MATH50003 Linear Algebra and Numerical Analysis Term 1

also known as

MATH50016 Linear Algebra 2 for JMC

Notes by Robert Weingart

Contents

General notes

Notation. Throughout this entire document, unless otherwise stated:

- F is an arbitrary field
- V is an arbitrary vector space over F
- $T : V \to V$ is an arbitrary linear map
- A is an arbitrary $n \times n$ matrix with entries in F

Furthermore,

- For any linear map or matrix S, c_S is the characteristic polynomial
- For any linear map or matrix S, m_S is the minimal polynomial (defined in section 9)
- To avoid ambiguity, : always means "such that" and | always means "divides"

1 Course Overview

Nothing to summarise here, I am including the section to maintain correct numbering.

2 Some revision from 1st Year Linear Algebra

Definition. Let V be finite-dimensional with a basis $B = \{v_1, \ldots, v_n\}$. Let a_{ij} be the jth component of $T(v_i)$ with respect to B. The **matrix of** T with respect to B is $[T]_B := (a_{ij})$.

Proposition 2.1. *For linear maps S and T and a basis B, $[ST]_B = [S]_B[T]_B$.*

Proposition. *For any polynomial $q(x)$, $[q(T)]_B = q([T]_B)$.*

Definition. Let V be finite-dimensional with two bases $E = \{e_1, \ldots, e_n\}$ and $F = \{f_1, \ldots, f_n\}$. Let p_{ij} be the ith component of f_j with respect to E. The **change of basis matrix** from E to F is (p_{ij}).

Proposition 2.2. *The change of basis matrix is invertible. If P is the change of basis matrix from E to F, then $[T]_F = P^{-1}[T]_E P$.*

Definition. The **characteristic polynomial** of T is $c_T(x) := \det xI_V - T$.

Proposition 2.3. *The eigenvalues of T are the roots of c_T. The eigenvectors to an eigenvalue λ are the non-zero vectors in $E_\lambda := \ker \lambda I_V - T$. $[T]_B$ is diagonal $\iff$ B consists of eigenvalues of T.*

Definition. The **λ-eigenspace** of T is E_λ as defined above.

Proposition 2.4. *If $F = \mathbb{C}$ and V is finite-dimensional then T has an eigenvalue in $\mathbb{C}$.*

Proposition 2.5. *Eigenvectors to different eigenvalues are linearly independent.*

Corollary 2.6. *If the characteristic polynomial of T has $\dim V$ distinct roots then T is diagonalisable.*

3 Algebraic and geometric multiplicities of eigenvalues

Definition. If λ is an eigenvalue of T, its **algebraic multiplicity** is $a(\lambda)$ such that $c_T(x) = (x - \lambda)^{a(\lambda)} q(x)$ where $q(\lambda) \neq 0$. Its **geometric multiplicity** is $g(\lambda) = \dim E_\lambda$.

Proposition 3.1. *$g(\lambda) \leq a(\lambda)$.*

Theorem 3.2. *If $\dim V = n$ and $\lambda_1, \ldots, \lambda_r$ are the distinct eigenvalues of $T : V \to V$, the following statements are equivalent:*

- *T is diagonalisable*
- $\sum_{i=1}^{r} g(\lambda_i) = n$
- $\forall i,\ g(\lambda_i) = a(\lambda_i)$

4 Direct sums

Definition. V is the **direct sum** of subspaces $V_1, \ldots, V_k$ iff

$$\forall v \in V,\ \exists! v_1 \in V_1, \ldots, v_k \in V_k : v = \sum_{i=1}^{k} v_i$$

that is to say, the selection of v_i is unique. In this case we write $V = V_1 \oplus \ldots \oplus V_k$.

