Linear Algebra

Shing Tak Lam

May 13, 2022

This set of notes assumes knowledge of the Groups, Rings and Modules; Vectors and Matrices courses. Furthermore, it assumes some basic category theory. In addition, we assume that all vector spaces have a basis, which is equivalent to the axiom of choice.

Contents

1 Vector spaces 1
- 1.1 Constructions . . . 2
- 1.2 Basis and dimension . . . 2
  - 1.2.1 Constructions . . . 3
  - 1.2.2 Extensionality . . . 4
  - 1.2.3 Rank-nullity . . . 4
- 1.3 Dual space . . . 4
  - 1.3.1 Double dual . . . 5

2 Matrices 6
- 2.1 Change of basis . . . 7
- 2.2 Elementary operations . . . 8
- 2.3 Trace . . . 9

3 Determinant 9

4 Eigenvectors and Eigenvalues 12
- 4.1 Minimal polynomial . . . 14
- 4.2 Jordan normal form . . . 16

5 Bilinear forms 17
- 5.1 Symmetric bilinear forms . . . 18
- 5.2 Sylvester's law of inertia . . . 19
- 5.3 Sesquilinear forms . . . 20
- 5.4 Skew-symmetric bilinear forms . . . 21

6 Inner product spaces 22
- 6.1 Orthogonality . . . 23
- 6.2 Adjoint . . . 24
- 6.3 Spectral theory . . . 25

1 Vector spaces

Definition 1.1 (Category of vector spaces)

The category $\mathbf{Vect}_F$ of all F-vector space has

- as objects, all F-vector space.
- as homs, all F-linear maps.

Definition 1.2 (Isomorphism, Endomorphism, Automorphism)
These are as defined in category theory.

Proposition 1.3. $f \in \mathrm{Hom}(V, W)$ is an isomorphism if and only if it is bijective.

1.1 Constructions

Definition 1.4 (Subspace)
Let V be a F-vector space. Then $W \subseteq V$ is a subspace of V, denoted by $W \leq V$, if and only if W is a F-vector space.

Definition 1.5 (Quotient)
Let V be a F-vector space, $W \leq V$. Then we have the quotient vector space V/W.

Definition 1.6 (Sum)
Let V be a F-vector space, $U_i \leq V$. Then we define the subspace

$$\sum U_i = \left\{ \sum u_j : u_j \in U_j \right\}$$

Definition 1.7 (Internal direct sum)
Let V be a F-vector space, $U, W \leq V$. Then we say V is the internal direct sum of $(U_i)_{i=1}^n$, $V = \bigoplus_{i=1}^n U_i$, if and only if for all $v \in V$, there exists unique $u_i \in U_i$ such that $v = \sum_{i=1}^n u_i$.

Proposition 1.8 (Kernels and cokernels). Kernels and cokernels exist in $\mathbf{Vect}_F$, and for any linear map $\alpha : V \to W$, we have the exact sequence.

$$0 \longrightarrow \ker(\alpha) \longrightarrow V \xrightarrow{\alpha} W \longrightarrow \mathrm{coker}(\alpha) \longrightarrow 0$$

with α injective if and only if $\ker(\alpha) = 0$, α surjective if and only if $\mathrm{coker}(\alpha) = 0$.

Theorem 1.9 (First isomorphism theorem). Any $\alpha : V \to W$ decomposes into

$$V \twoheadrightarrow V/\ker(\alpha) \xrightarrow{\cong} \mathrm{im}(\alpha) \hookrightarrow W$$

1.2 Basis and dimension

Definition 1.10 (Finite dimensional)
A F-vector space is finite dimensional if and only if it has a finite basis.

Theorem 1.11 (Steinitz exchange lemma)**.** Let V be a finite dimensional F-vector space, $v_1, \dots, v_m$ linearly independent, $w_1, \dots, w_n$ spanning set. Then $m \leq n$, and up to reordering, $v_1, \dots, v_m, w_1, \dots, w_n$ spans V.

Proof. By induction on m. $m = 0$ is trivial. Suppose we have $v_1, \dots, v_{m-1}, w_m, \dots, w_n$ spans V. Then $v_m \in \text{span}\{v_1, \dots, v_{m-1}, w_m, \dots, w_n\}$, so there exists α_i, β_i such that

$$v_m = \sum_{i<m} \alpha_i v_i + \sum_{i \geq m} \beta_i w_i$$

Since the v_i are linearly independent, we must have that some β_i is nonzero. Without loss of generality, $\beta_m \neq 0$. Then we have that

$$w_m = \frac{1}{\beta_m}\left(v_m - \sum_{i<m} \alpha_i v_i - \sum_{i>m} \beta_i w_i\right)$$

So $v_1, \dots, v_m, w_{m+1}, \dots, w_n$ spans V. Necessarily, $n \geq m$. □

Corollary 1.12. For a finite dimensional vector space V, all bases have the same size.

Definition 1.13 (Dimension)

Define $\dim(V)$ to be the size of any basis of V.

1.2.1 Constructions

Proposition 1.14. Let $U, W \leq V$. Then

$$\dim(U + W) = \dim(U) + \dim(W) - \dim(U \cap W)$$

Proof. Let $v_1, \dots, v_m$ be a basis of $U \cap W$. Extend to bases $v_1, \dots, v_m, u_1, \dots, u_n$ and $v_1, \dots, v_m, w_1, \dots, w_k$ of U, W respectively. Then $v_1, \dots, v_m, u_1, \dots, u_n, w_1, \dots, w_k$ is a basis for $U + W$. □

Lemma 1.15. The following are equivalent.

1. $$\sum_i V_i = \bigoplus_i V_i$$

2. For all i,
$$V_i \cap \left(\sum_{j \neq i} V_j\right) = 0$$

3. For any basis B_i of V_i, $B = \bigcup_i B_i$ is a basis for $\sum_i V_i$

Corollary 1.16.

$$\dim\left(\bigoplus_i V_i\right) = \sum_i \dim(V_i)$$

Proposition 1.17. Let $U \leq V$. Then

$$\dim(V/U) = \dim(V) - \dim(U)$$

Proof. Let $u_1, \dots, u_m$ be a basis for U. Extend to a basis $u_1, \dots, u_m, v_1, \dots, v_n$ for V. Then $v_1 + U, \dots, v_n + U$ is a basis for V/U. □

1.2.2 Extensionality

Lemma 1.18. Let V, W be F-vector space, B a basis for V. Then any $\alpha : B \to W$ can be extended uniquely to a linear map $V \to W$.

Theorem 1.19. If V is a F-vector space, $\dim(V) = n$, then V is isomorphic to F^n.

Proof. Map basis elements to basis elements. □

Lemma 1.20. If $\alpha : V \to W$ is an isomorphism, $v_1, \dots, v_n$ is a basis for V, then $\alpha(v_1), \dots, \alpha(v_n)$ is a basis for W.

Corollary 1.21. Two finite dimensional F-vector space are isomorphic if and only if they have the same dimension.

