}} \qquad \text{where the } g_{i,j} \text{ are independent } \mathcal{N}(0,1).$$

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

$$\mathbb{E}((A\mathbf{x})_i^2) = \mathbb{V}\Big(\sum a_{i,j} x_j\Big) = \sum x_j^2 \mathbb{V}(a_{i,j}) = \frac{\|\mathbf{x}\|_2^2}{m},$$
$$\mathbb{E}\big(\|A\mathbf{x}\|_2^2\big) = \|\mathbf{x}\|_2^2.$$

- In fact, $\|A\mathbf{x}\|_2^2$ concentrates around its mean: for $t \in (0,1)$,

(CI) $$\mathbb{P}\big(\big|\|A\mathbf{x}\|_2^2 - \|\mathbf{x}\|_2^2\big| > t\|\mathbf{x}\|_2^2\big) \leq 2\exp(-ct^2 m).$$

Covering Arguments

Covering Arguments

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

Covering Arguments

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

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

$$\mathbb{P}\big(\|A_S^* A_S - \mathrm{Id}\|_{2\to 2} > \delta\big) \leq 2\exp(-c\delta^2 m)$$

Covering Arguments

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

$$\mathbb{P}\big(\|A_S^* A_S - \mathrm{Id}\|_{2\to 2} > \delta\big) \leq 2\exp(-c\delta^2 m)$$

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 $\mathrm{card}(S) = s$. Then

$$\mathbb{P}\big(\|A_S^* A_S - \mathrm{Id}\|_{2\to 2} > \delta\big) \le 2\exp(-c\delta^2 m)$$

provided

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

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

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

Restricted Isometry Property

Restricted Isometry Property

- 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

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

Restricted Isometry Property

- 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.

Restricted Isometry Property

- 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

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

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

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

$$(1) \quad \|\mathbf{x} - \Delta(A\mathbf{x} + \mathbf{e})\|_2 \leq \frac{C}{\sqrt{s}} \sigma_s(\mathbf{x})_1 + D\|\mathbf{e}\|_2 \qquad \text{for all } \mathbf{x} \text{ and all } \mathbf{e}.$$

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,

$$(1) \quad \|\mathbf{x} - \Delta(A\mathbf{x} + \mathbf{e})\|_2 \leq \frac{C}{\sqrt{s}} \sigma_s(\mathbf{x})_1 + D\|\mathbf{e}\|_2 \qquad \text{for all } \mathbf{x} \text{ and all } \mathbf{e}.$$

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

$$(2) \qquad m \geq c\, s \ln(N/s).$$

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,

$$(1) \quad \|\mathbf{x} - \Delta(A\mathbf{x} + \mathbf{e})\|_2 \le \frac{C}{\sqrt{s}} \sigma_s(\mathbf{x})_1 + D\|\mathbf{e}\|_2 \qquad \text{for all } \mathbf{x} \text{ and all } \mathbf{e}.$$

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

$$(2) \qquad m \ge c\, s \ln(N/s).$$

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,

$$(1) \quad \|\mathbf{x} - \Delta(A\mathbf{x} + \mathbf{e})\|_2 \leq \frac{C}{\sqrt{s}} \sigma_s(\mathbf{x})_1 + D\|\mathbf{e}\|_2 \qquad \text{for all } \mathbf{x} \text{ and all } \mathbf{e}.$$

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

$$(2) \qquad m \geq c\, s \ln(N/s).$$

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 *for all* $\mathbf{x}$ in (1) by *for a fixed* $\mathbf{x}$.

Other types of random matrices

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.

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

  $$(1-\delta) /\!\!/ \mathbf{z} /\!\!/ \leq \|A\mathbf{z}\|_1 \leq (1+\delta) /\!\!/ \mathbf{z} /\!\!/ \qquad \text{for all } s\text{-sparse } \mathbf{z} \in \mathbb{R}^N,$$

  where the slanted norm is comparable to the ℓ_2-norm.

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

  $$(1-\delta) /\!\!/ \mathbf{z} /\!\!/ \leq \|A\mathbf{z}\|_1 \leq (1+\delta) /\!\!/ \mathbf{z} /\!\!/ \qquad \text{for all } s\text{-sparse } \mathbf{z} \in \mathbb{R}^N,$$

  where the slanted norm is comparable to the ℓ_2-norm.
- Adjacency matrices of lossless expanders (which exist with nonzero probability) satisfy

  $$(1-\theta)\|\mathbf{z}\|_1 \leq \|A\mathbf{z}\|_1 \leq \|\mathbf{z}\|_1 \qquad \text{for all } s\text{-sparse } \mathbf{z} \in \mathbb{R}^N.$$