EECS 16A
August 10, 2020
Lecture 7A
Grace Kuo

Today:
- Orthogonal Matching Pursuit (OMP)

Announcements:
- last lecture: tuesday
- last HW: 7A due wed., self grade due Thur.
- review sessions / OH
- exam format - free response

Overview:
- unique ← mod. 1, GE
- no solution ← least squares
- inf. solutions ← OMP
  - "not enough information"

Application of OMP: smart city

$n = 10{,}000$ big number

$$\vec{r} = \alpha_1 \vec{s}_1 + \alpha_2 \vec{s}_2 + \dots + \alpha_n \vec{s}_n$$

α_n ↑ messages

$\vec{s}_n$ ↖ device signature "gold codes" length 1023

$$\begin{bmatrix} \vec{r} \end{bmatrix} = \begin{bmatrix} \vec{s}_1 & \cdots \end{bmatrix}_{S} \begin{bmatrix} \vec{\alpha} \end{bmatrix}$$

$S \in \mathbb{R}^{1023 \times 10{,}000}$

→ we know that most devices transmit <u>zero = α</u>

↳ $\vec{\alpha}$ is sparse

→ $\langle s_i, s_j \rangle \approx 0$ if $i \neq j$

OMP helps us solve problems of the form

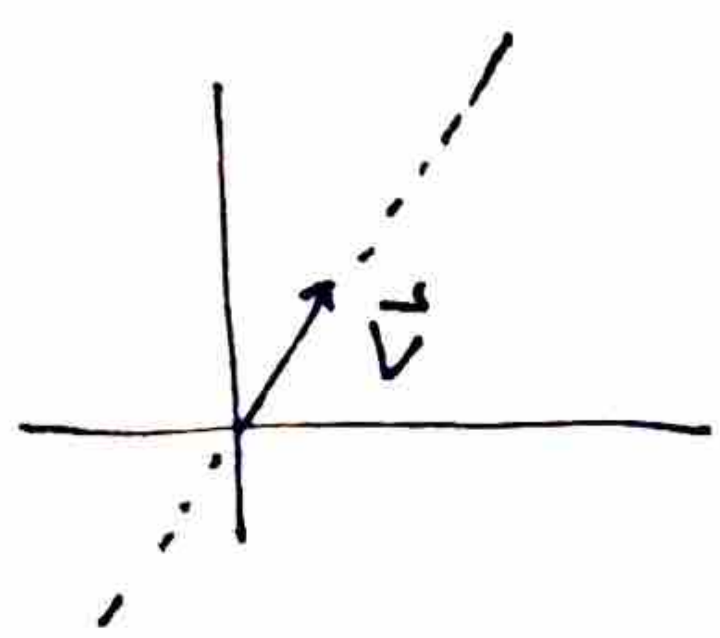
$$A\vec{x} = \vec{b} \qquad A = [\vec{a}_1 \ \vec{a}_2 \ \dots \ \vec{a}_n] \quad A \in \mathbb{R}^{m \times n}$$

where

- $m < n$ — not strict condition, case when most useful — more columns than rows
- $\vec{x}$ is spare (mosty zeros)
- $\langle a_i, a_j \rangle \approx 0$ columns are "nearly" orthogonal
- $\|a_1\| = \|a_2\| = \dots = \|\vec{a}_n\|$ — already taken care of with gold codes

---

Aside: "Normalizing vectors"



$$\alpha\vec{v}$$

$$\|\alpha\vec{v}\| = |\alpha| \, \|\vec{v}\|$$

$$\alpha = \frac{1}{\|\vec{v}\|} \qquad \|\alpha\vec{v}\| = \frac{\|\vec{v}\|}{\|\vec{v}\|} = 1$$

Back to smart city Example

$$\vec{r} = \alpha_{10}\vec{s}_{10} + \alpha_{200}\vec{s}_{200} + \alpha_5\vec{s}_5$$

Try $\langle \vec{r}, \vec{s}_1 \rangle = \langle \alpha_{10}\vec{s}_{10} + \alpha_{200}\vec{s}_{200} + \alpha_5\vec{s}_5, \vec{s}_1 \rangle$

$$= \langle \alpha_{10}\vec{s}_{10}, \vec{s}_1 \rangle + \langle \alpha_{200}\vec{s}_{200}, \vec{s}_1 \rangle + \langle \alpha_5\vec{s}_5, \vec{s}_1 \rangle$$

$$= \alpha_{10}\underbrace{\langle \vec{s}_{10}, \vec{s}_1 \rangle}_{\text{small}} + \alpha_{200}\underbrace{\langle \vec{s}_{200}, \vec{s}_1 \rangle}_{\text{small}} + \alpha_5\underbrace{\langle \vec{s}_5, \vec{s}_1 \rangle}_{\text{small}}$$

