Linear Analysis
Lecture 10

Finite Dimensional Spectral Theory

Let V be a finite dimensional vector space over $\mathbb{F} = \mathbb{C}$, and $L \in \mathcal{L}(V)$.
Definition. $\lambda \in \mathbb{C}$ is an **eigenvalue** of L if

$$\exists\, v \in V,\ v \neq 0,\ \ni Lv = \lambda v,$$

and v is called an eigenvector associated with the eigenvalue λ.

If (λ, v) is an eigenvalue-eigenvector pair for L, then $\mathrm{Span}\{v\}$ is a one-dimensional invariant subspace under L.

L acts on $\mathrm{Span}\{v\}$ by scalar multiplication by λ.

$$\begin{aligned}
E_\lambda &:= \mathcal{N}(\lambda I - L) = \text{ the } \lambda\text{-eigenspace of } L \\
\dim E_\lambda &:= \mathsf{m}_{\mathcal{G}}(\lambda) \ = \text{ geometric multiplicity of } \lambda \\
&= \text{ max number of lin. indep. eigenvectors for } \lambda \\
\Sigma(L) &:= \{\lambda \in \mathbb{C} : \det(\lambda I - L) = 0\} \ = \text{ the spectrum of } L \\
p_L(t) &:= \det(tI - L) = \text{ the characteristic polynomial for } L \\
\mathsf{m}_{\mathcal{A}}(\lambda) &:= \text{ algebraic multiplicity of } \lambda \\
&= \text{ the multiplicity of } \lambda \text{ as a root of } p_L.
\end{aligned}$$

Diagonalizability

Facts: Let $L \in \mathcal{L}(V)$.

(1) $\quad \forall\ \lambda \in \Sigma(L) \quad \mathsf{m}_{\mathcal{A}}(\lambda) \geq \mathsf{m}_{\mathcal{G}}(\lambda)$

(2) Eigenvectors corresponding to different eigenvalues are linearly independent.

Definition. $L \in \mathcal{L}(V)$ is **diagonalizable** if there is a basis $\mathcal{B} = \{v_1, \ldots, v_n\}$ of V consisting of eigenvectors of L.

Fact L is diagonalizable if there is a basis $\mathcal{B} = \{v_1, \ldots, v_n\}$ of V for which the matrix of L with respect to $\mathcal{B}$ is diagonal ($\in \mathbb{C}^{n \times n}$)

Diagonalizability

Fact $L \in \mathcal{L}(V)$ is diagonalizable if and only if

$$\forall\, \lambda \in \Sigma(L) \quad \mathsf{m}_{\mathcal{G}}(\lambda) = \mathsf{m}_{\mathcal{A}}(\lambda)\ .$$

In particular, if L has n distinct eigenvalues, then L is diagonalizable.

Proof Since $(\forall\, \lambda \in \Sigma(L))\ \mathsf{m}_{\mathcal{G}}(\lambda) \leq \mathsf{m}_{\mathcal{A}}(\lambda)$ and

$$\sum_{\lambda \in \Sigma(L)} \mathsf{m}_{\mathcal{A}}(\lambda) = n = \dim V,$$

we have

$$\sum_{\lambda \in \Sigma(L)} \mathsf{m}_{\mathcal{G}}(\lambda) \leq n$$

with equality iff

$$(\forall\, \lambda \in \Sigma(L))\ \mathsf{m}_{\mathcal{G}}(\lambda) = \mathsf{m}_{\mathcal{A}}(\lambda).$$

By Fact 2, $\sum_{\lambda \in \Sigma(L)} \mathsf{m}_{\mathcal{G}}(\lambda)$ is the maximum number of linearly independent eigenvectors of L which proves the result. $\square$

Diagonalizability

$A \in \mathbb{C}^{n\times n}$ is diagonalizable iff A is similar to a diagonal matrix:

$$\exists\ S \in C^{n\times n} \quad \text{such that} \quad S^{-1}AS = D \text{ is diagonal.}$$

Let $L \in \mathcal{L}(V)$ with V finite dimensional.