1.2.3 Rank-nullity

Definition 1.22 (Rank, nullity)
Let $\alpha : V \to W$ be a linear map. Then

$$\text{rank}(\alpha) = \dim(\text{im}(\alpha)) \quad \text{and} \quad \text{null}(\alpha) = \dim(\ker(\alpha))$$

Theorem 1.23 (Rank-nullity)**.** Let $\alpha : V \to W$, V finite dimensional. Then

$$\dim(U) = \text{rank}(\alpha) + \text{null}(\alpha)$$

Proof. Follows from first isomorphism theorem. □

Corollary 1.24. Let V, W be finite dimensional F-vector space, with $\dim(V) = \dim(U)$. Then α is injective if and only if it is surjective.

1.3 Dual space

Definition 1.25 (Dual space)
Let V be an F-vector space. Then we define the dual of V is the F-vector space

$$V^* = \text{Hom}(V, F)$$

Definition 1.26 (Dual basis)
Let V be an f.d. F-vector space. Say $e_1, \ldots, e_n$ be a basis for V, then we define the dual basis by

$$\varepsilon_i(e_j) = \delta_{ij}$$

Corollary 1.27. If V is finite dimensional, then $V \cong V^*$.

Definition 1.28 (Annihilitor)
If $U \leq V$, define the annihilitor of U by

$$U^0 = \mathrm{Ann}(U) = \{\alpha \in V^* : \alpha(U) = 0\} \leq V^*$$

Proposition 1.29. Suppose V is finite dimensional, $U \leq V$, then

$$\dim(V) = \dim(U) + \dim(\mathrm{Ann}(U))$$

Proof. Let $u_1, \ldots, u_m$ be a basis for U, extend to a basis $u_1, \ldots, u_m, v_1, \ldots, v_n$ of V. Let $\varepsilon_1, \ldots, \varepsilon_m, \xi_1, \ldots, \xi_n$ be the dual basis for V^*. Then $u_1, \ldots, u_m$ is a basis for U, and $\xi_1, \ldots, \xi_n$ is a basis for $\mathrm{Ann}(U)$. □

Definition 1.30 (Dual map)
Let $\alpha : V \to W$. Then we have the dual

$$\alpha^* : W^* \to V^*$$

by functoriality of the hom functor.

Lemma 1.31. Let $\alpha : V \to W$. Then

$$\ker(\alpha^*) = \mathrm{Ann}(\mathrm{im}(\alpha))$$

Lemma 1.32. Let $\alpha : V \to W$. Then

$$\mathrm{Im}(\alpha^*) \leq \mathrm{Ann}(\ker(\alpha))$$

Furthermore, if V, W are finite dimensional, then equality holds, and α^* is surjective if and only if α is injective.

1.3.1 Double dual

Definition 1.33 (Double dual)
Let V be an F-vector space. Then we define the double dual of V to be

$$V^{**} = (V^*)^*$$

Definition 1.34 (Canonical embedding)
Fix $v \in V$. Define $\hat{v} : V^* \to F$ by

$$\hat{v}(\varepsilon) = \varepsilon(v)$$

Theorem 1.35. $v \mapsto \hat{v}$ is an isomorphism.

Lemma 1.36. Let V be a finite dimensional F-vector space. Then

$$U \cong \hat{U} = \{\hat{u} : u \in U\} = \mathrm{Ann}(\mathrm{Ann}(U))$$

Lemma 1.37. Let V be a finite dimensional F-vector space. $U, W \leq V$. Then

$$\mathrm{Ann}(U + W) = \mathrm{Ann}(U) \cap \mathrm{Ann}(W) \quad \text{and} \quad \mathrm{Ann}(U \cap W) = \mathrm{Ann}(U) + \mathrm{Ann}(W)$$

2 Matrices

Definition 2.1 (Matrix)
Let R be a ring. Then the set of $m \times n$ R-matrices is denoted by

$$\mathrm{Mat}_{m,n}(R)$$

Definition 2.2 (Representation of a linear map)
Let $\alpha : V \to W$ be a linear map, $B = (v_1, \ldots, v_n)$ be a basis for V, $C = (w_1, \ldots, w_m)$ be a basis for W. Then define

$$[\alpha]_{B,C} = (a_{ij})$$

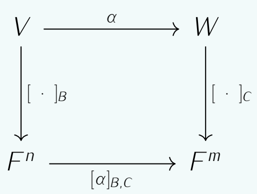
where $\alpha(v_j) = \sum_i a_{ij} w_i$.

Lemma 2.3. Fix bases B, C of V, W respectively, with $\dim(V) = n$, $\dim(W) = m$. Then

$$\alpha \mapsto [\alpha]_{B,C}$$

is an isomorphism $\mathrm{Hom}(V, W) \cong \mathrm{Mat}_{m,n}(F)$

Lemma 2.4.



commutes.

2.1 Change of basis

Definition 2.5 (Change of basis matrix)
Let B, B' be bases for V. The change of basis matrix from B to B' is

$$P = [\mathrm{id}]_{B',B}$$

Proposition 2.6.

$$[\alpha]_{B',C'} = [\mathrm{id}]_{C,C'}[\alpha]_{B,C}[\mathrm{id}]_{B',B}$$

or in matrix form,

$$A' = Q^{-1}AP$$

Definition 2.7 (Equivalent matrices)
A, B are equivalent if there exists invertible matrices P, Q such that

$$A = Q^{-1}BP$$

Proposition 2.8. Let V, W be f.d, $\alpha : V \to W$. Then we have bases B, C of V, W respectively such that

$$[\alpha]_{B,C} = \left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right)$$

where $r = \mathrm{rank}(\alpha)$.

Proof. Extend bases of kernel and image respectively. □

Definition 2.9 (Column, row rank)
The column (resp. row) rank of a matrix A is the dimension of the span of the columns (resp. rows) of A.

Proposition 2.10. Column rank of $[\alpha]_{B,C} = \mathrm{rank}(\alpha)$

Proposition 2.11. Two matrices are equivalent if and only if they have the same column rank.

Theorem 2.12. Column rank = row rank

Proof.

$$A = Q^{-1}\left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right)P$$

Taking the transpose gives the required result. □

Definition 2.13 (Similar, conjugate)
Two matrices A, B are similar, or conjugate if there is an invertible matrix P such that

$$B = P^{-1}AP$$

Proposition 2.14.

$$[a]_{B,C}^{\mathsf{T}} = [a^*]_{C^*,B^*}$$

Proposition 2.15. Let B, C be bases for V, P the change of basis matrix from B to C. Then the change of basis matrix from C^* to B^* is $(P^{-1})^{\mathsf{T}}$

2.2 Elementary operations

Definition 2.16 (Elementary operations)
The elementary column (row) operations are

- Swap column (row) i and j.
- Replace column (row) i with $\lambda\times$ column (row) i.
- Add $\lambda\times$ column (row) j to column (row) i.

Definition 2.17 (Elementary matrix)
For each elementary operation, let E be the result of performing the operation on the identity matrix. Then E is called an elementary matrix.

Proposition 2.18. Each column (row) operation corresponds to right (left) multiplication by the corresponding elementary matrix.

Corollary 2.19. Every matrix is equivalent to $\left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right)$.

Proof. By elementary operations. □

Proposition 2.20. If A is a square invertible matrix, then we can obtain I_n by column (row) operations only.

