Multilinear Algebra[1]

Tin-Yau Tam
Department of Mathematics and Statistics
221 Parker Hall
Auburn University
AL 36849, USA
tamtiny@auburn.edu

November 30, 2011

[1]Some portions are from B.Y. Wang's Foundation of Multilinear Algebra (1985 in Chinese)

Chapter 1

Review of Linear Algebra

1.1 Linear extension

In this course, U, V, W are finite dimensional vector spaces over $\mathbb{C}$, unless specified. All bases are **ordered** bases.

Denote by $\operatorname{Hom}(V, W)$ the set of all linear maps from V to W and $\operatorname{End} V := \operatorname{Hom}(V, V)$ the set of all linear operators on V. Notice that $\operatorname{Hom}(V, W)$ is a vector space under usual addition and scalar multiplication. For $T \in \operatorname{Hom}(V, W)$, the image and kernel are

$$\operatorname{Im} T = \{Tv : v \in V\} \subset W, \quad \operatorname{Ker} T = \{v : Tv = 0\} \subset V$$

which are subspaces. It is known that T is injective if and only if $\operatorname{Ker} T = 0$. The rank of T is $\operatorname{rank} T := \dim \operatorname{Im} T$.

Denote by $\mathbb{C}_{m\times n}$ the space of $m \times n$ complex matrices. Each $A \in \mathbb{C}_{m\times n}$ can be viewed as in $\operatorname{Hom}(\mathbb{C}^n, \mathbb{C}^m)$ in the obvious way. So we have the concepts, like rank, inverse, image, kernel, etc for matrices.

Theorem 1.1.1. Let $E = \{e_1, \ldots, e_n\}$ be a basis of V and let $w_1, \ldots, w_n \in W$. Then there exists a unique $T \in \operatorname{Hom}(V, W)$ such that $T(e_i) = w_i$, $i = 1, \ldots, n$.

Proof. For each $v = \sum_{i=1}^n a_i e_i$, define $Tv = \sum_{i=1}^n a_i T e_i = \sum_{i=1}^n a_i w_i$. Such T is clearly linear. If $S, T \in \operatorname{Hom}(V, W)$ such that $T(e_i) = w_i$, $i = 1, \ldots, n$, then $Sv = \sum_{i=1}^n a_i S e_i = \sum_{i=1}^n a_i T e_i = Tv$ for all $v = \sum_{i=1}^n a_i e_i \in V$ so that $S = T$. □

In other words, a linear map is completely determined by the images of basis elements in V.

A bijective $T \in \operatorname{Hom}(V, W)$ is said to be **invertible** and its inverse ($T^{-1} \circ T = I_V$ and $T \circ T^{-1} = I_W$) is linear, i.e., $T^{-1} \in \operatorname{Hom}(W, V)$: $T^{-1}(\alpha_1 w_1 + \alpha_2 w_2) = v$ means that

$$Tv = \alpha_1 w_1 + \alpha_2 w_2 = T(\alpha_1 T^{-1} w_1 + \alpha_2 T^{-1} w_2),$$

i.e., $v = T^{-1}w_1 + \alpha_2 T^{-1}w_2$ for all $w_1, w_2 \in W$ and $v \in V$. We will simply write ST for $S \circ T$ for $S \in \mathrm{Hom}\,(W, U)$, $T \in \mathrm{Hom}\,(V, W)$. Two vector spaces V and W are said to be **isomorphic** if there is an invertible $T \in \mathrm{Hom}\,(V, W)$.

Theorem 1.1.2. Let $T \in \mathrm{Hom}\,(V, W)$ and $\dim V = n$.

1. Then $\mathrm{rank}\, T = k$ if and only if there is a basis $\{v_1, \ldots, v_k, v_{k+1}, \ldots, v_n\}$ for V such that $Tv_1, \ldots, Tv_k$ are linearly independent and $Tv_{k+1} = \cdots = Tv_n = 0$.

2. $\dim V = \dim \mathrm{Im}\, T + \dim \mathrm{Ker}\, T$.

Proof. Since $\mathrm{rank}\, T = k$, there is a basis $\{Tv_1, \ldots, Tv_k\}$ for $\mathrm{Im}\, T$. Let $\{v_{k+1}, \ldots, v_{k+l}\}$ be a basis of $\mathrm{Ker}\, T$. Set $E = \{v_1, \ldots, v_k, v_{k+1}, \ldots, v_{k+l}\}$. For each $v \in V$, $Tv = \sum_{i=1}^k a_i Tv_i$ since $Tv \in \mathrm{Im}\, T$. So $T(v - \sum_{i=1}^k a_i v_i) = 0$, i.e., $v - \sum_{i=1}^k a_i v_i \in \mathrm{Ker}\, T$. Thus $v - \sum_{i=1}^k a_i v_i = \sum_{i=k+1}^{k+l} a_i v_i$, i.e., $v = \sum_{i=1}^{k+l} a_i v_i$ so that E spans V. So it suffices to show that E is linearly independent. Suppose $\sum_{i=1}^{k+l} a_i v_i = 0$. Applying T on both sides, $\sum_{i=1}^k a_i Tv_i = 0$ so that $a_1 = \cdots = a_k = 0$. Hence $\sum_{i=k+1}^{k+l} a_i v_i = 0$ and we have $a_{k+1} = \cdots = a_{k+l} = 0$ since $v_{k+1}, \ldots, v_{k+l}$ are linear independent. Thus E is linearly independent and hence a basis of V; so $k + l = n$. □

Theorem 1.1.3. Let $A \in \mathbb{C}_{m\times n}$.

1. $\mathrm{rank}\, A^*A = \mathrm{rank}\, A$.

2. For each $A \in \mathbb{C}_{m\times n}$, $\mathrm{rank}\, A = \mathrm{rank}\, A^* = \mathrm{rank}\, A^T$, i.e., column rank and row rank of A are the same.

