COMPSCI 690T Coding Theory and Applications	Jan 23, 2017
Lecture 16	
Instructor: Arya Mazumdar	*Scribe: Names Redacted*

1 Compressed Sensing

1.1 Randomized Construction of the sampling matrix A

One way of randomized construction is making A a random Gaussian matrix with m rows and n columns. Each cell of A is from $\mathcal{N}(0, 1/m)$. We can show that if $m = O(k \log \frac{n}{k})$, A will with high probability have δ_{2k}-RIP with $\delta_{2k} \leq \sqrt{2} - 1$.

The δ_{2k}-RIP means

$$(1 - \delta_{2k})||x||_{l_2}^2 \leq ||Ax||_{l_2}^2 \leq (1 + \delta_{2k})||x||_{l_2}^2$$

where x is $2k$-sparse vector. To simplify this, we can use a compact x' by taking the $2k$ non zero values from x, and a compact matrix B by the $2k$ columns from A at the non zero coordinates of x. So that

$$Bx' = Ax$$

and there is no constraint on x'. The δ_{2k}-RIP condition becomes

$$2 - \sqrt{2} \leq \frac{x'^T B^T B x'}{x^T x} \leq \sqrt{2}$$
$$0.586 \leq \frac{x'^T B^T B x'}{x'^T x'} \leq 1.414$$

By Rayleigh-Ritz theorem, we know for any symmetric matrix D

$$\max_{x \neq 0} \frac{x^T D x}{x^T x} = \text{max eigenvalue of } D$$
$$\min_{x \neq 0} \frac{x^T D x}{x^T x} = \text{min eigenvalue of } D$$

Now we reduce the sufficient condition for RIP into the bound of eigenvalues of $B^T B$ (or singular values of B) for any set of $2k$ columns B of A.

For an iid Gaussian $\mathcal{N}(0, 1/m)$ matrix of size $m \times 2k$, the probability that the singular value is greater than $1 + \epsilon$ or less than $1 - \epsilon$ is at most $e^{-\alpha n}$ for some $\alpha > 0$ that depends on the $\epsilon > 0$.

Hence for a Gaussian matrix A, the probability of δ_{2k}-RIP being violated is

$$\begin{aligned} P_e &\leq \binom{n}{2k} e^{-\alpha m} \\ &\leq (\frac{ne}{2k})^{2k} e^{-\alpha m} \\ &= e^{-\alpha m + 2k \log \frac{ne}{2k}} \end{aligned}$$

We can let $m = \frac{4}{\alpha} k \log \frac{n}{k}$ to make P_e bounded by a small value. Therefore $m = O(k \log \frac{n}{k})$ suffices.

1.2 Explicit Construction of A

If we have a binary code (m, N, d) with large d, we can construct matrix A expilicitly.

Let f be a bitwise mapping which maps 0 to $1/\sqrt{m}$ and 1 to $-1/\sqrt{m}$. We can construct A by using $f(x)$ as columns for every codeword x. If d is large enough, this matrix has RIP.

By the definition of f, we know the inner product between $f(x)$ and $f(x')$, where x, x' are two codewords, is

$$1 - \frac{2d(x, x')}{m}$$

We choose any $2k$ columns from A to form matrix B and compute $B^T B$. It has 1 on diagonal and $1 - 2d^*/m$ elsewhere. d^* is the distance between two different codewords. By Gershgorin circle theorem, we know the eigenvalue of A must be in at least one of the Gershgorin discs $D(a_{ii}, R_i)$. The discs are centered at the diagonal elements a_{ii} with radius $R_i = \sum_{j \neq i} |a_{ij}|$. In our case, a_{ii} is always 1 and there are $2k$ columns. We have

$$\begin{aligned} |\lambda_i - 1| &< \sum_{i \neq j} (1 - \frac{2d(c_i, c_j)}{m}) \\ &\leq (2k - 1)(1 - \frac{2d}{m}) \end{aligned}$$

We want $|\lambda_i - 1| \leq 0.414$, which means the sufficient condition for RIP is

$$\begin{aligned} (2k - 1)(1 - \frac{2d}{m}) &\leq 0.414 \\ 1 - \frac{2d}{m} &\leq \frac{0.414}{2k - 1} \\ d &\geq \frac{m}{2}(1 - \frac{0.414}{2k - 1}) \\ d &\geq \frac{m}{2}(1 - \frac{0.2}{k}) \end{aligned}$$

We can construct such code using the explicit construction for GV bound. We have

$$\begin{aligned} \frac{\log N}{m} &= 1 - h(\frac{1}{2} - \frac{0.1}{k}) \\ m &= \frac{\log N}{1 - h(\frac{1}{2} - \frac{0.1}{k})} \end{aligned}$$

The time complexity of such construction is in polynomial w.r.t. N. Note that in the context of coding theory, this complexity is bad because we care about the dimension of the code k, then it's in exponential w.r.t. k. However in compressed sensing, we care about the dimension of signal, which is N.

We can expand $h(\frac{1}{2} - x)$ to $1 - \Theta(x^2)$. Therefore,

$$m = \mathrm{O}(k^2 \log N)$$

There is no explicit construction of RIP matrix with $m = \mathrm{O}(k \log N)$.