Spectral Theory

Contents

Preliminary

Complex Numbers

Complex Numbers

In what follows, we assume all scalars, vectors, and matrices may be complex.

Notation.

- $\mathbb{R}$: the set of all real numbers
- $\mathbb{C}$: the set of all complex numbers, *i.e.*,

$$\{z = x + iy \,|\, x, y \in \mathbb{R}\} \quad \text{where } i = \sqrt{-1}.$$

Complex Numbers in MATLAB

Let $z = x + iy \in \mathbb{C}$.

MATLAB	Name	Notation
`real(z)`	real part of z	$\operatorname{Re} z$
`imag(z)`	imaginary part of z	$\operatorname{Im} z$
`conj(z)`	conjugate of z	$\overline{z}$
`abs(z)`	modulus of z	$\lvert z\rvert$
`angle(z)`	argument of z	$\arg(z)$

Euler's Formula

- Recall that the Maclaurin series for e^t is

$$e^t = 1 + t + \frac{t^2}{2} + \cdots + \frac{t^n}{n!} + \cdots = \sum_{n=0}^{\infty} \frac{t^n}{n!}, \quad -\infty < t < \infty.$$

- Replacing t by it and separating real and imaginary parts (using the cyclic behavior of powers of i), we obtain

$$e^{it} = \underbrace{\sum_{k=0}^{\infty} \frac{(-1)^k t^{2k}}{(2k)!}}_{\cos(t)} + i \underbrace{\sum_{k=0}^{\infty} \frac{(-1)^k t^{2k+1}}{(2k+1)!}}_{\sin(t)}$$

- The result is called the **Euler's formula**.

$$\boxed{e^{it} = \cos(t) + i\sin(t).}$$

Polar Representation and Complex Exponential

- **Polar representation:** A complex number $z = x + iy \in \mathbb{C}$ can be written as $z = re^{i\theta}$ where

$$r = |z|, \quad \tan\theta = \frac{y}{x}.$$

- **Complex exponentiation:**

$$e^z = e^{x+iy} = e^x e^{iy} = e^x \left(\cos y + i \sin y\right).$$

Complex Arrays

Complex Vectors

Denote by $\mathbb{C}^n = \mathbb{C}^{n\times 1}$ the space of all column vectors of n *complex* elements.

- The **hermitian** or **conjugate transpose** of $\mathbf{u} \in \mathbb{C}^n$ is denoted by $\mathbf{u}^*$:

$$\mathbf{u}^* \in \mathbb{C}^{1\times n}.$$

- The inner product of $\mathbf{u}, \mathbf{v} \in \mathbb{C}^n$ is defined by

$$\mathbf{u}^*\mathbf{v} = \sum_{k=1}^{n} \overline{u}_k v_k.$$

The 2-norm for complex vectors is defined in terms of this inner product:

$$\|\mathbf{u}\|_2^2 = \mathbf{u}^*\mathbf{u}.$$

Complex Matrices

Denote by $\mathbb{C}^{m \times n}$ the space of all complex matrices with m rows and n columns.

- The **hermitian** or conjugate transpose of $A \in \mathbb{C}^{m \times n}$ is denoted by A^*:

$$A^* = \left(\overline{A}\right)^{\mathrm{T}} = \overline{\left(A^{\mathrm{T}}\right)} \in \mathbb{C}^{n \times m}.$$

- A **unitary** matrix is a complex analogue of an orthogonal matrix. If $U \in \mathbb{C}^{n \times n}$ is unitary, then

$$U^* U = U U^* = I$$

and

$$\|U\mathbf{z}\|_2 = \|\mathbf{z}\|_2 , \quad \text{for any } \mathbf{z} \in \mathbb{C}^n.$$

Complex Matrices: Some Analogies

	Real	Complex
Norm	$\|\mathbf{v}\|_2 = \sqrt{\mathbf{v}^{\mathrm{T}}\mathbf{v}}$	$\|\mathbf{u}\|_2 = \sqrt{\mathbf{u}^*\mathbf{u}}$
Symmetry	$S^{\mathrm{T}} = S$ (symmetric matrix)	$S^* = S$ (hermitian matrix)
Orthonormality	$Q^{\mathrm{T}}Q = I$ (orthogonal matrix)	$U^*U = I$ (unitary matrix)
Householder	$H = I - \dfrac{2}{\mathbf{v}^{\mathrm{T}}\mathbf{v}}\mathbf{v}\mathbf{v}^{\mathrm{T}}$	$H = I - \dfrac{2}{\mathbf{u}^*\mathbf{u}}\mathbf{u}\mathbf{u}^*$

Eigenvalue Decomposition

Eigenvalue Decomposition

Eigenvalue Problem

Find a scalar **eigenvalue** λ and an associated nonzero **eigenvector** $\mathbf{v}$ satisfying

$$A\mathbf{v} = \lambda\mathbf{v}.$$

- The **spectrum** of A is the set of all eigenvalues; the **spectral radius** is $\max_j |\lambda_j|$.
- The problem is equivalent to
- An eigenvalue of A is a root of the **characteristic polynomial**

Eigenvalue Decomposition (cont')

Let $A \in \mathbb{C}^{n\times n}$ and suppose that $A\mathbf{v}_k = \lambda_k \mathbf{v}_k$ for $k \in \mathbb{N}[1, n]$.

