# MATH 593 - Multilinear Algebra

ARessegetes Stery

December 19, 2023

# Contents

# 1   The Tensor Algebra

**Definition 1.1** (Multilinear). *Let $R$ be a commutative ring, and $M_1, \cdots, M_n, N$ be $R$-modules. A map $\varphi : M_1 \times \cdots \times M_n \to N$ is **multilinear** if for all $i \in [\![1, n]\!]$, for all $x_j \in M_j$ for $j \neq i$, the map $\varphi(x_1, \cdots, x_{i-1}, -, x_{i+1}, \cdots, x_n) : M_i \to M$ is $R$-linear.*

**Remark 1.1.** Via performing induction on $n$, it can be shown that for a multilinear map $f : M_1 \times \cdots \times M_n \to N$, with the tensor map $\varphi : M_1 \times \cdots \times M_n \to M_1 \otimes_R \cdots \otimes_R M_n$ (which is multilinear), there exists a $R$-linear map $g : M_1 \otimes_R \cdots \otimes_R M_n \to P$ s.t. $g \circ \varphi = f$.

**Definition 1.2** (Tensor Algebra). *Let $M$ be a fixed $R$-module. Define $T^0(M) := R, T^1(M) = M$; and for $n \geq 2$, define $T^n(M) := \underbrace{M \otimes_R \cdots \otimes_R M}_{n \ times}$. Then the **tensor algebra** is defined as*

$$T(M) := \bigoplus_{i \geq 0} T^i(M) = R \oplus M \oplus (M \otimes_R M) \oplus \cdots$$

**Remark 1.2.** Since for all $i$, $T^i(M)$ has an $R$-module structure, $T(M)$ is also an $R$-module.

**Proposition 1.1.** $T(M)$ *also has an $R$-algebra structure.*

*Proof.* It suffices to define multiplication for each summand of $T(M)$ and check that it is well-defined. Define
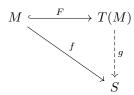
$$\alpha_{ij} : T^i(M) \times T^j(M) \to T^{i+j}(M), \quad (a_1 \otimes \cdots \otimes a_i, b_1 \otimes \cdots \otimes b_j) \mapsto (a_1 \otimes \cdots \otimes a_i \otimes b_1 \otimes \cdots \otimes b_j)$$

This is indeed well-defined, as by applying the universal property of tensor product for $i$ times gives the desired map. Notice that for the case where $i = 0$ or $j = 0$ this is just scalar multiplication, this is just scalar multiplication. $\square$

**Remark 1.3.** This can be extended to a map $T(M) \times T(M) \to T(M)$, which makes $T(M)$ a ring. The map is given by for $x = \oplus_{i \geq 0} x_i$, $y = \oplus_{j \geq 0} y_j$, the multiplication is defined as $x \cdot y = \oplus_{i,j \geq 0} \alpha_{i,j}(x_i, y_j)$, with $1 \in T^0(M) = R$. Moreover, the inclusion $R = T^0(M) \hookrightarrow T(M)$ is a ring homomorphism which makes $T(M)$ an $R$-algebra.

**Remark 1.4.** Notice that this differs from the polynomial ring in that it is not commutative (in terms of the direct summands). Therefore, the terms in $\oplus_{i,j} \alpha_{ij}$ cannot be collected into one. $T^n = M \otimes \cdots \otimes M$ has a basis given by $x_{i_1} \otimes \cdots \otimes x_{i_n}$, where $i_k \in [\![1, d]\!]$ for all $k \in [\![1, n]\!]$. Therefore, for $n \geq 2$, $T^n(M)$ is not commutative.

**Proposition 1.2** (Universal Property of $T(M)$). *Consider the (forgetful) functor $F :_R \underline{\mathrm{Alg}} \to {}_R\underline{\mathrm{Mod}}, M \mapsto T(M)$. Let $M$ be an $R$-module and $S$ an $R$-algebra. If $f : M \to S$ is a morphism of $R$-modules, then there exists a unique morphism of $R$-algebras $g : T(M) \to S$ s.t. $g|_{T^1(M)} = f$, i.e. $g \circ F = f$:*



*Proof.* Apply the universal property of tensor product and direct sum.