Proof. By recursion. Suppose after the k-th step we have

$$A^{(k)} = \left(\begin{array}{c|c} I_k & 0 \\ \hline * & * \end{array}\right)$$

We must have some $j > k$ such that $a_{k+1,j} \neq 0$. By column operations, without loss of generality $j = k+1$, $a_{k+1,j} = 1$. Then we can use this to clear out the rest of the $k+1$-th row. The result is

$$A^{(k+1)} = \left(\begin{array}{c|c} I_{k+1} & 0 \\ \hline * & * \end{array}\right)$$

□

Corollary 2.21. Any invertible square matrix is a product of elementary matrices.

2.3 Trace

Definition 2.22 (Trace)
For a square matrix A, define the trace

$$\operatorname{tr}(A) = \sum_i A_{ii}$$

Proposition 2.23.

$$\operatorname{tr}(AB) = \operatorname{tr}(BA)$$

Proposition 2.24. If A, B similar, then $\operatorname{tr}(A) = \operatorname{tr}(B)$.

Proposition 2.25.

$$\operatorname{tr}(A) = \operatorname{tr}\left(A^{\top}\right)$$

Definition 2.26 (Trace of an endomorphism)
For $\alpha \in \operatorname{End}(V)$, define $\operatorname{tr}(\alpha) = \operatorname{tr}\left([\alpha]_{B,B}\right)$ for any basis B.

Proposition 2.27.

$$\operatorname{tr}(\alpha) = \operatorname{tr}(\alpha^*)$$

3 Determinant

Definition 3.1 (Determinant)
Let $A \in \operatorname{Mat}_n(F)$. Define

$$\det(A) = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) A_{\sigma(1),1} \cdots A_{\sigma(n),n}$$

Lemma 3.2. If A is a (strict) upper triangular matrix, that is $a_{ij} = 0$ for $i > j$, then $\det(A) = \operatorname{tr}(A)$.

Proof. The only nonzero term in the sum is $\sigma = \operatorname{id}$. □

Lemma 3.3.

$$\det(A) = \det\left(A^{\mathsf{T}}\right)$$

Definition 3.4 (Volume form)
A volume form d on F^n is a multilinear map $d : (F^n)^n \to F$ such that

$$d(v_1, \ldots, v_i, \ldots, v_j, \ldots, v_n) = 0$$

if $v_i = v_j$ for some $i \neq j$[a].

[a] So d is alternating

Proposition 3.5. det as a function on the columns of a matrix, is a volume form.

Lemma 3.6. Let d be a volume form, $\sigma \in S_n$. Then

$$d(v_{\sigma(1)}, \ldots, v_{\sigma(n)}) = \operatorname{sign}(\sigma) d(v_1, \ldots, v_n)$$

Proof. Suffices to consider a transposition (ij) since S_n is generated by transpositions. Expanding the right hand side of

$$0 = d(v_1, \ldots, v_{i-1}, v_i + v_j, v_{i+1}, \ldots, v_{j-1}, v_i + v_j, v_{j+1}, \ldots, v_n)$$

gives the required result. □

Theorem 3.7. Let d be a volume form, A be a matrix, let $A^{(i)}$ be the columns of A. Then we have that

$$d(A^{(1)}, \ldots, A^{(n)}) = \det(A) d(e_1, \ldots, e_n)$$

Proof. Expand using multilinearity, to get that the LHS is

$$\sum_{i_1, \ldots, i_n} A_{i_1,1} \cdots A_{i_n,n} d(e_{i_1}, \ldots, e_{i_n})$$

and the volume form term forces $i_1, \ldots, i_n = 1, \ldots, n$, so we are summing over permutations, which gives the result required. □

Lemma 3.8.

$$\det(AB) = \det(A)\det(B)$$

Proof. Define the volume form d_A by

$$d_A(v_1, \ldots, v_n) = \det(Av_1, \ldots, Av_n)$$

Then

$$\det(AB) = d_A(B^{(1)}, \ldots, B^{(n)}) = d_A(e_1, \ldots, e_n)\det(B) = \det(A)\det(B)$$

□

Definition 3.9 (Singular)
A square matrix A is singular if $\det(A) = 0$, and it is nonsingular if $\det(A) \neq 0$.

Theorem 3.10. Let $A \in \mathrm{Mat}_n(F)$. Then the following are equivalent.

(i) A is invertible.

(ii) A is nonsingular.

(iii) $\mathrm{rank}(A) = n$.

Proof. (i) and (iii) are equivalent by rank-nullity. Suppose A is invertible. Then $1 = \det(A)\det\left(A^{-1}\right)$, so $\det(A) \neq 0$. Finally, suppose $\mathrm{rank}(A) < n$. Then $(A_1, \dots, A_n)$ is linearly dependent. So $\det(A) = 0$. □

Definition 3.11 (Determinant of an endomorphism)
For $\alpha : V \to V$, define

$$\det(\alpha) = \det([\alpha]_B)$$

for any basis B of V.

Lemma 3.12 (Determinant in block matrix form). Suppose $A \in M_k(F)$, $B \in M_l(F)$, $C \in M_{k,l}(F)$, and

$$M = \left(\begin{array}{c|c} A & C \\ \hline 0 & B \end{array}\right)$$

Then $\det(M) = \det(A)\det(B)$.

Proof.

$$\det(M) = \sum_{\sigma \in S_n} \mathrm{sign}(\sigma) \prod_{i=1}^{n} m_{\sigma(i),i}$$

For this to be nonzero, we must have that $\sigma(j) \leq k$ for $j \leq k$. So $\sigma = \sigma_1\sigma_2$, where σ_1 is a permutation of $1, \dots, k$, and σ_2 is a permutation of $k+1, \dots, n$. Splitting the sum and products gives the required result. □

Corollary 3.13. If

$$A = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

Then $\det(A) = \lambda_1 \cdots \lambda_n$.

Definition 3.14 (Minors)
Let A be a matrix, let $A_{\overline{i,j}}$ be A with the i-th row and j-th column removed. Then define the (i, j) minor to be $\det\left(A_{\overline{i,j}}\right)$.

Lemma 3.15 (Column/row expansion).

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} A_{ij} \det\left(A_{\overline{i,j}}\right) = \sum_{j=1}^{n} (-1)^{i+j} \det\left(A_{\overline{i,j}}\right)$$

Definition 3.16 (Adjugate)
Define the adjugate matrix of $A \in \mathrm{Mat}_n(F)$ by

$$(\mathrm{Adj}(A))_{ij} = (-1)^{i+j} \det\left(A_{\overline{j,i}}\right)$$

Proposition 3.17.

$$(\mathrm{Adj}(A))_{j,i} = \det\left(A^{(1)}, \dots, A^{(j-1)}, e_i, A^{(j+1)}, \dots, A^{(n)}\right)$$

Theorem 3.18.

$$\mathrm{Adj}(A)A = \det(A)I$$

Proof. Compute each entry of $\mathrm{Adj}(A)A$. □

4 Eigenvectors and Eigenvalues

Definition 4.1 (Eigenvector, eigenvalue)
Let $\alpha \in \mathrm{End}(V)$, $\lambda \in F$, $v \in V \setminus 0$. Then v is an eigenvector with eigenvalue λ if

$$\alpha(v) = \lambda v$$

Definition 4.2 (Eigenspace)
For an eignevalue λ, the eigenspace is

$$V_\lambda = \{\alpha(v) = \lambda v : v \in V\} \leq V$$

Lemma 4.3. λ is an eigenvalue if and only if $\det(\alpha - \lambda\,\mathrm{id}) = 0$.

