

Ultimately : Singular Value Decomposition (SVD)

Lec 35: Spectral Theory Eigenvalue Decomposition

Eigenvalue Decomposition (EVD)

Eigenvalue Problem

$n \times n$ matrix w/
complex entries

if $\lambda \in \mathbb{R}$,
stretch or compress

Give a square matrix $A \in \mathbb{C}^{n \times n}$,
find $\lambda \in \mathbb{C}$ and $\vec{v} \in \mathbb{C}^n$ such that
 $A\vec{v} = \lambda\vec{v}$ nonzero

Geometry

$$A\vec{v} =$$

"image of \vec{v}
under A "

$$\lambda\vec{v}$$

"scalar multiple
of \vec{v} "

- λ : eigenvalue of A
 - \vec{v} : eigenvector of A
- } eigenpair.

This eqn has a nontrivial soln \vec{v}
iff $A - \lambda I$ is singular. In other words,

$$A\vec{v} - \lambda\vec{v} = \vec{0}$$

$$(A - \lambda I)\vec{v} = \vec{0}$$

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

(characteristic
eqn)

poly. in λ (deg-n)

Terminology

- $A \in \mathbb{C}^{n \times n}$ has n eigenvalues

counting multiplicity:

$$\lambda_1, \lambda_2, \dots, \lambda_n$$

- The set of all eigenvalues of A is called the spectrum of A .

$$\text{Spec}(A) = \{\lambda_1, \dots, \lambda_n\}$$

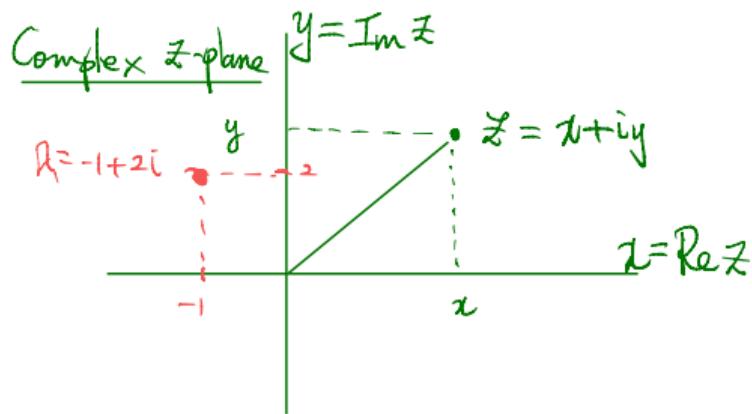
- $\max_{1 \leq j \leq n} |\lambda_j|$ = spectral radius

Digression: Geometry of Comp. numbers

$$z = x + iy \in \mathbb{R}$$

$$(x, y \in \mathbb{R})$$

$$i = \sqrt{-1}$$



Eigenvalue Decomposition (EVD)

- $A \in \mathbb{C}^{n \times n}$
- $\lambda_1, \lambda_2, \dots, \lambda_n$: eigenvalues
- $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$: eigenvectors

$$\Rightarrow AV = VD \text{ (general)}$$

- V : columns are eigenvectors
- D : diagonal matrix w/ eigenvalue on the diagonal.

$$A\vec{v}_j = \lambda_j \vec{v}_j \quad \text{for } j=1, \dots, n$$

$$\Rightarrow \begin{matrix} & \left[\begin{array}{c|c|c|c} A\vec{v}_1 & A\vec{v}_2 & \cdots & A\vec{v}_n \end{array} \right] \\ \begin{matrix} n \times n \text{ matrix} \\ V \end{matrix} & = \left[\begin{array}{c|c|c|c} \lambda_1 \vec{v}_1 & \lambda_2 \vec{v}_2 & \cdots & \lambda_n \vec{v}_n \end{array} \right] \end{matrix}$$

$\underbrace{\hspace{10em}}_{D}$

$$\Rightarrow A \left[\begin{array}{c|c|c|c} \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \end{array} \right] = \left[\begin{array}{c|c|c|c} \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \end{array} \right] \left[\begin{array}{cccc} \lambda_1 & & & 0 \\ & \lambda_2 & & \vdots \\ 0 & & \ddots & \lambda_n \end{array} \right]$$

In a special case where eigenvectors
of A are linearly independent, then \rightarrow rank (V) = n
(full rank)

$$V = [\vec{v}_1 \mid \vec{v}_2 \mid \cdots \mid \vec{v}_n]$$

$\nwarrow P \quad \nearrow$

Linearly independent

is nonsingular (or invertible).

Consequence: $AV = VD$ (true always for square A)

$$\Rightarrow A = V D V^{-1} \quad (\text{EVD})$$

Eigenvalue Decomposition (cont')

eigenvectors of A are lin-indep.

If A has an EVD, we say that A is **diagonalizable**; otherwise **nondiagonalizable**.

Theorem 1 (Diagonalizability)

If $A \in \mathbb{C}^{n \times n}$ has n distinct eigenvalues, then A is diagonalizable.

