LEC. 11: SPARSITY, EIGENANALYSIS

Prof. Piya Pal
ECE 269
FALL 2021

AGENDA

1. SPARSE SOLUTIONS TO LINEAR EQUATIONS
   - WHY THE PROBLEM IS DIFFICULT
   - MAIN IDEA OF OMP.

2. EIGENANALYSIS

SPARSE SOLUTIONS (CONTD..)

Since $m < n$, $\mathcal{N}(A) \neq \{0\}$ (think rank-nullity theorem)

$\Rightarrow$ $y_0 = Ax$ has "infinite solutions"

Impossible to tell which of these is x_0, UNLESS Ms. Y knows "something else" about x_0.

Typically in many scenarios x_0 turns out to be "sparse" (i.e. it has many "zero" entries).

Define:

$$S_x = \{ i \in [1, n], \text{ s.t. } x_i \neq 0 \}$$

S_x is called the support set of x.

$\|S_x\|_0$ = cardinality of S_x, also called "sparsity" of x_0.

x_0 has "sparsity" s if $\|S_{x_0}\|_0 = s$.

$\Leftrightarrow$ x_0 has "s" non-zero entries.

What if you know S_{x_0}?

Suppose you know S_{x_0}, $\|S_{x_0}\|_0 = s$

For ease of exposition, let $S_{x_0} = \{1, 2, \cdots, s\}$ (i.e. S_{x_0} comprises of the first "s" columns)

Then, is finding x_0 easy?

Since we know $S_{x_0} = \{1, 2, \cdots, s\}$, the non-zero elements of x_0 are $[x_0]_1, [x_0]_2, \cdots, [x_0]_s$ and they satisfy.

$$y = \underbrace{\begin{bmatrix} a_1 & a_2 & \cdots & a_s \end{bmatrix}}_{= A_{S_{x_0}}} \begin{bmatrix} [x_0]_1 \\ \vdots \\ [x_0]_s \end{bmatrix}$$

Hence, we can solve (in unknown z)

$$y = A_{S_{x_0}} z \qquad (1)$$

to recover $[x_0]_1, \cdots, [x_0]_s$.

This recovery will be successful if $\begin{bmatrix} [x_0]_1 \\ \vdots \\ [x_0]_s \end{bmatrix}$ is the unique solution to (1)

$$\boxed{\text{This happens if and only if the "s" columns of } A_{S_{x_0}} \text{ are linearly independent.}}$$

Since S_{x_0} can correspond to any subset (of size s) of columns of A.

We will need "every set of "s" columns" of A to be linearly independent for (1) to have unique solution for every possible choice of support S_{x_0}.

If we generate elements of A independently from a Gaussian distribution, every "s" columns will be linearly independent with probability 1, if $s \leq m$.

Hence for such random A, once we know S_{x_0}, we can indeed uniquely recover x_0.

KEY DIFFICULTY.

The key difficulty therefore is to determine the support S_{x_0} of x_0. Even if we know its size (s), there are still $\binom{n}{s}$ choices for S_{x_0}

$\Rightarrow$ No algorithm with polynomial complexity can solve this problem (NP hard)

$\Rightarrow$ Is all hope lost?

RECOVERING S_{x_0}

Can we develop a polynomial-time algorithm that recovers S_{x_0} for "most realizations" of the random matrix A ?

Specifically, if we generate $A \in \mathbb{R}^{m \times n}$ with iid entries can we design a polynomial-time algorithm that recovers S_{x_0} with high probability ?

Yes, Orthogonal Matching Pursuit (OMP) is such an algorithm that ensures that the probability of succesful recovery goes to 1, exponentially in n (i.e., with overwhelming probability S_{x_0} is recovered), provided the numbe of measurements, "m" satisfies

$$\boxed{m > \phi(s, n)} \quad \text{— } (2)$$

(2) is called the "sample complexity" of support recovery using OMP. It can be shown that m scales logarithmically with n. and "almost

linearly in s", which is "near-optimal" with respect to information-theoretic limits.