Definition 4.4 (Characteristic polynomial)
For $\alpha \in \mathrm{End}(V)$, the characteristic polynomial of α is

$$\chi_\alpha(t) = \det(\alpha - t\,\mathrm{id})$$

Definition 4.5 (Triangulable)
Let $\alpha \in \mathrm{End}(V)$. Then α is triangulable if there exists a basis B such that $[\alpha]_B$ is upper triangular. That is,

$$[\alpha]_B = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

Theorem 4.6. $\alpha \in \mathrm{End}(V)$ is triangulable if and only if χ_α can be written as a product of linear factors.

Proof. Suppose α is triangulable. Then we have that

$$[\alpha]_B = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

which means that $\chi_\alpha(t) = \det(\alpha - \lambda\,\mathrm{id}) = \prod(\lambda_i - \lambda)$.

Now suppose χ_α is a product of linear factors. By induction on n. If $n = 1$, it is trivial. On the other hand, suppose λ is a root of χ_α. Then let $v_1, \dots, v_j$ be a basis of the eigenspace V_λ. Complete this to a basis $B = (v_1, \dots, v_j, \dots, v_n)$ of V. Then

$$[\alpha]_B = \left(\begin{array}{c|c} \lambda I_k & * \\ \hline 0 & C \end{array}\right)$$

By computing the characteristic polynomial of this matrix, we find that

$$\chi_A(t) = (\lambda - t)^k \chi_C(t)$$

which means that χ_C is also a product of linear factors, and in a suitable basis is upper triangular. □

Lemma 4.7. Suppose $\alpha \in \mathrm{End}(V)$, where $F = \mathbb{R}$ or $\mathbb{C}$. Then say

$$\chi_\alpha(t) = (-1)^n t^n + c_{n-1} t^{n-1} + \dots + c_0$$

Then $c_0 = \det(\alpha)$ and $c_{n-1} = (-1)^{n-1} \mathrm{tr}(\alpha)$.

Proof. Since $\mathrm{Mat}_n(\mathbb{R})$ is a subring of $\mathrm{Mat}_n(\mathbb{C})$, without loss of generality assume $F = \mathbb{C}$.

$$c_0 = \chi_\alpha(0) = \det(\alpha)$$

and α is triangulable over $\mathbb{C}$, so we have that

$$\chi_\alpha(t) = \det \begin{pmatrix} a_1 - t & & * \\ & \ddots & \\ 0 & & a_n - t \end{pmatrix}$$

and $c_{n-1} = (-1)^{n-1}(a_1 + \dots + a_n) = (-1)^{n-1} \mathrm{tr}(\alpha)$. □

Definition 4.8 (Projection)
For each j, let

$$q_j(t) = \prod_{i \neq j} \frac{t - \lambda_i}{\lambda_j - \lambda_i}$$

be the j-th Lagrange basis polynomial. Then $q_j(\lambda_i) = \delta_{ij}$. Let $\Pi_j = q_j(\alpha)$ be the projection operator onto V_{λ_j}.

Lemma 4.9. Let $\alpha \in \mathrm{End}(V)$, and $\lambda_1, \ldots, \lambda_k$ be distinct eigenvalues. Then

$$\sum_{i=1}^{k} V_{\lambda_i} = \bigoplus_{i=1}^{k} V_{\lambda_i}$$

Proof. Let $v \in V_{\lambda_1} \cap (\sum_{i \geq 2} V_{\lambda_i})$. From $v \in V_{\lambda_1}$, we have that $\Pi_j(v) = v$. On the other hand, for $i \geq 2$, for any $w_i \in V_{\lambda_i}$, we have that $\Pi_j(w_i) = 0$. So $v = 0$. Thus the intersection is trivial and the sum is direct. □

Theorem 4.10 (Diagonalisability criterion)**.** Let $\alpha \in \mathrm{End}(V)$. Then α is diagonalisable if and only if there exists $p \in F[t]$ such that p is a product of distinct linear factors, and $p(\alpha) = 0$.

Proof. Suppose α is diagonalisable, with $\lambda_1, \ldots, \lambda_k$ being the distinct eigenvalues. Then $p(t) = \prod_{i=1}^{k}(t - \lambda_i)$ works.

Conversely, suppose such a p exists. Let $\lambda_1, \ldots, \lambda_k$ be the roots of p. Then note that these must be the eigenvalues of α. Let V_{λ_i} be the eigenspaces. Suffices to show that $V = \sum_i V_{\lambda_i}$, since the sum will be direct, and by considering a basis for each eigenspace we get the required result.

Note that by construction, $\sum_j \Pi_j = \mathrm{id}$, so for any vector $v \in V$, we have that

$$v = \sum_j \Pi_j(v) = \sum_j q_j(\alpha)(v)$$

Then

$$(\alpha - \lambda_j \,\mathrm{id})(q_j(\alpha)(v)) = \frac{1}{\prod_{i \neq j} \lambda_j - \lambda_i} p(\alpha)(v) = 0$$

So $\Pi_j(v) \in V_{\lambda_j}$, and V is a sum of the V_{λ_i}. □

Theorem 4.11 (Simultaneous diagonalization)**.** Let $\alpha, \beta \in \mathrm{End}(V)$ be diagonalisable. Then we have a basis B such that α, β are both diagonal if and only if $\alpha\beta = \beta\alpha$.

Proof. If such a basis exists, then diagonal matrices commute. On the other hand, suppose α and β commute. We have that

$$V = \sum_{i=1}^{k} V_{\lambda_i} \quad \text{where } \lambda_i \text{ eigenvalues of } \alpha$$

Then for $v \in V_{\lambda_i}$, $\alpha\beta(v) = \beta\alpha(v) = \beta(\lambda_i v) = \lambda_i \beta(v)$. So $(\alpha - \lambda \,\mathrm{id})(\beta(v)) = 0$, and $\beta(v) \in V_{\lambda_i}$. This means that for each λ_i we have an endomorphism $\beta_i = \beta|_{V_{\lambda_i}} \in \mathrm{End}(V_{\lambda_i})$.

Since β diagonalisable, we have $p \in F[t]$ such that $p(\beta) = 0$, p product of distinct linear factors. Then for each i, we have that $p(\beta_i) = p(\beta) = 0$, so each β_i is diagonalisable. Let B_i be a basis of V_{λ_i} such that $(\beta_i)_{B_i}$ is diagonal. Then $B = B_1 \cup \cdots \cup B_k$ works. □

4.1 Minimal polynomial

Definition 4.12 (Minimal polynomial)

Let V be a finite dimensional F-vector space, $\alpha \in \mathrm{End}(V)$. The minimal polynomial of α is $m_\alpha(t)$, which is the minimum degree (non-zero) polynomial such that $m_\alpha(\alpha) = 0$.