Notation. From now on I will use $\bigoplus$ to write a direct sum over a range, like $\sum$ for sums. This does not occur in the official notes, but I will use it here for brevity. For instance, in the situation above I will now write

$$V = \bigoplus_{i=1}^{k} V_i$$

Proposition 4.1. $V = V_1 \oplus V_2 \iff V_1 \cap V_2 = \{0\}$ *and* $\dim V_1 + \dim V_2 = \dim V$

Proposition 4.2. $V = \bigoplus_{i=1}^k V_i \iff \dim V = \sum_{i=1}^k \dim V_i$ *and for any bases* B_i *of the* V_i*,* $\bigcup_{i=1}^k B_i$ *is a basis of* V*.*

Definition 4.1. A subspace W of V is T**-invariant** $\iff T(W) \subseteq W$. In this case, the **restriction** of T to W is $T_W : W \to W$, where $T_W(w) = T(w)$.

Proposition 4.3. *Suppose there are* T*-invariant subspaces* V_i *with bases* B_i *such that* $V = \bigoplus V_i$*. Let* $A_i = [T_{V_i}]_{B_i}$ *and* $B = \bigcup B_i$*. Then*

$$[T]_B = \begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_k \end{pmatrix}$$

Notation. From now on, block diagonal matrices will be written using $\oplus$ or $\bigoplus$ whenever possible. For instance,

$$A \oplus B := \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$$

5 Quotient spaces

Definition. For a subspace W of V, the **quotient space** V/W consists of the cosets

$$W + v := \{w + v \mid w \in W\}$$

where $v \in V$. We define addition as

$$(W + v) + (W + v') := W + (v + v')$$

and scalar multiplication as

$$\lambda(W + v) := W + \lambda v$$

with $v, v' \in V$ and $\lambda \in F$ arbitrary.

Proposition 5.1. V/W *is a vector space over* F*.*

Proposition 5.2. *If* V *is finite-dimensional,* $\dim V/W = \dim V - \dim W$*.*

Proposition. *If* $B_W = \{w_1, \dots, w_r\}$ *is a basis of* V *and* $B = B_W \cup \{v_1, \dots, v_s\}$ *is a basis of* V *then* $\bar{B} := \{W + v_1, \dots, W + v_s\}$ *is a basis of* V/W*.*

Definition. If W is T-invariant, the **quotient map** $\bar{T} : V/W \to V/W$ is

$$\bar{T}(W + v) := W + T(v)$$

Proposition 5.3. *With* B*,* B_W *and* $\bar{B}$ *defined as above,*

$$[T]_B = \begin{pmatrix} [T_W]_{B_W} & Z \\ 0 & [\bar{T}]_{\bar{B}} \end{pmatrix}$$

for some $r \times s$ *matrix* Z*.*

Corollary 5.4. $c_T(x) = c_{T_W}(x) c_{\bar{T}}(x)$*.*

6 Triangularisation

Proposition 6.1. *The values along the main diagonal of an upper triangular matrix are its eigenvalues. The product of two upper triangular matrices is also upper triangular, and its diagonal entries are the products of the corresponding diagonal entries of the factors.*

Theorem 6.2. *(Triangularisation Theorem) If $c_T(x)$ factorises as a product of linear factors then there is a basis B of V such that $[T]_B$ is upper triangular.*

To triangularise T, choose an eigenvector w_1 of T and let $W_1 = \operatorname{span} w_1$. Let $W + w_2$ be an eigenvector for the quotient map with respect to W and let $W_2 = \operatorname{span}\{w_1, w_2\}$. Let $W + w_3$ be an eigenvector for the quotient map with respect to W_2 and let $W_3 = \operatorname{span}\{w_1, w_2, w_3\}$. Continue in this fashion until you have $B = \{w_1, \dots, w_n\}$. Then $[T]_B$ is upper triangular.

7 Cayley-Hamilton Theorem

Theorem 7.1. *(Cayley-Hamilton Theorem) $c_T(T) = 0$.*

Corollary 7.2. *$c_A(A) = 0$.*

Definition. The **companion matrix** of $p(x) = \sum_{i=0}^{n} a_i x^i$ where $a_n = 1$ is

$$C(p) = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 & -a_0 \\ 1 & 0 & 0 & \dots & 0 & -a_1 \\ 0 & 1 & 0 & \dots & 0 & -a_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & -a_{n-1} \end{pmatrix}$$

8 Polynomials

Definition. A **polynomial** over F is an expression of the form

$$p(x) := \sum_{i=1}^{n} a_i x^i$$

with all $a_i \in F$. The set of polynomials over F is $F[x]$.