BASICS OF OMP

- Start with empty set $S = \{\}$.
- Until $|S| = s$ (or appropriate stopping rule).
  - Greedily find an index (call it j) of the support
  - Append it to S, i.e
  $$S \longleftarrow S \cup \{j\}$$

KEY INSIGHT FROM JOEL TROPP.

  - Project y onto the subspace spanned by $\{a_i, i \in S\}$ and compute the projection error (or residual)
  - Repeat greedy step with residual.

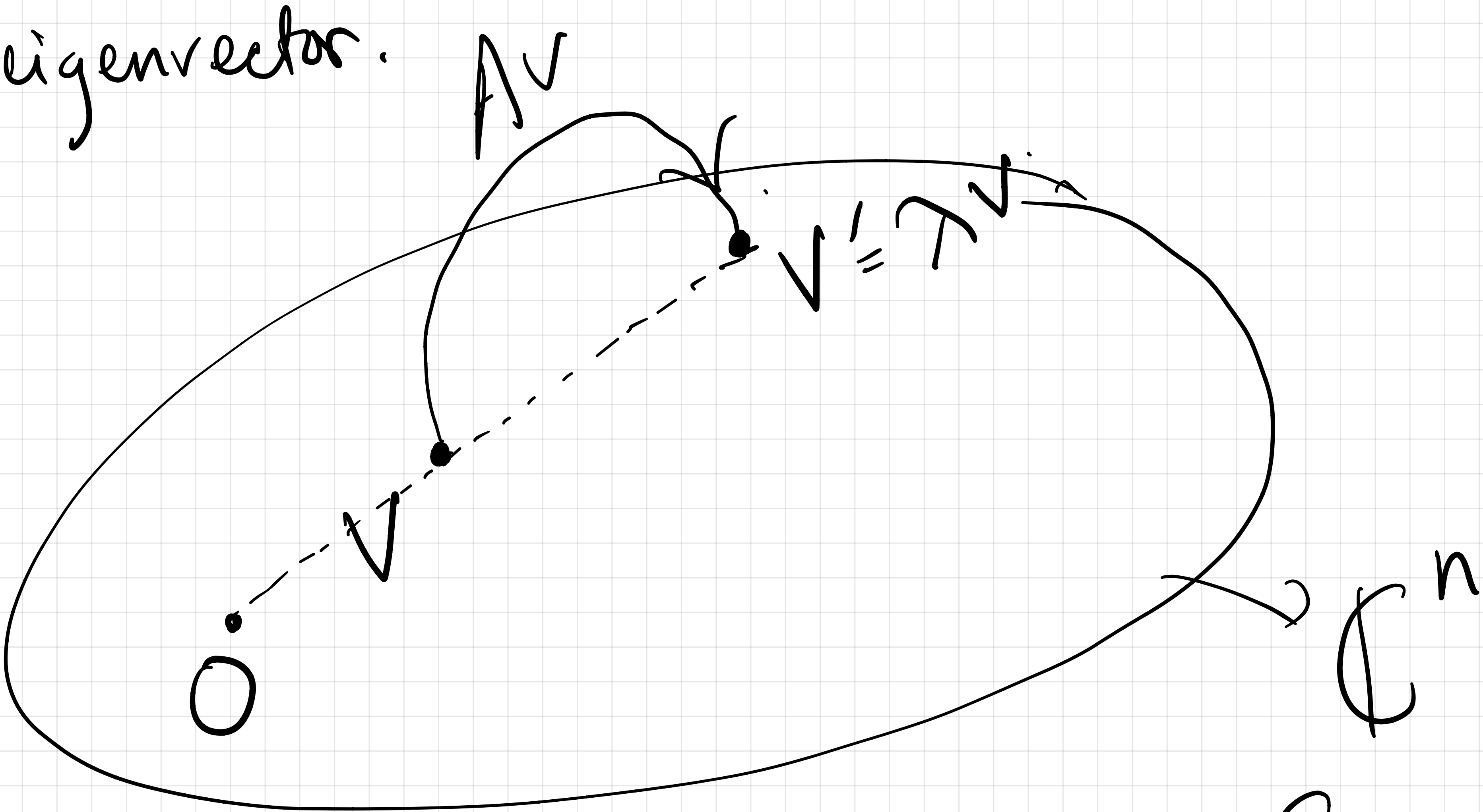
EIGENANALYSIS

Let $A \in \mathbb{C}^{n \times n}$, (complex square matrix)

If a scalar $\lambda \in \mathbb{C}$ and $v \in \mathbb{C}^n$ $(v \neq 0)$ such that

$$Av = \lambda v$$

then λ is called an eigenvalue of A and v is called the "corresponding" eigenvector.