≈ 0, small + close to zero

Try $\langle \vec{r}, \vec{s}_{10} \rangle = \alpha_{10} \underbrace{\langle \vec{s}_{10}, \vec{s}_{10} \rangle}_{\text{large, } 1023} + \alpha_{200} \underbrace{\langle \vec{s}_{200}, \vec{s}_{10} \rangle}_{\text{small}} + \alpha_5 \underbrace{\langle \vec{s}_5, \vec{s}_{10} \rangle}_{\text{small}}$

$= \text{large} \approx \alpha_{10}(1023)$

Let's design an algorithm

Plan:

$\Rightarrow$ consider

$$\left.\begin{array}{l} \langle \vec{r}, \vec{s}_1 \rangle \\ \vdots \\ \langle \vec{r}, \vec{s}_{10,000} \rangle \end{array}\right\} \text{Find max abs. value}$$

$\rightarrow$ say we find $\langle \vec{r}, \vec{s}_{200} \rangle$ has the largest inner product

$\downarrow$

Declare that device 200 is transmitting

what if we get a large negative value?

$\rightarrow$ transmitting $-\alpha$

$\Rightarrow$ Guess that only device 200 is transmitting

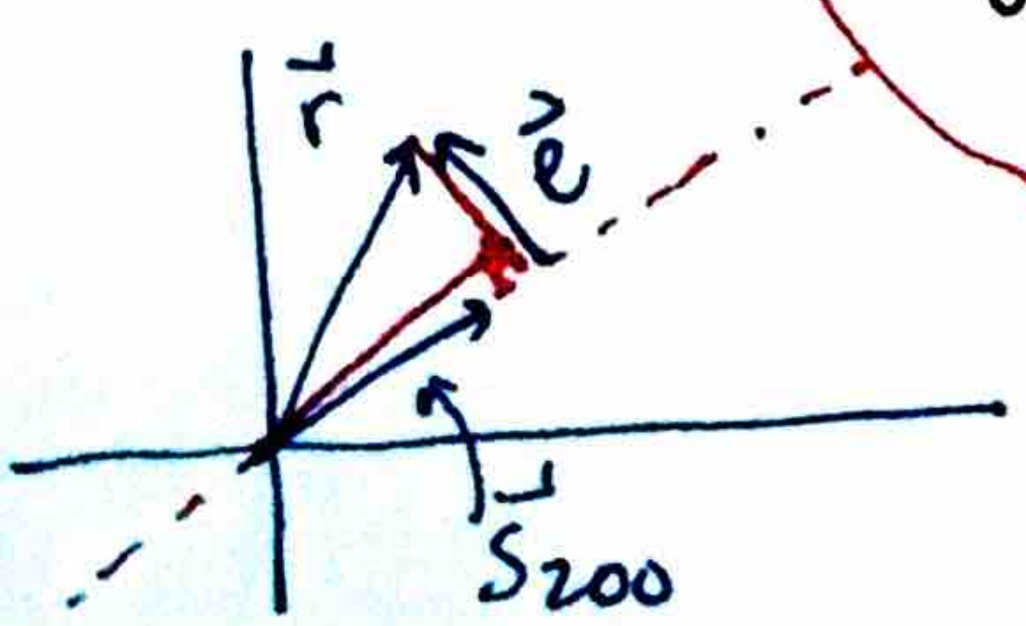
$$\vec{r} \approx \alpha_{200} \vec{s}_{200}$$

How do we find best fit α_{200} in this model?

A is 1023×1 $\longrightarrow$ $A = \begin{bmatrix} \vec{s}_{200} \end{bmatrix}$ $\quad \vec{x} = [\alpha_{200}]$ $\quad \vec{r} \approx A\vec{x}$