Define $f_n : \underbrace{M \times \cdots \times M}_{n \text{ times}} \to S$, where $f_n(x_1, \cdots, x_n) = f(x_1) \cdots f(x_n)$. This is clearly multilinear, which implies that there exists a unique $R$-linear map $g_n : T^n(M) \to S$ s.t. $g_n(x_1 \otimes \cdots \otimes x_n) = f(x_1) \cdots f(x_n)$. Apply the universal property of direct sum on the superscript gives that there exists a unique $R$-linear map $g : T(M) \to S$ s.t. $g|_{T^n(M)} = g_n$. Check the followings:

- $g$ is a morphism of $R$-algebras. Since it is already a morphism of $R$-modules, it suffices to check that this definition is compatible with multiplication. For $x = \oplus_{i=0}^n x_i, y = \oplus_{j=0}^m y_j$, this gives

$$g(x \cdot y) = g(x_0 \otimes \cdots \otimes x_n \otimes y_0 \otimes \cdots \otimes y_m) = \Pi_{i=0}^n f(x_i) \cdot \Pi_{j=0}^m f(y_j) = g(x_0 \otimes \cdots \otimes x_n) \cdot g(y_0 \otimes \cdots \otimes y_m) = g(x) \cdot g(y)$$

- $g$ is the unique morphism of $R$-algebras $T(M) \to S$, s.t. $g|_{T^1(M)} = f$. This is clear as defining $g|_{T^1(M)}$ gives the map on $g|_{T^n(M)}$ for all $n$, as by the definition of the multiplication. Furthermore, the map restricted to $T^0(M)$ is given by the associated morphism with the $R$-algebra $S$. Both of which are uniquely determined.

$\square$

**Remark 1.5.** This makes $T$ a functor, which maps from $R$-modules to $R$-algebras. For all $R$-linear maps $f : M \to N$, there exists a unique morphism of $R$-algebras $T(f)$ s.t. the following diagram commutes: This makes $T$ a functor as it preserves

$$
\begin{array}{ccc}
M & \xrightarrow{\quad f \quad} & N \\
\downarrow & & \downarrow \\
T(M) & \xrightarrow{\quad T(f) \quad} & T(N)
\end{array}
$$

compositions. Further this is the left adjoint of the forgetful functor $G$ which only regards $S$ as an $R$-module, i.e. we have the following isomorphism

$$\mathrm{Hom}_{R\underline{\mathrm{Alg}}}(T(M), S) \simeq \mathrm{Hom}_R(M, G(S))$$

as by the universal property of tensor algebra the morphism from $T(M)$ to $S$ is uniquely defined by the map $f : M \to S$.

**Definition 1.3** (Graded Ring). *A ring $R$ is a **graded ring** if it comes with a decomposition $R = \oplus_{i \geq 0} R_i$ as abelian groups; and multiplication satisfies the relation $R_p \cdot R_q \subseteq R_{p+q}$ for all $p, q \geq 0$.*

**Remark 1.6.** Consider the subring $R \subseteq R_0$. If $R_0$ lies in the center of $R$, i.e. $R_0 \subseteq \{a \in R \mid ab = ba \forall b \in R\}$, then $R$ becomes an $R_0$-algebra; and the decomposition $R = \oplus_{i \geq 0} R_i$ is a direct sum of $R_0$-modules.

**Example 1.1.** Consider the following examples of graded rings:

1. The tensor algebra $T(M)$ is a graded ring, where $R_0 = T^0(M) = R$

2. The multivariate polynomials $S = R[x_1, \cdots, x_n]$ is a graded ring, where $S_d = \oplus_{\{i_1, \cdots, i_d \mid \sum_k i_k = d\}} R x_1^{i_1} \cdots x_d^{i_d}$.

**Definition 1.4** (Homogeneous). *If $R = \oplus_{i \geq 0} R_i$ is a graded ring, the elements of $R_n$ are **homogeneous of degree** $n$.*

**Definition 1.5** (Morphism of Graded Rings). *If $R$ and $S$ are graded rings, then a morphism of graded rings $f : R \to S$ is a ring homomorphism s.t. $f(R_i) \subseteq S_i$ for all $i$. Such definition gives the result that graded rings form a category.*

**Definition 1.6** (Homogeneous Ideal). *If $R$ is a graded ring, and $I \subseteq R$ an ideal. $I$ is a **homogeneous ideal** if $I = \oplus_{i \geq 0}(I \cap R_i)$. Equivalently, for all $f \in I$, for all $f_i \in R_i$ s.t. $f = \sum_{i=0}^d f_i$, then $f_j \in I$ for all $j$.*

**Remark 1.7.** If further $I$ is two-sided, then $R/I = \oplus_{i\geq 0}(R_i/(R_i \cap I))$ as a direct sum of abelian groups. In this case, $R/I$ is a graded ring, and the quotient $\pi : R \to R/I$ is a morphism of graded rings.

**Proposition 1.3.** *Let $R$ be a graded ring, and $I \subseteq R$ an ideal. Then $I$ is homogeneous if and only if it can be generated by homogeneous elements.*

*Proof.* Show implication in two directions:

$\Rightarrow$: Since $I$ is homogeneous, there exists ideals $I_k \subseteq R_k$ s.t. $I = \oplus_{k\geq 0} I_k$. Then it is generated by the generating sets of $I_k$, which are all homogeneous.