- Then

- If V is nonsingular, we can further write

  which is called an **eigenvalue decomposition (EVD)** of A. If $\mathbf{v}$ is an eigenvector of A, then so is $c\mathbf{v}$, $c \neq 0$. Thus an EVD is not unique.

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.

Calculating EVD in MATLAB

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

Understanding EVD: Change of Basis

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

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

$$\mathbf{z} = X\mathbf{u} = u_1\mathbf{x}_1 + u_2\mathbf{x}_2 + \cdots + u_n\mathbf{x}_n.$$

  - The scalars $u_1, \ldots, u_n$ 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')

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

Interpretation

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

What Is EVD Good For?

Suppose $A \in \mathbb{C}^{n\times n}$ has an EVD $A = VDV^{-1}$.

- Economical computation of powers A^k:

- Analyzing convergence of iterates $(\mathbf{x}_1, \mathbf{x}_2, \ldots)$ constructed by

$$\mathbf{x}_{j+1} = A\mathbf{x}_j, \quad j = 1, 2, \ldots$$

Conditioning of Eigenvalues

Theorem 2 (Bauer-Fike)

Let $A \in \mathbb{C}^{n\times n}$ be diagonalizable, $A = VDV^{-1}$, with eigenvalues $\lambda_1, \ldots, \lambda_n$. If μ is an eigenvalue of $A + \delta A$ for a complex matrix δA, then

$$\min_{1\leqslant j\leqslant n} \left|\mu - \lambda_j\right| \leqslant \kappa_2(V) \left\|\delta A\right\|_2 .$$

Singular Value Decomposition

Singular Value Decomposition: Overview

Singular Value Decomposition

Theorem 3 (SVD)

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

$$A = U\Sigma V^*, \tag{SVD}$$

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 $\sigma_1, \sigma_2, \ldots, \sigma_r$, for $r = \min\{m, n\}$, are called the **singular values** of A and they are nonnegative numbers ordered as

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

Singular Value Decomposition (cont')

Thick vs Thin SVD

Suppose that $m > n$ and observe that:

SVD in MATLAB

- Thick SVD: `[U,S,V] = svd(A);`
- Thin SVD: `[U,S,V] = svd(A, 0);`

Understanding SVD

Geometric Perspective

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

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

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

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

> 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 itself[1] 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.

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

Properties of SVD

Properties of SVD

SVD and the 2-Norm

Theorem 4

Let $A \in \mathbb{C}^{m \times n}$ *have an SVD* $A = U\Sigma V^*$. *Then*

1. $\|A\|_2 = \sigma_1$ *and* $\|A\|_F = \sqrt{\sigma_1^2 + \sigma_2^2 + \cdots + \sigma_r^2}$.
2. *The rank of* A *is the number of nonzero singular values.*
3. *Let* $r = \min\{m, n\}$. *Then*

$$\kappa_2(A) = \|A\|_2 \|A^+\|_2 = \frac{\sigma_1}{\sigma_r}.$$

Connection to EVD

Let $A = U\Sigma V^* \in \mathbb{C}^{m\times n}$ and $B = A^*A$. Observe that

- $B \in \mathbb{C}^{n\times n}$ is a *hermitian matrix*[2], *i.e.*, $B^* = B$.
- B has an EVD:

  - The squares of singular values of A are eigenvalues of B.
  - An EVD of $B = A^*A$ reveals the singular values and a set of right singular vectors of A.

[2]This is the $\mathbb{C}$-extension of real symmetric matrices.

Theorem 5

The nonzero singular values of $A \in \mathbb{C}^{m \times n}$ are the square roots of the nonzero eigenvalues of A^*A or AA^*.

Unitary Diagonalization and SVD

Unitary Diagonalization of Hermitian Matrices

The previous discussion is relevant to hermitian matrices constructed in a specific manner. For a generic hermitian matrix, we have the following result.

Theorem 6 (Spectral Decomposition)

Let $A \in \mathbb{C}^{n\times n}$ be hermitian. Then A has a unitary diagonalization

$$A = VDV^{-1},$$

where $V \in \mathbb{C}^{n\times n}$ is unitary and $D \in \mathbb{R}^{n\times n}$ is diagonal.

In words, a hermitian matrix (or symmetric matrix) has a complete set of orthonormal eigenvectors and all its eigenvalues are real.

Notes on Unitary Diagonalization and Normal Matrices

- A unitarily diagonalizable matrix $A = VDV^{-1}$ with $D \in \mathbb{C}^{n \times n}$, is called a **normal matrix**[3]. All hermitian matrices are normal.
- Let $A = VDV^{-1} \in \mathbb{C}^{n \times n}$ be normal. Since $\kappa_2(V) = 1$ (why?), Bauer-Fike implies that eigenvalues of A can be changed by no more than $\|\delta A\|_2$.