Lemma 4.13. Let $\alpha \in \mathrm{End}(V)$. Then $p(\alpha) = 0$ if and only if $m_\alpha \mid p$.

Proof. $\Longleftarrow$ is obvious. Suppose $p(\alpha) = 0$. Then we have that $\deg(p) \geq \deg(m_\alpha)$, and we have that $p = qm_\alpha + r$, with $\deg(r) < \deg(m_\alpha)$. By minimality of $\deg(m_\alpha)$, we must have $r = 0$. □

Corollary 4.14. Minimum polynomial is unique up to a constant.

Theorem 4.15 (Cayley-Hamilton).

$$\chi_\alpha(\alpha) = 0$$

Remark 4.16. In the following proof, we will need to take the determinant and adjugate of matrices defined over general rings, and not just fields. This is well defined, and the same formulae will work. Furthermore, we will use the isomorphism $\mathrm{Mat}_n(F[t]) \cong (\mathrm{Mat}_n(F))[t]$ throughout the proof without explicitly mentioning it.

Proof. Let $A \in \mathrm{Mat}_n(F)$, $\det(t\,\mathrm{id} - A) = (-1)^n \chi_\alpha(t) = t^n + a_{n-1}t^{n-1} + \cdots + a_0$. Let $B = t\,\mathrm{id} - A \in \mathrm{Mat}_n(F[t])$. Then we have that

$$\mathrm{Adj}(B) = B_{n-1}t^{n-1} + \cdots + B_1 t + B_0 \in (\mathrm{Mat}_n(F))[t]$$

Then note that

$$(t\,\mathrm{id} - A)(B_{n-1}t^{n-1} + \cdots + B_1 t + B_0) = (t\,\mathrm{id} - A)\,\mathrm{Adj}(B) = B\,\mathrm{Adj}(B) = \det(B)\,\mathrm{id} = (t^n + a_{n-1}t^{n-1} + \cdots + a_0)\,\mathrm{id}$$

Equating coefficients, we have that

$$B_{n-1} = \mathrm{id}, B_{n-2} - AB_{n-1} = a_{n-1}\,\mathrm{id}, \ldots, -AB_0 = a_0\,\mathrm{id}$$

Rearranging, we find that

$$A^n B_{n-1} = A^n, A^{n-1}B_{n-2} - A^n B_{n-1} = a_{n-1}A^{n-1}, \ldots, -AB_0 = a_0\,\mathrm{id}$$

Summing these equations, we find that

$$A^n + a_{n-1}A^{n-1} + \cdots + a_0\,\mathrm{id} = 0$$

□

Definition 4.17 (Algebraic, geometric multiplicity)

For $\alpha \in \mathrm{End}(V)$, the algebraic multiplicity is a_λ, which is the multiplicity of $t - \lambda$ in $\chi_\alpha(t)$. The geometric multiplicity is $g_\lambda = \dim(V_\lambda)$.

Lemma 4.18.

$$g_\lambda \leq a_\lambda$$

Proof. Let $v_1, \ldots, v_{g_\lambda}$ be a basis for v_λ. Extend to a basis $B = (v_1, \ldots, v_{g_\lambda}, \ldots, v_n)$ for V. Then we have that

$$[\alpha]_B = \left(\begin{array}{c|c} \lambda\,\mathrm{id}_{g_\lambda} & * \\ 0 & A \end{array}\right)$$

Then $\det(\alpha - t\,\mathrm{id}) = (\lambda - t)^{g_\lambda}\chi_A(t)$, so $g_\lambda \mid \chi_\alpha$, and $g_\lambda \leq a_\lambda$. □

Lemma 4.19. Let c_λ be the multiplicity of $t - \lambda$ in $m_\alpha(t)$. Then $c_\lambda \leq a_\lambda$.

Proof. By Cayley-Hamilton $m_\alpha \mid \chi_\alpha$. □

Lemma 4.20 (Diagonalisation over $\mathbb{C}$)**.** Let $\alpha \in \text{End}(V)$, where V is a finite dimensional $\mathbb{C}$-vector space. Then the following are equivalent.

(i) α is diagonalisable.

(ii) For all eigenvalues λ of α, $a_\lambda = g_\lambda$.

(iii) For all eigenvalues λ of α, $c_\lambda = 1$.

Proof. We have already shown that (i) $\iff$ (iii). Note that

$$V \geq \bigoplus_{i=1}^{k} V_{\lambda_i}$$

By considering the dimensions, we get that (i) $\iff$ (ii). □

4.2 Jordan normal form

Definition 4.21 (Jordan block)

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

Definition 4.22 (Jordan normal form)

A matrix $A \in \text{Mat}_n(\mathbb{C})$ is in Jordan normal form if it is a block diagonal matrix where each block is a Jordan block.

Theorem 4.23 (Jordan normal form)**.** Any $A \in \text{Mat}_n(\mathbb{C})$ is similar to a matrix in Jordan normal form, which is unique up to reordering of the Jordan blocks.

Proof. See GRM. □

Proposition 4.24. From the JNF of a matrix, we can find

- $a_\lambda = \sum$ size of blocks with eigenvalue λ
- $g_\lambda =$ number of blocks with eigenvalue λ
- $c_\lambda =$ size of largest block with eigenvalue λ.

Theorem 4.25 (Generalised eigenspace decomposition)**.** Let $\alpha \in \text{End}(V)$, where V is a finite dimensional $\mathbb{C}$-vector space. Define the generalised eigenspace

$$V_j = \ker((\alpha - \lambda_j \,\text{id})^{c_j})$$

Then

$$V = \bigoplus_{j=1}^{k} V_j$$

Proof. Let $p_j(t) = \prod_{i \neq j} (t - \lambda_i)^{c_i}$. Consider the ideal $(p_1, \ldots, p_k) \trianglelefteq \mathbb{C}[t]$. Since $\mathbb{C}$ is a field, $\mathbb{C}[t]$ is a PID. So we have $f \in \mathbb{C}[t]$ such that $(p_1, \ldots, p_k) = (f)$. In this case, since $f \mid p_i$ for all i, we must in fact have $f \in \mathbb{C}^\times$. So $(p_1, \ldots, p_k) = (1)$. This means that we have $q_1, \ldots, q_k$ such that

$$q_1 p_1 + \cdots + q_k p_k = 1$$

Define the projection $\Pi_j = q_j p_j(\alpha)$. Then $\sum_j \Pi_j = \mathrm{id}$. Furthermore,

$$(a - \lambda_j \,\mathrm{id})^{c_j} \Pi_j(v) = (\alpha - \lambda_j \,\mathrm{id})^{c_j} q_j p_j(\alpha)(v) = q_j m_\alpha(\alpha(v)) = 0$$

So $\Pi_j(v) \in V_j$, which means that $V = \sum_j V_j$. Furthermore, note that for $i \neq j$, $\Pi_i \Pi_j = 0$, since $m_\alpha \mid p_i p_j$, which also means that $\Pi_i = \Pi_i \,\mathrm{id} = \Pi_i \sum_j \Pi_j = \Pi_i^2$. So Π_i idempotent.