Definition. The **zero polynomial** is $p(x) = 0$.

Definition. The degree of a non-zero polynomial is the highest $\deg p$ such that $x^{\deg p}$ occurs in $p(x)$ with a non-zero coefficient. $\deg 0$ is undefined.

Definition. $p(x)$ **divides** $q(x) \iff \exists r(x) \in F[x] : q(x) = p(x)r(x)$. In this case we write $p(x) \mid q(x)$.

Notation. From now on I will usually write p rather than $p(x)$.

Proposition 8.1. *(Euclidean Algorithm) For $f, g \in F[x]$ non-constant,*

$$\exists q, r \in F[x] : f = qg + r$$

with $r = 0$ or $\deg r < \deg g$.

Definition 8.1. $d \in F[x]$ is a **greatest common divisor** of $f, g \in F[x]$ both non-zero iff

- $d \mid f$
- $d \mid g$
- For all $e \in F[x]$, $e \mid f$ and $e \mid g \implies e \mid d$.

In this case we write $\gcd(f, g) = d$.

Proposition 8.2. *For $f, g \in F[x]$ non-zero, $\gcd(f, g)$ exists and is unique up to scalar multiplication.*

Definition. $f, g \in F[x]$ are **co-prime** $\iff \gcd(f, g) = 1$.

Proposition 8.3. *For $f, g \in F[x]$ non-zero, $\exists r, s \in F[x] : \gcd(f, g) = rf + sg$.*

Definition. An **irreducible** polynomial is one which is non-constant and cannot be factorised as a product of polynomials of smaller degree.

Proposition 8.4. *For $p(x) \in \mathbb{Q}[x]$ monic with integer coefficients, all roots of p in $\mathbb{Q}$ are integers. If $p(x)$ is reducible, then $p = ab$ where a and b are also monic with integer coefficients (this statement is called Gauss's Lemma).*

Proposition 8.5. *If p is irreducible and $p \mid ab$, either $p \mid a$ or $p \mid b$.*

Corollary 8.6. *If p is irreducible and $p \mid \prod g_i$, then $\exists i : p \mid g_i$.*

Theorem 8.7. *(Unique Factorization Theorem) Any non-constant polynomial can be written as a product of irreducible factors. This factorisation is unique up to scalar multiplication of the factors.*

9 The minimal polynomial of a linear map

Definition. $m_T(x) \in F[x]$ is a **minimal polynomial** for T if $m_T(T) = 0$, m_T is monic and there is no polynomial of smaller degree for which the other two conditions hold.

Proposition 9.1. *m_T is unique. For all $p \in F[x]$, $p(T) = 0 \iff m_T \mid p$.*

Proposition 9.2. *$m_T \mid c_T$. For all $\lambda \in F$, $c_T(\lambda) = 0 \implies m_T(\lambda) = 0$.*

Theorem 9.3. *All irreducible factors of c_T divide m_T.*

Proposition 9.4. *For T_W and $\bar{T}$ as in sections 4 and 5, $m_{T_W} \mid m_T$ and $m_{\bar{T}} \mid m_T$.*

10 Primary Decomposition

Theorem 10.1. *(Primary Decomposition Theorem) Suppose $m_T(x) = \prod_{i=1}^k f_i(x)^{n_i}$ with the f_i distinct and irreducible. Let $V_i := \ker f_i(T)^{n_i}$. Then $V = \bigoplus_{i=1}^k V_i$, each V_i is T-invariant, and $m_{T_{V_i}} = f_i(x)^{n_i}$.*

Corollary 10.2. *T is diagonalisable $\iff$ m_T is a product of distinct linear factors.*

Proposition 10.3. *Suppose $g_1, g_2 \in F[x]$ are coprime and $g_1(T)g_2(T) = 0$. Let $V_i = \ker g_i(T)$. Then $V = V_1 \oplus V_2$ and the V_i are T-invariant. If additionally $m_T = g_1 g_2$, then $m_{T_{V_i}} = g_i$.*