Notes.

- Let $A, B \in \mathbb{C}^{n \times n}$. We say that B is **similar** to A if there exists a nonsingular matrix X such that

$$B = XAX^{-1}.$$

- So **diagonalizability is similarity to a diagonal matrix**.
- Similar matrices share the same eigenvalues.

↳ Exercise : Confirm this .

(Review your 2568 notes / text.)

Calculating EVD in MATLAB

- $E = \text{eig}(A)$
produces a column vector E containing the eigenvalues of A .
- $[V, D] = \text{eig}(A)$
produces V and D in an EVD of A , $A = VDV^{-1}$.

Understanding EVD: Change of Basis

syn. to an invertible matrix.

Let $X \in \mathbb{C}^{n \times n}$ be a nonsingular matrix.

In part., columns of X are linearly independent.

- The columns $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ of X form a basis of \mathbb{C}^n .
- Any $\mathbf{z} \in \mathbb{C}^n$ is uniquely written as

$$\begin{array}{c} X \text{ nonsingular} \\ \hline X^{-1} \vec{z} = \vec{u} \end{array} \quad \boxed{\mathbf{z} = X\mathbf{u} = u_1\mathbf{x}_1 + u_2\mathbf{x}_2 + \cdots + u_n\mathbf{x}_n. \text{ (Linear comb. of basis vectors)}} \\ = \left[\begin{array}{c|c|c|c} \vec{x}_1 & \vec{x}_2 & \cdots & \vec{x}_n \end{array} \right] \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$$

- The scalars u_1, \dots, u_n are called the coordinates of \mathbf{z} with respect to the columns of X .
- The vector $\mathbf{u} = X^{-1}\mathbf{z}$ is the representation of \mathbf{z} with respect to the basis consisting of the columns of X .



Upshot

Left-multiplication by X^{-1} performs a **change of basis** into the coordinates associated with the columns of X .

Understanding EVD: Change of Basis (cont')

V -basis = basis consisting of columns of V
 $= \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$

Suppose $A \in \mathbb{C}^{n \times n}$ has an EVD $A = VDV^{-1}$. Then, for any $\mathbf{z} \in \mathbb{C}^n$, $\mathbf{y} = A\mathbf{z}$ can be written as (i.e., A is diagonalizable)

$$V^{-1}\mathbf{y} = DV^{-1}\mathbf{z}$$

(coord. of \vec{y}
w.r.t. V -basis)

(coord. of \vec{z}
w.r.t. V -basis)

$$\begin{aligned}\vec{y} &= A\vec{z} \\ &= VDV^{-1}\vec{z} \\ V^{-1}\vec{y} &= V^{-1}VDV^{-1}\vec{z}\end{aligned}$$

Interpretation

The matrix A is a diagonal transformation in the coordinates with respect to the V -basis.

basis consisting of eigenvectors

* diagonal transformation : coordinates are rescaled independently

For example

$$D : \vec{e}_1 \mapsto \vec{e}_1 \quad (D\vec{e}_1 = \vec{e}_1)$$

$$D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

$$D : \vec{e}_2 \mapsto 5\vec{e}_2$$

$$D : \vec{e}_3 \mapsto 2\vec{e}_3$$

What is EVD good for?

$$[V, D] = \text{eig}(A);$$

Q. Why do we like diagonalizable matrices?

$$A = VDV^{-1}, \quad V \text{ nonsingular, } D \text{ diagonal.}$$

Situation: Matrix powers.

$$\bullet A^2 = (VDV^{-1})(VDV^{-1})$$

$$= V D \underbrace{(V^{-1} V)}_{\text{I}} D V^{-1}$$

$$= V D^2 V^{-1}$$

$$\begin{aligned} \bullet A^3 &= (VDV^{-1}) \underbrace{(VDV^{-1})}_{\text{I}} \underbrace{(VDV^{-1})}_{\text{I}} \\ &= V D^3 V^{-1} \end{aligned}$$

• In general,

$$A^k = V D^k V^{-1}$$

$$\left| \begin{array}{l} D = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} \\ D^k = \begin{bmatrix} \lambda_1^k & & 0 \\ & \ddots & \\ 0 & & \lambda_n^k \end{bmatrix} \end{array} \right.$$

Lec 36: Spectral Theory Singular Value Decomposition (SVD)

Singular Value Decomposition: Overview

Singular Value Decomposition

$$A = LU, \quad A = QR, \quad A = VDV^{-1}$$

Theorem 1 (SVD)

Let $A \in \mathbb{C}^{m \times n}$. Then A can be written as

$$A = U\Sigma V^*, \quad (\text{SVD})$$

where $U \in \mathbb{C}^{m \times m}$ and $V \in \mathbb{C}^{n \times n}$ are unitary and $\Sigma \in \mathbb{R}^{m \times n}$ is diagonal. If A is real, then so are U and V .

