LAG 2 Notes

Francesco Chotuck

Contents

1 Linear Operators

1.1 Linear operators introduction

Definition 1.1. Let V be a vector space. A linear transformation $T : V \to V$ is called a **linear operator** on V. The set of linear operators on V is denoted $\operatorname{End}(V)$.

Remark. *The set of linear operators on V is denoted $\operatorname{End}(V)$ because these linear operators are endomorphisms.*

Definition 1.2. Let $M_n(\mathbb{F})$ denote the set of $n \times n$ matrices with entries in $\mathbb{F}$.

Note. $V =$ *finite dimensional vector space over $\mathbb{F}$; sometimes $\mathbb{F}$ is called the **ground field** of V.*

1.2 Eigenvalues, Eigenspaces and Spectrum

Definition 1.3. (Eigenvectors and eigenvalues). If $T : V \to V$ is a linear operator, then we call a vector $\mathbf{v} \in V$ an **eigenvector** for T if

$$\mathbf{v} \neq \mathbf{0} \quad \text{and} \quad T(\mathbf{v}) = \lambda\mathbf{v} \text{ for some } \lambda \in \mathbb{F}.$$

Remark. *Standard basis vectors are eigenvectors if the matrix is a diagonal matrix.*

Definition 1.4. (Spectrum). The set of all eigenvalues of a linear operator T is called the **spectrum** of T and is denoted by $\sigma(T)$.

Definition 1.5. (Identity). The identity transformation is the linear operator $I : V \to V$ that is defined by $I(\mathbf{v}) = \mathbf{v}$ for all $\mathbf{v} \in V$. The linear operator is well defined for any vector space over any field.

Lemma 1.1. *Let T be a linear operator on a vector space V and $\mathbf{v} \in V$ and $\lambda \in \mathbb{F}$. Then $\mathbf{v}$ is an eigenvector of T with eigenvalue λ if and only if*

$$\mathbf{v} \neq \mathbf{0} \quad \textit{and} \quad \mathbf{v} \in \ker(T - \lambda I).$$

Definition 1.6. (Eigenspace). Let $T : V \to V$ be a linear operator and $\lambda \in \mathbb{F}$ an eigenvalue of T. Then we define a linear subspace V_λ of V by setting

$$V_\lambda = \ker(T - \lambda I)$$

1.3 Similarity and Diagonalizability

Definition 1.7. (Similarity). If A and B are $n \times n$ matrices over $\mathbb{F}$, then we say that A is similar to B, if there exists an $n \times n$ invertible matrix Q over $\mathbb{F}$ such that $B = QAQ^{-1}$. If A is similar to B, then we write $A \simeq B$. If we want to emphasise the dependence on $\mathbb{F}$ we may say instead that A is similar to B over $\mathbb{F}$.

Corollary 1.1.1. *If two matrices $A, B \in M_n(\mathbb{F})$ are similar then they have the same spectrum, i.e. $\sigma(A) = \sigma(B)$*

1.4 The characteristic polynomial

Definition 1.8. The **characteristic polynomial** of an $n \times n$ matrix A is the polynomial $p_A(x) = \det(A - xI)$ (in the variable x). The characteristic equation of A is the equation $\det(A - xI) = 0$.

Lemma 1.2. *Let $\lambda \in \mathbb{F}$. Then λ is an eigenvalue for A if and only if $p_A(\lambda) = 0$.*

$$\begin{aligned} \text{An } n \times n \text{ matrix } M \text{ is invertible} &\iff \operatorname{rank}(M) = n \\ &\iff \operatorname{nullity}(M) = 0 \\ &\iff \det(M) \neq 0. \end{aligned}$$

2 Diagonalisability

Definition 2.1. (Diagonalizability).

1. We say that a linear operator $T : V \to V$ is **diagonalizable** if there is a matrix $[T]_{\mathcal{B}}$ representing T that is diagonal (for some choice basis $\mathcal{B}$ of V).

2. An $n \times n$ matrix A with entries in $\mathbb{F}$ is said to be **diagonalizable over** $\mathbb{F}$ if A is similar over $\mathbb{F}$ to a diagonal matrix. We call A **diagonalizable** for short if A is diagonalizable over $\mathbb{F}$ and the field $\mathbb{F}$ is clear from context.

2.1 Criteria for diagonalizability

Theorem 2.1. A linear map $T : V \to V$ is diagonalizable if and only V has a basis consisting of eigenvectors for T . A matrix A in $M_n(\mathbb{F})$ is diagonalizable over $\mathbb{F}$ if and only if $\mathbb{F}^n$ has a basis consisting of eigenvectors for A.

2.2 How do we diagonalise in practice?

If an $n \times n$ matrix A is diagonalizable so, it must have a basis of eigenvectors $\mathcal{B} := \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ with corresponding (possibly repeating) eigenvalues $\lambda_1, \dots, \lambda_n$. Let

$$D = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots 0 & \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}.$$

Our goal is to find a matrix P so that

$$A = PDP^{-1}.$$

> We get P by stacking the eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_n$ into columns next to one another.

Example. Let

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

- **Step 1:** Finding the eigenvalues.

$$\begin{aligned} p_A(x) = \det(A - xI) &= \det \begin{pmatrix} -x & 1 \\ 1 & -x \end{pmatrix} \\ &= (-x)(-x) - (1)(1) \\ &= x^2 - 1 \\ &= (x+1)(x-1) \end{aligned}$$

So, the eigenvalues are $\lambda_1 = 1$ and $\lambda_2 = -1$.

- **Step 2:** Finding eigenvectors. Let us find the eigenvectors for $\lambda_1 = 1$,

$$\begin{gathered} (A - (1)I)\mathbf{v}_1 = \mathbf{0} \\ \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \mathbf{v}_1 = \mathbf{0} \end{gathered}.$$

To find $\mathbf{v}_1$ let us form the augmented matrix and perform row reduction.

$$\left(\begin{array}{cc|c} -1 & 1 & 0 \\ 1 & -1 & 0 \end{array}\right) \xrightarrow{R2 \to R1 + R2} \left(\begin{array}{cc|c} -1 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right) \xrightarrow{R1 \to (-1)R1} \left(\begin{array}{cc|c} 1 & -1 & 0 \\ 0 & 0 & 0 \end{array}\right).$$

Therefore, if $\mathbf{v}_1 = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ then $x_1 - x_2 = 0 \Rightarrow x_1 = x_2$. Set $x_1 = 1$ and it follows that $x_2 = 1$ so, $\mathbf{v}_1 = (1, 1)$.