11 Jordan Canonical Form

11.1 Definition and properties

Definition. A **Jordan block** is an $n \times n$ matrix of the form

$$J_n(\lambda) := \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & \dots & 0 \\ 0 & 0 & \lambda & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda \end{pmatrix}$$

Proposition 11.1. *Let $J = J_n(\lambda)$.*

- $c_J(x) = m_J(x) = (x-\lambda)^n$
- *λ is the only eigenvalue of J*
- $a(\lambda) = n$
- $g(\lambda) = 1$
- $J - \lambda I = J_n(0)$
- $(J-\lambda I)e_1 = 0$ *and* $(J-\lambda I)e_k = e_{k-1}$
- $(J-\lambda I)^n = 0$
- $\operatorname{rank}(J-\lambda I)^i = n - i$
- $(J-\lambda I)^i e_k = e_{k-i}$

Proposition 11.2. *Let $A = \bigoplus A_i$.*

- $c_A = \prod c_{A_i}$
- $m_A = \operatorname{lcm}\{m_{A_i}\}$
- *For any eigenvalue λ of A,* $\dim E_\lambda(A) = \sum \dim E_\lambda(A_i)$
- *For any $q(x) \in F[x]$,* $q(A) = \bigoplus q(A_i)$

Theorem 11.3. *If c_A is a product of linear factors, then A is similar to a matrix of the form*

$$J = \bigoplus_{i=1}^{k} J_{n_i}(\lambda_i)$$

where the λ_i are the eigenvalues of A, not necessarily distinct. J is unique up to reordering the Jordan blocks.

Definition. The **Jordan Canonical Form** of A is J as defined above.

Proposition 11.4. *Suppose J has a blocks associated with an eigenvalue λ, with sizes $n_1, \dots, n_a$. Then:*

- $\sum_{i=1}^{a} n_i = a(\lambda)$
- $a = g(\lambda)$
- $\max\{n_i\} = \max\{r : (x-\lambda)^r \mid m_A(x)\}$ *(the size of the largest block is the exponent of $(x-\lambda)$ as a factor of m_A)*

11.2 Steps in the proof of Theorem 11.3 that are numbered as though they were important even though they are just special cases of Theorem 11.3

Theorem 11.5. *The JCF is unique up to reordering the blocks.*

Theorem 11.6. *If c_T is a product of linear factors, then there exists a basis B of V such that $[T]_B$ is in JCF.*

Definition. T is **nilpotent** $\iff \exists k : T^k = 0$.

Theorem 11.7. *If T is nilpotent, then there exists a basis B of V such that $[S]_B = \bigoplus J_{n_i}(0)$*

Corollary 11.8. *In the situation above, $[T + \lambda I_V]_B = \bigoplus J_{n_i}(\lambda)$.*

12 Cyclic Decomposition and Rational Canonical Form

12.1 Cyclic Decomposition

Definition. Let V be finite-dimensional. The T**-cyclic subspace** of V generated by $v \in V$ is

$$\begin{aligned} Z(v,T) &:= \{f(T)(v) : f(x) \in F[x]\} \\ &= \operatorname{span}\{v, T(v), T^2(v), \ldots\} \end{aligned}$$

Note that $Z(v,T)$ is T-invariant. We abbreviate the restriction $T_{Z(v,T)}$ to T_v.

Definition. The T**-annihilator** of v and $Z(v,T)$ is m_v, the monic polynomial of smallest degree such that $m_v(T)(v) = 0$. It can be constructed as follows: Let k be as small as possible such that $\{v, T(v), T^2(v), \ldots, T^k(v)\}$ is not linearly independent. Then $T^k(v) = -a_0 v - a_1 T(v) - \ldots - a_{k-1}T^{k-1}(v)$, and m_v is the polynomial with coefficients a_i.