Proof. (1) Notice $Ax = 0$ if and only if $A^*Ax = 0$. It is because $A^*Ax = 0$ implies that $(Ax)^*(Ax) = x^*A^*Ax = 0$. So $Ax_1, \ldots, Ax_k$ are linearly independent if and only if $A^*Ax_1, \ldots, A^*Ax_k$ are linearly independent. Hence $\mathrm{rank}\, A^*A = \mathrm{rank}\, A$.
(2) Since $\mathrm{rank}\, A = \mathrm{rank}\, A^*A \le \mathrm{rank}\, A^*$ from Problem 3 and thus $\mathrm{rank}\, A^* \le \mathrm{rank}\, A$ since $(A^*)^* = A$. Hence $\mathrm{rank}\, A = \mathrm{rank}\, A^T$ (why?). □

Problems

1. Show that $\dim \mathrm{Hom}\,(V, W) = \dim V \dim W$.

2. Let $T \in \mathrm{Hom}\,(V, W)$. Prove that $\mathrm{rank}\, T \le \min\{\dim V, \dim W\}$.

3. Show that if $T \in \mathrm{Hom}\,(V, U)$, $S \in \mathrm{Hom}\,(U, W)$, then $\mathrm{rank}\, ST \le \min\{\mathrm{rank}\, S, \mathrm{rank}\, T\}$.

4. Show that the inverse of $T \in \mathrm{Hom}\,(V, W)$ is unique, if exists. Moreover T is invertible if and only if $\mathrm{rank}\, T = \dim V = \dim W$.

5. Show that V and W are isomorphic if and only $\dim V = \dim W$.

6. Show that if $T \in \mathrm{Hom}(V, U)$, $S \in \mathrm{Hom}(U, W)$ are invertible, then ST is also invertible. In this case $(ST)^{-1} = T^{-1}S^{-1}$.

7. Show that $A \in \mathbb{C}_{m\times n}$ is invertible only if $m = n$, i.e., A has to be square. Show that $A \in \mathbb{C}_{n\times n}$ is invertible if and only if the columns of A are linearly independent.

8. Prove that if $A, B \in \mathbb{C}_{n\times n}$ such that $AB = I_n$, then $BA = I_n$.

Solutions to Problems 1.1

1. Let $\{e_1, \ldots, e_n\}$ and $\{f_1, \ldots, f_m\}$ be bases of V and W respectively. Then $\xi_{ij} \in \mathrm{Hom}(V, W)$ defined by $\xi_{ij}(e_k) = \delta_{ik} f_j$, $i, j, k = 1, \ldots, n$, form a basis of $\mathrm{Hom}(V, W)$. Thus $\dim \mathrm{Hom}(V, W) = \dim V \dim W$.

2. From definition $\mathrm{rank}\, T \leq \dim W$. From Theorem 1.1.2 $\mathrm{rank}\, T \leq \dim V$.

3. Since $\mathrm{Im}\, ST \subset \mathrm{Im}\, S$, $\mathrm{rank}\, ST \leq \mathrm{rank}\, S$. By Theorem 1.1.2 $\mathrm{rank}\, ST = \dim V - \dim \mathrm{Ker}\, ST$. But $\mathrm{Ker}\, T \subset \mathrm{Ker}\, ST$ so that $\mathrm{rank}\, ST \leq \dim V - \dim \mathrm{Ker}\, T = \mathrm{rank}\, T$.

4. Suppose that S and $S' \in \mathrm{Hom}(V, W)$ are inverses of $T \in \mathrm{Hom}(V, W)$. Then $S = S(TS') = (ST)S' = S'$; inverse is unique. $T \in \mathrm{Hom}(V, W)$ invertible $\Leftrightarrow$ T is bijective (check!). Now injective $T \Leftrightarrow \mathrm{Ker}\, T = 0$, and surjective $T \Leftrightarrow \mathrm{rank}\, T = \dim W$

5. One implication follows from Problem 4. If $\dim V = \dim W$, let $E = \{e_1, \ldots, e_n\}$ and $\{f_1, \ldots, f_n\}$ be bases of V and W, define $T \in \mathrm{Hom}(V, W)$ by $Te_i = f_i$ for all i which is clearly invertible since $T^{-1} f_i = e_i$.

6. $(T^{-1}S^{-1})(ST) = I_V$ and $(ST)(T^{-1}S^{-1}) = I_W$.

7. From Problem 4, a matrix A is invertible only if A is square. The second statement follows from the fact the $\mathrm{rank}\, A$ is the dimension of the column space of A.

8. If $AB = I$, then $\mathrm{rank}\, AB = n$ so $\mathrm{rank}\, A = n (= \mathrm{rank}\, B)$ by Problem 3. So $\mathrm{Ker}\, A = 0$ and $\mathrm{Ker}\, B = 0$ by Theorem 1.1.2. Thus A and B is invertible. Then $I = A^{-1}ABB^{-1} = A^{-1}B^{-1} - (BA)^{-1}$ so that $BA = I$. (or, without using Theorem 1.1.2, note that $B = B(AB) = (BA)B$ so that $(I - BA)B = 0$. Since $\mathrm{Im}\, B = \mathbb{C}^n$, $I - BA = 0$, i.e., $BA = I$).

1.2 Matrix representations of linear maps

In this section, U, V, W are finite dimensional vector spaces.

Matrices $A \in \mathbb{C}_{m\times n}$ can be viewed as elements in $\mathrm{Hom}(\mathbb{C}^n, \mathbb{C}^m)$. On the other hand, each $T \in \mathrm{Hom}(V, W)$ can be realized as a matrix once we fix bases for V and W.

Let $T \in \mathrm{Hom}\,(V, W)$ with bases $E = \{e_1, \ldots, e_n\}$ for V and $F = \{f_1, \ldots, f_m\}$ for W. Since $Te_j \in W$, we have

$$Te_j = \sum_{i=1}^{m} a_{ij} f_i, \quad j = 1, \ldots, n.$$

The matrix $A = (a_{ij}) \in \mathbb{C}_{m\times n}$ is called the **matrix representation** of T with respect to the bases E and F, denoted by $[T]_E^F = A$.

From Theorem 1.1.1 $[S]_E^F = [T]_E^F$ if and only if $S = T$, where $S, T \in \mathrm{Hom}\,(V, W)$.

The **coordinate vector** of $v \in V$ with respect to the basis $E = \{e_1, \ldots, e_n\}$ for V is denoted by $[v]_E := (a_1, \ldots, a_n)^T \in \mathbb{C}^n$ where $v = \sum_{i=1}^n a_i v_i$. Indeed we can view $v \in \mathrm{Hom}\,(\mathbb{C}, V)$ with 1 as a basis of $\mathbb{C}$. Then $v \cdot 1 = \sum_{i=1}^n a_i e_i$ so that $[v]_E$ is indeed $[v]_1^E$.

