What to do if syndroms are corrupted also?

Grigory Kabatiansky, IITP RAS, Russia;
Serge Vladuts, Aix-Marseille University and IITP RAS

Albena, September 8, 2013

The Compressed Sensing (CS)Problem is formulated as a problem of reconstructing of n-dimensional t-sparse vector $x \in \mathbb{R}^n$ by a few, namely, r linear measurements $s_i = (h_i, x)$. Denote by H an $r \times n$ matrix, whose rows are $h_1, \ldots, h_r$, and syndrom vector $s = (s_1, \ldots, s_r)$. Then CS problem is how to find vector x such that

$$s = Hx^T$$

and its Hamming weight, equals to the number of nonzero coordinates of x and denoted by $wt(x)$ or $||x||_0$, is at most t. All we know how to solve this problem - just take Reed-Solomon code over reals. Then $r = 2t$ and decoding algorithm is polynomial!

The main achievement of Compressed Sensing is that the corresponding algorithm(s) can recover a sparse vector x even if measurements are not exact, i.e. if we know (h_i, x) with some errors $e = (e_1, \ldots, e_r)$, what is formulated as solving equation

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

A discrete version of this problem leads us to a new problem of coding theory Now let us consider a new problem of coding theory: **To reconstruct the vector x of the Hamming weight $wt(x) \leq t$ if some but not more than l of its syndrom coordinates are incorrect**.

The corresponding matrix H we call q -ary (t, l)-compressed sensing (CS) matrix.

Definition

A q-ary $r \times n$ matrix H called a (t, l)-CS matrix iff

$$d(Hx^T, Hy^T) \geq 2l + 1. \quad (2)$$

for any two distinct vectors $x, y \in \mathbb{F}_q^n$ such that $wt(x) \leq t$ and $wt(y) \leq t$.

Proposition A q -ary $r \times n$ matrix H is a (t, l)-CS matrix iff $wt(Hz^T) \geq 2l + 1$ for any nonzero vector $z \in \mathbb{F}_q^n$ such that $wt(z) \leq 2t$.

In particular case $t = 1,\ l = 1$ we have nonadaptive (or deterministic) version of famous Ulam's problem on searching with lie, and for $t = 1$ and l arbitrary - a nonadaptive version of Ulam's problem on searching with l lies.

Let us denote:
by $V_q(n,d) = \sum_{i=0}^{i=d} C_n^i (q-1)^i = |\{x \in \mathbb{F}_q^n : wt(x) \leq d\}|$ the volume of q-ary n dimensional ball of radius d;
by $A_q(n, 2t+1)$ the maximal possible cardinality of q-ary code correcting t errors.

Theorem

(Hamming bound). For any q -ary (t,l)-CS $r \times n$-matrix

$$A_q(r, 2l+1) \geq V_q(n,t). \tag{3}$$

Proof. According to definition the set of syndroms $\{Hx^T : \ wt(x) \leq t\}$ forms a code of length r capable to correct l errors, in particular, all syndroms are different, and, on the other hand, the number of syndroms cannot be larger than $A_q(r, 2l+1)$.

Hamming bound-II

For binary case Plotkin bound gives that

$$(2l+1)2^{r-4l} \geq A_2(r, 2l+1) \geq V_2(n, t) \tag{4}$$

or, in asymptotic form,

$$\frac{r}{n} \geq 4\frac{l}{n} + H(t/n) + o(1) \tag{5}$$

Theorem

(G-V bound). A (t,l)-CS $r \times n$-matrix exists if

$$V_q(n, 2t-1)V_q(r, 2l) \leq q^r. \quad (6)$$

Hence $r(n,l) \leq \log_q V_q(n, 2t-1) + \log_q V_q(r, 2l)$ for $2\tau < \frac{q-1}{q}, 2\lambda < \frac{q-1}{q}$.
Ulam's problem with many (l) lies corresponds to the case $t = 1,\ q = 2$ and both bounds asymptotically coincides, namely, the minimal number of questions is $l(4 + o(1))$. For arbitrary q the answer is $l(\frac{2q}{q-1} + o(1))$.

Restricted δ-Isometry Property(RIP) matrix H if

$$(1-\delta_D)||x||_2 \leq ||Hx^T||_2 \leq (1+\delta_D)||x||_2, \tag{7}$$

for any vector $x \in \mathbb{R}^n : ||x||_0 \leq D$, where $0 < \delta_D < 1$. Typical result looks like this

"if $\delta_{3t} + 3\delta_{4t} < 2$ then the solution x^* of linear programming problem is unique and equal to x"

or, additionally *" and $||x^* - x||_2 \leq C\varepsilon$ for any perturbation e with $||e||_2 \leq \varepsilon$ "*

We can rewrite it as $\delta_{4t} < 1/2$ or that for any z s.t. $||z||_0 \leq 4t$

$$\frac{1}{2}||z||_2 \leq ||Hz^T||_2 \leq \frac{3}{2}||z||_2$$

It was proved that such matrices exist if $r = O(t\log(n/t))$

Two major questions:
What is O in $r = O(tlog(n/t))$ for $t/n = const$, i.e., what is the *capacity* of this channel??? RIP probably is too strong since enough

$$\mu_{2t}||x||_2 \leq ||Hx^T||_2, \tag{8}$$

If error-correcting codes can "fight" with errors in syndrom? For instance, RS codes?

1. D.L. Donoho, "Compressed sensing", IEEE Trans on Information Theory, v. 52 (4), pp. 1289-1306, 2006
2. E.J. Candes, T. Tao, " Near-Optimal Signal Recovery From Random Projections: Universal Encoding Strategies? ", IEEE Trans Information Theory, v. 52 (12), pp. 5406 - 5425, 2006
3. E.J. Candes, T. Tao, "Decoding by linear programming", IEEE Trans Information Theory, v. 51 (12), pp. 4203 - 4215, 2005
4, A. Pelc, "Solution of Ulam's Problem on Searching with a Lie", Journal of Comb.Theory, Ser.A 44, pp129-140, 1987.