Proposition 12.1. *With k defined as above, $B = \{v, T(v), T^2(v), \ldots, T^{k-1}(v)\}$ is a basis of $Z(v,T)$. $[T_v]_B = C(m_v)$. $m_{T_v} = m_v$.*

Theorem 12.2. *(Cyclic Decomposition Theorem) Let V be finite-dimensional. If $m_T(x) = f(x)^k$ with f irreducible, then there exist unique numbers r and $k_1, \ldots, k_r$ such that*

$$\begin{aligned} k = k_1 \geq k_2 \geq \ldots \geq k_r \\ V = \bigoplus_{i=1}^{r} Z(v_i, T) \\ m_{v_i} = f(x)^{k_i} \quad i = 1, \ldots, r \end{aligned}$$

for some $v_1, \ldots, v_r \in V$

Corollary 12.3. *If T is as above, there exists a basis B of V such that*

$$[T]_B = \bigoplus_{i=1}^{r} C(f(x)^{k_i})$$

with $k = k_1 \geq k_2 \geq \ldots \geq k_r$ uniquely determined by T.

Corollary 12.4. *If $m_A(x) = x^k$, then*

$$A \sim \bigoplus_{i=1}^{r} C(x^{k_i})$$

with $k = k_1 \geq k_2 \geq \ldots \geq k_r$ uniquely determined by A.

12.2 RCF

Theorem 12.5. *(Rational Canonical Form) Let V be finite-dimensional. If*

$$m_T(x) = \prod_{i=1}^{t} f_i(x)^{k_i}$$

with the f_i distinct and irreducible, then there are unique numbers

$$\begin{gathered} r_1, \ldots, r_t \\ k_{11}, \ldots, k_{1r_1} \\ k_{22}, \ldots, k_{2r_2} \\ \vdots \\ k_{t1}, \ldots, k_{tr_t} \end{gathered}$$

such that

$$\begin{aligned} k_i = k_{i1} \geq k_{i2} \geq \ldots \geq k_{ir_i} \quad i = 1, \ldots, t \\ [T]_B = \bigoplus_{i=1}^{t} \left(\bigoplus_{j=1}^{r_i} C\left(f_i(x)^{k_{ij}}\right) \right) \end{aligned}$$

for some basis B of V.

Corollary 12.6. *If the minimal polynomial of A is of the form of m_T above, then A is similar to a unique matrix of the form of $[T]_B$ above.*

Definition. The **rational canonical form** of A is the matrix described in the corollary above.

Proposition 12.7. *This proposition is not in the official notes, but it should be. Let f_i be one of the factors of m_A. Consider the blocks of the RCF of A associated with f_i, which are*

$$C(f_i(x)^{k_{i1}}) \oplus \ldots \oplus C(f_i(x)^{k_{ir_i}})$$

with the exponents defined as in Theorem 12.5. Then:

- $\sum_{j=1}^{r_i} k_{ij} = \max \{k : f^k \mid c_A\}$
- $r_i = \frac{n - \operatorname{rank}(f(A))}{\deg f}$
- $k_{i1} = k_i = \max \{k : f^k \mid m_A\}$

The first and third statements are trivial; I included them to emphasise the parallels to Proposition 11.4. The second statement was proved in a live lecture. (TODO: add link)

13 The dual space

Definition. A **linear functional** on V is a linear map $\phi : V \to F$.

Definition. The **dual space** of V is V^*, the set of linear functionals on V. Addition and scalar multiplication are defined as usual using the equivalent operations on V. This is a vector space over F.

Definition. The **Kronecker delta** is

$$\delta_{ij} := \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

Proposition 13.1. *Let $n = \dim V$ and $B = \{v_1, \ldots, v_n\}$ be a basis of V. Define*

$$\phi_i(v_j) = \delta_{ij} \quad i = 1, \ldots, n, j = 1, \ldots, n$$

meaning that for any $v = \sum \alpha_j v_j$

$$\phi_i(v) = \alpha_i$$

Then $\{\phi_1, \ldots, \phi_n\}$ is a basis of V^*, and therefore $\dim V^* = \dim V$.

Definition. Let V be finite-dimensional. The **annihilator** of $X \subseteq V$ is

$$X^0 := \{\phi \in V^* : \phi(x) = 0 \quad \forall x \in X\}$$

This is a subspace of V^*.