- The columns of U are called the **left singular vectors** of A ;
- The columns of V are called the **right singular vectors** of A ;
- The diagonal entries of Σ , written as $\sigma_1, \sigma_2, \dots, \sigma_r$, for $r = \min\{m, n\}$, are called the **singular values** of A and they are nonnegative numbers ordered as

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r \geq 0.$$

$$A = V D V^{-1}$$

SVD (real)

$A \in \mathbb{R}^{m \times n}$, $m > n$ (more rows than cols; tall rect.)

$$A = U \Sigma V^T$$

where

- $U \in \mathbb{R}^{m \times m}$ orthogonal ($U^T U = I \in \mathbb{R}^{m \times m}$)
- $\Sigma \in \mathbb{R}^{m \times n}$ diagonal
- $V \in \mathbb{R}^{n \times n}$ orthogonal ($V^T V = I \in \mathbb{R}^{n \times n}$)

Orthogonal matrices $\rightarrow QR$.

$$\begin{matrix} V^{-1} \\ \parallel \\ V^T \end{matrix}$$

Cartoon View

$$A_{m \times n} = U_{m \times m} \left(\begin{array}{c|c|c|c} \sigma_1 & & & 0 \\ & \ddots & & \\ & & \sigma_n & 0 \end{array} \right) V_{n \times n}^T$$

Terms

- $\vec{u}_1, \dots, \vec{u}_m$: left singular vectors
- $\vec{v}_1, \dots, \vec{v}_n$: right singular vectors
- $\sigma_1, \dots, \sigma_n$: singular values

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n \geq 0$$

Generalization to Complex SVD

$A \in \mathbb{C}^{m \times n}$, $m > n$

$$A = U \Sigma V^*$$

where

- $U \in \mathbb{C}^{m \times m}$ unitary
- $\Sigma \in \mathbb{R}^{m \times n}$ diagonal
- $V \in \mathbb{C}^{n \times n}$ unitary

Note 1

$$V^* = (\bar{V})^T$$

- Conjugate transpose of V
- a.k.a. the hermitian of V .

Note 2

A complex $U \in \mathbb{C}^{m \times m}$ is unitary

$$\text{if } U^* U = I.$$

(columns of U are orthonormal.)

$\vec{u}^* \vec{v}$ inner product in complex vectors

Singular Value Decomposition (cont')

Thick vs Thin SVD

Suppose that $m > n$ and observe that:

$$A = U \Sigma V^* \quad (\text{thick SVD})$$

Note that $V^* V = I$,

$$AV = U \Sigma$$

$$= \begin{array}{c|c} \vec{u}_1 \dots \vec{u}_n & \vec{u}_{n+1} \dots \vec{u}_m \\ \hline \overbrace{\quad}^U & \overbrace{\quad}^{U^\perp} \end{array} \begin{array}{c} \sum \\ \left\{ \begin{array}{l} n \text{ row} \\ 0 \end{array} \right\} \end{array} = \widehat{U} \widehat{\Sigma}$$

Thus, $A = \widehat{U} \widehat{\Sigma} V^* \quad (\text{thin SVD}).$

SVD in MATLAB

- Thick SVD: $[U, S, V] = \text{svd}(A);$
- Thin SVD: $[U, S, V] = \text{svd}(A, 0);$

Understanding SVD

Geometric Perspective

Write $A = U\Sigma V^*$ as $AV = U\Sigma$:

$$A\mathbf{v}_k = \sigma_k \mathbf{u}_k, \quad k = 1, \dots, r = \min\{m, n\}.$$

The image of the unit sphere under any $m \times n$ matrix is a hyperellipse.

Algebraic Perspective

Alternately, note that $\mathbf{y} = A\mathbf{z} \in \mathbb{C}^m$ for any $\mathbf{z} \in \mathbb{C}^n$ can be written as

$$(U^*\mathbf{y}) = \Sigma(V^*\mathbf{z}).$$

Any matrix $A \in \mathbb{C}^{m \times n}$ can be viewed as a diagonal transformation from \mathbb{C}^n (source space) to \mathbb{C}^m (target space) with respect to suitably chosen orthonormal bases for both spaces.

SVD vs. EVD

Recall that a diagonalizable $A = VDV^{-1} \in \mathbb{C}^{n \times n}$ satisfies

$$\mathbf{y} = A\mathbf{z} \quad \longrightarrow \quad \left(V^{-1}\mathbf{y}\right) = D\left(V^{-1}\mathbf{z}\right).$$

This allowed us to view any diagonalizable square matrix $A \in \mathbb{C}^{n \times n}$ as a diagonal transformation from \mathbb{C}^n to itself¹ with respect to the basis formed by a set of eigenvector of A .

Differences.

- **Basis:** SVD uses two ONBs (left and right singular vectors); EVD uses one, usually non-orthogonal basis (eigenvectors).
- **Universality:** all matrices have an SVD; not all matrices have an EVD.
- **Utility:** SVD is useful in problems involving the behavior of A or A^+ ; EVD is relevant to problems involving A^k .

¹The source and the target spaces of the transformation coincide.