Theorem 1.2.1. Let $T \in \mathrm{Hom}\,(V, W)$ and $S \in \mathrm{Hom}\,(W, U)$ with bases $E = \{e_1, \ldots, e_n\}$ for V, $F = \{f_1, \ldots, f_m\}$ for W and $G = \{g_1, \ldots, g_l\}$ for U. Then

$$[ST]_E^G = [S]_F^G [T]_E^F$$

and in particular

$$[Tv]_F = [T]_E^F [v]_E, \quad v \in V.$$

Proof. Let $A = [T]_E^F$ and $B = [S]_F^G$, i.e., $Te_j = \sum_{i=1}^m a_{ij} f_i$, $j = 1, \ldots, n$. and $Tf_i = \sum_{k=1}^l b_{ki} g_k$, $i = 1, \ldots, m$. So

$$STe_j = \sum_{i=1}^{m} a_{ij} Sf_i = \sum_{i=1}^{m} a_{ij} (\sum_{k=1}^{l} b_{ki}) g_k = \sum_{k=1}^{l} (BA)_{kj} g_k.$$

So $[ST]_E^G = BA = [S]_F^G [T]_E^F$. □

Consider $I := I_V \in \mathrm{End}\, V$. The matrix $[I]_E^{E'}$ is called the **transitive matrix** from E to E' since

$$[v]_{E'} = [Iv]_{E'} = [I]_E^{E'} [v]_E, \quad v \in V,$$

i.e., $[I]_E^{E'}$ transforms the coordinate vector $[v]_E$ with respect to E to $[v]_{E'}$ with respect to E'. From Theorem 1.2.1, $[I]_{E'}^E = ([I]_E^{E'})^{-1}$.

Two operators $S, T \in \mathrm{End}\, V$ are said to be **similar** if there is an invertible $P \in \mathrm{End}\, V$ such that $S = P^{-1}AP$. Similarity is an equivalence relation and is denoted by $\sim$.

Theorem 1.2.2. Let $T \in \mathrm{End}\, V$ with $\dim V = n$ and let E and E' be bases of V. Then $A := [T]_{E'}^{E'}$ and $B := [T]_E^E$ are similar, i.e., there is an invertible $P \in \mathbb{C}_{n\times n}$ such that $PAP^{-1} = B$. Conversely, similar matrices are matrix representations of the same operator with respect to different bases.

Proof. Let $I = I_V$. By Theorem 1.2.1, $[I]^E_{E'}[I]^{E'}_E = [I]^E_E = I_n$ where I_n is $n \times n$ identity matrix. Denote by $P := [I]^E_{E'}$ (the transitive matrix from E' to E) so that $P^{-1} = [I]^{E'}_E$. Thus

$$[T]^{E'}_{E'} = [ITI]^{E'}_{E'} = [I]^{E'}_E [T]^E_E [I]^E_{E'} = P^{-1}[T]^E_E P.$$

Suppose that A and B are similar, i.e., $B = R^{-1}BR$. Let E be a basis of V. By Theorem 1.1.1 A uniquely determines $T \in \operatorname{End} V$ such that $[T]^E_E = A$. By Theorem 1.2.1, an invertible $R \in \mathbb{C}_{n\times n}$ uniquely determines a basis E' of V such that $[T]^E_{E'} = R$ so that

$$[T]^{E'}_{E'} = [I]^{E'}_E [T]^E_E [I]^E_{E'} = R^{-1}[T]^E_E R = B.$$

□

So functions $\varphi : \mathbb{C}_{n\times n} \to \mathbb{C}$ that take constant values on similarity orbits of $\mathbb{C}_{n\times n}$, i.e., $\varphi(A) = P^{-1}AP$ for all invertible $P \in \mathbb{C}_{n\times n}$, are defined for operators, for examples, determinant and trace of a operator.

Problems

1. Let $E = \{e_1, \dots, e_n\}$ and $F = \{f_1, \dots, f_m\}$ be bases for V and W. Show that the matrix representation $\varphi := [\,\cdot\,]^F_E : \operatorname{Hom}(V, W) \to \mathbb{C}_{m\times n}$ is an isomorphism.

2. Let E and F be bases of V and W respectively. Show that if $T \in \operatorname{Hom}(V, W)$ is invertible, then $[T^{-1}]^E_F = ([T]^F_E)^{-1}$.

3. Let E and F be bases of V and W respectively. Let $T \in \operatorname{Hom}(V, W)$ and $A = [T]^F_E$. Show that $\operatorname{rank} A = \operatorname{rank} T$.

Solutions to Problems 1.2

1. It is straightforward to show that φ is linear by comparing the (ij) entry of $[\alpha S+\beta T]^F_E$ and $\alpha[S]^F_E+\beta[T]^F_E$. Since $\dim \operatorname{Hom}(V, W) = mn = \dim \mathbb{C}_{m\times n}$, it suffices to show that φ is injective. From Theorem 1.1.1, φ is injective.

2. From Theorem 1.2.1, $[T^{-1}]^E_F[T]^F_E = [T^{-1}T]^E_E = [I_V]^E_E = I_n$ where $\dim V = n$. Thus use Problem 1.8 to have $[T^{-1}]^E_F = ([T]^F_E)^{-1}$, or show that $[T]^F_E[T^{-1}]^E_F = I_n$ similarly.