Proposition 13.2. $\dim W^0 = \dim V - \dim W$.

14 Inner Product Spaces

14.1 Definition and matrix representation

Notation. From now on let F be either $\mathbb{R}$ or $\mathbb{C}$, such that there is a conjugate map $\bar{\cdot} : F \to F$.

Definition. An **inner product** on V is a map $(\cdot, \cdot) : V \times V \to F$ where

- $(\lambda u + \mu v, w) = \lambda(u, w) + \mu(v, w)$
- $\overline{(v, w)} = (w, v)$
- $v \neq 0 \implies (v, v) > 0$

V together with $(\cdot, \cdot)$ is an **inner product space**.

Definition. A is **Hermitian** $\iff$ $A^T = \bar{A}$.

Definition. For $\dim V = n$, the **matrix of an inner product** for some basis $B = \{v_1, \ldots, v_n\}$ has entries $a_{ij} = (v_i, v_j)$. This matrix A is Hermitian, and for any $v, w \in V$, $(v, w) = [v]_B^T A[\bar{w}]_B$, so A defines $(\cdot, \cdot)$ uniquely.

Definition. A Hermitian matrix A is **positive-definite** iff

$$\forall x \in F^n \setminus \{0\},\ x^T A \bar{x} > 0$$

14.2 Geometric functions

Definition. The **length** of $u \in V$ is $||u|| := \sqrt{(u, u)}$. The **distance** between $u, v \in V$ is $d(u, v) := ||u - v||$. u is a **unit vector** $\iff$ $||u|| = 1$.

Proposition 14.1. *For any u, v, w in an inner product space:*

- $|(u, v)| \leq ||u||||v||$
- $||u + v|| \leq ||u|| + ||v||$
- $||u - v|| \leq ||u - w|| + ||w - v||$

14.3 Dual space

Definition. $f_v \in V^*$ is $f_v(w) = (w, v)$.

Definition. $\bar{V}$ is a vector space over F with the same vectors and addition as V, but scalar addition defined as $\lambda * v := \bar{\lambda} v$, where the operation on the right hand side is scalar multiplication from V.

Proposition 14.2. *$\pi(v) := f_v$ is a vector space isomorphism $\pi : \bar{V} \to V^*$.*

Corollary 14.3. $\forall f \in V^*,\ \exists! v \in V : f = f_v$.

14.4 Orthogonality

Definition. u and v are **orthogonal** $\iff$ $(u, v) = 0$. A set of vectors is **orthogonal** iff any pair of distinct vectors is orthogonal. An orthogonal set of unit vectors is **orthonormal**.

Definition. For $W \subseteq V$, the **orthogonal complement** of W is

$$W^\perp := \{u \in V : (u, w) = 0 \quad \forall w \in W\}$$

This is a subspace of V. Note: the term "orthogonal complement" does not appear in the notes, but this definition does and this is the standard term for this construction.

Proposition 14.4. *If V is finite-dimensional and W is a subspace of V, then $V = W \oplus W^\perp$.*

Theorem 14.5. *Any finite-dimensional inner product space has an orthonormal basis, and any orthonormal set can be extended to an orthonormal basis.*

Proposition. *(Gram-Schmidt Process) Given a basis $\{v_1, \ldots, v_n\}$, define*

$$u_1 := \frac{v_1}{||v_1||}$$
$$w_i := v_i - \sum_{j=1}^{i-1} (v_i, u_j) u_j$$
$$u_i := \frac{w_i}{||w_i||}$$

Then $\{u_1, \ldots, u_n\}$ is an orthonormal basis.

Proposition 14.6. *For $v \in V$ and an orthonormal basis $\{u_1, \ldots, u_n\}$ for V, $v = \sum_{i=1}^n (v, u_i)u_i$ and $||v||^2 = \sum_{i=1}^n |(v, u_i)|^2$.*

Definition. The **projection** of $v \neq 0$ along $w \neq 0$ is $\frac{(v,w)}{(w,w)}v$

Definition. The **orthogonal projection map** along a subspace W is $\pi_W(w + w') = w$ where $w \in W$ and $w' \in W^\perp$ (recall that for any v there are unique $w \in W, w' \in W^\perp$ such that $v = w + w'$, so this is well-defined).