$\Leftarrow$: If $I$ can be generated by homogeneous elements, then for all $x \in R$ there exists a decomposition

$$x = \sum_{r \in R} c_r r = \sum_{i\geq 0} \sum_{r \in R_i} c_{ri} r_i$$

where only finitely many $c_{ri}$s can be non-zero, and $r_i \in I$. Collecting all the terms in the inner summation gives $x = \sum_{i\geq 0} c_i r_i$ for $r_i \in I_i \subseteq R_i$, which satisfies the definition of homogeneous ideals.

$\square$

**Remark 1.8.** It is not necessary (and also not true) that all the homogeneous elements must (can) have the same degree. For example, it is completely valid to have $R \cdot (I \cap R_0) \subsetneq I \cap R_1$, which prevents any homogeneous generating set of the same degree from existing.

# 2   Exterior and Symmetric Algebra

**Definition 2.1** (Symmetric, Alternating Maps). *Let $R$ be a commutative ring, with $M$ and $N$ $R$-modules. Let $\varphi : M^n \to N$ be a multilinear map. It is defined to be:*

i) ***Symmetric***, *if for all $\sigma \in S_n$, and for all $x_1, \cdots, x_n \in M$, $\varphi(x_1, \cdots, x_n) = \varphi(x_{\sigma(1)}, \cdots, \varphi(x_{\sigma(n)}))$.*

ii) ***Alternating***, *if $\varphi(x_1, \cdots, x_n) = 0$ whenever $x_i = x_j$ for $i \neq j$. This is equivalent to stating that $\varphi$ is **skew-symmetric**, where $\varphi(x_1, \cdots, x_i, \cdots, x_j, \cdots, x_n) = -\varphi(x_1, \cdots, x_j, \cdots, x_i, \cdots, x_n)$ for all $i < j$.*

**Remark 2.1.** Since all elements in $S_n$ (symmetric group) are generated by transpositions, to show that a map is symmetric it suffices to show the equality for all transpositions.

**Definition 2.2** (Exterior Algebra). *Let $T$ be the tensor algebra, with $J_n$ an $T^n$-submodule generated by $\{x_1 \otimes \cdots \otimes x_n \mid x_i = x_j \ \forall i \neq j\}$. Define $\Lambda^n(M) := T^n(M)/J_n$; and the **exterior algebra** $\Lambda(M) := \oplus_{n\geq 0} \Lambda^n(M)$. The algebra structure is inherited from that of the tensor algebra.*

**Remark 2.2.** Consider the ideal $J := \oplus_{i\geq 2} J_n \subseteq T(M)$. Here only summands with degree greater than or equal to 2 are taken into consideration as the definition "alternating" makes no sense for the lower degree cases. This is a two-sided ideal as tensor product is $R$-balanced; and each summand is an element in $J_n$ for some $n$. It is further homogeneous, as by definition. In this notation the exterior algebra can also be expressed as $\Lambda(M) := T(M)/J$.

**Remark 2.3.** The equivalence class of $x_1 \otimes \cdots \otimes x_n$ in $\Lambda(M)$ is often denoted as $x_1 \wedge \cdots \wedge x_n$, i.e. the <u>wedge product</u>.

**Proposition 2.1** (Universal Property of $\Lambda^n(M)$). *The map $\varphi : M^n \to \lambda^n(M)$, $(x_1, \cdots, x_n) \mapsto x_1 \wedge \cdots \wedge x_n$ is an alternating multilinear map; and for all alternating multilinear map $\psi : M^n \to P$, there exists a unique $R$-linear map $f : \Lambda^n(M) \to P$ s.t. $f \circ \varphi = \psi$*

*Proof.* This follows directly from the definition of the exterior algebra, and the universal property of tensor product. $\square$

**Example 2.1.** If $g : M \to P$ is an $R$-linear map, then this gives an $R$-linear map $\Lambda^n g : \Lambda^n(M) \to \Lambda^n(P)$ for all $n$, s.t. $x_1 \wedge \cdots \wedge x_n \mapsto g(x_1) \wedge \cdots \wedge g(x_n)$. The construction implies that combining all $\Lambda^n g$s gives a morphism of graded $R$-algebra.

