CIS 700: “algorithms for Big Data”

Lecture 9: Compressed Sensing

Slides at http://grigory.us/big-data-class.html

Grigory Yaroslavtsev

http://grigory.us

Penn
University of Pennsylvania

Compressed Sensing

- Given a sparse signal $x \in \mathbb{R}^n$ can we recover it from a small number of measurements?
- Goal: design $A \in \mathbb{R}^{d \times n}$ which allows to recover any s-sparse $x \in \mathbb{R}^n$ from Ax.
- A = matrix of i.i.d. Gaussians $N(0,1)$
- Application: signals are usually sparse in some Fourier domain

Reconstruction

- Reconstruction:

$$\min ||x||_0, \text{ subject to: } Ax = b$$

- Uniqueness: If there are two s-sparse solutions x_1, x_2:

$$A(x_1 - x_2) = 0$$

then A has $2s$ linearly dependent columns

- If $d = \Omega(s^2)$ and A is Gaussian then unlikely to have linearly dependent columns
- $||x||_0$ not convex, NP-hard to reconstruct
- $||x||_0 \rightarrow ||x||_1$: $\min ||x||_1$, subject to: $Ax = b$
- When does this give sparse solutions?

Subgradient

- $\min ||x||_1$, subject to: $Ax = b$
- $||x||_1$ is convex but not differentiable
- Subgradient ∇f:
  - equal to gradient where f is differentiable
  - any linear lower bound where f is not differentiable

$$\forall x_0, \Delta x:\ f(x_0 + \Delta x) \geq f(x_0) + (\nabla f)^T \Delta x$$

- Subgradient for $||x||_1$:
  - $\nabla\left(||x||_1\right)_i = sign(x_i)$ if $x_i \neq 0$
  - $\nabla\left(||x||_1\right)_i \in [-1,1]$ if $x_i = 0$
- For all Δx such that $A\Delta x = 0$ satisfies $\nabla^T \Delta x \geq 0$
- Sufficient: $\exists w$ such that $\nabla = A^T w$ so $\nabla^T \Delta x = wA\Delta x = 0$

Exact Reconstruction Property

- **Subgradient Thm.** If $Ax_0 = b$ and there exists a subgradient ∇ for $||x||_1$ such that $\nabla = A^T w$ and columns of A corresponding to x_0 are linearly independent then x_0 minimizes $||x||_1$ and is unique.
- (Minimum): Assume $Ay = b$. Will show
$$||y||_1 \geq ||x_0||_1$$
- $z = y - x_0 \Rightarrow Az = Ay - Ax_0 = 0$
- $\nabla^T z = 0 \Rightarrow$
$$||y||_1 = ||x_0 + z|| \geq ||x_0|| + \nabla^T z = ||x_0||_1$$

Exact Reconstruction Property

- (Uniqueness): assume $\tilde{x}_0$ is another minimum
- ∇ at x_0 is also a subgradient at $\tilde{x}_0$
- $\forall \Delta x: A\Delta x = 0$:

$$\begin{aligned} \left|\left|\tilde{x}_0 + \Delta x\right|\right|_1 &= \left|\left|x_0 + \tilde{x}_0 - x_0 + \Delta x\right|\right| \\ &\geq \left|\left|x_0\right|\right|_1 + \nabla^T(\tilde{x}_0 - x_0 + \Delta x) \\ &= \left|\left|x_0\right|\right|_1 + \nabla^T(\widetilde{x_0} - x_0) + \nabla^{\mathrm{T}}\Delta x \end{aligned}$$

- $\nabla^T(\widetilde{x_0} - x_0) = w^T A(\widetilde{x_0} - x_0) = w^T(b - b) = 0$
- $\left|\left|\tilde{x}_0 + \Delta x\right|\right|_1 \geq \left|\left|x_0\right|\right|_1 + \nabla^{\mathrm{T}}\Delta x$
- $(\nabla)_{\mathrm{i}} = \mathrm{sign}((\mathrm{x}_0)_{\mathrm{i}}) = \mathrm{sign}((\tilde{x}_0)_{\mathrm{i}})$ if either is non-zero, otherwise equal to 0
- $\Rightarrow x_0$ and $\tilde{x}_0$ have same sparsity pattern
- By linear independence of columns of A: $x_0 = \widetilde{x_0}$

Restricted Isometry Property

- Matrix A satisfies restricted isometry property (RIP), if for any s-sparse x there exists δ_s:

$$(1\ -\delta_s)\left||x|\right|_2^2 \le \left||Ax|\right|_2^2 \le (1+\delta_s)\left||x|\right|_2^2$$