This implies that the sum is direct, since if we have $v \in V_1 \cap (\sum_{i \geq 2} V_i)$. Since $v \in V_1$, by construction $\Pi_j(v) = 0$ for $j \geq 2$, which means that $\Pi_1(v) = v$. On the other hand, for $w_i \in V_i$ for $i \geq 2$, $\Pi_1(w_i) = 0$ by the same argument. So $v = \Pi_1(v) = 0$. □

5 Bilinear forms

Definition 5.1 (Bilinear form)

A bilinear form $\phi : U \times V \to F$ is a function which is linear in each of its arguments.

Definition 5.2 (Representation of bilinear forms)

Let $B = (v_1, \ldots, v_m)$ and $C = (w_1, \ldots, w_n)$, then

$$[\phi]_{B,C} = \big(\phi(v_i, w_j)\big)_{i,j}$$

Lemma 5.3.

$$\phi(u, v) = [u]_B^\top [\phi]_{B,C} [v]_C$$

Definition 5.4 (Left (right) map)

Given a bilinear form $\phi : U \times V \to F$, define $\phi_L : U \to V^*$ by

$$\phi_L(u) = \phi(u, \cdot)$$

and define ϕ_R similarly.

Lemma 5.5. Let B, C be bases for U, V, and B^*, C^* the dual bases. Let $A = [\phi]_{B,C}$. Then $[\phi_L]_{B,C^*} = A^\top$, and $[\phi_R]_{C,B^*} = A$.

Definition 5.6 (Non-degenerate)

ϕ is non-degenerate if $\ker(\phi_L) = 0$ and $\ker(\phi_R) = 0$.

Lemma 5.7. ϕ is non-degenerate if and only if A is invertible, and $\dim(U) = \dim(V)$.

Proof. By rank-nullity. □

Corollary 5.8. If $\dim U = \dim V$, then choosing a non-degenerate bilinear form $\phi : U \times V \to F$ is equivalent to choosing an isomorphism $\phi_L : U \to V^*$.

Definition 5.9 (Orthogonal)
For $T \subseteq U$, define

$$T^\perp = \{v \in V : \phi(t, v) = 0 \forall t \in T\}$$

and for $S \subseteq V$, define

$$^\perp S = \{u \in U : \phi(u, s) = 0 \forall s \in S\}$$

Proposition 5.10 (Change of basis). Let B, B' be bases for U, C, C' bases for V. Then P, Q be the respective change of bases matrices. Then

$$[\phi]_{B',C'} = P^\mathsf{T} [\phi]_{B,C} Q$$

Definition 5.11 (Rank)
The rank of a bilinear form ϕ is the rank of any matrix representing it.

Definition 5.12 (Congruent)
$A, B \in \mathrm{Mat}_n(F)$ are congruent if there exists $P \in \mathrm{Mat}_n(F)$ invertible such that

$$A = P^\mathsf{T} B P$$

5.1 Symmetric bilinear forms

Definition 5.13 (Symmetric)
A bilinear form $\phi : V \times V \to F$ is symmetric if $\phi(u, v) = \phi(v, u)$.

Proposition 5.14. ϕ is symmetric if and only if for any basis B, $[\phi]_B$ is symmetric.

Proposition 5.15. $[\phi]_B$ is diagonal only if ϕ is symmetric.

Definition 5.16 (Quadratic form)
A map $Q : V \to F$ is a quadratic form if there exists a bilinear form ϕ such that $Q(v) = \phi(v, v)$.

Remark 5.17. The matrix representing a quadratic form with respect to a given basis need not be unique. If A is one, then so is $\frac{1}{2}(A + A^{\mathsf{T}})$[a].

[a]This is the symmetric part of A.

Proposition 5.18 (Polarisation identity)**.** If ϕ is a symmetric bilinear form, $Q(v) = \phi(v, v)$ a quadratic form, then

$$\phi(u, v) = \frac{1}{2}\left(Q(u + v) - (Q(u) + Q(v))\right)$$

Proposition 5.19. If $Q : V \to F$ is a quadratic form, then there exists a unique symmetric bilinear form $\phi : V \times V \to F$ such that $Q(v) = \phi(v, v)$.

Proof. Follows by the polarisation identity. □

Theorem 5.20 (Diagonalisation of symmetric bilinear forms)**.** Let $\phi : V \times V \to F$ be a symmetric bilinear form, with $\dim(V) = n$. Then there exists a basis B for V such that $[\phi]_B$ is diagonal.

Proof. By induction on n. $n = 1$ is trivial. If $\phi(u, u) = 0$ for all u, then by the polarisation identity $\phi = 0$. So if ϕ is nonzero, then there must be e_1 such that $\phi(e_1, e_1) \neq 0$.

Let $U = (\operatorname{span}\{e_1\})^{\perp} = \{v \in V : \phi(e_1, v) = 0\} = \ker(\phi(e_1, \cdot))$. Then $V = \operatorname{span} e_1 \oplus U$, which means that $\dim(U) = n - 1$. By the induction hypothesis, we can choose a basis $e_2, \dots, e_n$ for U such that $\phi|_{U \times U} : U \times U \to F$ is diagonal. Then $e_1, \dots, e_n$ is a basis for V such that ϕ is diagonal. □

Corollary 5.21. Over an algebraically closed field F (e.g. $\mathbb{C}$), we have a basis for V such that

$$[\phi] = \left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right)$$

where $r = \operatorname{rank}(\phi)$.

Proof. By rescaling and reordering the basis elements. □

Corollary 5.22. Over $\mathbb{R}$, we have a basis for V such that

$$[\phi] = \left(\begin{array}{c|c|c} I_p & 0 & 0 \\ \hline 0 & -I_q & 0 \\ \hline 0 & 0 & 0 \end{array}\right)$$

where $p, q \geq 0$, $p + q = \operatorname{rank}(\phi)$.

Proof. Rescale and reorder the basis elements. □

5.2 Sylvester's law of inertia

Theorem 5.23 (Sylvester's law of inertia)**.** Suppose we have that

$$[\phi]_B = \left(\begin{array}{c|c|c} I_p & 0 & 0 \\ \hline 0 & -I_q & 0 \\ \hline 0 & 0 & 0 \end{array}\right) \quad \text{and} \quad [\phi]_{B'} = \left(\begin{array}{c|c|c} I_{p'} & 0 & 0 \\ \hline 0 & -I_{q'} & 0 \\ \hline 0 & 0 & 0 \end{array}\right)$$

Then $p = p'$, $q = q'$.

Remark 5.24. In the course we use positive definite to mean $\phi(u, u) > 0$, and positive semidefinite to mean $\phi(u, u) \geq 0$.

Proof. Suffices to show $p \geq p'$ by symmetry. Furthermore, the argument for p and q are similar so we only prove one case.