- **Step 3:** Finding eigenvectors. Repeat Step 2 for $\lambda_2 = -1$ and obtain that $\mathbf{v}_2 = (1, -1)$.
- **Step 4:** Forming P. The matrix P is formed by adjoining the eigenvectors in the order in which they appear in the diagonal matrix. If

$$A = P \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} P^{-1}.$$

Then $P = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$.

Remark. *To find the eigenvectors it is also sufficient to analyse one of the rows instead of performing row reduction. For example,* $\begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \mathbf{v}_1 = \mathbf{0}$ *using row one we have that* $-x_1 + x_2 = 0 \Rightarrow x_2 = x_1$.

Remark. *An easy way to check if a vector is an eigenvector is by using the definition of eigenvectors i.e.* $A\mathbf{v} = \lambda\mathbf{v}$. *Following from the previous example, we have that* A *has eigenvalue* 1 *and* -1. *Let us check if* $\mathbf{v} = (1, 1)$ *is an eigenvector. So,* $A\mathbf{v} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} = (1) \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \lambda\mathbf{v}$.

2.3 Multiplicity

First of all let us recall the definition of linear independence.

> A collection of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k \in V$ is called linearly independent if, whenever
>
> $$\alpha_1\mathbf{v}_1 + \ldots + \alpha_k\mathbf{v}_k = \mathbf{0} \quad \text{for } \alpha_1, \ldots, \alpha_k \in \mathbb{F},$$
>
> then $\alpha_1 = \alpha_2 = \ldots = \alpha_k = 0$. In other words,there is only one way to express $\mathbf{0}$ as a linear combination of the $\mathbf{v}_i$, namely with all coefficients α_i equal to 0.

Lemma 2.2. *If* $T : V \to V$ *is a linear map, and* $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k \in V$ *are eigenvectors for* T *with distinct eigenvalues* $\lambda_1, \lambda_2, \ldots, \lambda_k$, *then* $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ *is linearly independent.*

Corollary 2.2.1. *Suppose* V *is an* n*–dimensional vector space and* $T : V \to V$ *a linear map. If* T *has* n *distinct eigenvalues then it is diagonalizable.*

Equally this result can be formulated for matrices:

Any $n \times n$ matrix over $\mathbb{F}$ with n distinct eigenvalues in $\mathbb{F}$ is diagonalizable over $\mathbb{F}$.

Theorem 2.3. (The Fundamental Theorem of Algebra). Consider the (monic) polynomial $f(x) = x_n + a_{n-1}x^{n-1} + \ldots + a_0$ with coefficients $a_0, a_1, \ldots, a_{n-1} \in \mathbb{C}$, then

$$f(x) = (x - \alpha_1)(x - \alpha_2) \ldots (x - \alpha_n)$$

for some $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{C}$.

Note. *Monic means the leading coefficient is* 1.

Definition 2.2. The **algebraic multiplicity** m_λ of an eigenvalue λ of $A \in M_n(\mathbb{C})$ is the multiplicity of λ as a root of the characteristic equation. That is, if

$$p_A(x) = (\lambda_1 - x)^{m_1}(\lambda_2 - x)^{m_2} \ldots (\lambda_k - x)^{m_k},$$

where $\lambda_1, \ldots, \lambda_k \in \mathbb{C}$ are the distinct eigenvalues for A (without repetition), then the algebraic multiplicity of each λ_i is m_i.

Definition 2.3. The **geometric multiplicity** n_λ of λ is defined to be the dimension of the associated eigenspace V_λ. Equivalently, the geometric multiplicity is the nullity of $(A - \lambda I)$. We may write $n_i = \dim(\ker(A - \lambda_i I))$ for the geometric multiplicity of λ_i.

Remark. *Note that for an $n \times n$ matrix A the characteristic polynomial $p_A(x) = \det(A - xI)$ has degree n in x. Thus if $m_1, \ldots, m_k$ are the algebraic multiplicities of the eigenvalues of A, then $\sum_{i=1}^{k} m_i = n$.*

Proposition 1. Suppose that B is similar to A. Then

1. A and B have the same characteristic polynomial, $p_A(x) = p_B(x)$;
2. the algebraic multiplicity of an eigenvalue λ of A is the same as its algebraic multiplicity as an eigenvalue of B;
3. the geometric multiplicity of an eigenvalue λ of A is the same as its geometric multiplicity as an eigenvalue of B.

Theorem 2.4. A square matrix A is diagonalizable over $\mathbb{C}$ if and only if $m_\lambda = n_\lambda$ for all eigenvalues λ of A.

2.4 Upper triangular matrices

Theorem 2.5. (Schur triangulation). Every square matrix is similar over $\mathbb{C}$ to an upper-triangular matrix.

Remark. *For a fixed square $n \times n$ matrix Q the map $M_n(\mathbb{C}) \to M_n(\mathbb{C})$ that sends A to QAQ^{-1} is a linear operator that is also known as* ***"conjugation by** Q**"**. *Therefore if two matrices are similar, they are sometimes also referred to as conjugate. We may say* ***"conjugating** A **by** Q **gives** B**"** *to express the equation* $B = QAQ^{-1}$.

Characteristic polynomial of a triangular matrix. If T is an upper-triangular matrix,

$$T = \begin{pmatrix} t_{11} & t_{12} & \dots & t_{1n} \\ 0 & t_{22} & \dots & t_{2n} \\ \vdots & 0 & \ddots & \vdots \\ 0 & 0 & \dots & t_{nn} \end{pmatrix},$$

then the characteristic polynomial of T is $p_T(x) = (t_{11} - x)(t_{22} - x) \dots (t_{nn} - x)$. Moreover if A is similar to the upper-triangular matrix T then also $p_A(x) = p_T(x) = (t_{11} - x)(t_{22} - x) \dots (t_{nn} - x)$.

Definition 2.4. For any matrix $A \in M_n(\mathbb{F})$, the **trace** of A denoted $trace(A) \in \mathbb{F}$ be defined to be the sum of the diagonal entries of A.

