An introduction to compressive sampling

Zhang Liu

02/01/2008

A new sampling paradigm

- Nyquist/Shannon: sample at twice the bandwidth of the signal

- An digital camera example
  - Acquire a full 10 megapixel image
  - Compression retains the largest transform coefficients, say 100K
  - Motivation: design a < 1 MP camera to achieve the same quality

- Compressive sampling

$$y = Ax$$

- signal $x \in \mathbf{R}^n$, samples $y \in \mathbf{R}^m$, $m << n$

- Assume x is compressible, meaning sparse in some fixed basis

Under-sample of discrete fourier coefficents

$$y = Ax$$

A is arbitrary m rows of a $n \times n$ DFT matrix

$$y_k = \frac{1}{\sqrt{n}} \sum_{t=0}^{n-1} x_t e^{-j2\pi\omega_k t/n}, \quad k = 1, \ldots, m$$

Theorem 1 Suppose $\mathbf{card}(x) \leq S$, we can almost always recover x exactly by solving LP

$$\begin{array}{rll} x^* = & \text{arg.min.} & \|x\|_1 \\ & \text{subject to} & Ax = y \end{array}$$

With overwhelming probability if $m \geq CS\log(n)$

Uniform uncertainty principle (UUP)

UUP states the sensing matrix A obeys a "'restricted isometry hypothesis"'

S-restricted isometry constant δ_S of A is defined as

$$\text{min. } \delta_S, \quad \text{s.t. } (1-\delta_S)\|c\|_2^2 \leq \|A_t c\|_2^2 \leq (1+\delta_S)\|c\|_2^2, \quad \forall c\ \forall A_t$$

A_t represents all submatrix of A of dimension $m \times t$, $t \leq S$

Theorem 2 Suppose $\mathbf{card}(x) \leq S$ and $\delta_{2S} + \delta_{3S} < 1$, then x^* recovers x exactly

Theorem 3 Suppose $\delta_{3S} + \delta_{4S} < 2$, then

$$\|x^* - x\|_2 \leq C\frac{\|x - x_S\|_1}{\sqrt{S}}.$$

x_S is a vector keeping S largest entries of x and setting others 0

C is well behaved, i.e. $C \leq 8.77$ for $\delta_{4S} = 1/5$.

Random matrices

Design of the sensing matrice $A \in \mathbf{C}^{m\times n}$ is usually difficult

Fortunately, many random matrice realization works with high probability

- *Gaussian matrix*: each entry is independent sample of $(0, 1/m)$ Gaussian distribution,
  $$S \leq \frac{Cm}{\log(n/m)}$$
  Theorem 2 and 3 hold with probability $1 - O(e^{-\gamma n})$.

- *Binary matrix*: each entry is independent sample of symmetric Bernoulli distribution,
  $$S \leq \frac{Cm}{\log(n/m)}$$
  Theorem 2 and 3 hold with probability $1 - O(e^{-\gamma n})$.

- *Fourier matrix*: random rows of DFT matrix,

$$S \le \frac{Cm}{\log(n)^4}$$

Theorem 2 holds with overwhelming probability

- *Incohenrent matrix*: random rows of $n \times n$ unitary matrix U, then renormalize column,

$$S \le \frac{Cm}{n \log(n)^4 (\max_{i,j} |U_{i,j}|)^2}$$

Theorem 2 holds with overwhelming probability

Robust compressive sampling

$$y = Ax + e$$

e is measurement error with bounded energy $\|e\|_2 \leq \epsilon$

We can solve a SOCP for this problem

$$\begin{array}{ll} \text{minimize} & \|x\|_1 \\ \text{subject to} & \|y - Ax\|_2 \leq \epsilon \end{array}$$

theorem 4 Suppose $\delta_{3S} + \delta_{4S} < 2$, the solution x^* to the SOCP obeys

$$\|x^* - x\|_2 \leq C_{1,S}\epsilon + C_{2,S}\frac{\|x_0 - x_{0,S}\|_1}{\sqrt{S}}$$

- The first term is proportional to the size of the measurement error
- The second term is the error one would obtain from the noiseless case