[3]Usual defintion: $A \in \mathbb{C}^{n \times n}$ is normal if $AA^* = A^*A$.

Unitary Diagonalization and SVD

Theorem 7

Let $A \in \mathbb{C}^{n\times n}$ be hermitian. Then the singular values of A are the absolute values of the eigenvalues of A.

Precisely, if $A = VDV^{-1}$ is a unitary diagonalization of A, then

$$A = \big(V\,\mathrm{sign}(D)\big)\,|D|\,V^*$$

is an SVD, where

$$\mathrm{sign}(D) = \begin{bmatrix} \mathrm{sign}(d_1) & & \\ & \ddots & \\ & & \mathrm{sign}(d_n) \end{bmatrix}, \qquad |D| = \begin{bmatrix} |d_1| & & \\ & \ddots & \\ & & |d_n| \end{bmatrix}.$$

When Do Unitary EVD and SVD Coincide?

Theorem 8

If $A = A^*$, then the following statements are equivalent:

1. *Any unitary EVD of A is also an SVD of A.*
2. *The eigenvalues of A are positive numbers.*
3. $\mathbf{x}^* A\mathbf{x} > 0$ *for all nonzero* $\mathbf{x} \in \mathbb{C}^n$. *(HPD)*

- The equivalence of 1 and 2 is immediate from Theorem 7
- The property in 3 is called the **hermitian positive definiteness**, *c.f.*, symmetric positive definiteness.
- The equivalence of 2 and 3 can be shown conveniently using **Rayleigh quotient**; see next slide.

Note: Rayleigh Quotient

Let $A \in \mathbb{R}^{n \times n}$ be fixed. The **Rayleigh quotient** is the map $R_A : \mathbb{R}^n \to \mathbb{R}$ given by

$$R_A(\mathbf{x}) = \frac{\mathbf{x}^{\mathrm{T}} A \mathbf{x}}{\mathbf{x}^{\mathrm{T}} \mathbf{x}}.$$

- R_A maps an eigenvector of A into its associated eigenvalue, *i.e.*, if $A\mathbf{v} = \lambda \mathbf{v}$, then $R_A(\mathbf{v}) = \lambda$.
- If $A = A^{\mathrm{T}}$, then $\nabla R_A(\mathbf{v}) = \mathbf{0}$ for an eigenvector $\mathbf{v}$, and so

  $$R_A(\mathbf{v} + \epsilon \mathbf{z}) = R_A(\mathbf{v}) + 0 + O(\epsilon^2) = \lambda + O(\epsilon^2), \quad \text{as } \epsilon \to 0.$$

  The Rayleigh quotient is a quadratic approximation of an eigenvalue.

Reduction of Dimensions

Low-Rank Approximations

Let $A \in \mathbb{C}^{m\times n}$ with $m \geqslant n$. Its thin SVD $A = \widehat{U}\widehat{\Sigma}V^*$ can be written as

$$
\begin{aligned}
A &= \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_n \end{bmatrix} \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{bmatrix} \begin{bmatrix} \mathbf{v}_1^* \\ \vdots \\ \mathbf{v}_n^* \end{bmatrix} \\
&= \begin{bmatrix} \sigma_1\mathbf{u}_1 & \cdots & \sigma_n\mathbf{u}_n \end{bmatrix} \begin{bmatrix} \mathbf{v}_1^* \\ \vdots \\ \mathbf{v}_n^* \end{bmatrix} = \sum_{j=1}^{r} \sigma_j \mathbf{u}_j \mathbf{v}_j^*,
\end{aligned}
$$

where r is the rank of A.

- Each outer product $\mathbf{u}_j\mathbf{v}_j^*$ is a rank-1 matrix.
- Since $\sigma_1 \geqslant \sigma_2 \geqslant \cdots \geqslant \sigma_r > 0$, important contributions to A come from terms with small j.

Low-Rank Approximations (cont')

For $1 \leqslant k \leqslant r$, define

$$A_k = \sum_{j=1}^{k} \sigma_j \mathbf{u}_j \mathbf{v}_j^* = U_k \Sigma_k V_k^*,$$

where

- U_k is the first k columns of U;
- V_k is the first k columns of V;
- Σ_k is the upper-left $k \times k$ submatrix of Σ.

This is a rank-k approximation of A.

Best Rank-k Approximation

Theorem 9 (Eckart-Young)

Let $A \in \mathbb{C}^{m\times n}$. Suppose A has rank r and let $A = U\Sigma V^*$ be an SVD. Then

- $\|A - A_k\|_2 = \sigma_{k+1}$, *for* $k = 1, \ldots, r-1$.
- *For any matrix B with* $\operatorname{rank}(B) \leqslant k$, $\|A - B\|_2 \geqslant \sigma_{k+1}$.