**Proposition 2.2.** *If $(x_1, \cdots, x_d)$ is a system of generators of $M$, then $\Lambda^n(M)$ is generated as an $R$-module by $\{x_{i_1} \wedge \cdots \wedge x_{i_n} \mid 1 \le i_1 < \cdots < i_n \le d\}$. In particular, for all $n > d$, $\Lambda^n(M) = 0$, which implies that $\Lambda(M)$ is a finitely generated $R$-module.*

*Proof.* For the cases where $n \le d$, the result follows from the fact that the tensor of several modules is generated by the tensor of the generators of the corresponding modules; and multilinear maps into the tensor product is alternating. $\square$

**Proposition 2.3.** *$M$ is a free, finitely generated $R$-module, with basis $(x_1, \cdots, x_d)$. Then for all $n \le d$, $\Lambda^n(M)$ is free with basis given by $\{x_{i_1} \wedge \cdots \wedge x_{i_n} \mid 1 \le i_1 < \cdots < i_n \le d\}$. Its rank is $\binom{d}{n}$.*

*Proof.* By the previous proposition, we only need to show that all the elements are linearly independent over $R$. Since the tensor product is commutative, fix the representation for $e_I$ for $I \subseteq \{1, \cdots, d\}$ to be $e_I = x_{i_1} \wedge \cdots \wedge x_{i_n}$ for $i_1 < \cdots < i_n$.

Now consider the $I$s respectively. For a fixed $I$, define $\bar{I} := \{1, \cdots, d\} \setminus I$. By the fact that maps into the exterior algebra is alternating, we have

$$0 = e_{\bar{I}} \wedge \sum_{|J|=n} a_J e_J = \sum_{|J|=n} a_J (e_{\bar{I}} \wedge e_J) = a_I (x_1 \wedge \cdots \wedge x_n) \tag{$*$}$$

We now seek to prove that $a_I = 0$ for all $I$, via apply a transformation into $R$. Consider the map $\psi : M^d \to R$ s.t. $\psi(u_1, \cdots, u_d) = \det(a_{ij})$, where $u_i = \sum_{i=1}^n a_{ij} x_j$, which is multilinear and alternating by construction. Then, by the universal property of exterior algebra, there exists a map $f : \Lambda^d(M) \to R$ s.t. $\psi(u_1, \cdots, u_d) = f(u_1 \wedge \cdots \wedge u_d)$. In particular, $f(x_1 \wedge \cdots \wedge x_d) = 1$. Applying $f$ to $(*)$ gives $a_I = 0$, which gives as a consequence the linear independence. $\square$

**Definition 2.3** (Symmetric Algebra). *Define $I$ as the two-sided ideal generated by elements in the form of $\{x \otimes y - y \otimes x \mid x, y \in M\}$. The **symmetric algebra** $S(M) := T(M)/I$. This is a commutative $R$-algebra.*

**Remark 2.4.** By construction $I$ is generated by homogeneous elements of degree 2, which is a homogeneous ideal (with $I_0 = I_1 = \{0\}$). This gives an alternative expression for the symmetric algebra

$$S(M) = \bigoplus_{n \ge 0} \frac{T^n(M)}{I \cap T^n(M)}$$

which indicates that this is a graded ring. The denominator $I \cap T^n(M)$ is often denoted $S^n(M)$ or $\mathrm{Sym}^n(M)$.

**Proposition 2.4** (Universal Property of $S(M)$). *$S(M)$ is a commutative $R$-algebra; and we have an inclusion $M \hookrightarrow S(M)$ which gives an isomorphism $M \simeq S^1(M)$. If $S$ is a commutative $R$-algebra, and $\beta : M \to S$ is a $R$-linear map, then there exists a unique $R$-algebra homomorphism $f : S(M) \to S$ s.t. $f \circ \alpha = \beta$.*

*Proof.* By the universal property of $T(M)$, there exists a unique $R$-algebra homomorphism $\tilde{\beta} : T(M) \to S$ s.t. $\tilde{\beta} \mid_M = \alpha$. Since $S$ is commutative, $I \subseteq \ker(\tilde{\beta})$. By the universal property of quotient, there exists a unique morphism $\beta : S(M) \to S$ s.t. $\beta \mid_M = \alpha$. $\qquad\square$

**Example 2.2.** Let $M$ be a free $R$-module, with basis $x_1, \cdots, x_n$. The above universal property gives the isomorphisms where $S$ is a commutative $R$-algebra:

$$\{R\text{-algebra homomorphisms } S(M) \to S\} \simeq \{R\text{-linear maps } M \to S\} \simeq \{\text{maps } \{x_1, \cdots, x_n\} \to S\}$$

which implies that $S(M)$ satisfies the universal property of multivariate polynomials, i.e. we have the isomorphism $S(M) \simeq R[x_1, \cdots, x_n]$.

**Example 2.3.** Let $\varphi : M^p \to T^p(M) \to S^p(M)$ be a symmetric multilinear map. For every symmetric multilinear map $\psi : M^p \to N$, there exists a unique $R$-linear map $f : S^p(M) \to N$ s.t. $\psi = f \circ \varphi$. This can be proved similarly using the universal property of quotient rings.