3. (Roy) Let $\operatorname{rank} T = k$ and let $E = \{v_1, \dots, v_n\}$ and $F = \{w_1, \dots, w_m\}$ be bases for V and W. When we view $v \in V$ as an element in $\operatorname{Hom}(\mathbb{C}, V)$, by Problem 1, the coordinate maps $[\,\cdot\,]_E : V \to \mathbb{C}^n$ and $[\,\cdot\,]_F : W \to \mathbb{C}^m$ are isomorphisms. So

$$\begin{aligned} \operatorname{rank} T &= \dim \operatorname{Im} T = \dim\langle Tv_1, \dots, Tv_n\rangle = \dim\langle [Tv_1]_F, \dots, [Tv_n]_F\rangle \\ &= \dim\langle [T]^F_E[v_1]_E, \dots, [T]^F_E[v_n]_E\rangle = \dim \operatorname{Im} A = \operatorname{rank} A. \end{aligned}$$

1.3 Inner product spaces

Let V be a vector space. An **inner product** on V is a function $(\cdot,\cdot) : V \times V \to \mathbb{C}$ such that

1. $(u, v) = \overline{(v, u)}$ for all $u, v \in V$.
2. $(\alpha_1 v_1 + \alpha_2 v_2, u) = \alpha_1(v_1, u) + \alpha_2(v_2, u)$ for all $v_1, v_2, u \in V$, $\alpha_1, \alpha_2 \in \mathbb{C}$.
3. $(v, v) \geq 0$ for all $v \in V$ and $(v, v) = 0$ if and only if $v = 0$.

The space V is then called an inner product space. The **norm** induced by the inner product is defined as

$$\|v\| = \sqrt{(v, v)}, \quad v \in V.$$

Vectors v satisfying $\|v\| = 1$ are called **unit vectors**. Two vectors $u, v \in V$ are said to be **orthogonal** if $(u, v) = 0$, denoted by $u \perp v$. A basis $E = \{e_1, \ldots, e_n\}$ is called an **orthogonal basis** if the vectors are orthogonal. It is said to be **orthonormal** if

$$(e_i, e_j) = \delta_{ij}, \quad i, j = 1, \ldots, n.$$

where

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

is the Kronecker delta notation.

Theorem 1.3.1. (Cauchy-Schwarz inequality) Let V be an inner product space. Then

$$|(u, v)| \leq \|u\|\|v\|, \quad u, v \in V.$$

Equality holds if and only if u and v are linearly dependent, i.e., one is a scalar multiple of the other.

Proof. It is trivial when $v = 0$. Suppose $v \neq 0$. Let $w = u - \frac{(u,v)}{\|v\|^2} v$. Clearly $(w, v) = 0$ so that

$$0 \leq (w, w) = (w, u) = (u, u) - \frac{(u, v)}{\|v\|^2}(v, u) = \|u\|^2 - \frac{|(u, v)|^2}{\|v\|^2}.$$

Equality holds if and only if $w = 0$, i.e. u and v are linearly dependent. □

Theorem 1.3.2. (Triangle inequality) Let V be an inner product space. Then

$$\|u + v\| \leq \|u\| + \|v\|, \quad u, v \in V.$$

Proof.

$$\|u + v\|^2 = (u + v, u + v) = \|u\|^2 + 2\mathrm{Re}\,(u, v) + \|v\|^2 \leq \|u\|^2 + 2|(u, v)| + \|v\|^2$$

By Theorem 1.3.1, we have

$$\|u + v\|^2 \leq \|u\|^2 + 2\|u\|\|v\| + \|v\|^2 = (\|u\| + \|v\|)^2.$$

□

Theorem 1.3.3. Let $E = \{e_1, \ldots, e_n\}$ be an orthonormal basis of V. For any $u, v \in V$,

$$\begin{aligned} u &= \sum_{i=1}^{n} (u, e_i) e_i, \\ (u, v) &= \sum_{i=1}^{n} (u, e_i)(e_i, v). \end{aligned}$$

Proof. Let $u = \sum_{j=1}^{n} a_j e_j$. Then $(u, e_i) = (\sum_{j=1}^{n} a_j e_j, e_i) = a_i$, $i = 1, \ldots, n$. Now $(u, v) = (\sum_{i=1}^{n} (u, e_i) e_i, v) = \sum_{i=1}^{n} (u, e_i)(e_i, v)$. □

Denote by $\langle v_1, \ldots, v_k \rangle$ the **span** of the vectors $v_1, \ldots, v_k$.

Theorem 1.3.4. (Gram-Schmidt orthogonalization) Let V be an inner product space with basis $\{v_1, \ldots, v_n\}$. Then there is an orthonormal basis $\{e_1, \ldots, e_n\}$ such that

$$\langle v_1, \ldots, v_k \rangle = \langle e_1, \ldots, e_k \rangle, \quad k = 1, \ldots, n.$$

Proof. Let

$$\begin{aligned} e_1 &= \frac{v_1}{\|v_1\|}, \\ e_2 &= \frac{v_2 - (v_2, e_1)e_1}{\|v_2 - (v_2, e_1)e_1\|} \\ &\cdots \\ e_n &= \frac{v_n - (v_n, e_1)e_1 - \cdots - (v_n, e_{n-1})e_{n-1}}{\|v_n - (v_n, e_1)e_1 - \cdots - (v_n, e_{n-1})e_{n-1}\|} \end{aligned}$$

It is direct computation to check that $\{e_1, \ldots, e_n\}$ is the desired orthonormal basis. □

The inner product on $\mathbb{C}^n$: for any $x = (x_1, \ldots, x_n)^T$ and $y = (y_1, \ldots, y_n)^T \in \mathbb{C}^n$,

$$(x, y) := \sum_{i=1}^{n} x_i \overline{y_i}$$

is called the standard inner product on $\mathbb{C}^n$. The induced norm is called the **2-norm** and is denoted by

$$\|x\|_2 := (x, x)^{1/2}.$$

Problems

1. Let $E = \{e_1, \ldots, e_n\}$ be a basis of V and $v \in V$. Prove that $v = 0$ if and only if $(v, e_i) = 0$ for all $i = 1, \ldots, n$.

2. Show that each orthogonal set of nonzero vectors is linearly independent.

3. Let $E = \{e_1, \dots, e_n\}$ be an orthonormal basis of V. If $u = \sum_{i=1}^n a_i e_i$ and $n = \sum_{i=1}^n b_i e_i$, then $(u, v) = \sum_{i=1}^n a_i \overline{b_i}$.

4. Show that an inner product is completely determined by an orthonormal basis, i.e., if the inner products $(\cdot, \cdot)$ and $\langle \cdot, \cdot \rangle$ have the same orthonormal basis, then $(\cdot, \cdot)$ and $\langle \cdot, \cdot \rangle$ are the same.

5. Show that $(A, B) = \operatorname{tr}(B^* A)$ defines an inner product on $\mathbb{C}_{m\times n}$, where B^* denotes the complex conjugate transpose of B.