- Exact isometry:
  - all eigenvalues are ± 1
  - for orthogonal x, y: $x^T A^T A y = 0$
- Let A_S be the set of columns of A in set S
- **Lem:** If A satisfies RIP and $\delta_{s_1+s_2} \le \delta_{s_1} + \delta_{s_2}$:
  - For S of size s singular values of A_S in $[1-\delta_s, 1+\delta_s\ ]$
  - For any orthogonal x, y with supports of size s_1, s_2:

$$|x^T A^T A y| \le \left||x|\right|\left||y|\right|(\delta_{s_1} + \delta_{s_2})$$

Restricted Isometry Property

- **Lem:** If A satisfies RIP and $\delta_{s_1+s_2} \leq \delta_{s_1} + \delta_{s_2}$:
  - For S of size s singular values of A_S in $[1-\delta_s, 1+\delta_s\ ]$
  - For any orthogonal x, y with supports of size s_1, s_2:
    $$|x^T A^T A y| \leq 3/2\big||x|\big|\big||y|\big|(\delta_{s_1} + \delta_{s_2})$$
- W.l.o.g $\big||x|\big| = \big||y|\big| = 1$ so $\big||x+y|\big|^2 = 2$
  $$2(1-\delta_{s_1+s_2}) \leq \big||A(x+y)|\big|^2 \leq\ 2(1+\delta_{s_1+s_2})$$
  $$2\left(1-(\delta_{s_1}+\delta_{s_2})\right) \leq \big||A(x+y)|\big|^2 \leq\ 2\left(1+(\delta_{s_1}+\delta_{s_2})\right)$$
- $(1\ -\delta_{s_1}) \leq \big||Ax|\big|^2 \leq (1+\delta_{s_1})$
- $(1\ -\delta_{s_2}) \leq \big||Ay|\big|^2 \leq (1+\delta_{s_2})$

Restricted Isometry Property

- $2x^T A^T Ay$
$= (x+y)^T\, A^T A(x+y) - x^T A^T Ax - y^T A^T Ay$
$= \left|\left|A(x+y)\right|\right|^2 - \left|\left|Ax\right|\right|^2 - \left|\left|Ay\right|\right|^2$
- $2x^T A^T Ay \leq 2\left(1 + \left(\delta_{s_1} + \delta_{s_2}\right)\right) - \left(1\ - \delta_{s_1}\right) - \left(1\ - \delta_{s_2}\right) = 3(\delta_{s_1} + \delta_{s_2})$
- $x^T A^T Ay \leq \frac{3}{2}\left|\left|x\right|\right| \cdot \left|\left|y\right|\right| \cdot (\delta_{s_1} + \delta_{s_2})$

Reconstruction from RIP

- **Thm.** If A satisfies RIP with $\delta_{s+1} \leq \frac{1}{10\sqrt{s}}$ and x_0 is s-sparse and satisfies $Ax_0 = b$. Then a $\nabla(\left|\left|\cdot\right|\right|_1)$ exists at x_0 which satisfies conditions of the "subgradient theorem".
- Implies that x_0 is the unique minimum 1-norm solution to $Ax = b$.
- $S = \{i | (x_0)_i \neq 0\}, \bar{S} = \{i | (x_0)_i = 0\}$
- Find subgradient u search for w: $u = A^T w$
  - for $i \in S$: $u_i = sign(x_0)$
  - 2-norm of the coordinates in $\bar{S}$ is minimized

Reconstruction from RIP

- Let z be a vector with support S:

$$z_i = \text{sign}((x_0)_i)$$

- Let $w = A_S\left(A_S^T A_S\right)^{-1} z$
- A_S has independent columns by RIP
- For coordinates in S:

$$(A^T w)_S = A_S^T A_S\left(A_S^T A_S\right)^{-1} z = z$$

- For coordinates in $\bar{S}$:

$$(A^T w)_{\bar{S}} = A_{\bar{S}}^T A_S\left(A_S^T A_S\right)^{-1} z$$

- Eigenvalues of $A_S^T A_S$ are in $[(1-\delta_s)^2, (1+\delta_s)^2]$
- $||\left(A_S^T A_S\right)^{-1}|| \le \frac{1}{(1-\delta_s)^2}$, let $p = \left(A_S^T A_S\right)^{-1} z$, $||p|| \le \frac{\sqrt{s}}{(1-\delta_s)^2}$
- $A_S p = Aq$ where q has all coordinates in $\bar{S}$ equal 0
- For $j \in \bar{S}$: $(A^T w)_j = e_j^T A^T A q$ so $|(A^T w)_j| \le \frac{\frac{3}{2}(\delta_s+\delta_1)\sqrt{s}}{(1-\delta_s)^2} \le \frac{\frac{3}{2}(\delta_{s+1})\sqrt{s}}{(1-\delta_s)^2} \le \frac{1}{2}$