Remark. *Let A be an $n \times n$ matrix, and let $\lambda_1, \lambda_2, \dots, \lambda_n$ be its (complex) eigenvalues (counting multiplicities). Then*

1. $trace(A) = \lambda_1 + \lambda_2 + \dots + \lambda_n$.
2. $\det(A) = \lambda_1 \lambda_2 \dots \lambda_n$.

Corollary 2.5.1. *Let A be an $n \times n$ matrix over $\mathbb{C}$. For any eigenvalue λ of A, the geometric multiplicity is less than or equal to the algebraic multiplicity: $n_\lambda \leq m_\lambda$.*

2.5 The Cayley-Hamilton Theorem

Given a polynomial $p(x) = \sum_{k=0}^{N} a_k x^k$, we define the evaluation of the polynomial on a square matrix A by setting

$$p(A) := \sum_{k=0}^{N} a_k A^k$$

where we agree, by convention, that $A^0 = I$.

Theorem 2.6. (Cayley-Hamilton Theorem). Every square matrix A satisfies its characteristic equation; i.e.,$p_A(A) = \mathbf{0}$.

2.6 Minimal polynomial

Lemma 2.7. *For any complex $n \times n$ matrix A there is a unique lowest degree monic (coefficient of leading term is 1) polynomial $m_A(x)$ over C, such that $m_A(A) = \underline{\mathbf{0}}$. The polynomial $m_A(x)$ has the following properties:*

1. *If $q(x)$ is a polynomial such that $q(A) = \underline{\mathbf{0}}$, then $m_A(x)$ divides $q(x)$.*
2. *The roots of $m_A(x) = 0$ are precisely the eigenvalues λ of A.*
3. *If A and B are similar, then $m_A(x) = m_B(x)$.*

The polynomial $m_A(x)$ defined in this way is called the ***minimal polynomial*** *of A.*

Remark. *Property* (1) *tells us that $m_A(x)|p_A(x)$, as $p_A(x) = 0$ by Cayley-Hamilton.*

Theorem 2.8. An $n \times n$ matrix A is diagonalizable over $\mathbb{C}$ if and only if its minimal polynomial $m_A(x)$ has no repeated roots.

Example. From Skills 5: how to calculate minimal polynomial.

Definition of the minimal polynomial:

Lemma 1.10.3 (The minimal polynomial). *For any complex $n \times n$ matrix A there is a unique lowest degree monic polynomial $m_A(x)$ over $\mathbb{C}$, such that $m_A(A) = \underline{\mathbf{0}}$. The polynomial $m_A(x)$ has the following properties:*

1. *If $q(x)$ is a polynomial such that $q(A) = \underline{\mathbf{0}}$, then $m_A(x)$ divides $q(x)$.* $\rightarrow m_A(x) \mid p_A(x)$ as $p_A(A) = 0$ by Cayley-Hamilton
2. *The roots of $m_A(x) = 0$ are precisely the eigenvalues λ of A.*
3. *If A and B are similar, then $m_A(x) = m_B(x)$.*

The polynomial $m_A(x)$ defined in this way is called the **minimal polynomial** *of A.*

3. ⋆ Find the minimal polynomial for $A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 1 & -1 & 2 \end{pmatrix}$.

We know (from sheet 3) $p_A(x) = (1-x)(x-2)^2$

1) Since $p_A(A) = 0$ (Cayley-Hamilton), $m_A(x)$ divides $p_A(x)$.

2) The roots of $m_A(x)$ are 1 and 2

$\Rightarrow m_A(x) = (x-1)(x-2)$ or $(x-1)(x-2)^2$

(remember: $m_A(x)$ has to be monic).

Try the smaller degree first ($m_A(x)$ is the <u>lowest</u> degree poly such that $m_A(A) = 0$)

If $m_A(x) = (x-1)(x-2)$:

$A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 1 & -1 & 2 \end{pmatrix}$

$m_A(A) = (A - I)(A - 2I)$

$= \begin{pmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & -1 & 1 \end{pmatrix}\begin{pmatrix} -1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix}$

$= \begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \\ -1 & 0 & -1 \end{pmatrix} \neq 0 \Rightarrow m_A(x)$ is <u>not</u> $(x-1)(x-2)$

Therefore $m_A(x) = (x-1)(x-2)^2$

2.7 Spectral mapping theorem

Theorem 2.9. For any matrix $A \in M_n(\mathbb{C})$ and any polynomial p the spectrum of $p(A)$ is related to the spectrum of A by the identity

$$\sigma(p(A)) = p(\sigma(A)),$$

where $p(\sigma(A)) := \{p(\lambda) : \lambda \in \sigma(A)\}$. In other words, μ is an eigenvalue of $p(A)$ if and only if there exists an eigenvalue λ of A such that $p(\lambda) = \mu$.

Corollary 2.9.1. *Let p be a polynomial and let A be a square matrix with eigenvalues $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$. Then $p(A)$ is invertible if and only if*

$$p(\lambda_k) \neq 0, \quad l = 1, 2, \ldots, n.$$

2.8 Jordan Canonical Form

Theorem 2.10. Every square matrix A is similar over C to a partitioned matrix of the form:

$$\begin{pmatrix} T_1 & \underline{\mathbf{0}} & \ldots & \underline{\mathbf{0}} \\ \underline{\mathbf{0}} & T_2 & \ldots & \underline{\mathbf{0}} \\ \vdots & \underline{\mathbf{0}} & \ddots & \vdots \\ \underline{\mathbf{0}} & \underline{\mathbf{0}} & \ldots & T_k \end{pmatrix},$$

where each T_i is a square matrix of some dimension $r_i \times r_i$ which has the general form

$$T_i = J_{\lambda_i, r_i} \begin{pmatrix} \lambda_i & 1 & 0 & \ldots & 0 \\ 0 & \lambda_i & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_i & 1 \\ 0 & 0 & \ldots & 0 & \lambda_i \end{pmatrix}$$

for some eigenvalue λ_i of A. Moreover the matrix T as above is unique, up to permuting the order of the blocks $T_1, \ldots, T_k$.