6. Prove $|\operatorname{tr}(A^* B)| \le \operatorname{tr}(A^* A) \operatorname{tr}(B^* B)$ for all $A, B \in \mathbb{C}_{m\times n}$.

Solutions to Problems 1.3

1. $v = 0 \Leftrightarrow (v, u) = 0$ for all $u \in V \Rightarrow (v, e_i) = 0$ for all i. Conversely each $u \in V$ is of the form $\sum_{i=1}^n a_i e_i$ so that $(v, e_i) = 0$ for all i implies $(v, u) = 0$ for all $u \in V$.

2. Suppose that $S = \{v_1, \dots, v_n\}$ is an orthogonal set of nonzero vectors. With $\sum_{i=1}^n a_i v_i = 0$, $a_j (v_j, v_j) = (\sum_{i=1}^n a_i v_i, v_j) = 0$ so that $a_j = 0$ for all j since $(v_j, v_j) \ne 0$. Thus S is linearly independent.

3. Follows from Theorem 1.3.3.

4. Follows from Problem 3.

5. Straightforward computation.

6. Apply Cauchy-Schwarz inequality on the inner product defined in Problem 5.

1.4 Adjoints

Let V, W be inner product spaces. For each $T \in \operatorname{Hom}(V, W)$, the **adjoint** of T is $S \in \operatorname{Hom}(W, V)$ such that $(Tv, w)_W = (v, Sw)_V$ for all $v \in V, w \in W$ and is denoted by T^*. Clearly $(T^*)^* = T$.

Theorem 1.4.1. Let W, V be inner product spaces. Each $T \in \operatorname{Hom}(V, W)$ has a unique adjoint.

Proof. By Theorem 1.3.4, let $E = \{e_1, \dots, e_n\}$ be an orthonormal basis of V. For any $w \in W$, define $S \in \operatorname{Hom}(W, V)$ by

$$Sw := \sum_{i=1}^n (w, Te_i)_W e_i.$$

For any $v \in V$, by Theorem 1.3.3, $v = \sum_{i=1}^n (v, e_i)_V e_i$ so that

$$\begin{aligned}(v, Sw)_V &= (v, \sum_{i=1}^n (w, Te_i)_W e_i)_V = \sum_{i=1}^n (Te_i, w)_W (v, e_i)_V \\ &= (T \sum_{i=1}^n (v, e_i)_V e_i, w)_W = (Tv, w)_W.\end{aligned}$$

Uniqueness follows from Problem 1. □

Theorem 1.4.2. Let $E = \{e_1, \dots, e_n\}$ and $F = \{f_1, \dots, f_m\}$ be orthonormal bases of the inner product spaces V and W respectively. Let $T \in \operatorname{Hom}(V, W)$. Then $[T^*]_F^E = ([T]_E^F)^*$, where the second $*$ denotes the complex conjugate transpose.

Proof. Let $A := [T]_E^F$, i.e., $Te_j = \sum_{k=1}^m a_{kj} f_k$, $j = 1, \dots, n$. So $(Te_j, f_i)_W = a_{ij}$. By the proof of Theorem 1.4.1

$$T^* f_j = \sum_{i=1}^n (f_j, Te_i)_W e_i = \sum_{i=1}^n \overline{(Te_i, f_j)}_W e_i = \sum_{i=1}^n \overline{a}_{ji} e_i.$$

So $[T^*]_F^E = A^* = ([T]_E^F)^*$. □

Notice that if E and F are not orthonormal, then $[T^*]_F^E = ([T]_E^F)^*$ may not hold.

Problems

1. Let $S, T \in \operatorname{Hom}(V, W)$ where V, W are inner product spaces. Prove that
   (a) $(Tv, w) = 0$ for all $v \in V$ and $w \in W$ if and only if $T = 0$.
   (b) $(Sv, w) = (Tv, w)$ for all $v \in V$ and $w \in W$ if and only if $S = T$.
2. Show that $(ST)^* = T^* S^*$ where $T \in \operatorname{Hom}(V, W)$ and $S \in L(W, U)$ and U, V, W are inner product spaces.
3. Let E and F be orthonormal bases of inner product space V. Prove that $([I]_E^F)^* = ([I]_E^F)^{-1} = [I]_F^E$.
4. Let G be a basis of the inner product space V. Let $T \in \operatorname{End} V$. Prove that $[T^*]_G^G$ and $([T]_G^G)^*$ are similar.
5. Let V, W be inner product spaces. Prove that if $T \in \operatorname{Hom}(V, W)$, then $\operatorname{rank} T = \operatorname{rank} T^*$.
6. The adjoint $* : \operatorname{Hom}(V, W) \to \operatorname{Hom}(W, V)$ is an isomorphism.

Solutions to Problems 1.4

1. (a) $(Tv, w) = 0$ for all $v \in V$ and $w \in W \Rightarrow Tv = 0$ for all $v \in V$, i.e., $T = 0$.

   (b) Consider $S - T$.

2. Since $(v, T^*S^*w)_V = (Tv, S^*w)_V = (STv, w)_V$, by the uniqueness of adjoint, $(ST)^* = T^*S^*$.

3. Notice that $I^* = I$ where $I := I_V$. So from Theorem 1.4.2 $([I]_E^F)^* = [I^*]_F^E = [I]_E^F$. Then by Problem 2.2, $([I]_E^F)^{-1} = [I]_F^E$.

4. (Roy and Alex) Let E be an orthonormal basis of V. Let $P = [I_V]_G^E$ (the transition matrix from G to E). Then $[T^*]_G^G = P^{-1}[T^*]_E^E P$ and $[T]_E^E = P[T]_G^G P^{-1}$. By Theorem 1.4.2 $[T^*]_E^E = ([T]_E^E)^*$. Together with Problem 2

$$\begin{aligned}[T^*]_G^G &= P^{-1}[T^*]_E^E P = P^{-1}([T]_E^E)^* P \\ &= P^{-1}(P[T]_G^G P^{-1})^* P = (P^*P)^{-1}([T]_G^G)^*(P^*P),\end{aligned}$$

   i.e. $([T]_G^G)^*$ and $([T^*]_G^G)$ are similar.

   Remark: If G is an orthonormal basis, then P is unitary and thus $([T]_G^G)^* = ([T^*]_G^G)$, a special case of Theorem 1.4.2.