# 3   Symmetric, Alternating and Hermitian Forms

**Definition 3.1** (Hermitian). *Let $E$ be a finite-dimensional $\mathbb{C}$-vector space. $g : E \times E \to \mathbb{C}$ is **Hermitian** if:*

*i)* $g(-, v)$ *is linear for all $v \in E$.*

*ii)* $\overline{g(v, -)}$ *is linear for all $v \in E$.*

*iii)* $g(v, w) = \overline{g(w, v)}$ *for all $v, w \in E$.*

**Remark 3.1.** By the definition, if $v = w$ this gives $g(v, v) = \overline{g(v, v)}$, which implies that $g(v, v) \in \mathbb{R}$ if $g$ is Hermitian.

**Definition 3.2.** *The **kernel** of a bilinear (Hermitian) map $g$ is defined as*

$$\ker(g) := \{v \in E \mid g(v, w) = 0 \; \forall w \in E\}$$

*which is a linear subspace of $E$.*

**Definition 3.3** (Orthogonality). *Let $F \subseteq E$ be a linear subspace. Then its **orthogonal complement** $F^\perp := \{v \in E \mid g(v, w) = 0 \; \forall w \in F\}$. which is a linear subspace of $E$. Two vector subspaces $F_1, F_2 \subseteq E$ are **orthogonal** if $F_1 \subseteq F_2^\perp$, which is equivalently $g(v, w) = 0$ for all $v \in F_1, w \in F_2$.*

**Example 3.1.** $E^\perp$ w.r.t. $g$ is the kernel of $g$.

**Remark 3.2.** For orthogonal vector subspaces $F_1$ and $F_2$, it is often denoted $F_1 \perp F_2$ for $F_1 \subseteq F_2^\perp$; and $E = F_1 \perp F_2$ if $E = F_1 \oplus F_2$ for $F_1 \perp F_2$.

**Definition 3.4.** *A bilinear (Hermitian) form $g$ is **nondegenerate** if $\ker(g) = 0$.*

**Remark 3.3.** If $W \subseteq E$ is a subspace s.t. $E = \ker g \perp W$, then $g|_W$ is nondegenerate.

Suppose that there exists $x \in W$ s.t. $x \in \ker g|_W$. By definition $g(x, y) = 0$ for all $y \in \ker g$. This implies that $x \in \ker g$, which contradicts with the hypothesis that $E$ is the internal direct sum of $\ker g$ and $W$.

**Proposition 3.1.** *Let $F \subseteq E$ be a linear subspace. Suppose that $g$ is nondegenerate. Then*

  *i)* $\dim F^{\perp} + \dim F = \dim E$; *and* $(F^{\perp})^{\perp} = F$.

  *ii) The following statements are equivalent:*

    *(a)* $g|_{F \times F}$ *is nondegenerate.*

    *(b)* $E = F \perp F^{\perp}$.

    *(c)* $g|_{F^{\perp} \times F^{\perp}}$ *is nondegenerate.*

*Proof.* Consider $V^*$ to be the dual vector space of $V$, which is $\text{Hom}_K(V, K)$ for $g$ bilinear, and conjugate linear maps for $G$ Hermitian. Since to specify the linear maps it suffices to specify where the basis is mapped to, $\dim_K(V) = \dim_K(V^*)$. This gives a natural map $\varphi : E \to E^*$ defined as $\varphi(v) := g(v, -)$. $\varphi$ is further an isomorphism, as since $g$ is nondegenerate $\ker g = 0$; and by construction $\ker \varphi = \ker g$.

Define $K := \ker(E^* \to F^*)$. This gives an exact sequence

$$0 \longrightarrow K \longrightarrow E^* \longrightarrow F^* \longrightarrow 0$$

where since $F^{\perp} = \varphi^{-1}(K)$ we have $\dim F^{\perp} = \dim K = \dim E - \dim F$. By definition $F \subseteq (F^{\perp})^{\perp}$; and applying the above argument again gives $\dim F = \dim(F^{\perp})^{\perp}$.

For the second argument, $g|_{F \times F}$ being nondegenerate implies that $F \cap F^{\perp} = \{0\}$, indicating that $E = F \perp F^{\perp}$. Applying symmetric argument gives the other equivalence. $\qquad\square$

**Definition 3.5.** *If $g$ is symmetric of Hermitian, an **orthogonal basis** of $E$ is a basis $(e_1, \cdots, e_n)$ s.t. for all $i \neq j$, $g(e_i, e_j) = 0$.*

**Remark 3.4.** This definition is meaningless for $g$ being alternating, as by definition $g(e_i, e_i) = 0$, which implies that the whole space is the kernel of $g$.