If A is the matrix for L in some basis, then

$$L \text{ is diagonalizable iff } A \text{ is diagonalizable.}$$

Matrices and Complex Numbers

Consider $\mathbb{C}^n$ with the Euclidean inner product $\langle \cdot, \cdot \rangle$, and $\mathbb{C}^{n\times n}$ with the Frobenius inner product, $\langle A, B\rangle = \langle A, B\rangle_F = \operatorname{tr} B^H A$.
Here $\|\cdot\|$ will denote the norm induced by $\langle \cdot, \cdot \rangle$, and $\|A\|$ the associated operator norm induced on $\mathbb{C}^{n\times n}$. $\|A\|_F$ will denote the Frobenius norm.
Many objects and operations in $\mathbb{C}^{n\times n}$ correspond to similar objects and operations in $\mathbb{C}$.
To see this we begin by identifying the conjugation operation in $\mathbb{C}$ with the Hermitian-transpose operation in $\mathbb{C}^{n\times n}$:

$$\left[\, z \mapsto \bar{z} \text{ in } \mathbb{C}\,\right] \quad \sim \quad \left[\, A \mapsto A^H \text{ in } \mathbb{C}^{n\times n}\,\right] \qquad (*)$$

Given $z \in \mathbb{C}$,

$$z = \bar{z} \quad \Longleftrightarrow \quad z \in \mathbb{R}.$$

Thus, by the identification (*), the Hermitain matrices ($A = A^H$) correspond to the real numbers. Correspondingly,

$$\bar{z} = -z \quad \Longleftrightarrow \quad z \in i\mathbb{R}.$$

The matrices for which $A^H = -A$ are called the skew-Hermitian matrices, and they correspond to the purely imaginary numbers.

$$\begin{array}{lcll} \mathcal{S}_n & = & n \times n \text{ Hermitian matrices} & \text{(real numbers)} \\ \bar{\mathcal{S}}_n & = & n \times n \text{ skew-Hermitian matrices} & \text{(purely imaginary numbers)} \end{array}$$

Facts

(1) A is Hermitian iff iA is skew-Hermitian.

(2) Every $A \in \mathbb{C}^{n\times n}$ has a unique representation of the form $A = X + iY$, where X, $Y \in \mathcal{S}_n$:

$$X = \frac{1}{2}(A + A^H) \quad \text{and} \quad Y = \frac{1}{2i}(A - A^H) \ .$$

(3) $A \in \mathbb{C}^{n\times n}$ is Hermitian iff $(\forall\, x \in \mathbb{C}^n) \quad \langle Ax, x\rangle \in \mathbb{R}$.

(4) $A \in \mathbb{C}^{n\times n}$ is skew-Hermitian iff $(\forall\, x \in \mathbb{C}^n) \quad \langle Ax, x\rangle \in i\mathbb{R}$.

Caution: If $A \in \mathbb{R}^{n\times n}$ and $(\forall\, x \in \mathbb{R}^n) \quad \langle Ax, x\rangle \in \mathbb{R}$, A need not be symmetric. For example, if

$$A = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix},$$

then

$$\langle Ax, x\rangle = \langle x, x\rangle \ \forall\, x \in \mathbb{R}^n \ .$$

Now consider $\mathbb{C}^{n\times n}$ as a vector space over $\mathbb{R}$.

This is also an inner product space with inner product

$$\langle A, B\rangle = \mathcal{R}e\langle A, B\rangle_F \ .$$

In this setting we have the following facts:

(1) $\mathcal{S}_n$ and $\bar{\mathcal{S}}_n$ are subspaces of $\mathbb{C}^{n\times n}$.

(2) $\quad \mathcal{S}_n^{\perp} \ = \ \bar{\mathcal{S}}_n \ .$

(3) Given $A \in \mathbb{C}^{n\times n}$, the orthogonal projections onto $\mathcal{S}_n$ and $\bar{\mathcal{S}}_n$ are

$$\frac{1}{2}(A + A^H) \quad \text{and} \quad \frac{1}{2}(A - A^H),$$

respectively.

Matrices and Complex Numbers

The analogue of the nonnegative reals are the positive semi-definite matrices.

Definition. $A \in \mathbb{C}^{n\times n}$ is called **positive semi-definite** (or nonnegative) if
$$(\forall\, x \in \mathbb{C}^n) \quad \langle Ax, x\rangle \geq 0\ .$$

Since $A \in \mathbb{C}^{n\times n}$ is Hermitian iff $(\forall\, x \in \mathbb{C}^n)$ $\langle Ax, x\rangle \in \mathbb{R}$, a positive semi-definite $A \in \mathbb{C}^{n\times n}$ is automatically Hermitian, but one often says Hermitian positive semi-definite anyway.
Consider $A^{\mathrm{H}}A$ as the analogue of $|z|^2$ for $z \in \mathbb{C}$. Observe that $A^{\mathrm{H}}A$ is positive semi-definite:
$$\langle A^{\mathrm{H}}Ax, x\rangle = \langle Ax, Ax\rangle = \|Ax\|^2 \geq 0.$$