5. Follows from Theorem 1.4.2, Problem 2.3 and $\operatorname{rank} A = \operatorname{rank} A^*$ where A is a matrix.

6. It is straightforward to show that $* : \operatorname{Hom}(V, W) \to \operatorname{Hom}(W, V)$ is a linear map. Since $\dim(V, W) = \dim(W, V)$ it remains to show that $*$ is injective. First show that $(T^*)^* = T$: for all $v \in V$, $w \in W$,

$$(T^*w, v) = \overline{(v, T^*w)}_V = \overline{(Tv, w)}_V = (w, Tv)$$

   Then for any $S, T \in \operatorname{Hom}(V, W)$, $S^* = T^* \Leftrightarrow S = T$.

1.5 Normal operators and matrices

An operator $T \in \operatorname{End} V$ on the inner product space V is **normal** if $T^*T = TT^*$; **Hermitian** if $T = T^*$, **positive semi-definite**, abbreviated as psd or $T \geq 0$ if $(Tx, x) \geq 0$ for all $x \in V$; **positive definite**, abbreviated as pd or $T > 0$ if $(Tx, x) > 0$ for all $0 \neq x \in V$; **unitary** if $T^*T = I$. Unitary operators form a group.

When $V = \mathbb{C}^n$ is equipped with the standard inner product and orthonormal basis, linear operators are viewed as matrices in $\mathbb{C}_{n\times n}$ and the adjoint is simply the complex conjugate transpose. Thus a matrix $A \in \mathbb{C}_{n\times n}$ is said to be normal if $A^*A = AA^*$; Hermitian if $A = A^*$, psd if $(Ax, x) \geq 0$ for all $x \in \mathbb{C}^n$; pd if $(Ax, x) > 0$ for all $0 \neq x \in \mathbb{C}^n$; **unitary** if $A^*A = I$. Unitary matrices in

$\mathbb{C}_{n\times n}$ form a group and is denoted by $\mathrm{U}_n(\mathbb{C})$. One can see immediately that A is unitary if A has orthonormal columns (rows since $AA^* = I$ as well from Problem 1.8).

Theorem 1.5.1. (Schur triangularization theorem) Let $A \in \mathbb{C}_{n\times n}$. There is $U \in \mathrm{U}_n(\mathbb{C})$ such that U^*AU is upper triangular.

Proof. Let λ_1 be an eigenvalue of A with unit eigenvector x_1, i.e., $Ax_1 = \lambda_1 x_1$ with $\|x_1\|_2 = 1$. Extend via Theorem 1.3.4 to an orthonormal basis $\{x_1, \dots, x_n\}$ for $\mathbb{C}^n$ and set $Q = [x_1 \cdots x_n]$ which is unitary. Since $Q^*Ax_1 = \lambda_1 Q^* x_1 = \lambda_1(1, \dots, 0)^T$,

$$\hat{A} := Q^*AQ = \begin{pmatrix} \lambda_1 & * \cdots & * \\ 0 & & \\ \vdots & A_1 & \\ 0 & & \end{pmatrix}$$

where $A_1 \in \mathbb{C}_{(n-1)\times(n-1)}$. By induction, there is $U_1 \in \mathrm{U}_{n-1}(\mathbb{C})$ such that $U_1^* A_1 U_1$ is upper triangular. Set

$$P := \begin{pmatrix} 1 & 0 \cdots & 0 \\ 0 & & \\ \vdots & U_1 & \\ 0 & & \end{pmatrix}$$

which is unitary. So

$$P^* \hat{A} P = \begin{pmatrix} \lambda_1 & * \cdots & * \\ 0 & & \\ \vdots & U_1^* A_1 U_1 & \\ 0 & & \end{pmatrix}$$

is upper triangular and set $U = QP$. □

Theorem 1.5.2. Let $T \in \operatorname{End} V$, where V is an inner product space. Then there is an orthonormal basis E of V such that $[T]_E^E$ is upper triangular.

Proof. For any orthonormal basis $E' = \{e_1', \dots, e_n'\}$ of V let $A := [T]_{E'}^{E'}$. By Theorem 1.5.1 there is $U \in \mathrm{U}_n(\mathbb{C})$, where $n = \dim V$, such that U^*AU is upper triangular. Since U is unitary, $E = \{e_1, \dots, e_n\}$ is also an orthonormal basis, where $e_j = \sum_{i=1}^n u_{ij} e_i'$, $j = 1, \dots, n$ (check!) and $[I]_E^{E'} = U$. Hence $[T]_E^E = [I]_{E'}^E [T]_{E'}^{E'} [I]_E^{E'} = U^*AU$ since $U^* = U^{-1}$. □

Lemma 1.5.3. Let $A \in \mathbb{C}_{n\times n}$ be normal and let r be fixed. Then $a_{rj} = 0$ for all $j \neq r$ if and only if $a_{ir} = 0$ for all $i \neq r$.

Proof. From $a_{rj} = 0$ for all $j \neq r$ and $AA^* = A^*A$,

$$|a_{rr}|^2 = \sum_{j=1}^n |a_{rj}|^2 = (AA^*)_{rr} = (A^*A)_{rr} = |a_{rr}|^2 + \sum_{i\neq r} |a_{ir}|^2.$$

So $a_{ir} = 0$ for all $i \neq r$. Then apply it on the normal A^*. □

Theorem 1.5.4. Let $A \in \mathbb{C}_{n\times n}$. Then

1. A is normal if and only if A is unitarily similar to a diagonal matrix.
2. A is Hermitian (psd, pd) if and only if A is unitarily similar to a real (nonnegative, positive) diagonal matrix.
3. A is unitary if and only if A is unitarily similar to a diagonal unitary matrix.

Proof. Notice that if A is normal, Hermitian, psd, pd, unitary, so is U^*AU for any $U \in \mathrm{U}_n(\mathbb{C})$. By Theorem 1.5.1, there is $U \in \mathrm{U}_n(\mathbb{C})$ such that U^*AU is upper triangular and also normal. By Lemma 1.5.3, the result follows. The rest follows similarly. □

From Gram-Schmidt orthgonalization (Theorem 1.3.4), one has the QR decomposition of an invertible matrix.