Let $U \leq V$ be any subspace such that $\phi|_{U \times U}$ is positive definite. Let $B = (v_1, \dots, v_p, \dots, v_{p+q}, \dots, v_n)$. Let $W = \text{span}\{v_{p+1}, \dots, v_n\}$. Then ϕ is negative semidefinite on W, so $U \cap W = 0$. This means that

$$n \geq \dim(U) + \dim(W) = \dim(U) + (n - p) \implies p \geq \dim(U)$$

Let $B' = (v'_1, \dots, v'_p, \dots, v_{p+q'}, \dots, v'_n)$. Then setting $U = \text{span}\{v'_1, \dots, v'_p\}$ we get that $p \geq p'$. □

Definition 5.25 (Signature)

For a symmetric bilinear form $\phi : V \times V \to \mathbb{R}$, define the signature to be

$$S(\phi) = p - q$$

Definition 5.26 (Kernel)

For a symmetric bilinear form ϕ, define the kernel

$$\ker(\phi) = \{v \in V : \forall u \in U, \phi(u, v) = 0\}$$

Lemma 5.27 (Rank-nullity).

$$\dim(\ker(\phi)) + \text{rank}(\phi) = 0$$

Proposition 5.28. There exists a subspace T such that

$$\phi|_T = 0 \quad \text{and} \quad \dim(T) = n - (p + q) + \min(p, q)$$

Proof. Suppose $B = (v_1, \dots, v_p, \dots, v_{p+q}, \dots, v_n)$ is a basis which makes ϕ into the form in Sylvester's law of inertia. Without loss of generality, assume $p \geq q$. Then $T = \text{span}\{v_1 + v_{p+1}, \dots, v_q + v_{p+q}, v_{p+q+1}, \dots, v_n\}$ works. □

Proposition 5.29. $\dim(T)$ in the previous proposition is maximal.

5.3 Sesquilinear forms

Definition 5.30 (Sesquilinear form)

Let V, W be $\mathbb{C}$-vector space. Then $\phi : V \times W \to \mathbb{C}$ is a sesquilinear form if $\phi(\cdot, w)$ is linear, and $\phi(v, \cdot)$ is conjugate linear.

Definition 5.31 (Representation of a sesquilinear form)
Let $B = (v_1, \ldots, v_m)$ and $C = (w_1, \ldots, w_n)$, then

$$[\phi]_{B,C} = \big(\phi(v_i, w_j)\big)_{i,j}$$

Lemma 5.32.

$$\phi(v, w) = [u]_B^{\mathsf{T}}[\phi]_{B,C}\overline{[v]_C}$$

Definition 5.33 (Change of basis)

$$[\phi]_{B',C'} = P^{\mathsf{T}}[\phi]_{B,C}\overline{Q}$$

Definition 5.34 (Hermitian)
Let $\phi : V \times V \to \mathbb{C}$ be a sesquilinear form. Then ϕ is Hermitian if $\phi(u, v) = \overline{\phi(v, u)}$.

Proposition 5.35. For a Hermitian form ϕ, $\phi(u, u)$ is real.

Remark 5.36. This means that we can refer to positive/negative (semi)definite Hermitian forms.

Lemma 5.37. ϕ is Hermitian if and only if for any basis B, we have that

$$[\phi]_B = \overline{[\phi]_B^{\mathsf{T}}}$$

Lemma 5.38 (Polarisation)**.** Let $\phi : V \times V \to \mathbb{C}$ be sesquilinear, $Q(v) = \phi(v, v)$. Then

$$\phi(u, v) = \frac{1}{4}\left(Q(u + v) - Q(u - v)\right) + \frac{i}{4}\left(Q(u + iv) - Q(u - iv)\right)$$

Theorem 5.39 (Sylvester's law of inertia)**.** Let ϕ be a Hermitian form. Then we have a basis such that

$$[\phi] = \left(\begin{array}{c|c|c} I_p & 0 & 0 \\ \hline 0 & -I_q & 0 \\ \hline 0 & 0 & 0 \end{array}\right)$$

Furthermore, p, q are independent of the choice of basis.

Proof. Existence and uniqueness follows from similar proofs for the real case. □

5.4 Skew-symmetric bilinear forms

Definition 5.40 (Skew-symmetric)
Let $\phi : V \times V \to R$ be a bilinear form. Then it is skew-symmetric if $\phi(u, v) = -\phi(v, u)$.

Proposition 5.41. $\phi(u, u) = 0$.

Proposition 5.42.

$$[\phi]_B = -[\phi]_B^{\mathsf{T}}$$

Proposition 5.43. Every square matrix is the sum of a symmetric and an antisymmetric matrix.

Proof.

$$A = \frac{1}{2}\left(A + A^{\mathsf{T}}\right) + \frac{1}{2}\left(A - A^{\mathsf{T}}\right)$$

□

Theorem 5.44 (Sylvester's law of inertia). Let ϕ be a skew-symmetric form. Then we have a basis $B = (v_1, w_1, \ldots, v_m, w_m, u_{2m+1}, \ldots, u_n)$ such that

$$[\phi]_B = \begin{pmatrix} T & & & \\ & \ddots & & \\ & & T & \\ & & & 0 \end{pmatrix} \quad \text{where} \quad T = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

Proof. By induction on $\dim(V)$. If $\phi = 0$ then we are done. Otherwise, we have v_1, w_1 such that $\phi(v_1, w_1) \neq 0$. By rescaling, $\phi(v_1, w_1) = 1$ and $\phi(w_1, v_1) = -1$. Then let $U = \operatorname{span}\{v_1, w_1\}$, and define

$$W = \{v \in V : \phi(v_1, v) = \phi(w_1, v) = 0\}$$

Then $V = U \oplus W$. □

Corollary 5.45. The rank of a skew-symmetric form is even.

6 Inner product spaces

Definition 6.1 (Inner product)

Let V be a vector space over $\mathbb{C}$, an inner product on V is a positive definite Hermitian form ϕ on V.

Remark 6.2. For real vector spaces, we have symmetric bilinear forms instead. However, most of the time we won't have to worry about the difference.

Remark 6.3. We write $\langle u, v\rangle = \phi(u, v)$ to denote the inner product of u and v.

Theorem 6.4 (Cauchy-Schwarz inequality).

$$|\langle u, v\rangle| \leq \|u\|\|v\|$$

Corollary 6.5 (Triangle inequality).

$$||u+v|| \leq ||u|| + ||v||$$

6.1 Orthogonality

Definition 6.6 (Orthogonal)
A set $\{e_1, \ldots, e_k\}$ of vectors are orthogonal if $i \neq j \implies \langle e_i, e_j \rangle = 0$.

Definition 6.7 (Orthonormal)
A set $\{e_1, \ldots, e_k\}$ of vectors is orthonormal if $\langle e_i, e_j \rangle = \delta_{ij}$.

Lemma 6.8. A set of orthogonal nonzero vectors is linearly independent.

Lemma 6.9 (Parseval). If $e_1, \ldots, e_n$ is an orthonormal basis, then

$$\langle u, v \rangle = \sum_{i=1}^{n} u_i \overline{v_i}$$

Theorem 6.10 (Gram-Schmidt orthogonalisation). Let V be an inner product space, v_i ($i \in I \subseteq \mathbb{N}$) be a collection of linearly independent vectors. Then there exists e_i orthonormal such that for all k,

$$\operatorname{span}\{e_1, \ldots, e_k\} = \operatorname{span}\{v_1, \ldots, v_k\}$$

Proof. Define $e_1 = v_1/\|v_1\|$, and

$$u_{k+1} = v_{k+1} - \sum_{i=1}^{k} \langle v_{k+1}, e_i \rangle e_i$$

and $e_{k+1} = u_{k+1}/\|u_{k+1}\|$. □

Corollary 6.11. Any finite dimensional inner product space has an orthonormal basis.