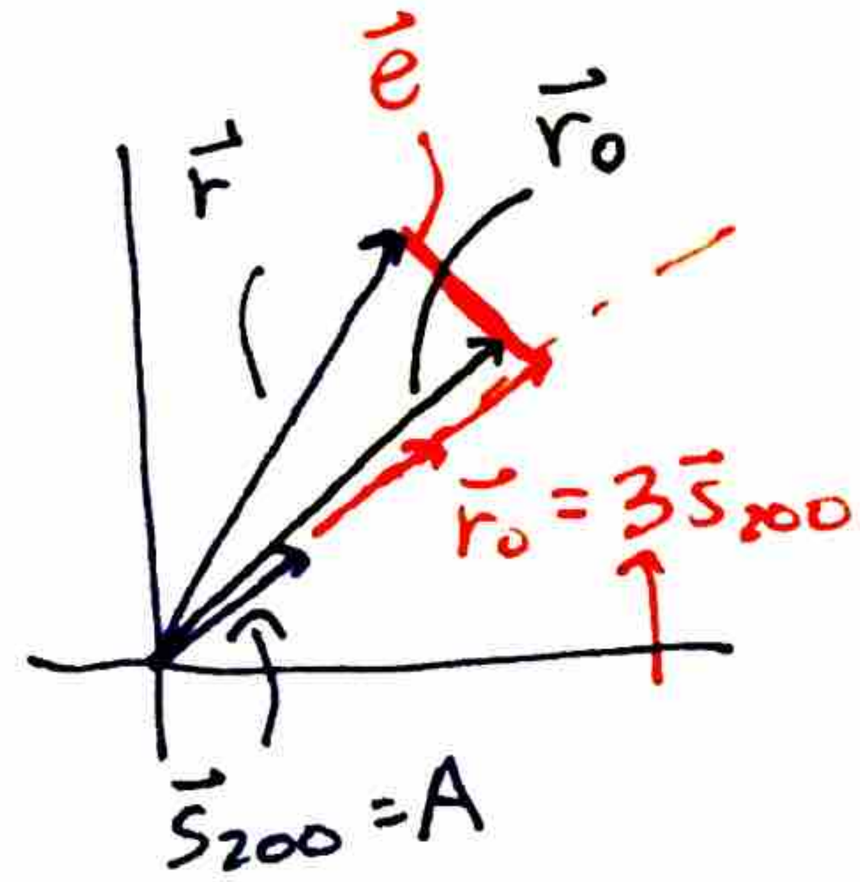
$\hat{x} = (A^T A)^{-1} A^T \vec{r} = \alpha_{200}$

$\vec{r}_0 = A\hat{x} = A(A^T A)^{-1} A^T \vec{r}$



$\langle \vec{s}_{200}, \vec{s}_{200} \rangle^{-1} \langle \vec{s}_{200}, \vec{r} \rangle$

$= \dfrac{\langle \vec{s}_{200}, \vec{r} \rangle}{\| \vec{s}_{200} \|^2}$



$\Rightarrow$ How well did we do?

look at error:

$$\vec{e} = \vec{r} - \vec{r}_0$$

$\alpha_{10}\vec{s}_{10} + \alpha_{200}\vec{s}_{200} + \alpha_5\vec{s}_5$

$\vec{r}_0 \approx \alpha_{200}\vec{s}_{200}$
$\quad\quad\quad x \quad A$

expect: $\vec{e} \approx \alpha_{10}\vec{s}_{10} + \alpha_5\vec{s}_5$

go back and repeat using error vector!

$\boxed{\text{ITERATION 2}}$

$\Rightarrow$ consider

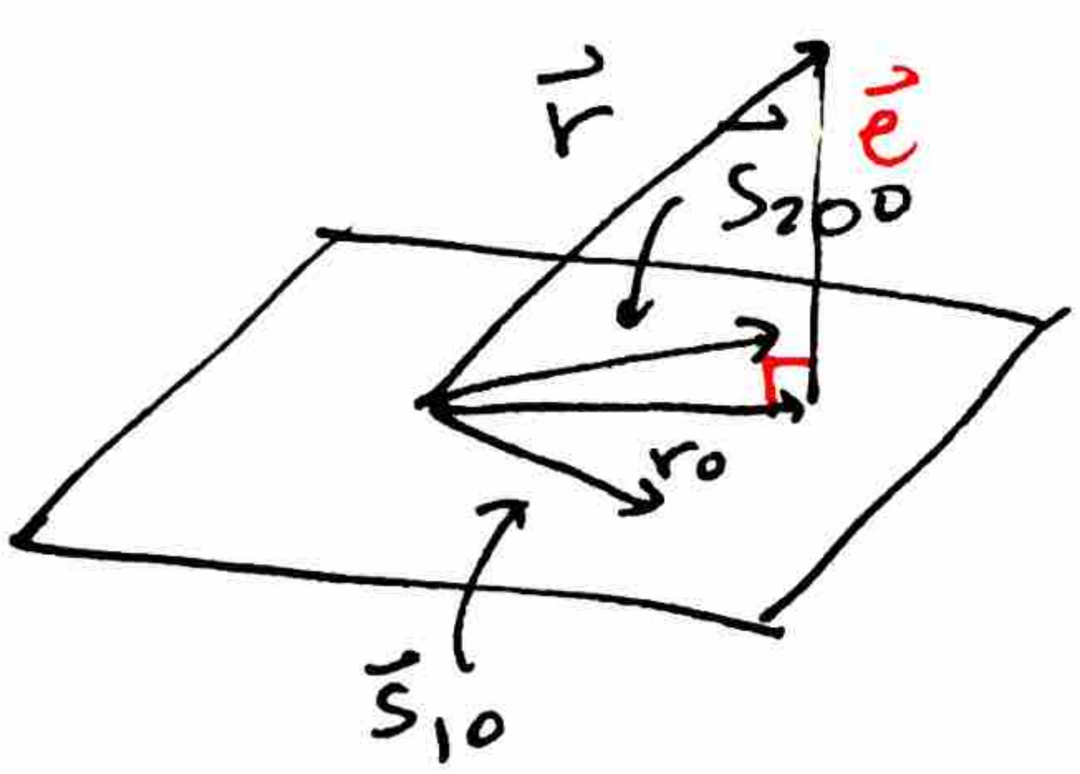
$$\left.\begin{matrix} \langle \vec{e}, \vec{s}_1 \rangle \\ \langle \vec{e}, \vec{s}_2 \rangle \\ \vdots \\ \langle \vec{e}, \vec{s}_{10,000} \rangle \end{matrix}\right\}$$