Given any vector $u = \alpha v$, $\alpha \in \mathbb{C}$

If v is eigenvector of A, then for $\alpha \neq 0$,

$u = \alpha v$ is also an eigenvector

of A corresponding to λ

ie if $Av = \lambda v$

then $A(\alpha v) = \alpha A v = \alpha \lambda v$

$= \lambda(\alpha v)$

$\Rightarrow$ There are infinite eigenvectors corresponding to same eigenvalue.

EIGENVALUES & CHARACTERISTIC POLYNOMIAL

λ is an eigenvalue of A if and only if

$$Av = \lambda v, \text{ for some } v \neq 0.$$

$$\Leftrightarrow (A - \lambda I)v = 0 \text{ for some } v \neq 0$$

$\Leftrightarrow (A-\lambda I)$ has a non-trivial Null space.

Since $A - \lambda I$ is a "<u>square</u>" matrix

$(A - \lambda I)$ has a non-trivial Null space

$$\Leftrightarrow \det(A - \lambda I) = 0$$

$$\boxed{\lambda \text{ is an eigenvalue of } A \in \mathbb{C}^{n \times n} \text{ if and only if } \det(A - \lambda I) = 0}$$

$\det(A-\lambda I)$ is a polynomial in λ, of degree "n"

$$\Rightarrow \boxed{\det(A-\lambda I) = c_0 + c_1\lambda + c_2\lambda^2 + \cdots + c_n\lambda^n}$$

(Here $c_0, c_1, \ldots, c_n$ depend on A).

$\Rightarrow \det(A-\lambda I)=0$ has n "complex" roots.

These "n" roots are the "n" eigenvalues of A.

$$\boxed{\begin{array}{l}\text{A matrix } A \in \mathbb{C}^{n\times n} \text{ has } n \text{ (complex) eigenvalues, given by the roots of} \\ \text{its "characteristic polynomial", i.e roots of} \\ \qquad \det(A-\lambda I) = 0\end{array}}$$

EIGENSPACE

Let λ be an eigenvalue of $A \in \mathbb{C}^{n\times n}$.

Consider the following set :-

$$E_\lambda = \{ v \neq 0, \text{ s.t. } Av = \lambda v \}$$

= set of all eigenvectors associated with the <u>same</u> eigenvalue λ.

Suppose $v_1 \in E_\lambda$, $v_2 \in E_\lambda \Rightarrow \begin{cases} Av_1 = \lambda v_1 \\ Av_2 = \lambda v_2 \end{cases}$

Let $V_3 = V_1 + V_2$

$$Av_3 = A(v_1 + v_2) = Av_1 + Av_2$$
$$= \lambda v_1 + \lambda v_2$$
$$= \lambda(v_1 + v_2) = \lambda v_3$$

$\Rightarrow$ v_3 is also an eigenvector of A.

Similarly show that $\alpha_1 v_1 + \alpha_2 v_2$ is also an eigenvector of A corresponding to λ (provided $\alpha_1, \alpha_2 \in \mathbb{C}$ are chosen so that $\alpha_1 v_1 + \alpha_2 v_2 \neq 0$)

Note:

- $E_\lambda \cup \{0\}$ is a subspace. (convince yourselves)

$$\boxed{\bullet\ E_\lambda \cup \{0\} = \mathcal{N}(A - \lambda I)}$$

$\mathcal{N}(A-\lambda I)$ is called the eigenspace corresponding to the eigenvalue λ.

dim$(\mathcal{N}(A-\lambda I))$ is called the "geometric multiplicity" of the eigenvalue λ, and is denoted by $g_A(\lambda)$.

(Note $g_A(\lambda) \geq 1$)