ECE 68000: MODERN AUTOMATIC CONTROL

Professor Stan Żak

Modeling sparse malicious packet drop attacks

Modeling sparse malicious packet drop attacks

- Estimating disturbances of the communication network such as noise, delays, and packet drops formulated as a sparse vector recovery problem
- Sparse $\boldsymbol{e}$—more zero entries than non-zero entries in the vector $\boldsymbol{e}$

Definition (Sparse vector recovery problem)

Estimate an unknown vector $\boldsymbol{x}$ in the linear system, $\boldsymbol{A}\boldsymbol{x} + \boldsymbol{e} = \boldsymbol{b}$, where the vector $\boldsymbol{b}$ and the matrix $\boldsymbol{A}$ are known and $\boldsymbol{e}$ models the unknown disturbances

Analysis of $\boldsymbol{Ax} + \boldsymbol{e} = \boldsymbol{b}$

Assumptions:

1. $\boldsymbol{b}$ and the full column rank matrix $\boldsymbol{A}$ are known;
2. Only a "small" number of entries of $\boldsymbol{b}$ corrupted by $\boldsymbol{e}$

Justifying the second assumption

Candes and Tao's observation: if the number of nonzero entries of the error vector is "large", then it is in general impossible to reconstruct $\boldsymbol{x}$ from $\boldsymbol{Ax} + \boldsymbol{e} = \boldsymbol{b}$ for a given $\boldsymbol{A}$ and $\boldsymbol{b}$

E. J. Candes and T. Tao, *Decoding by linear programming*, IEEE Transactions on Information Theory, Vol. 51, No. 12, pp. 4203–4215, 2005

Reconstructing $\boldsymbol{x}$ from $\boldsymbol{A}\boldsymbol{x} + \boldsymbol{e} = \boldsymbol{b}$ for a given $\boldsymbol{A}$ and $\boldsymbol{b}$

- Cannot have too many non-zero entries in $\boldsymbol{e}$ to reconstruct $\boldsymbol{x}$!
- Indeed, let $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ and suppose $m = 2n$
- Consider two distinct fixed vectors $\boldsymbol{x}$ and $\hat{\boldsymbol{x}}$
- Suppose the vector $\boldsymbol{b} \in \mathbb{R}^m$ is constructed by setting n coefficients of $\boldsymbol{b}$ equal to those of $\boldsymbol{A}\boldsymbol{x}$ and n coefficients of $\boldsymbol{b}$ equal to those of $\boldsymbol{A}\hat{\boldsymbol{x}}$
- Then we have $\boldsymbol{b} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{e} = \boldsymbol{A}\hat{\boldsymbol{x}} + \hat{\mathbf{e}}$ for some $\boldsymbol{e}$ and $\hat{\mathbf{e}}$
- In sum, the maximum number of nonzero coefficients in $\boldsymbol{e}$ should be smaller than $n = m/2$ if we are to be able to reconstruct $\boldsymbol{x}$

Cannot have too many non-zero elements in $\boldsymbol{e}$ to recover $\boldsymbol{x}$ in $\boldsymbol{A}\boldsymbol{x} + \boldsymbol{e} = \boldsymbol{b}$

Example

$$\boldsymbol{A} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}, \quad \boldsymbol{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \hat{\boldsymbol{x}} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

Then two coefficients of $\boldsymbol{b}$ equal to those of $\boldsymbol{A}\boldsymbol{x}$ and two equal to those of $\boldsymbol{A}\hat{\boldsymbol{x}}$.

Example—Contd.

Solving the equations $\boldsymbol{e} = \boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}$ and $\hat{\boldsymbol{e}} = \boldsymbol{b} - \boldsymbol{A}\hat{\boldsymbol{x}}$, we obtain

$$\boldsymbol{e} = \begin{bmatrix} 0 \\ 0 \\ -1 \\ -1 \end{bmatrix} \quad \text{and} \quad \hat{\boldsymbol{e}} = \begin{bmatrix} -1 \\ -1 \\ 0 \\ 0 \end{bmatrix}.$$

Note that both $\boldsymbol{e}$ and $\hat{\boldsymbol{e}}$ have $m = n/2$ nonzero components

- We do not know if $\boldsymbol{e}$ or $\hat{\boldsymbol{e}}$ corrupts the system
- We cannot recover $\boldsymbol{x}$ in $\boldsymbol{A}\boldsymbol{x} + \boldsymbol{e} = \boldsymbol{b}$
- We have to have less than $m = n/2$ nonzero components in $\boldsymbol{e}$ to start talking about recovering $\boldsymbol{x}$ in $\boldsymbol{A}\boldsymbol{x} + \boldsymbol{e} = \boldsymbol{b}$

How to recover $\boldsymbol{x}$ from $\boldsymbol{A}\boldsymbol{x} + \boldsymbol{e} = \boldsymbol{b}$?