In fact, $\|A^{\mathrm{H}}A\| = \|A\|^2$ since

$$\|A\| = \|A^H\|,\ \|A^{\mathrm{H}}A\| \leq \|A^H\| \cdot \|A\| = \|A\|^2$$

and
$$\begin{aligned} \|A^{\mathrm{H}}A\| &= \sup_{\|x\|=1} \|A^{\mathrm{H}}Ax\| &&= \sup_{\|x\|=1} \sup_{\|y\|=1} |\langle A^{\mathrm{H}}Ax, y\rangle| \\ &\geq \sup_{\|x\|=1} \langle A^{\mathrm{H}}Ax, x\rangle &&= \sup_{\|x\|=1} \|Ax\|^2 \ = \ \|A\|^2\ . \end{aligned}$$

The analogue of complex numbers of modulus 1 are the **unitary matrices**.

Definition. $A \in \mathbb{C}^{n\times n}$ is **unitary** if $A^{\mathrm{H}} A = I$.

Since injectivity is equivalent to surjectivity for $A \in \mathbb{C}^{n\times n}$ unitary, it follows that $A^H = A^{-1}$ and $AA^H = I$ (or equivalently, $A^{\mathrm{H}} A = I$).

Proposition. For $A \in \mathbb{C}^{n\times n}$, the following are equivalent:

(1) A is unitary.
(2) The columns of A form an orthonormal basis of $\mathbb{C}^n$.
(3) The rows of A form an orthonormal basis of $\mathbb{C}^n$.
(4) A preserves the Euclidean norm: $(\forall\, x \in \mathbb{C}^n)\ \|Ax\| = \|x\|$.
(5) A preserves the Euclidean inner product: $(\forall\, x, y \in \mathbb{C}^n)\ \langle Ax, Ay\rangle = \langle x, y\rangle$.
(6) $\kappa(A) = 1$ (perfectly conditioned in operator 2-norm).

Proof Sketch. Let $a_1, \ldots, a_n$ be the columns of A. Then

$$A^{\mathrm{H}} A = I \Leftrightarrow a_i^H a_j = \delta_{ij}.$$

So (1) $\Leftrightarrow$ (2). Similarly

$$(1) \;\Leftrightarrow\; AA^H = I \;\Leftrightarrow\; (3).$$

Since $\|Ax\|^2 = \langle Ax, Ax\rangle = \langle A^{\mathrm{H}} Ax, x\rangle$ and $A^{\mathrm{H}} A$ is Hermitian,

$$(4) \;\Leftrightarrow\; \langle (A^{\mathrm{H}} A - I)x, x\rangle = 0\,\forall\, x \in \mathbb{C}^n \;\Leftrightarrow\; A^{\mathrm{H}} A = I \Leftrightarrow\; (1).$$

Finally, clearly (5) $\Rightarrow$ (4), and (4) $\Rightarrow$ (5) by polarization. □

The correspondences between $\mathbb{C}$ and $\mathbb{C}^{n\times n}$, although rich, can only take one so far. There are important objects in $\mathbb{C}^{n\times n}$ for which it seems that there is no suitable corresponding notion in $\mathbb{C}$. For example, *normal matrices* don't really have an analogue in $\mathbb{C}$.

Definition. $A \in \mathbb{C}^{n\times n}$ is **normal** if $AA^H = A^{\mathrm{H}}A$.

Proposition. For $A \in \mathbb{C}^{n\times n}$, the following are equivalent:

(1) A is normal.

(2) The Hermitian and skew-Hermitian parts of A commute, i.e.,

$$\text{if } A = B + iC \;\; B, C \in \mathcal{S}_n \;\Rightarrow\; BC = CB \;.$$

(3) $(\forall\, x \in \mathbb{C}^n)\ \|Ax\| = \|A^H x\|$.

Proof Sketch. (1) ⇔ (2) (exercise).
Since

$$\|Ax\|^2 = \langle A^{\mathrm{H}}Ax, x\rangle \quad \text{and} \quad \|A^H x\|^2 = \langle AA^H x, x\rangle,$$

and since $A^{\mathrm{H}}A$ and AA^H are Hermitian,

$$(3) \quad \Leftrightarrow \quad (\forall\, x \in \mathbb{C}^n)\ \langle (A^{\mathrm{H}}A - AA^H)x, x\rangle = 0 \quad \Leftrightarrow \quad (1).$$

□

Observe that Hermitian, skew-Hermitian, and unitary matrices are all normal.