Note. *The notation $\underline{\mathbf{0}}$ denotes the zero matrix.*

The form of the matrix in the theorem is called *Jordan canonical form* or *Jordan normal form.* The matrix T also called a *Jordan normal form matrix* (JNF matrix), and the individual blocks, the matrices T_i, are called the *Jordan blocks* of the JNF matrix T. We remark that there can be repetition among the eigenvalues $\lambda_1, \ldots, \lambda_k$ appearing in the expression, and the sizes of the Jordan blocks $T_1, T_2, \ldots, T_k$ may differ from each other; also some of the T_i may be 1×1, in which case we have $T_i = (\lambda_i)$.

Example.

If T is a 3×3 matrix over $\mathbb{C}$ it can have one, two or three distinct eigenvalues. This leads to the following three cases.

- If $p_A(x)$ has distinct roots $\lambda_1, \lambda_2, \lambda_3$, then A is diagonalizable and

$$\begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}.$$

  So in this case $T_1 = (\lambda_1), T_2 = (\lambda_2), T_3 = (\lambda_3)$.

- If $p_A(x) = (\lambda_1 - x)(\lambda_2 - x)^2$ with $\lambda_1 \neq \lambda_2$, then

$$\begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_2 \end{pmatrix} \text{ or } \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 1 \\ 0 & 0 & \lambda_2 \end{pmatrix}.$$

  In the first case $k = 3, T_1 = (\lambda_1)$ and $T_2 = T_(\lambda_2)$; in the second case $k = 2$, $T_1 = (\lambda_1)$ and $T_2 = \begin{pmatrix} \lambda_2 & 1 \\ 0 & \lambda_2 \end{pmatrix}$.

- If $p_A(x) = (\lambda - x)^3$, then

$$\begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix}, \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}, \text{ or } \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}.$$

  In the first case $k = 3, T_1 = T_2 = T_3 = (\lambda)$; in the second case $k = 2, T1 = (\lambda)$ and $T_2 = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$; in the last case $k = 1$ and $T = T_1$.

If A is similar to the JNF matrix T, then it has the same eigenvalues, multiplicities, characteristic and minimal polynomials, since these agree for matrices that are similar.

Example. From the lecture scans:

Example

$$
\begin{array}{ccccccccccc}
\multicolumn{11}{c}{J = \left(\begin{array}{ccccccccccc}
\boxed{i} & & & & & & & & & & \\
& \boxed{i} & & & & & & & & & \\
& & i & 1 & & & & & & & \\
& & & i & & & & & & & \\
& & & & 2 & 1 & & & & & \\
& & & & & 2 & 1 & & & & \\
& & & & & & 2 & & & & \\
& & & & & & & 2 & 1 & & \\
& & & & & & & & 2 & & \\
& & & & & & & & & \boxed{3} & \\
& & & & & & & & & & \boxed{3}
\end{array}\right)} \\
ie_1 & ie_2 & ie_3 & & 2e_5 & & & 2e_8 & & 3e_{10} & 3e_{11}
\end{array}
$$

$\sigma(J) = \{i, 2, 3\}$

$$\left.\begin{array}{l} m_i = 4 \\ m_2 = 5 \\ m_3 = 2 \end{array}\right\} \text{algebraic multiplicities}$$

$P_J(x) = (i-x)^4(2-x)^5(3-x)^2$

geometric multiplicity:

$n_\lambda = \dim(V_\lambda)$

$Je_1 = ie_1$

$Je_2 = ie_2$

$Je_3 = ie_3$

$Je_5 = 2e_5$

$Je_8 = 2e_8$

$Je_{10} = 3e_{10}$

$Je_{11} = 3e_{11}$

$V_i = \mathrm{Span}(e_1, e_2, e_3)$

$V_2 = \mathrm{Span}(e_5, e_8)$

$V_3 = \mathrm{Span}(e_{10}, e_{11})$

$$\left.\begin{array}{l} n_i = 3 \\ n_2 = 2 \\ n_3 = 2 \end{array}\right\} \text{geometric multiplicities}$$

Minimal polynomial

$$J = \begin{pmatrix} \boxed{J_{\mu_1,k_1}} & & & \\ & \boxed{J_{\mu_2,k_2}} & & \\ & & \ddots & \\ & & & \boxed{J_{\mu_r,k_r}} \end{pmatrix}$$

$$\boxed{m_J(x) = \prod_{\lambda \in \sigma(J)} (x-\lambda)^{d_\lambda} .}$$

<u>Claim</u>: For $\lambda \in \sigma(J)$ let J_{μ_s,k_s} be the largest Jordan block with eigenvalue $\mu_s = \lambda$. Then $\boxed{d_\lambda = k_s}$.

The proof of this Claim relies on the fact that the minimal polynomial of a single Jordan block $J_{\mu,k}$ is $(x-\mu)^k$. See Problem Sheet 4!

<u>Example</u> The minimal polynomial for our 11×11 matrix J is

$$m_J(x) = (x-i)^2 (x-2)^3 (x-3)$$

$$J = \begin{pmatrix}
i & & & & & & & & & & \\
& i & & & & & & & & & \\
& & i & 1 & & & & & & & \\
& & & i & & & & & & & \\
& & & & 2 & 1 & & & & & \\
& & & & & 2 & 1 & & & & \\
& & & & & & 2 & & & & \\
& & & & & & & 2 & 1 & & \\
& & & & & & & & 2 & & \\
& & & & & & & & & 3 & \\
& & & & & & & & & & 3
\end{pmatrix}$$

Remark. *Jordan blocks must have the same number on the diagonal and next to each diagonal entry there must be a* 1.

From Skill session 5:

- The eigenvalues of T are the diagonal entries of T so, m_λ is number of times λ appears on the diagonal.
- If λ is an eigenvalue of T, n_λ is equal to the number of Jordan blocks with λ as its diagonal entries.
- The minimal polynomial of T is:

$$m_T = \prod_{\lambda\in\sigma(T)} (x-\lambda)^{r_\lambda},$$

where r_λ is the size of the largest Jordan block with eigenvalue λ.