Definition 6.12 (Orthogonal matrix)
$A \in \mathrm{Mat}_n(\mathbb{R})$ is orthogonal if $A^{\mathsf{T}} A = \mathrm{id}$.

Proposition 6.13 (QR decomposition). Let $A \in \mathrm{Mat}_n(\mathbb{R})$ be nonsingular. Then Q can be written uniquely as

$$A = QR$$

where Q orthogonal, R upper triangular. If $A \in \mathrm{Mat}_n(\mathbb{C})$, then Q can be chosen to be unitary.

Proof. Gram-Schmidt on the columns of A. Or see IB Numerical Analysis for more computationally stable methods. □

Definition 6.14 (Orthogonal direct sum)

Let $V_1, V_2 \leq V$. Then we say that V is the orthogonal direct sum of V_1, V_2, written $V = V_1 \overset{\perp}{\oplus} V_2$, if $V = V_1 \oplus V_2$ and for all $v_1 \in V_1, v_2 \in V_2$, $\langle v_1, v_2 \rangle = 0$.

Definition 6.15 (Orthogonal)
For $U \leq V$, define

$$U^{\perp} = \{v \in V : \langle u, v \rangle = 0\} \leq V$$

Proposition 6.16.

$$V = U \overset{\perp}{\oplus} U^{\perp}$$

Definition 6.17 (Projection)
Suppose $V = U \oplus W$. Then define the projection onto W by $\Pi : V \to W$ by $v = u + w \mapsto w$.

Remark 6.18. In general, Π depends on U. However, we usually take $U = W^{\perp}$.

Lemma 6.19. Let $W \leq V$, $e_1, \ldots, e_k$ be a basis for W, Π projection onto W. Then

$$\Pi(v) = \sum_{i=1}^{k} \langle v, e_i \rangle e_i$$

Proposition 6.20 (Least squares)**.** For all $v \in V$, $w \in W$,

$$\left\| v - \Pi(v) \right\| \leq \|v - w\|$$

with equality holding if and only if $w = \Pi(v)$.

6.2 Adjoint

Proposition 6.21. Let $\alpha : V \to W$ be a linear map. Then there is a unique linear map $\alpha^* : W \to V$ such that

$$\langle \alpha(v), w \rangle = \langle v, \alpha^*(w) \rangle$$

Moreover, if B, C are orthonormal bases of V, W respectively, then

$$[\alpha^*]_{C,B} = \overline{[\alpha]_{B,C}^{\mathsf{T}}}$$

Definition 6.22 (Adjoint)
Define α^* to be the adjoint of α.

Proposition 6.23. By identifying $V \cong V^*$ and $W \cong W^*$, the dual and adjoint are the same maps.

Definition 6.24 (Self adjoint)
Let $\alpha \in \text{End}(V)$. Then if $\alpha^* = \alpha$, we say α is self adjoint.

Proposition 6.25. α self-adjoint if and only if $\langle \alpha(u), v \rangle = \langle u, \alpha(v) \rangle$.

Proposition 6.26. If $F = \mathbb{R}$, then self adjoint is equivalent to symmetric. If $F = \mathbb{C}$, then self adjoint is equivalent to Hermitian.

Definition 6.27 (Isometry)
Let $\alpha : V \to W$. If $\langle \alpha(v), \alpha(w) \rangle = \langle v, w \rangle$, then we say α is an isometry.

Proposition 6.28. α is an isometry if and only if $\alpha^* = \alpha^{-1}$.

Proposition 6.29. For endomorphisms over $\mathbb{R}$, isometry is equivalent to orthogonal. For endomorphisms over $\mathbb{C}$, isometry is equivalent to unitary.

6.3 Spectral theory

Lemma 6.30. Let $\alpha \in \text{End}(V)$ be self adjoint. Then all eigenvalues of α are real, and eigenvectors with distinct eigenvalues are orthogonal.

Theorem 6.31. Let $\alpha \in \text{End}(V)$ be self adjoint. Then V has an orthonormal basis of eigenvectors.

Proof. By induction on $n = \dim(V)$. $n = 1$ is trivial. By FTA, χ_α has a root over $\mathbb{C}$. Since α is self adjoint, the root is real. Let $\lambda \in \mathbb{R}$, and let v be an eigenvector with eigenvalue λ and norm 1. Let $U = (\text{span}\{v\})^\perp$. Then we have that $\alpha|_U : U \to U$ is a self-adjoint endomorphism. □

Corollary 6.32. All self adjoint operators are diagonalisable by unitary operators.

Corollary 6.33. V is a direct sum of the eigenspaces.

Proposition 6.34. Let V be a complex inner product space, $\alpha \in \mathrm{End}(V)$ isometry (unitary). Then all eigenvalues of α have modulus 1, and eigenvectors with distinct eigenvalues are orthogonal.

Theorem 6.35. Let V be a complex inner product space, $\alpha \in \mathrm{End}(V)$ isometry (unitary). Then V has an orthonormal basis of eigenvectors.

Proof. By induction as for self-adjoint maps. □

Proposition 6.36. A symmetric (Hermitian) matrix is diagonalisable by an orthogonal (unitary) matrix.

Proof. Matrix with eigenvectors as columns. □

Proposition 6.37. Let ϕ be a symmetric (Hermitian) form, then there is an orthonormal basis of V such that ϕ is diagonal.

Proof. Basis of eigenvectors. □

Remark 6.38. The diagonal entries are the eigenvalues. Furthermore, $S(\phi) = p - q$, where p is the number of positive eigenvalues and q is the number of negative eigenvalues.

Proposition 6.39 (Simultaneous diagonalisation)**.** Let ϕ, ψ be symmetric (Hermitian) forms, ϕ is positive definite. Then there exists a basis of V such that ϕ, ψ are diagonal.

Proof. (V, ϕ) is an inner product space. So we have an orthonormal basis of (V, ϕ) such that ψ is diagonal. Furthermore, $[\phi] = \mathrm{id}$. □

Proposition 6.40 (Simultaneous diagonalisation of matrices)**.** Ley A, B be square symmetric (Hermitian) matrices, A positive definite. Then there exists $Q \in \mathrm{Mat}_n(\mathbb{R})$ $(Q \; in \, \mathrm{Mat}_n(\mathbb{C}))$ invertible such that

$$Q^{\mathsf{T}}AQ \quad (Q^{\mathsf{T}}A\overline{Q}) \quad \text{and} \quad Q^{\mathsf{T}}BQ \quad (Q^{\mathsf{T}}B\overline{Q})$$

are diagonal.

Proof. $\phi(u, v) = u^{\mathsf{T}}Av$ is a positive symmetric (Hermitian) form, so apply previous result. □