Proposition 14.7. $||v - \pi_W(v)|| = \min\{||v - w|| \mid w \in W\}$ *i.e. $\pi_W(v)$ is the closest vector to v in W. If $\{u_1, \ldots, u_r\}$ is an orthonormal basis of W, then $\pi_W(v) = \sum_{i=1}^r (v, v_i)v_i$.*

Proposition 14.8. *Let $E = \{e_1, \ldots, e_n\}$ and $F = \{f_1, \ldots, f_n\}$ be two orthonormal bases of V. Let $P = (p_{ij})$ be the change of basis matrix, i.e. the matrix such that $f_i = \sum_{j=1}^n p_{ji}e_j$. Then $P^T\bar{P} = I$.*

Definition. P is **orthogonal** $\iff$ P is real and $P^TP = I$. P is **unitary** $\iff$ P is complex and $P^T\bar{P} = I$. Each of these types is a group, called a **classical group**.

15 Linear maps on inner product spaces

15.1 Definition and adjoints

Proposition 15.1. *Let V be finite-dimensional For any linear map $T : V \to V$ there is a unique linear map $T^* : V \to V$ such that*

$$\forall u, v \in V,\ (T(u), v) = (u, T^*(v))$$

Definition. T^* is the **adjoint** of T. T is **self-adjoint** $\iff$ $T = T^*$.

Proposition 15.2. *If B is an orthonormal basis then $[T^*]_B = [\bar{T}]_B^T$*

Theorem 15.3. *(Spectral Theorem) If T is self-adjoint then V has an orthonormal basis of T-eigenvectors.*

Corollary 15.4. *If A is real and symmetric, there exists an orthogonal P such that $P^{-1}AP$ is diagonal. If A is complex and Hermitian, there exists a unitary P such that $P^{-1}AP$ is diagonal.*

Lemma 15.5. *If T is self-adjoint, all its eigenvalues are real, eigenvectors to distinct eigenvalues are orthogonal, and if W is T-invariant then so is $W^\perp$.*

15.2 How to find an orthonormal basis

If T is self-adjoint, we can find the basis described in Theorem 15.3 as follows: For each eigenvalue of T, use the Gram-Schmidt Process to find an orthonormal basis of the eigenspace. Then take the union of all these bases.

16 Bilinear and Quadratic Forms

16.1 Bilinear forms

Notation. F is now back to being arbitrary.

Definition. A **bilinear form** on V is $(\cdot,\cdot) : V \times V \to F$ where

$$(\alpha v_1 + \beta v_2, w) = \alpha(v_1, w) + \beta(v_2, w)$$
$$(v, \alpha w_1 + \beta w_2) = \alpha(v, w_1) + \beta(v, w_2)$$

Definition. For $\dim V = n$, the **matrix of a bilinear form** for some basis $B = \{v_1, \ldots, v_n\}$ has entries $a_{ij} = (v_i, v_j)$. Then for any $v, w \in V$, $(v, w) = [v]_B^T A[w]_B$, so A defines $(\cdot,\cdot)$ uniquely.

16.2 Symmetry and skew-symmetry

Definition. $(\cdot,\cdot)$ is **symmetric** $\iff$ $\forall v, w \in V,\ (v, w) = (w, v)$. $(\cdot,\cdot)$ is **skew-symmetric** $\iff \forall v, w \in V,\ (v, w) = -(w, v)$.

Definition. The **characteristic** of F is the smallest $\operatorname{char} F$ such that $\operatorname{char} F = 0$ in F, or 0 if there is no such number. $\operatorname{char}\mathbb{R} = \operatorname{char}\mathbb{C} = 0$. $\operatorname{char}\mathbb{F}_p = p$.

Lemma 16.1. $\operatorname{char} F \neq 2 \implies \forall v \in V,\ (v, v) = 0$

Theorem 16.2. *$(\cdot,\cdot)$ is symmetric or skew-symmetric* $\iff \forall v, w \in V,\ ((v, w) = 0 \iff (w, v) = 0)$.