Lemma 2.11. *(Rank-nullity lemma). Suppose that $T : V \to W$ is a linear map, $\{\mathbf{w}_1, \dots, \mathbf{w}_r\}$ is a basis for the image,* $\mathrm{Im}(T)$, *and $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is a basis for the kernel,* $\ker(T)$. *For $i = 1, ..., r$, let $\mathbf{u}_i \in V$ be such that $T(\mathbf{u}_i) = \mathbf{w}_i$. Then $S = \{\mathbf{u}_1, \dots, \mathbf{u}_r, \mathbf{v}_1, \dots, \mathbf{v}_k\}$ is a basis for V.*

3 Inner product spaces

All vector spaces here are finite-dimensional.

3.1 Inner products in $\mathbb{R}^n$ and $\mathbb{C}^n$. Inner product spaces

Definition 3.1. For any two vectors $\mathbf{z} = (z_1, \dots, z_n) \in \mathbb{C}^n$ and $\mathbf{w} = (w_1, \dots, w_n) \in \mathbb{C}^n$ we define

- the *norm* of $\mathbf{z}$ by

$$||\mathbf{z}|| := \sqrt{\sum_{k=1}^{n} |z_k|^2} = \sqrt{\sum_{k=1}^{n} \left((\mathrm{Re}\{z_k\})^2 + (\mathrm{Im}\{z_k\})^2\right)}.$$

- the *dot product* of $\mathbf{x}$ and $\mathbf{y}$ given by

$$\mathbf{x}\cdot\mathbf{y} := \sum_{k=1}^{n} z_k \overline{w_k}.$$

Definition 3.2. Let V be a vector space over $\mathbb{F}$. An inner product on V is a map

$$\langle \cdot, \cdot \rangle : \underbrace{V \times V}_{\text{Cartesian product}} \to \mathbb{F}$$

such that for all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$ and all scalars $\alpha, \beta \in \mathbb{F}$,

(i) Linearity in the first vector:

$$\langle \alpha\mathbf{x} + \beta\mathbf{y}, \mathbf{z} \rangle = \alpha\langle \mathbf{x}, \mathbf{z} \rangle + \beta\langle \mathbf{y}, \mathbf{z} \rangle;$$

(ii) Conjugate symmetry:

$$\langle \mathbf{x}, \mathbf{y} \rangle = \overline{\langle \mathbf{y}, \mathbf{x} \rangle};$$

(iii) Non-negativity:

$$\langle \mathbf{x}, \mathbf{x} \rangle \geq 0;$$

(iv) Non-degeneracy:

$$\langle \mathbf{x}, \mathbf{x} \rangle = 0 \iff \mathbf{x} = \mathbf{0} \in V;$$

are satisfied.

Remark. *If we use property (i) and (ii):*

$$\begin{aligned}
\langle \mathbf{x}, \alpha\mathbf{y} + \beta\mathbf{z} \rangle &= \overline{\langle \alpha\mathbf{y} + \beta\mathbf{z}, \mathbf{x} \rangle} \\
&= \overline{\alpha\langle \mathbf{y}, \mathbf{x} \rangle + \beta\langle \mathbf{z}, \mathbf{x} \rangle} \\
&= \overline{\alpha\langle \mathbf{y}, \mathbf{x} \rangle} + \overline{\beta\langle \mathbf{z}, \mathbf{x} \rangle} \\
&= \overline{\alpha}\overline{\langle \mathbf{y}, \mathbf{x} \rangle} + \overline{\beta}\overline{\langle \mathbf{x}, \mathbf{z} \rangle} \\
&= \overline{\alpha}\langle \mathbf{x}, \mathbf{y} \rangle + \overline{\beta}\langle \mathbf{x}, \mathbf{z} \rangle.
\end{aligned}$$

This is called 'conjugate linear in the second entry'.

Definition 3.3. Let $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{F}$ be an inner product. We call

$$||\mathbf{x}|| := \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} \geq 0$$

the *norm* of $\mathbf{x} \in V$ with respect to the inner product. We will also say that $|| \cdot ||$ is the norm *associated* to the inner product $\langle \cdot, \cdot \rangle$.

Definition 3.4. An *inner product* space is a vector space over $\mathbb{F}$ together with a specified inner product on it.

Lemma 3.1. *Let V be an inner product space and $\mathbf{x} \in V$. Then $\mathbf{x} = \mathbf{0}$ if and only if*

$$\langle \mathbf{x}, \mathbf{y} \rangle = 0 \quad \forall \mathbf{y} \in V.$$

Corollary 3.1.1. *Let V be an inner product space and $\mathbf{x}, \mathbf{y} \in V$. Then $\mathbf{x} = \mathbf{y}$ holds if and only if*

$$\langle \mathbf{x}, \mathbf{z} \rangle = \langle \mathbf{y}, \mathbf{z} \rangle \quad \forall \mathbf{z} \in V.$$

Corollary 3.1.2. *Let V be a vector space, W an inner product space and $A, B \to W$ two linear transformations such that*

$$\langle A\mathbf{x}, \mathbf{y} \rangle = \langle B\mathbf{x}, \mathbf{y} \rangle \quad \forall \mathbf{x} \in V, \mathbf{y} \in W.$$

Then $A = B$.

Theorem 3.2. Let V be an inner product space over $\mathbb{F}$ with associated norm $||\cdot||$. Then for all $\mathbf{x}, \mathbf{y} \in V$ we have

1. $|\langle \mathbf{x}, \mathbf{y} \rangle| \leq ||\mathbf{x}|| \, ||\mathbf{y}||$ (Cauchy-Schwartz inequality),
2. $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ (Triangle inequality)

We have equality in the Cauchy-Schwartz inequality if and only if one of $\mathbf{x}, \mathbf{y}$ is a scalar multiple of the other.

Theorem 3.3. Let V be an inner product space over $\mathbb{F}$. Then for any $\mathbf{x}, \mathbf{y} \in V$,

1. $\langle \mathbf{x}, \mathbf{y} \rangle = \frac{1}{4}\|\mathbf{x} + \mathbf{y}\|^2 - \frac{1}{4}\|\mathbf{x} - \mathbf{y}\|^2$, for $\mathbb{R}$.
2. $\langle \mathbf{x}, \mathbf{y} \rangle = \frac{1}{4}\|\mathbf{x} + \mathbf{y}\|^2 - \frac{1}{4}\|\mathbf{x} - \mathbf{y}\|^2 + \frac{i}{4}\|\mathbf{x} + i\mathbf{y}\|^2 - \frac{i}{4}\|\mathbf{x} - i\mathbf{y}\|^2$, for $\mathbb{C}$.