- Candes and Tao's† idea:
- We can recover $\boldsymbol{x}$ if we have $\boldsymbol{e}$
- Plan: Reconstruct $\boldsymbol{e}$ and then compute $\boldsymbol{x}$

† E. J. Candes and T. Tao, *Decoding by linear programming*, IEEE Transactions on Information Theory, Vol. 51, No. 12, pp. 4203–4215, 2005

Reconstructing $\boldsymbol{e}$ from $\boldsymbol{Ax} + \boldsymbol{e} = \boldsymbol{b}$?

- Find a matrix $\boldsymbol{F} \in \mathbb{R}^{(m-n)\times m}$ such that $\boldsymbol{FA} = \boldsymbol{O}$
- Premultiply both sides of $\boldsymbol{Ax} + \boldsymbol{e} = \boldsymbol{b}$ by $\boldsymbol{F}$ to obtain, $\boldsymbol{FAx} + \boldsymbol{Fe} = \boldsymbol{Fb}$
- Let $\boldsymbol{z} = \boldsymbol{Fb}$
- Then, since $\boldsymbol{FAx} = \boldsymbol{0}$, we obtain

  $$\boldsymbol{Fe} = \boldsymbol{z},$$

  where $\boldsymbol{z}$ is known
- Thus the original problem has been reduced to reconstructing the sparse error vector $\boldsymbol{e}$ from under-determined system of equations

Finding the sparsest solution to $\boldsymbol{F}\boldsymbol{e} = \boldsymbol{z}$

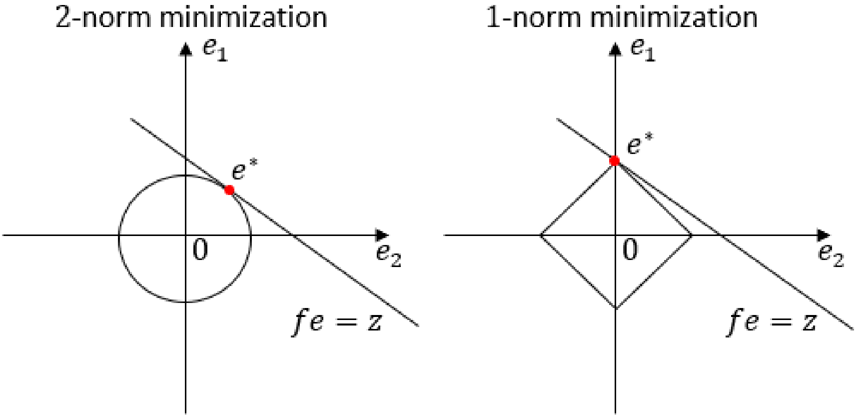
Definition (0-norm of a vector)

The 0-norm of a finite dimensional vector $\boldsymbol{x}$, denoted $\|\boldsymbol{x}\|_0$, is the number of nonzero entries in $\boldsymbol{x}$

Definition (Finding the sparsest solution problem)

$$\begin{aligned} &\min \|\boldsymbol{e}\|_0, && \boldsymbol{e} \in \mathbb{R}^m \\ &\text{subject to} && \boldsymbol{F}\boldsymbol{e} = \boldsymbol{z} \end{aligned}$$

Minimizing $\|\boldsymbol{e}\|$ subject to $\boldsymbol{F}\boldsymbol{e} = \boldsymbol{z}$



The minimal 1-norm solution is the sparsest

- D. L. Donoho and M. Elad, *For most large underdetermined systems of linear equations the minimal l_1-norm solution is also the sparsest solution*, SIAM Review, Vol. 56, No. 6, pp. 797–829, 2006
- Therefore, instead of minimizing $\|\boldsymbol{e}\|_0$, we consider an optimization problem where we minimize the 1-norm of a solution subject to the constraint, $\boldsymbol{F}\boldsymbol{e} = \boldsymbol{z}$

Finding the minimal 1-norm solution

- Since $\|\boldsymbol{e}\|_1 = \sum_{i=1}^{m} |e_i|$ is a convex function, we have a convex optimization problem,

$$\begin{array}{ll} \min \|\boldsymbol{e}\|_1, & e \in \mathbb{R}^m \\ \text{subject to} & \boldsymbol{F}\boldsymbol{e} = \boldsymbol{z} \end{array}$$

- Our objective: Find the unique solution $\boldsymbol{e}$ to the above problem
- Once we find $\boldsymbol{e}$, we can then recover $\boldsymbol{x}$

Sparse Vectors

Definition (i-sparse vector)