Theorem 1.5.5. (QR decomposition) Each invertible $A \in \mathbb{C}_{n\times n}$ can be decomposed as $A = QR$, where $Q \in \mathrm{U}_n(\mathbb{C})$ and R is upper triangular. The diagonal entries of R may be chosen positive; in this case the decomposition is unique.

Problems

1. Show that "upper triangular" in Theorem 1.5.1 may be replaced by "lower triangular".
2. Prove that $A \in \operatorname{End} V$ is unitary if and only if $\|Av\| = \|v\|$ for all $v \in V$.
3. Show that $A \in \mathbb{C}_{n\times n}$ is psd (pd) if and only if $A = B^*B$ for some (invertible) matrix B. In particular B may be chosen lower or upper triangular.
4. Let V be an inner product space. Prove that (i) $(Av, v) = 0$ for all $v \in V$ if and only if $A = 0$, (ii) $(Av, v) \in \mathbb{R}$ for all $v \in V$ if and only if A is Hermitian. What happens if the underlying field $\mathbb{C}$ is replaced by $\mathbb{R}$?
5. Show that if A is psd, then A is Hermitian. What happens if the underlying field $\mathbb{C}$ is replaced by $\mathbb{R}$?
6. Prove Theorem 1.5.5.
7. Prove that if $T \in \operatorname{Hom}(V, W)$ where V and W are inner product spaces, then $T^*T \ge 0$ and $TT^* \ge 0$. If T is invertible, then $T^*T > 0$ and $TT^* > 0$.

Solutions to Problems 1.5

1. Apply Theorem 1.5.1 on A^* and take complex conjugate transpose back.

2. $A \in \operatorname{End} V$ is unitary $\Leftrightarrow A^*A = I_V$ So unitary A implies $\|Av\|^2 = (Av, Av) = (A^*Av, v) = (v, v) = \|v\|^2$. Conversely, $\|Av\| = \|v\|$ for all $v \in V$ implies $((A^*A - I)v, v) = 0$ where $A^*A - I$ is Hermitian. Apply Problem 5.

3. If $A = B^*B$, then $x^*Ax = xB^*Bx = \|Bx\|_2^2 \geq 0$ for all $x \in \mathbb{C}^n$. Conversely if A is psd, we can define a psd $A^{1/2}$ via Theorem 1.5.4. Apply QR on $A^{1/2}$ to have $A^{1/2} = QR$ where Q is unitary and R is upper triangular. Hence $A = A^{1/2}A^{1/2} = (A^{1/2})^*A^{1/2} = R^*Q^*QR = R^*R$. Set $B := R$.

   Let $A^{1/2} = PL$ where P is unitary and L is lower triangular (apply QR decomposition on $(A^*)^{-1}$). Then $A = L^*L$.

4. (i) It suffices to show for Hermitian A because of Hermitian decomposition $A = H+iK$ where $H := (A+A^*)/2$ and $K = (A-A^*)/(2i)$ are Hermitian and
$$(Av, v) = (Hv, v) + i(Kv, v)$$
   and $(Hv, v), (Kv, v) \in \mathbb{R}$. So $(Av, v) = 0$ for all v if and only if $(Hv, v) = (Kv, v) = 0$ for all $v \in V$. Suppose A is Hermitian and $(Av, v) = 0$ for all $v \in V$. Then there is an orthonormal basis of eigenvectors $\{v_1, \ldots, v_n\}$ of A with corresponding eigenvalues $\lambda_1, \ldots, \lambda_n$, but all eigenvalues are zero because $\lambda_i(v_i, v_i) = (Av_i, v_i) = 0$. When $\mathbb{C}$ is replaced by $\mathbb{R}$, we cannot conclude that $A = 0$ since for any real skew symmetric matrix $A \in \mathbb{R}_{n\times n}$, $x^TAx = 0$ (why?).
   (Brice) Notice that A is psd (because $(Av, v) = 0$). From Problem 3 via matrix-operator approach, $A = B^*B$ for some $B \in \operatorname{End} V$. Thus from Problem 4.2, $A = A^*$.
   (ii)
$$2[(Av, w) + (Aw, v)] = (A(v + w), v + w) - (A(v - w), v - w) \in \mathbb{R}$$
   so that $\operatorname{Im}[(Av, w) + (Aw, v)] = 0$ for all v, w. Similarly
$$2i[-(Av, w) + (Aw, v)] = (A(v + iw), v + iw) - (A(v - iw), v - iw) \in \mathbb{R}$$
   so that $\operatorname{Re}[-(Av, w) + (Aw, v)] = 0$ for all v, w. So
$$(Av, w) = \overline{(Aw, v)} = (v, Aw),$$
   i.e., $A = A^*$. When $\mathbb{C}$ is replaced by $\mathbb{R}$, we cannot conclude that A is real symmetric since for any real skew symmetric matrix $A \in \mathbb{R}_{n\times n}$, $x^TAx = 0$.

5. Follow from Problem 4.

6. Express the columns (basis of $\mathbb{C}^n$) of A in terms of the (orthonormal) columns of Q.

7. $(T^*Tv, v) = (Tv, Tv) \geq 0$ for all $v \in V$. So $T^*T \geq 0$ and similar for $TT^* \geq 0$. When T is invertible, $Tv \neq 0$ for all $v \neq 0$ so $(T^*Tv, v) = (Tv, Tv) > 0$.

1.6 Inner product and positive operators

Theorem 1.6.1. Let V be an inner product space with inner product $(\cdot,\cdot)$ and let $T \in \operatorname{End} V$. Then $\langle u, v\rangle := (Tu, v)$, $u, v \in V$, defines an inner product if and only if T is pd with respect to $(\cdot,\cdot)$.

Proof. Suppose $\langle u, v\rangle := (Tu, v)$, $u, v \in V$, defines an inner product, where $(\cdot,\cdot)$ is an inner product for V. So

$$(Tu, v) = \langle u, v\rangle = \overline{\langle v, u\rangle} = \overline{(Tv, u)} = (T^* u, v)$$

for all $u, v \in V$. So $T = T^*$, i.e., self-adjoint. For any $v \neq 0$, $0 < \langle v, v\rangle = (Tv, v)$ so that T is pd with respect to $(\cdot,\cdot)$. The other implication is trivial. □

The next theorem shows that in the above manner inner products are in one-one correspondence with pd matrices.