These are known as the **polarisation identities.**

Lemma 3.4. *Let V be an inner product space with norm $\|\cdot\|$. The for any $\mathbf{x}, \mathbf{y} \in V$,*

$$\|\mathbf{x} + \mathbf{y}\|^2 + \|\mathbf{x} - \mathbf{y}\|^2 = 2(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2).$$

This is known as the ***triangle inequality.***

Definition 3.5. Let V be a vector space over $\mathbb{F}$. A *norm* on V is a map $\|\cdot\| : V \to \mathbb{R}$ such that for all $\mathbf{x}, \mathbf{y} \in V$ and all $\alpha \in \mathbb{F}$ we have

(i) Homogeneity:

$$\|\alpha \mathbf{x}\| = |\alpha| \, \|\mathbf{x}\|,$$

(ii) Triangle inequality:

$$\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|,$$

(iii) Non-negativity:
$$\|\mathbf{x}\| \geq 0 \quad \forall \mathbf{x} \in V,$$

(iv) Non-degeneracy:
$$\|\mathbf{x}\| = 0 \text{ if and only if } \mathbf{x} = \mathbf{0}.$$

Definition 3.6. A vector space equipped with a norm is called a *normed space.*

Theorem 3.5. (Jordan-von Neumann). A norm in a normed space is obtained from some inner product if and only if it satisfies the parallelogram identity.

3.2 Orthogonality

An inner product space $(V, \langle \cdot, \cdot \rangle)$ has an associated norm $|| \cdot ||$.

Definition 3.7. Let V be an inner product space. Two vectors $\mathbf{x}, \mathbf{y} \in V$ are called orthogonal if
$$\langle \mathbf{x}, \mathbf{y} \rangle = 0 \quad (\text{shorthand } \mathbf{x} \perp \mathbf{y}).$$

Definition 3.8. Let V be an inner product space and $E \subset V$ a subspace. We say that a vector $\mathbf{x} \in V$ is orthogonal to E if $\mathbf{x}$ is orthogonal to all vectors $\mathbf{y} \in E$. We also say that two subspaces $E, F \subset V$ are orthogonal if all vectors in E are orthogonal to F (and vice versa). From the lecture scans:

- $\mathbf{x} \in V$ is orthogonal to $E \subset V$ if $\langle \mathbf{x}, \mathbf{y} \rangle = 0 \; \forall \mathbf{y} \in E$ i.e. $\mathbf{x} \perp E$.
- Two subspaces $E, F \subset V$ are orthogonal if $\langle \mathbf{x}, \mathbf{y} \rangle = 0 \; \forall \mathbf{x} \in E$ and $\forall \mathbf{y} \in F$.

Lemma 3.6. *Let V be an inner product space and $E = span\, \{\mathbf{v}_1, \ldots, \mathbf{v}_r\} \subset V$. Then $x \perp E$ if and only if*
$$\mathbf{x} \perp \mathbf{v}_j \quad \forall j = 1, \ldots, r.$$

Definition 3.9. Let V be an inner product space. We say that a subset of $V, S = \mathbf{x}_1, \ldots, x_{\mathbf{x}} \subset V$, is an *orthogonal set* if $\langle \mathbf{x}_i, \mathbf{x}_j \rangle = 0$ for all $i \neq j$. If in addition $||\mathbf{x}_i|| = 1$ for all $i = 1, \ldots, n$, we say that S is an *orthonormal set.*

Lemma 3.7. *(Generalized Pythagorean identity). Let V be an inner product space and $S = \mathbf{x}_1, \ldots, \mathbf{x}_n$ an orthogonal set. Then for any $\alpha_1, \ldots, \alpha_n \in \mathbb{F}$,*
$$\left\| \sum_{j=1}^{n} \alpha_j \mathbf{x}_j \right\|^2 = \sum_{j=1}^{n} |\alpha_j|^2 \|\mathbf{x}_j\|^2.$$

Corollary 3.7.1. *Let V be an inner product space. Any orthogonal set of non-zero vectors is linearly independent.*

Definition 3.10. Let V be an inner product space. An orthogonal (or orthonormal) set $S \subset V$ which is also a basis of V is called an *orthogonal* (or *orthonormal*) basis.

3.3 Writing arbitrary vectors as linear combinations of orthogonal basis vectors

Suppose $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is an orthogonal basis for V (in this case $\mathbf{e}_n$ is not representing the standard basis). Let $\mathbf{x} \in V$.Then there exists $\alpha_1, \ldots, \alpha_n \in \mathbb{F}$ such that $\mathbf{x} = \alpha_1 \mathbf{e}_1 + \alpha_2 \mathbf{e}_2 + \ldots + \alpha_n \mathbf{e}_n$.

Question: If we know x, how do we find the α_i?

$$
\begin{aligned}
\langle \mathbf{x}, \mathbf{e}_i \rangle &= \left\langle \sum_{j=1}^{n} \alpha_j, \mathbf{e}_i \right\rangle \\
&= \sum_{j=1}^{n} \alpha_j \quad \underbrace{\langle \mathbf{e}_j, \mathbf{e}_i \rangle}_{=0 \text{ unless } j=k} \\
&= \alpha_i \|\mathbf{e}_i\|^2 \\
&\Rightarrow \alpha_i = \frac{\langle \mathbf{x}, \mathbf{e}_i \rangle}{\|\mathbf{e}_i\|^2}.
\end{aligned}
$$

Thus

$$
\boxed{\mathbf{x} = \sum_{j=1}^{n} \frac{\langle \mathbf{x}, \mathbf{e}_i \rangle}{\|\mathbf{e}_i\|^2} \mathbf{e}_j.}
$$

3.4 Orthogonal projections and the Gram-Schmidt process

The only non-trivial subspaces E in $V = \mathbb{R}^2$ are straight lines through the origin, and thus the *orthogonal projection* of a vector $\mathbf{x}$ on E is visualized as in Figure 2.2.