Decoding by LP

$x \in \mathbf{R}^m$ is the information

$A \in \mathbf{R}^{n \times m}$ is the coding matrix, $n > m$

$$y = Ax + e$$

e is sparse and completely corrupt y_i if e_i nonzero

$$\text{minimize} \quad \|y - Ax\|_1$$

Let any $F \in \mathbf{R}^{n-m \times n}$ with $FA = 0$, this is the same as solving

$$\begin{array}{ll} \text{minimize} & \|e\|_1 \\ \text{subject to} & Fy = Fe \end{array}$$

Coding matrix

Popular choice of coding matrix is a random matrix

Theorem 5: suppose coding matrix A is i.i.d. $N(0,1)$, corruption e is sparse, then exact decoding all x with probability exceeding $1 - O(e^{-\gamma m})$.

Numerical experiment of exact recovery of x occur with high probability

– if corruption is less than 17% when $n = 2m$

– if corruption is less than 14% when $n = 4m$

Connection to statistical estimation

$$y = Ax + e,$$

$e \in \mathbf{R}^m$ is Gaussian i.i.d with zero mean and variance σ^2

$$\text{Classical result:} \quad \|x^* - x\|_2^2 \leq Cm\sigma^2$$

Suppose x is sparse, we can do better

Dantzig selector (DS)

$$\begin{array}{ll} \text{minimize} & \|x\|_1 \\ \text{subject to} & \|A^T(y - Ax)\|_\infty \leq \lambda\sigma \end{array}$$

Theorem 6: suppose $\mathbf{card}(x) \leq S$ and $\delta_{2S} + \theta_{S,2S} < 1 - t$, set $\lambda = (1 - t^{-1}1)\sqrt{s \log p}$. With very high probability, solution of DS obeys

$$\|x^* - x\|^2 \leq O(\log p) \left(\sigma^2 + \sum_i \min(x_i^2, \sigma^2) \right)$$

Extension to matrix

To obtain a sparse vector x

$$\text{minimize} \quad \|x\|_1,$$

where $\|x\|_1 = \sum_i |x_i|$

To obtain a low-rank matrix X

$$\text{minimize} \quad \|X\|_*,$$

where $\|X\|_*$ denotes the nuclear norm and is the sum of singular values

$$\|X\|_* = \sum_i \sigma_i(X)$$

Nuclear norm minimization

- $\|x\|_1$ is the convex envelope of $\mathbf{card}(x)$
- $\|X\|_*$ is the convex envelope of $\mathbf{rank}(x)$

Nuclear norm minimization

$$\begin{array}{ll} \text{minimize} & \|X\|_* \\ \text{subject to} & \mathcal{A}(X) = b, \end{array}$$

where $\mathcal{A} : \mathbf{R}^{m \times n} \rightarrow \mathbf{R}^p$

Exact low-rank matrix recovery

r-restricted isometry constant: smalles number $\delta_r(\mathcal{A})$ such that

$$(1-\delta_r(\mathcal{A}))\|X\|_F \leq \|\mathcal{A}(X)\| \leq (1+\delta_r(\mathcal{A}))\|X\|_F$$

holds for all matrices X of rank at most r

Theorem 7: $\delta_{5r} < 1/10$, minimum rank solution is obtained

Nearly isometric families

Nearly isometric random mapping:

— Gaussian matrix: $A_{i,j}$ is $N(0, 1/p)$

— Bernoulli matrix: $A_{i,j} = \left\{ \begin{array}{ll} \sqrt{1/p} & \text{with probability } 1/2 \\ -\sqrt{1/p} & \text{with probability } 1/2 \end{array} \right.$

Theorem 8: fix $0 < \delta < 1$, $\mathcal{A}$ is nearly isometric random mapping,

$$p \geq c_0 r(m+n)\log(mn),$$

$\delta_r(\mathcal{A}) \leq \delta$ with probability at least $1 - e^{-c_1 p}$.

References

- E. J. Candès.
  Compressive sampling.
  Proceedings of the International Congress of Mathematicians,
  Madrid, Spain, 2006

- Benjamin Recht, Maryam Fazel, and Pablo A. Parrilo.
  Guaranteed Minimum Rank Solutions to Linear Matrix Equations via Nuclear Norm Minimization.
  Submitted to *SIAM Review*