**Proposition 3.2.** *If $g$ is Hermitian; or $g$ is symmetric with $char(K) \neq 2$, then there exists an orthogonal basis.*

*Proof.* Perform induction on the dimension of the vector space. Since the base case is vacuous, it suffices to show the inductive step.

Suppose that the above proposition is true for vector spaces whose dimension is no greater than $d$. Consider a vector space $E$ over $K$ s.t. $\dim E = d + 1$. Let $F \subseteq E$ to be a vector subspace with dimension $d$, and $v \in E \smallsetminus F$ s.t. $g(v, v) \neq 0$. Assume that $g$ is non-degenerate (or otherwise consider $W$ where $E = \ker(g) \oplus W$; and set up the dimensions accordingly). Now consider the cases separately:

- $g$ is symmetric. Since 2 is invertible, for all $v, w \in E$, $g(v, w) = 2^{-1}(g(v + w, v + w) - g(v, v) - g(w, w))$. which can be set to 0 for suitable choice of $v$.

- $g$ is Hermitian. Consider separately the real part and the imaginary part. The above construction gives

$$g(v, w) + g(w, v) = g(v + w, v + w) - g(v, v) - g(w, w)$$

which indicates that $\text{Re}(g(v, w)) = 0$. Further consider $g(iv, w)$, which gives $\text{Im}(g(v, w)) = 0$.

Since $g_{F \times F}$ is nondegenerate (by the choice of $v$), $E = F \perp F^{\perp}$, which indicates that the basis of $F$ and $F^{\perp}$ combined gives an orthogonal basis. $\square$

**Theorem 3.1** (Sylvester)**.** *If $g$ is a symmetric and nondegenerate bilinear form over $K = \mathbb{R}$, then there exists $r \in \mathbb{N}$ s.t. for all orthogonal basis $(e_1, \cdots, e_n)$, from the sets $g(e_k, e_k)$ for $k \in [\![1, n]\!]$ exactly $r$ out of them are positive, and the rest are negative since $g$ is nondegenerate. For the $r$ defined as above $(r, \dim E - r)$ is defined as the **signature** of $g$.*

*Proof.* Let $(e_i)$ and $(e'_i)$ be two orthogonal bases. Define $a_i = g(e_i, e_i)$, and $b_j = g(e'_j, e'_j)$ for all $i$ and $j$. Reorder the basis s.t. $a_i > 0$ if and only if $i \le r$; and $b_j > 0$ if and only if $j \le s$. We would like to show that $r = s$, via showing $r \le s$ through the fact that $(e_1, \cdots, e_r, e'_{s+1}, \cdots, e'_n)$ are linearly independent; and by symmetric argument.

Suppose
$$\sum_{i=1}^{r} \alpha_i e_i - \sum_{j=1}^{n-s} \beta_j e'_{s+j} = 0$$

define $u$ as such and notice the follows:
$$u = \sum_{i=1}^{r} \alpha_i e_i = \sum_{j=1}^{n-s} \beta_j e'_{s+j} \implies g(u, u) = \underbrace{\sum_{i=1}^{r} \alpha_i^2 a_i}_{\ge 0} = \underbrace{\sum_{j=1}^{r} \beta^2 b_{s+j}}_{\le 0}$$

which implies that $\alpha_i = \beta_j = 0$ for all $i \le r, j \le s$. $\square$

**Remark 3.5.** For the Hermitian case, since $g(r, r) \in \mathbb{R}$ for all $r$, the proof is exactly the same via replacing $\alpha_i^2$ by $\alpha_i \cdot \overline{\alpha_i}$ and same for $\beta_j$s.

**Remark 3.6.** If $e_i, \cdots, e_n$ is an orthogonal basis of $g$, one can find $\lambda_i$ s.t. $e'_i = \lambda_i e_i$ satisfying that $g(e'_i, e'_i) = \pm 1$ for all $i$. Such $(e'_i)$ gives an **orthonormal basis**.

**Definition 3.6.** *A symmetric (Hermitian) bilinear form $g$ is **positive definite** if it is nondegenerate with signature $(\dim E, 0)$. Equivalently, $g$ is positive definite if and only if $g(v, v) > 0$ for all $v \in E$ s.t. $v \ne \{0\}$.*

**Example 3.2.** If $H$ is a positive definite Hermitian form, with $(e_1, \cdots, e_n)$ an orthonormal basis, then for $v = \sum_i \alpha_i e_i$ and $w = \sum_j \beta_j e_j$, $H(v, w) = \sum_i \alpha_i \overline{\beta_i}$.