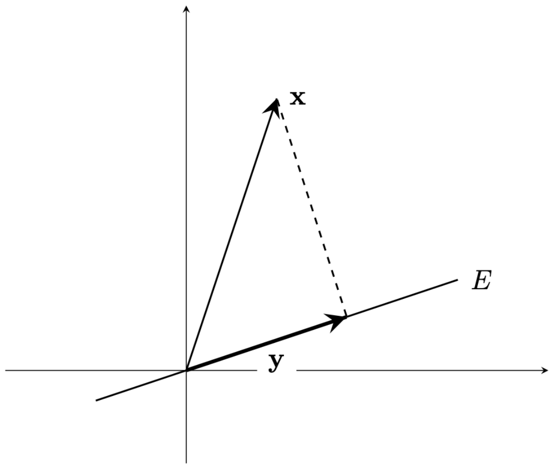

Figure 2.2: The orthogonal projection $\mathbf{y} = P_E\mathbf{x}$ of a given vector $\mathbf{x} \in \mathbb{R}^2$ on the subspace $E \subset \mathbb{R}^2$.

Definition 3.11. Let V be an inner product space and $E \subset V$ a subspace. For a vector $\mathbf{x} \in V$, its orthogonal projection $P_E\mathbf{x}$ on the subspace E is a vector $\mathbf{y}$ such that

(i) $\mathbf{y} \in E$,

(ii) $(\mathbf{x} - \mathbf{y}) \perp E$.

We write $\mathbf{y} = P_E\mathbf{x}$ for the orthogonal projection.

Theorem 3.8. Let V be an inner product space and $E \subset V$ a subspace. The orthogonal projection $\mathbf{y} = P_E\mathbf{x}$ minimizes the distance from $x \in V$ to E, i.e.

$$\forall \mathbf{z} \in E : \quad \|\mathbf{x} - \mathbf{y}\| \leq \|\mathbf{x} - \mathbf{z}\|.$$

Moreover, if for some $\mathbf{z} \in E$ we have $\|\mathbf{x} - \mathbf{y}\| = \mathbf{x} - \mathbf{z}$, then $\mathbf{y} = \mathbf{z}$.

Theorem 3.9. Let V be an inner product space, $E \subset V$ a subspace with orthogonal basis $\{\mathbf{x}_1, \ldots, \mathbf{x}_r\}$. Then the orthogonal projection $P_E\mathbf{x}$ of a vector $\mathbf{x} \in V$ on E is given by the formula

$$\boxed{P_E\mathbf{x} = \sum_{j=1}^{r} \frac{\langle \mathbf{x}, \mathbf{x}_j \rangle}{\|\mathbf{x}_j\|^2} \mathbf{x}_j.}$$

Theorem 3.10. (Gram-Schmidt orthogonalization). Let V be an inner product space and $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ linearly independent vectors in V. Then one may construct orthogonal vectors $\mathbf{y}_1, \ldots, \mathbf{y}_n \in V$ such that for each $1 \leq r \leq n$ we have

$$\operatorname{span}\{\mathbf{x}_1, \ldots, \mathbf{x}_r\} = \operatorname{span}\{\mathbf{y}_1, \ldots, \mathbf{y}_r\}.$$

So by the theorem to construct an orthogonal basis we follow this algorithm:

$$\boxed{\mathbf{y}_{r+1} := \mathbf{x}_{r+1} - P_{E_r}\mathbf{x}_{r+1} - \sum_{k=1}^{r} \frac{\langle \mathbf{x}_{r+1}, \mathbf{y}_k \rangle}{\|\mathbf{y}_k\|^2}\mathbf{y}_k.}$$

From the lecture scans:

Example: Sheet 7, Q3.

$\mathbb{R}^3$ with dot product.

$$\{\underset{\underline{x}_1}{(1,2,-2)},\ \underset{\underline{x}_2}{(1,-1,4)},\ \underset{\underline{x}_3}{(2,1,1)}\}$$

1. $y_1 = \underline{x}_1 = (1,2,-2)$

2. $y_2 = \underline{x}_2 - \dfrac{\langle \underline{x}_2, y_1 \rangle}{\|y_1\|^2} y_1 = (1,-1,4) - \dfrac{(-9)}{9}(1,2,-2) = (2,1,2)$

$$\langle \underline{x}_2, y_1 \rangle = \underline{x}_2 \cdot y_1 = \begin{pmatrix} 1 \\ -1 \\ 4 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 2 \\ -2 \end{pmatrix} = -9.$$

$$\|y_1\|^2 = \|(1,2,-2)\|^2 = 9$$

3. $y_3 = \underline{x}_3 - \dfrac{\langle \underline{x}_3, y_1 \rangle}{\|y_1\|^2} y_1 - \dfrac{\langle \underline{x}_3, y_2 \rangle}{\|y_2\|^2} y_2$

$\langle \underline{x}_3, y_1 \rangle = 2$, $\|y_1\|^2 = 9$, $\langle \underline{x}_3, y_2 \rangle = 7$, $\|y_2\|^2 = 9$.

$$y_3 = (2,1,1) - \frac{2}{9}(1,2,-2) - \frac{7}{9}(2,1,2)$$

$$= (2/9, -2/9, -1/9)$$

3.5 Adjoints of linear operators

Let $A \in M_{m,n}(\mathbb{C})$. We define the adjoint matrix of A as:

$$A^* = \overline{A^T} \in M_{m,n}(\mathbb{C}).$$

Recall the dot the product on $\mathbb{C}^n$: $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n \Rightarrow \mathbf{x} \cdot \mathbf{y} = \sum_{j=1}^n \mathbf{x}_k \overline{\mathbf{y}_j}$; then we can think of $\mathbf{x}, \mathbf{y} \in M_{m,n}(\mathbb{C}^n)$. Then

$$\mathbf{x} \cdot \mathbf{y} = \underbrace{\mathbf{y}^*}_{\substack{1\times n \\ \text{matrix}}} \underbrace{\mathbf{x}}_{\substack{n\times 1 \\ \text{matrix}}} .$$

Lemma 3.11. *Let $A \in M_n(\mathbb{C})$ and $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$. Then $(A\mathbf{x}) \cdot \mathbf{y} = \mathbf{x} \cdot (A^*\mathbf{y})$*

Theorem 3.12. Let V be an inner product space and $T : V \to V$ a linear operator on it. Then there exists a unique linear operator $T^* : V \to V$ such that

$$\langle T\mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, T^*\mathbf{y} \rangle \quad \forall \mathbf{x}, \mathbf{y} \in V.$$

Definition 3.12. The operator T^* in the last theorem is called the **adjoint** of T.

