Big Data Algorithms **Fall 2017**

Lecture 11: Compressed Sensing at Nov. 29

Lecturer: Xiaotie Deng *Scribes: Hongxiao Bai*

11.1 Overview

Given a vector $\mathbf{x} = (x_1, x_2, \cdots, x_n)^T$, we want to find a linear transformation $\mathbf{Ax}$ to reduce its dimention. After reduction, we also hope that we have the recover operation of recovering from short data to long. The main reason that we can reduce its dimention is the sparsity of data.

For compression, roughly speaking, we want to have an operation on dimention $n \to k = f(n)$. What do we expect $f(\cdot)$? Informally, we want to keep data at $log(n)$ size in mind. This ability has been ensured by *JL Lemma*.

Sensing is signals of different forms of different applications, and we can use length, distance, signal strength and weight of importance in system to measure it.

11.2 Modeling of Signal

11.2.1 Norm

Signal can be seen as a vector $\mathbf{x} = (x_1, x_2, \cdots, x_n)^T$. Norm is a measurement of vector

$$\|x\|_p = (\sum_i |x_i|^p)^{1/p}.$$

And we have

$$L_0 = \|x\|_0 = \lim_{p \to 0} (\sum_i |x_i|^p)^{1/p} = \lim_{n \to \infty} (\sum_i |x_i|^{1/n})^n = \# \text{ of non-zeros}$$
$$L_\infty = \|x\|_\infty = \lim_{p \to \infty} (\sum_i |x_i|^p)^{1/p} = max|x_i|$$

Is $L_{1/2}$ a norm? Can we use it for distance? (It's a norm but it doesn't obey the triangle inequality.)

11.2.2 Bases

Bases are independent vectors $\phi_1, \phi_2, \cdots, \phi_n$. Dependent means $\exists (\alpha_1, \alpha_2, \cdots, \alpha_n) \neq 0, \sum_i \alpha_i \phi_i = 0$.
Orthogonal bases: $\phi_i^T \phi_j = 0, i \neq j$.
Any vector $\mathbf{x}$ can be uniquely represented by a basis: $\mathbf{x} = \mathbf{\Phi c} = \sum_{i=1}^n c_i \phi_i$, and $c_i = \phi_i^T x$.

11.2.3 Frames

Frames is a subset of orthogonal independent vectors.
Idea: use frames to represent vector **x**. We have errors however data size is reduced.

Bases: square matrix
Frames: $d \times n$ matrix

$\Phi = (\phi_1, \phi_2, \cdots, \phi_n)_{(d\times n)}$ is a frame if $\exists 0 < A < B < \infty, \forall x : A\|x\|_2^2 \leq \|\Phi x\|_2^2 \leq B\|x\|_2^2$.
The optimum frame bound A and B are the minimum and maximum eigenvalue of $\Phi\Phi^T$.

Since $\|\Phi^T x\|_2^2 = (\Phi^T x)^T(\Phi^T x) = x^T\Phi\Phi^T x$, so $\Phi\Phi^T$ is positive definite, and we can write

$$\Phi\Phi^T = P^T diag(\lambda_1, \lambda_2, \cdots, \lambda_k, 0, \cdots, 0)P$$

and $\lambda_1 > \cdots > \lambda_k > 0$, and $x^T\Phi\Phi^T x = \sum_i \lambda_i x^T x$, so

$$\lambda_k \sum x^T x \leq \|\Phi^T x\|_2^2 \leq \lambda_1 \sum x^T x$$

.
Dual frame $\tilde{\Phi}$: $\Phi\tilde{\Phi}^T = \tilde{\Phi}\Phi^T = I$.
So we can set $\tilde{\Phi} = (\Phi\Phi^T)^{-1}\Phi$, it is *Moore Penrose Inverse*.

11.3 Low-Dimention Signal Models

Norm is non-linear.
Sparse: # of non-zeors is small.
k-sparse: # of non-zeros $= k$.
Linear combination of two k-sparse signals may not be k-sparse any more.

For a compression $y = Ax$, we want to recover x. Suppose $Ax = y$, $Ax' = y$ and $x, x' \in \Sigma_k$,then $A(x - x') = 0$.
We want $x = x'$, i.e. unique recovering, so we want $Az = 0$ only have zero solution.
and A is a $d \times n$ mathix, $d \ll n$, so $rank(A) = d$, and $d + 1$ vectors will become dependent.

11.4 Sensing Matrix

$\Lambda \subset \{1, 2, \cdots, n\}$, and $\Lambda^c = \{1, 2, \cdots, n\} - \Lambda$. (c means complement.)
So

$$h_\Lambda = \begin{cases} h_i, & i \in \Lambda \\ 0, & i \in \Lambda^c \end{cases}$$

Null space $N(A) = \{z : Az = 0\}$.
Spark: the smallest number of columns of A that are linearly dependent.

Theorem 11.1 *$\forall y \in R^m$,there exists at most one signal $x \in \Sigma_k$ such that $y = Ax$ iff $spark(A) > 2k$.*