A vector $\boldsymbol{e}$ is i-sparse if it has at most i non-zero components, that is, $\|\boldsymbol{e}\|_0 \leq i$

Example

Let

$$\boldsymbol{e} = \begin{bmatrix} 0 \\ 0 \\ -1 \\ -1 \end{bmatrix}.$$

Then, $\|\boldsymbol{e}\|_0 = 2$

A Very Important Technical Result

- Consider an under-determined system, $\boldsymbol{F}\boldsymbol{e} = \boldsymbol{z}$, where $\boldsymbol{F}$ and $\boldsymbol{z}$ are given
- Let $\Sigma_i = \{\boldsymbol{e} : \|\boldsymbol{e}\|_0 \leq i\}$ be the set of all i-sparse vectors
- Let $\mathcal{N}(\boldsymbol{F})$ denote the null space of the matrix $\boldsymbol{F}$

Lemma

If $\Sigma_{2i} \cap \mathcal{N}(\boldsymbol{F}) = \{\mathbf{0}\}$, then any i-sparse solution of the under-determined system $\boldsymbol{F}\boldsymbol{e} = \boldsymbol{z}$ is unique

Proof of Lemma

- By contradiction: $S_1 \implies S_2 \iff \text{NOT}\,(S_1 \text{ AND NOT } S_2)$
- Suppose $\boldsymbol{e}^{(1)}$ and $\boldsymbol{e}^{(2)}$ are two different i-sparse solutions of the under-determined system $\boldsymbol{F}\boldsymbol{e} = \boldsymbol{z}$
- Then $\boldsymbol{F}(\boldsymbol{e}^{(1)} - \boldsymbol{e}^{(2)}) = \mathbf{0}$ and thus $\boldsymbol{e}^{(1)} - \boldsymbol{e}^{(2)} \in \mathcal{N}(F)$
- Since $\boldsymbol{e}^{(1)}$ and $\boldsymbol{e}^{(2)}$ are in Σ_i, we also have $\boldsymbol{e}^{(1)} - \boldsymbol{e}^{(2)} \in \Sigma_{2i}$
- Therefore $\boldsymbol{e}^{(1)} - \boldsymbol{e}^{(2)} \in \Sigma_{2i} \cap \mathcal{N}(F) = \{\mathbf{0}\}$
- It follows that we must have $\boldsymbol{e}^{(1)} = \boldsymbol{e}^{(2)}$, a contradiction, and thus an i-sparse solution of the under-determined system $\boldsymbol{F}\boldsymbol{e} = \boldsymbol{z}$ must be unique

$\square$

The Spark of a Matrix

Definition

The spark of the matrix $\boldsymbol{F}$ is the smallest number of linearly dependent columns in $\boldsymbol{F}$, that is,

$$\operatorname{spark}(\boldsymbol{F}) = \min\{\|\boldsymbol{d}\|_0 : \boldsymbol{F}\boldsymbol{d} = \boldsymbol{0}, \boldsymbol{d} \neq \boldsymbol{0}\}$$

The Spark of a Matrix—Example

Example

$$\operatorname{spark}\begin{bmatrix} 1 & 1 & 3 & 0 \\ 1 & 1 & 2 & 0 \\ 1 & 1 & 4 & 0 \\ 0 & 1 & 3 & -1 \end{bmatrix} = 3$$

Indeed

- No zero column so no set of one columns linearly dependent
- No set of two columns that are linearly dependent
- There is a set of three columns that are linearly dependent; the first, the second, and the fourth columns are linearly dependent

Some Properties of the Spark of a Matrix

Let $\boldsymbol{A}$ be an $m \times n$ matrix, where $m \geq n$.

- Then, $\text{spark}(\boldsymbol{A}) = n + 1 \iff \text{rank}(\boldsymbol{A}) = n$, that is, the spark of $\boldsymbol{A}$ equals $n + 1$ if and only if $\boldsymbol{A}$ is a full column rank matrix
- $\text{spark}(\boldsymbol{A}) = 1 \iff \boldsymbol{A}$ has a zero column
- If $\text{spark}(\boldsymbol{A}) \neq n + 1$, then

$$\text{spark}(\boldsymbol{A}) \leq \text{rank}(\boldsymbol{A}) + 1$$

Restatement of the Very Important Technical Result

Corollary:

$\text{spark}(\boldsymbol{F}) > 2i$ is equivalent to $\Sigma_{2i} \cap \mathcal{N}(\boldsymbol{F}) = \{\boldsymbol{0}\}$. Therefore, $\text{spark}(\boldsymbol{F}) > 2i$ implies that the i-sparse solution to $\boldsymbol{F}\boldsymbol{e} = \boldsymbol{z}$ is unique