Theorem 3.13. Let V be an inner product space and $T, U : V \to V$ linear operators on V. If $c \in \mathbb{F}$, then

1. $(T + U)^* = T^* + U^*$;
2. $(cT)^* = \overline{c}T^*$;
3. $(TU)^* = U^*T^*$;
4. $(T^*)^* = T$.

Theorem 3.14. Let V be an inner product space and $T : V \to V$ a linear operator on V. Then

1. $\ker T^* = (\operatorname{Im} T)^\perp$;
2. $\ker T = (\operatorname{Im} T^*)^\perp$;
3. $\operatorname{Im} T^* = (\ker T)^\perp$;
4. $\operatorname{Im} T = (\ker T^{*\perp})$.

Definition 3.13. Let V be an inner product space. A linear operator $T : V \to V$ such that $T = T^*$ is called **self-adjoint** (or *Hermitian*).

3.6 Isometries and unitary operators

Definition 3.14. Let V and W be inner product spaces over $\mathbb{F}$. A linear map $U : V \to W$ is called an **isometry** (or *norm preserving*) if

$$\|U\mathbf{x}\|_W = \|\mathbf{x}\|_V \quad \forall \mathbf{x} \in V.$$

A norm preserving linear operator $U : V \to V$ on an inner product space V is also called a **unitary operator.**

Lemma 3.15. *Eigenvalues of unitary operators have absolute value* 1.

Theorem 3.16. Let V, W be inner product spaces over $\mathbb{F}$. Then $U : V \to W$ preserves norms if and only if it preserves inner products, i.e.

$$\langle U\mathbf{x}, U\mathbf{y}\rangle_W = \langle \mathbf{x}, \mathbf{y}\rangle_V \quad \forall \mathbf{x}, \mathbf{y} \in V.$$

Lemma 3.17. *Let V be an inner product space and $U : V \to V$ a linear map. Then U is norm preserving if and only if $U^*U = UU^* = I$. That is, $U^{-1} = U^*$.*

Corollary 3.17.1.

1. *$U : V \to V$ is a unitary operator. Let $\{\mathbf{x}_1, \ldots, \mathbf{x}_r\} \subset V$ be an orthonormal set. Then $\{U\mathbf{x}_1, \ldots, U\mathbf{x}_r\}$ is an orthonormal set.*

2. *Let $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ be an orthonormal basis of an inner product space V. An operator $U : V \to V$ is unitary if and only if $\{U\mathbf{x}_1, \ldots, U\mathbf{x}_n\}$ is an orthonormal basis of V.*

Definition 3.15. A matrix $U \in M_n(\mathbb{C})$ is called unitary if $U^*U = I_{n\times n}$ where $U^* = \overline{U^T}$. A matrix $O \in M_n(\mathbb{R})$ is called orthogonal if $O^TO = I_{n\times n}$.

Proposition 2. Let $U \in Mn(C)$ be a unitary matrix. Then

1. $|\det U| = 1$, so in particular for orthogonal matrices $O \in M_n(\mathbb{R}) : \det O = \pm 1$.

2. If λ is an eigenvalue of U, then $|\lambda| = 1$.

Definition 3.16. Let $A, B \in M_n(\mathbb{C})$. We say that B is unitarily equivalent to A if there exists a unitary matrix $U \in M_n(\mathbb{C})$ such that $A = UBU^{-1}$. If the same holds true for an orthogonal matrix $O \in M_n(\mathbb{R})$, i.e. $A = OBO^{-1}$, then B is orthogonally equivalent to A.

Theorem 3.18. A matrix $A \in M_n(\mathbb{C})$ is unitarily equivalent to a diagonal matrix (i.e. can be unitarily diagonalized) if and only if it has n orthonormal eigenvectors (in other words, there exists an orthonormal basis of $\mathbb{C}^n$ consisting of eigenvectors for A).

4 Structure of operators in inner product

5 Bilinear and quadratic forms

6 Appendix

6.1 Functions of linear operators

Example. Say $f(x) = x^2$ and $A \begin{pmatrix} a & b \\ c & d \end{pmatrix}$: then,

$$f(A) = A^2 = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a^2+bc & ab+bd \\ ac+cd & bc+d^2 \end{pmatrix}.$$

Example. $f(x) = x^2 + 1$. Then we set $f(A) = A^2 + I$ so,

$$\begin{pmatrix} a^2+bc & ab+bd \\ ac+cd & bc+d^2 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a^2+bc+1 & ab+bd \\ ac+cd & bc+d^2+1 \end{pmatrix}$$

Definition 6.1. Let $p(x) = a_k x^k + a_{k-1} x^{k-1} + \ldots + a_1 x + a_0(1)$ where $a_0, a_1, \ldots a_k \in \mathbb{F}$ and $A \in M_n(\mathbb{F})$ then

$$p(A) = a_k A^k + a_{k-1} A^{k-1} + \ldots + a_1 A + a_0 I.$$

Lemma 6.1. *If $A \simeq B$ then $p(A) \simeq p(B)$ for any polynomial p. In fact if $B = QAQ^{-1}$ then $p(B) = Qp(A)Q^{-1}$.*

6.1.1 Exponential function

Replace $p(x)$ by $\exp(x) = e^x$. Recall that

$$\exp(x) = 1 + x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \frac{1}{4!}x^4 + \ldots = \sum_{n=0}^{\infty} \frac{1}{n!} x^n.$$

Lemma 6.2. *Suppose A is diagonalizable (so $A = QDQ^{-1}$ for some diagonal matrix D).*

- *Then $\sum_{n=0}^{\infty} \frac{1}{n!} A^n$ converges and we call the limit* $\exp(A)$.
- *Moreover if* $D = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix}$ *the limit is* $\exp(A) = Q \begin{pmatrix} e^{\lambda_1} & 0 & \ldots & 0 \\ 0 & e^{\lambda_2} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & e^{\lambda_n} \end{pmatrix} Q^{-1}$.

Remark. *If* $v(t) = \begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix}$ *satisfy* $\dot{v}(t) = Av(t)$ *and* $v(0) = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$. *Then* $\mathbf{v}(t) = \exp(tA) \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$.