**Proposition 3.3.** *If $g$ is a nondegenerate alternating form, then $E = E_1 \perp \cdots \perp E_m$; and each $E_i$ has a basis $(u_i, v_i)$ s.t. $g(u_i, v_i) = 1$ for all $i$.*

*Proof.* For $v, u \in E$ which are linearly independent, $g$ being alternating implies $g(v, u) = 0$. Further if there exists $u_1, u_2 \in E$ s.t. $g(v, u_1) \ne 0, g(v, u_2) \ne 0$, this gives that there exists $c \in K$ s.t. $g(v, u_1 - cu_2) = 0$ which implies that $(v, u_1, u_2)$ is not linearly independent. $\square$

# 4   The Spectral Theorem

**Remark 4.1.** Fix the notation. For $g$ being symmetric bilinear or Hermitian, $g(-, -)$ is often denoted as $\langle -, - \rangle$.

In this part two kinds of maps are considered:

1. $E$ is a finite dimensional vector space over $K = \mathbb{R}$, with a positive-definite symmetric bilinear form.

2. $E$ is a finite dimensional vector space over $K = \mathbb{C}$, with a positive-definite Hermitian form.

**Remark 4.2.** Any symmetric or Hermitian bilinear form can be described using a matrix w.r.t. any basis $(e_1, \cdots, e_n)$ of $V$: $A = (g(e_i, e_j))$, $1 \le i, j \le n$. $g$ being symmetric gives ${}^tA = A$; and $g$ being Hermitian gives ${}^tA = \bar{A}$. If the basis is orthonormal, then it has $\pm 1$ on the diagonal.

**Definition 4.1** (Adjoint of linear operators). *Fix $y \in E$. Consider $E \ni x \mapsto \langle y, u(x) \rangle \in K$. This map is in $E^*$, which is* $\mathrm{Hom}_{\mathbb{R}}(E, \mathbb{R})$ *in the first case, and conjugate linear maps to $\mathbb{C}$ in the second. Suppose further that $\langle -, - \rangle$ is nondegenerate, there exists a unique map $u^*$ s.t. $\langle u^*(y), x \rangle = \langle y, u(x) \rangle$.*

**Remark 4.3.** $u^*$ is also linear. Consider the following identification:

$$u^*(\lambda y) = \langle \lambda y, u(x) \rangle = \lambda \langle y, u(x) \rangle = \lambda \langle u^*(y), x \rangle = \lambda u^*(y)$$

**Definition 4.2.** *For $u$ symmetric or Hermitian, it is **self-adjoint** if $u = u^*$, i.e. for all $x, y \in E$, $\langle u(x), y \rangle = \langle x, u(y) \rangle$.*

**Proposition 4.1.** *A linear map is self-adjoint if and only if its matrix representation is symmetric or Hermitian.*

*Proof.* Let $(e_1, \cdots, e_n)$ be an orthonormal basis of $E$. By linearity it suffices to consider the evaluation of the map on the basis. Let the matrix representations of $u$ and $u^*$ be $(a_{ij})$ and $(b_{ij})$, respectively. Then the evaluation can be expressed via

$$u(e_k) = \sum_{i=1}^n a_{ik} e_i, \quad u^*(e_\ell) = \sum_{j=1}^n b_{j\ell} e_j$$

Now consider the criterion for judging whether an operator is self-adjoint: for $\langle e_\ell, u(e_k) \rangle = \langle u^*(e_\ell), e_k \rangle$ we have

$$\text{LHS} = \left\langle e_\ell, \sum_{i=1}^n a_{ik} e_k \right\rangle = \overline{a_{\ell k}}, \qquad \text{RHS} = \left\langle \sum_{j=1}^n b_{j\ell} e_j, e_k \right\rangle = b_{kl}$$

which implies that $B = {}^t\bar{A}$. $\qquad\square$

**Theorem 4.1** (Spectral Theorem). *Let $u : E \to E$ be a self-adjoint linear map. Then:*

i) *All eigenvalues of $u$ are in $\mathbb{R}$.*

ii) *If $u_1, u_2$ are eigenvectors corresponding to distinct eigenvalues $\lambda_1, \lambda_2$, then $\langle u_1, u_2 \rangle = 0$.*

iii) *There exists an orthonormal basis of $E$ consisting of eigenvectors of $u$, which implies that $u$ is diagonalizable.*

**Remark 4.4.** In particular, this implies that symmetric linear maps on $\mathbb{R}$, and Hermitian linear maps on $\mathbb{C}$, are diagonalizable.