Find max abs. value $\rightarrow$ we find $\langle \vec{e}, \vec{s}_{10} \rangle$ is max

Declare that $\vec{s}_{10}$ and $\vec{s}_{200}$ are transmitting

$\Rightarrow$ Guess:

$$\vec{r} \approx \alpha_{200}\vec{s}_{200} + \alpha_{10}\vec{s}_{10}$$



$$\vec{r} \approx \underbrace{\begin{bmatrix} | & | \\ \vec{s}_{200} & \vec{s}_{10} \\ | & | \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} \alpha_{200} \\ \alpha_{10} \end{bmatrix}}_{\vec{x}}$$

$$\hat{x} = (A^T A)^{-1} A^T \vec{r} \quad \text{not } \vec{e}$$

$$\vec{r}_0 = A\hat{x}$$

$\Rightarrow$ How did we do?

$$\vec{e} = \vec{r} - \vec{r}_0 \approx \alpha_5\vec{s}_5$$

ITERATION 3

⇒ take all inner products

$\langle \vec{e}, \vec{s}_i \rangle$ → find ~~aa~~ $\vec{s}_5$ is largest $\langle \vec{e}, \vec{s}_5 \rangle$

⇒ Guess:

$$\vec{r} \approx \alpha_{200} \vec{s}_{200} + \alpha_{10} \vec{s}_{10} + \alpha_5 \vec{s}_5$$

$$\vec{r} = \underbrace{\begin{bmatrix} | & | & | \\ \vec{s}_{200} & \vec{s}_{10} & \vec{s}_5 \\ | & | & | \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} \alpha_{200} \\ \alpha_{10} \\ \alpha_5 \end{bmatrix}}_{\vec{x}}$$

$$\hat{x} = (A^T A)^{-1} A^T \vec{r}$$

$$\vec{r}_0 = A \hat{x}$$

⇒ How did we do?

$\vec{e} = \vec{r} - \vec{r}_0$ ← close to zero!
(exactly zero w/ no noise)

OMP Algorithm:

Problem: $S\vec{\alpha} = \vec{r}$ $\quad$ $S = [\vec{s}_1 \dots \vec{s}_n]$

$\vec{\alpha}$ is sparse $\quad$ $\|\vec{s}_1\| = \|\vec{s}_2\| = \dots = \|\vec{s}_n\|$

Algorithm:

initialize $\vec{e} = \vec{r}$, $A = [\ ]$, $\vec{x} = [\ ]$

on each iteration: (k)

① Take inner product of $\vec{e}$ with every column of $\vec{S}$.
Find i such that $|\langle \vec{s}_i, \vec{e} \rangle|$ is maximized.

② Update matrix A and vector $\vec{x}$

$$A = [\underset{\substack{\uparrow \\ \text{prior } A}}{A} \mid \underset{\substack{\uparrow \\ \text{new column}}}{\vec{s}_i}] \qquad \vec{x} = \begin{bmatrix} \vec{x} \\ \alpha_i \end{bmatrix} \qquad \vec{x} \in \mathbb{R}^k$$

A ~~■~~ $1023 \times k$

③ Find the best estimate of $\vec{r}$ by projecting $\vec{r}$ onto the column space of A

$$\vec{r} \approx A\vec{x} \qquad \hat{x} = (A^T A)^{-1} A^T \vec{r}$$

$$\vec{r}_0 = A\hat{x}$$

④ Update the error:

$$\vec{e} = \vec{r} - \vec{r}_0$$

Repeat until
- ~~either~~ K iterations have passed
- ~~noise is~~ error is small enough $\|e\| <$ threshhold