Double Sparse Compressed Sensing Problem

Grigory Kabatiansky, IITP RAS, Russia
Cédric Tavernier, Assystem, France.

ACCT-XIV, Svetlogorsk
Sept 12, 2014

The Compressed Sensing (CS) theme was born in papers by Donoho,"Compressed sensing", and Candes & Tao, "Near-Optimal Signal Recovery From Random Projections: Universal Encoding Strategies? ", both in IEEE-IT, 2006.
The CS problem is to reconstruct an n-dimensional t-sparse vector $x \in \mathbb{R}^n$ by a few linear measurements $s_i = (h_i, x)$ even if measurements (h_i, x) are known with some errors e_i, where $i = 1, \ldots, r$. Saying in other words, the goal is to find a solution x the following equation

$$s = Hx^T + e, \tag{1}$$

such that its l_0 norm (or Hamming weight), denoted by $||x||_0$, is at most t, if Euclidean length $||e||_2$ of *syndrom error vector* $e = (e_1, \ldots, e_r)$ is small enough, i.e. $||e||_2 \leq \varepsilon$. Here H denotes an $r \times n$ matrix, whose rows are $h_1, \ldots, h_r$.

CS uses the following popular "trick" replacing hard problem of finding solution of Eq. (1) with minimal l_0 norm on finding solution with minimal l_1 norm. I.e., find *arg* min $\sum |x_i|$ such that $||s - Hx^T|| \leq \varepsilon$.
This problem is LP problem. Moreover it was proved that if matrix H is *RIP-matrix* then the solution x^* of LP problem is a good *approximation* to the solution x_0 of the original problem and the corresponding number of measurements r has the same order of grow as the minimal possible number of measurements, namely,

$$r_{min} = O(t \log \frac{n}{t}) \quad (2)$$

The compressed sensing problem is usually investigated under assumption that the error vector $e = (e_1, \ldots, e_r)$ has relatively small Euclidean norm (length) $||e||_2$. We consider another assumption, namely, that the error vector e is also sparse, say $||e||_0 \leq t$, but in return its Euclidean norm can be arbitrary large and we find not an approximation but exact solution of (1)!
We call these assumptions: $||x||_0 \leq l$ and $||e||_0 \leq t$, as *double sparse*.
Our main result is that for double sparse CS problem

$$r_{min} = 2(t + l)$$

Definition

A real $r \times n$ matrix H called a (t, l)-double sparse compressed sensing (DSCS) matrix if

$$||Hx^T - Hy^T||_0 \geq 2l + 1 \quad (3)$$

for any two distinct vectors $x, y \in \mathbb{R}^n$ such that $||x||_0 \leq t$ and $||y||_0 \leq t$.

This definition immediately leads to the following
Proposition A real $r \times n$ matrix H is a (t, l)-DSCS matrix iff

$$||Hz^T||_0 \geq 2l + 1 \quad (4)$$

for any nonzero vector $z \in \mathbb{R}^n$ such that $||z||_0 \leq 2t$.

Construction of MDS codes for DSCS

$$H = G^T \tilde{H}, \tag{5}$$

where a real $\tilde{r} \times n$ matrix $\tilde{H}$ be a parity-check matrix of an $(n, n - \tilde{r})$-code code over $\mathbb{R}$, correcting t errors, and G be a generator (systematic) matrix of an $(r, \tilde{r})$-code over $\mathbb{R}$ of length r, correcting l errors.
Saying in words, we encode columns of parity-check matrix $\tilde{H}$, which already capable to correct t errors, by a code, correcting l errors, in order to restore correctly syndrom of $\tilde{H}$.

Theorem

Matrix $H = G^T \tilde{H}$ is a (t, l)-DSCS matrix.

Proof. According to Proposition it is enough to prove that $||Hz^T||_0 \geq 2l + 1$ for any nonzero vector $z \in \mathbb{R}^n$ such that $||z||_0 \leq 2t$. Indeed, $u = \tilde{H}z^T \neq 0$ since any $2t$ columns of $\tilde{H}$ are linear independent. Then $Hz^T = G^T\tilde{H}z^T = G^T u = (u^T G)^T$ and $u^T G$ is a nonzero vector of a code over $\mathbb{R}$, correcting l errors. Hence $||Hz^T||_0 = ||u^T G||_0 \geq 2l + 1$. □

How to decode? First we decode vector $\hat{s} = s + e$ by a decoding algorithm of the code with generator matrix G. Since $||e||_0 \leq l$ this algorithm outputs the correct syndrome s. Then we form a syndrome $\tilde{s}$ by selecting first $\tilde{r}$ coordinates of s and apply syndrom decoding algorithm for the corresponding syndrom equation

$$\tilde{s} = \tilde{H}x^T. \tag{6}$$

Why not just MDS??

It is a very natural question why do not use just a single RS-code with extra redudancy in order to correct possible errors in measurements, i.e., in its syndrom?!
It is rather old question, which goes back to time of the French Revolution, when R.Prony (J. de Ecole Polytechnique 1, pp.24-76, 1795) asked how to reconstruct a polynomial of a given degree by its value in some points, when at most l of these values could be incorrect. The modern solution was given in M.T. Comer, E.L. Kaltofen, C.Pernet, "Sparse Polynomial Interpolationb and Berlekamp-Massey Algorithms That Correct Outlier Errors in Input Values", namely, it was shown that it is possible to solve equation (1) by RS-code *iff* its redudancy $r \geq 2t(2l+1)$.
We see that it is too much expensive solution for double sparse CS-problem.

Capacity problem of CS

The minimal possible number of measurements to solve equation

$$s = Hx^T + e, \tag{7}$$

has the order

$$r_{min} = O(t \log \frac{n}{t}) \tag{8}$$

if the solution is enough good approximation, namely, if

$$||x^* - x_0|| \leq C||e||_2 \tag{9}$$

For t-constant it is OK, but for $t = \lambda n$ it gives $r = O(n)$ and we want to know constant which stands in this $O(n)$!