*Proof of Theorem 4.1.* Consider the statements sequentially:

i) Let $\lambda \in \mathbb{C}$ be a root of the characteristic polynomial of $u$:

- First consider $E$ as a vector space over $\mathbb{C}$. Let $y$ be an eigenvector associated to $\lambda$. This gives

$$\lambda \langle y, y \rangle = \langle \lambda y, y \rangle = \langle u(y), y \rangle = \langle y, u(y) \rangle = \langle y, \lambda y \rangle = \bar{\lambda} \langle y, y \rangle$$

Since $y \neq 0$, $\langle y, y \rangle \neq 0$, which implies that $\lambda = \overline{\lambda}$.

- Now let $E$ be a vector space over $\mathbb{R}$. Seek to reduce to the case above. Consider the extension of scalars $E_{\mathbb{C}} := E \otimes_{\mathbb{R}} \mathbb{C}$, with the inclusion $E \hookrightarrow E_{\mathbb{C}}$, $x \mapsto x \otimes 1$. This gives an extension of $\langle -, - \rangle$ to $E_{\mathbb{C}}$, which is also Hermitian. This implies that orthogonality in $E_{\mathbb{C}}$ is well-defined, and any orthonormal basis in $E$ is also orthonormal in $E_{\mathbb{C}}$.

  Now consider the corresponding map after the extension of scalar $i_{\mathbb{C}} := u \otimes \mathrm{Id}_{\mathbb{C}}$, which is given by the same matrix representation as that in $E$. By the result in the first case, this is self-adjoint; and since the characteristic polynomial of these two maps are the same, all roots are real numbers.

ii) Let $\lambda_1, \lambda_2$ be the eigenvalues to which the eigenvectors $v_1, v_2$ are associated to. Since $u$ is self-adjoint, this gives

$$\lambda_1 \langle v_1, v_2 \rangle = \langle u(v_1), v_2 \rangle = \langle v_1, u(v_2) \rangle = \overline{\lambda_2} \langle v_1, v_2 \rangle = \lambda_2 \langle v_1, v_2 \rangle$$

By hypothesis $\lambda_1 \neq \lambda_2$, and by definition they are non-zero. This gives $\langle v_1, v_2 \rangle = 0$.

iii) Show this via induction on the dimension of $E$:

- *Base case.* Then for any element in the vector space it is linear in some basis vector that is normalized.

- *Inductive step.* Let $v \in E$ be an eigenvector of $u$ associated to the eigenvalue $\lambda$. Define $F = (K \cdot v)^{\perp}$ where $K$ is the underlying field. Since we are working over a vector space, this gives $E = F \oplus F^{\perp}$, indicating that $\langle -, - \rangle|_{F^{\perp} \times F^{\perp}}$ is positive definite (since the basis is orthonormalized). Let $y \in F^{\perp}$. By construction $\langle u(y), v \rangle = \langle \sum_i \alpha_i \lambda_i y_i, v \rangle = 0$ where $y_i$ is associated to the eigenvalue $\lambda_i$. Further $u|_{F^{\perp} \times F^{\perp}}$ is self-adjoint, as taking a perspective on its matrix representation a matrix with one column and one row deleted remains to be Hermitian if itself is. Then $v$ together with an orthonormal basis of $F^{\perp}$ gives an orthonormal basis of $E$.

$\square$

**Definition 4.3** (Unitary Operator). *A linear map $u : E \to E$ is **unitary** if it is a bijection; and $u^* = u^{-1}$.*

**Remark 4.5.** Equivalently, this gives $\langle u^{-1}(x), y \rangle = \langle x, u(y) \rangle$, which implies that

$$\langle x, y \rangle = \langle u(x), u(y) \rangle \implies \langle x, x \rangle = \langle u(x), u(x) \rangle \implies x = 0 \text{ if and only if } u(x) = 0$$

**Remark 4.6.** Translating this into the matrix representation $A$ of $u$, $u$ is unitary if and only if $A \cdot {}^t\overline{A} = \mathrm{Id}$. This implies that a matrix is unitary if and only if its inverse is (assuming that it is invertible).

**Proposition 4.2.** *If $B$ is an orthonormal basis, and $A$ is a change of basis matrix between $B$ and $B'$. Then $B'$ is orthonormal if and only if $A$ is unitary.*

*Proof.* By the Remark 4.6 and symmetric argument it suffices to prove that for one direction. Remark 4.5 $\langle u(v_i), u(v_j) \rangle = \langle v_i, v_j \rangle$ which is $\pm 1$ for $i = j$; and $0$ otherwise. This is exactly the definition of an orthonormal basis. $\square$

**Remark 4.7.** By Spectral Theorem, there exists an orthonormal basis consists of eigenvectors. Further the matrix $U$ with columns elements of the orthonormal basis is unitary. Then for $A$ Hermitian, $U \cdot A \cdot U^{-1}$ is diagonal.