Definition. For $X \subseteq V$

$$X^\perp := \{u \in V : (u, w) = 0 \quad \forall w \in W\}$$

This is a subspace of V.

Definition. $(\cdot,\cdot)$ is **non-degenerate** $\iff V^\perp = \{0\}$.

Proposition 16.3. *Let V be finite-dimensional and $(\cdot,\cdot)$ be non-degenerate and symmetric or skew-symmetric. Define $f_v(u) := (v, u)$. Then $\phi(v) := f_v$ is an isomorphism $\phi : V \to V^*$. Furthermore, if W is a subspace of V, then* $\dim W^\perp = \dim V - \dim W$.

Definition. A and B both $n \times n$ over F are **congruent** $\iff$ there exists P invertible over F such that $B = P^T A^P$. Two bilinear forms are **equivalent** $\iff$ their matrices are congruent.

Theorem 16.4. *Let V be finite-dimensional,* $\operatorname{char} F \neq 2$, *and $(\cdot,\cdot)$ non-degenerate and skew-symmetric. Then $\dim V$ is even, and there is a basis $B = e_1, f_1, \ldots, e_m, f_m$ for $2m = n$ such that the matrix of $(\cdot,\cdot)$ with respect to B is*

$$\bigoplus_{i=1}^{m} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

Equivalently,

$$(e_i, f_i) = -(f_i, e_i) = 1$$

and if $i \neq j$

$$(e_i, e_j) = (f_i, f_j) = (e_i, f_j) = (f_j, e_i) = 0$$

Corollary 16.5. *If A is invertible and skew-symmetric and* $\operatorname{char} F \neq 2$*, then A is congruent to a matrix of the form above.*

Theorem 16.6. *Let V be finite-dimensional,* $\operatorname{char} F \neq 2$*, and $(\cdot,\cdot)$ non-degenerate and symmetric. Then there is an orthogonal basis $B = \{u_1, \ldots, u_n\}$ such that $(u_i, u_j) = 0$ when $i \neq j$ and $(u_i, u_i) \neq 0$. The matrix of $(\cdot,\cdot)$ with respect to this B is* $\operatorname{diag}(u_1, u_1), \ldots, (u_n, u_n)$.

Corollary 16.7. *If A is invertible and symmetric and* $\operatorname{char} F \neq 2$*, then A is congruent to a diagonal matrix.*

16.3 Quadratic forms

Definition. A **quadratic form** on V is $Q : V \to F$ with $Q(v) := (v, v)$, where $(\cdot,\cdot)$ is symmetric. Q is **non-degenerate** $\iff$ $(\cdot,\cdot)$ is non-degenerate.

Definition. Q, Q' are **equivalent** $\iff$ there is an invertible matrix P such that $Q(Py) = Q'(y)$ $\iff$ their matrices are congruent. Then letting $x = Py$, we have $Q(x) = (Py)^T A(Py) = y^T P^T APy = Q'(y)$ where A is the matrix of Q.

Theorem 16.8. *Let $V = F^n$ and Q be non-degenerate. If $F = \mathbb{C}$ then Q is equivalent to $Q_0(x) := \sum_{i=1}^n x_i^2$, whose matrix is I_n. If $F = \mathbb{R}$ then Q is equivalent to a unique $Q_{p,q} := \sum_{i=1}^p x_i^2 - \sum_{i=p+1}^{p+q} x_i^2$ where $p + q = n$, whose matrix is $I_p \oplus -I_q$. If $F = \mathbb{Q}$ then there are infinitely many inequivalent non-degenerate quadratic forms.*

16.4 Applications

Definition. Let $(\cdot,\cdot)$ be non-degenerate and symmetric or skew-symmetric. T is an **isometry** of $(\cdot,\cdot)$ $\iff$ $\forall u, v \in V,\ (T(u), T(v)) = (u, v)$. The set of isometries $I(V, (\cdot,\cdot))$ is a subgroup of $GL(V)$.