Compressed Sensing

Feb 18, 2016

Course Webpage:
http://www.cs.cornell.edu/Courses/cs4786/2016sp/
\( d \) and \( n \) so large we can’t even store in memory

Only have time to be linear in \( \text{size}(X) = n \times d \)
\[ y_t = x_t^T W \]
\[ Y = X \times \begin{bmatrix}
+1 & \ldots & -1 \\
-1 & \ldots & +1 \\
+1 & \ldots & -1 \\
\vdots & & \\
+1 & \ldots & -1 \\
K & & -1
\end{bmatrix}
\]

\[ d / \sqrt{K} \]

**Pick a Random W**
JL Lemma:
For any $\epsilon > 0$, for $K$ large enough, with high probability over draw of $W$, for all pairs of data points $i, j \in \{1, \ldots, n\}$,

$$(1 - \epsilon) \|y_i - y_j\|_2 \leq \|x_i - x_j\|_2 \leq (1 + \epsilon) \|y_i - y_j\|_2$$

$$K \approx \frac{\log(n)}{\epsilon^2}$$
Can we always recover $x_t$’s form $y_t$’s?

Answer: In general no. When $d > n$ we have an underdetermined system of linear equations.
Can we always recover $x_t$’s form $y_t$’s if $x_t$’s are sparse?

Answer: Yes!
\( \ell_0 \) (norm) of a vector \( \mathbf{x} \in \mathbb{R}^d \) measures its “sparsity” and is given by

\[
\| \mathbf{x} \|_0 = \# \text{ non-zero entries of } \mathbf{x}
\]

Examples:
When $x_t$’s are sparse, recovery is possible through random projections.

Random matrix transformations preserve distances of all sparse vectors!

This is referred to as restricted isometry property.

With this property one can successfully perform sparse recovery.
A projection matrix $W$ of size $K \times d$ possesses $(\epsilon, s)$-RIP, if for all pairs of $\mathbf{x}, \mathbf{x}' \in \mathbb{R}^d$ with $\|\mathbf{x}\|_0, \|\mathbf{x}'\|_0 \leq s$,

$$(1 - \epsilon) \left\| y - y' \right\|_2 \leq \left\| \mathbf{x} - \mathbf{x}' \right\|_2 \leq (1 + \epsilon) \left\| y - y' \right\|_2$$

where $\mathbf{y} = \mathbf{x}^\top W$ and $\mathbf{y}' = \mathbf{x}'^\top W$.

- When $K > \frac{s \log d}{\epsilon^2}$, random matrix $W$ satisfies $(\epsilon, s)$-RIP with high probability.
Algorithm for Recovery:

\[ \hat{x}_t = \arg \min_{x: y_t = x^T W} \|x\|_0 \]

Recall definition of RIP:

\( W \) possesses \((\varepsilon, s)\)-RIP, if for all pairs of \( x, x' \in \mathbb{R}^d \) with \( \|x\|_0, \|x'\|_0 \leq s \),

\[
(1 - \varepsilon) \|y - y'\|_2 \leq \|x - x'\|_2 \leq (1 + \varepsilon) \|y - y'\|_2
\]

This algorithm is computationally expensive!
$B_0(1) = \{ \mathbf{x} \in \mathbb{R}^d : \| \mathbf{x} \|_0 \leq 1, \forall i \leq d, |\mathbf{x}[i]| \leq 1 \}$
$B_1(1) = \left\{ \mathbf{x} \in \mathbb{R}^d : \sum_{i=1}^{d} |\mathbf{x}[i]| \leq 1 \right\}$
Most volume in the center with protruding tentacles reaching out.

\[ \|x\|_1 = \sum_{i=1}^{d} |x[i]| \]

Replace \( \ell_0 \) by \( \ell_1 \).
Perform random projections with large enough $K$

For recovery compute the following:

$$\tilde{x}_t = \arg\min_{x} \|x\|_1$$

subject to $y_t = x^t W$

This can be computed efficiently: linear programming problem

With high probability for all $t$'s, $\tilde{x}_t = x_t$
If $W$ has $(\epsilon, s)$-RIP then matrix

$$\Phi W$$

has $(\epsilon', s)$-RIP for invertible matrices $\Phi$.

So if data is likely to be sparse under transformation $\Phi$, i.e. $z_t = x_t^\top \Phi$ and $z_t$ is the image we see,

- Compressed sensing part is the same, Simply project using random projection
- While reconstructing, use $\Phi W$ instead

Eg. JPEG we use Fourier Transformation, JPEG 2000 Discrete wavelet transformation. If golden standard changes, only minor change in reconstruction, sensing is the same.
Compressed Sensing

- Used for image compression, instead of capturing image in large file and then compressing, directly capture low dimensional representation through random transform
- Allows fast sensing of signals without processing delays
- Random projection can be pushed to hardware level
- JPEG, JPEG 2000 techniques can be applied during sparse recovery.