Real Matrices

The above definitions can all be specialized to the real case.

• Real Hermitian matrices are (real) **symmetric** matrices: $A^T = A$.

Every $A \in \mathbb{R}^{n\times n}$ can be written uniquely as $A = B + C$ where $B = B^T$ (symmetric) and $C = -C^T$ (skew-symmetric) :

$$B = \frac{1}{2}(A + A^T) \text{ and } C = \frac{1}{2}(A - A^T).$$

• Real unitary matrices are called **orthogonal matrices**, and are characterized by $A^T A = I$ or $A^T = A^{-1}$.

Since

$$(\forall\, A \in \mathbb{R}^{n\times n})(\forall\, x \in \mathbb{R}^n)\ \langle Ax, x\rangle \in \mathbb{R},$$

there is no characterization of symmetric matrices analogous to that given above for Hermitian matrices.

Also, unlike the complex case, the values of the quadratic form $\langle Ax, x\rangle$ for $x \in \mathbb{R}^n$ only determine the symmetric part of A, not A itself.

Matrices and Complex Numbers

The analogy with the complex numbers clarifies when considering eigenvalues.

Let $A \in \mathbb{C}^{n\times n}$ have characteristic polynomial $p_A(t) = \det(tI - A)$.

Since $\overline{p_A(t)} = p_{A^H}(\bar{t})$, we have $\lambda \in \Sigma(A) \Leftrightarrow \bar{\lambda} \in \Sigma(A^H)$.

If A is Hermitian and $\lambda \in \Sigma(A)$ with $Ax = \lambda x$, then

$$\lambda\langle x, x\rangle = \langle Ax, x\rangle = \langle x, Ax\rangle = \bar{\lambda}\langle x, x\rangle \ .$$

so $\lambda = \bar{\lambda}$, that is, $\Sigma(A) \subset \mathbb{R}$.

Also eigenvectors corresponding to different eigenvalues of Hermitian matrix are orthogonal: if $Ax = \lambda x$ and $Ay = \mu y$, then

$$\lambda\langle x, y\rangle = \langle Ax, y\rangle = \langle x, Ay\rangle = \mu\langle x, y\rangle,$$

so $\langle x, y\rangle = 0$ if $\lambda \neq \mu$.

Any eigenvalue λ of a unitary matrix satisfies $|\lambda| = 1$ since

$$|\lambda| \cdot \|x\| = \|Ax\| = \|x\|.$$

Again, eigenvectors corresponding to different eigenvalues of a unitary matrix are orthogonal: if $Ax = \lambda x$ and $Ay = \mu y$, then

$$\begin{aligned} \lambda\langle x, y\rangle &= \langle Ax, y\rangle = \langle x, A^* y\rangle \\ &= \langle x, A^{-1} y\rangle = \langle x, \mu^{-1} y\rangle \\ &= \bar{\mu}^{-1}\langle x, y\rangle = \mu\langle x, y\rangle. \end{aligned}$$

Unitary Hermitian Matrices

Matrices that are both Hermitian and unitary, i.e., $A = A^H = A^{-1}$, satisfy $A^2 = I$ and can be thought of as generalizations of reflections. For example, the matrix $A = -I$ has this property.

Reflections: Householder Transformations

These are the reflections across a hyperplane $\mathcal{H} \subset \mathbb{C}^{n\times n}$ passing through the origin. If $\mathcal{H}$ has normal vector $y \in \mathbb{C}^{n\times n} \backslash \{0\}$ (i.e. $\mathcal{H} = \{y\}^\perp$), then reflection across this hyperplane is given by the linear transformation

$$R = I - \frac{2}{\langle y, y\rangle} yy^H \ .$$

Note that

$$\frac{\langle x, y\rangle}{\langle y, y\rangle} y \quad \text{and} \quad x - \frac{\langle x, y\rangle}{\langle y, y\rangle} y$$

are the orthogonal projections of x onto $\mathrm{Span}\{y\}$ and $\{y\}^\perp$, resp.'ly. Therefore,

$$Rx = x - 2\frac{\langle x, y\rangle}{\langle y, y\rangle} y$$

is the reflection of x across $\{y\}^\perp$.

These transformations are called **Householder transformations** or **Householder reflections**.