Theorem 1.6.2. Let $(\cdot,\cdot)$ and $\langle\cdot,\cdot\rangle$ be inner products of V. Then there exists a unique $T \in \operatorname{End} V$ such that $\langle u, v\rangle := (Tu, v)$, $u, v \in V$. Moreover, T is positive definite with respect to both inner products.

Proof. Let $E = \{e_1, \dots, e_n\}$ be an orthonormal basis of V with respect to $(\cdot,\cdot)$. So each $v \in V$ can be written as $v = \sum_{i=1}^n (v, e_i)e_i$. Now for each $v \in V$ defines $T \in \operatorname{End} V$ by

$$Tv := \sum_{i=1}^n \langle v, e_i\rangle e_i$$

Clearly

$$(Tu, v) = \sum_{i=1}^n \langle u, e_i\rangle (e_i, v) = \langle u, \sum_{i=1}^n (v, e_i)e_i\rangle = \langle u, v\rangle.$$

Since $\langle\cdot,\cdot\rangle$ is an inner product, from Theorem 1.6.1 T is pd with respect to $(\cdot,\cdot)$. When $v \neq 0$, $\langle Tv, v\rangle = (T^2 v, v) = (Tv, Tv) > 0)$ so that T is pd with respect to $\langle\cdot,\cdot\rangle$. The uniqueness follows from Problem 4.1. □

Theorem 1.6.3. Let $F = \{f_1, \dots, f_n\}$ be a basis of V. There exists a unique inner product $(\cdot,\cdot)$ on V such that F is an orthonormal basis.

Proof. Let $(\cdot,\cdot)$ be an inner product with orthonormal basis $E = \{e_1, \dots, e_n\}$. By Problem 1.4 $S \in \operatorname{End} V$ defined by $Sf_i = e_i$ is invertible. Set $T := S^*S > 0$ (* and $T > 0$ from Problem 5.7 are with respect to $(\cdot,\cdot)$). So $\langle u, v\rangle = (Tu, v)$ is an inner product by Theorem 1.6.1 and $f_1, \dots, f_n$ are orthonormal with respect to $\langle\cdot,\cdot\rangle$. It is straightforward to show the uniqueness. □

Problems

1. Let $E = \{e_1, \ldots, e_n\}$ be a basis of V. For any $u = \sum_{i=1}^n a_i e_i$ and $v = \sum_{i=1}^n b_i e_i$, show that $(u, v) := \sum_{i=1}^n a_i \bar{b}_i$ is the unique inner product on V so that E is an orthonormal basis.

2. Find an example that there are inner products $(\cdot, \cdot)$ on $\langle \cdot, \cdot \rangle$ on V and $T \in \operatorname{End} V$ so that T is pd with respect to one but not to the other.

3. Let V be an inner product space. Show that for each $A \in \mathbb{C}_{n\times n}$ there are $u_1, \ldots, u_n; v_1, \ldots, v_n \in \mathbb{C}^n$ such that $a_{ij} = (u_i, v_j)$ $1 \le i, j \le n$.

4. Suppose that $(\cdot, \cdot)$ on $\langle \cdot, \cdot \rangle$ are inner products on V, $T \in \operatorname{End} V$ and let $T^{(*)}$ and $T^{\langle * \rangle}$ be the corresponding adjoints. Show that $T^{(*)}$ and $T^{\langle * \rangle}$ are similar.

Solutions to Problems 1.6

1. Uniqueness follows from Theorem 1.6.3.

2.

3.

4. By Theorem 1.6.1 there is a pd $S \in \operatorname{End} V$ with respect to both inner products such that $\langle u, v \rangle = (Su, v)$ for all $u, v \in V$ and thus $\langle S^{-1}u, v \rangle = (u, v)$. So

$$\langle u, T^{\langle * \rangle} v \rangle = \langle Tu, v \rangle = (STu, v) = (u, T^{(*)} S^{(*)} v) = \langle S^{-1}u, T^{(*)} S^{(*)} v \rangle$$

Since $S^{(*)} = S^{\langle * \rangle} = S^*$, say, and $(A^{-1})^* = (A^*)^{-1}$ for any $A \in \operatorname{End} V$ ($(A^{-1})^* A^* = (AA^{-1})^* = I$ by Problem 4.2). Then

$$\langle u, T^{\langle * \rangle} v \rangle = \langle u, (S^{-1})^* T^{(*)} S^* v \rangle = \langle u, (S^*)^{-1} T^{(*)} S^* v \rangle.$$

Hence $T^{\langle * \rangle} = (S^*)^{-1} T^{(*)} S^*$.

1.7 Invariant subspaces

Let W be a subspace of V. For any $T \in \operatorname{Hom}(V, U)$, the **restriction** of T, denote by $T|_W$, is the unique $T_1 \in \operatorname{Hom}(W, U)$ such that $T_1(w) = T(w)$ for all $w \in W$.

Let W be a subspace of V and $T \in \operatorname{End} V$. Then W is said to be **invariant** under T if $Tw \in W$ for all $w \in W$, i.e., $T(W) \subset W$. The trivial invariant subspaces are 0 and V. Other invariant subspaces are called **nontrivial** or **proper** invariant subspaces. If W is invariant under T, then the restriction $T|_W \in \operatorname{Hom}(W, V)$ of T induces $T|_W \in \operatorname{End} W$ (we use the same notation $T|_W$).

Theorem 1.7.1. Let V be an inner product space over $\mathbb{C}$ and $T \in \operatorname{End} V$. If W is an invariant subspace under T and T^*, then

(a) $(T|_W)^* = T^*|_W$.

(b) If T is normal, Hermitian, psd, pd, unitary, so is $T|_W$.

Proof. (a) For any $x, y \in W$, $T|_W x = Tx$, $T^*|_W y = T^* y$. By the assumption,

$$(T|_W x, y)_W = (Tx, y)_V = (x, T^* y)_V = (x, T^*|_W y)_W.$$

So $(T|_W)^* = T^*|_W$.
(b) Follows from Problem 7.3. $\square$

Problems

1. Prove that if $W \subset V$ is a subspace and $T_1 \in \operatorname{Hom}(W, V)$, then there is $T \in \operatorname{Hom}(V, U)$ so that $T|_W = T_1$. Is T unique?

2. Show that the restriction on $W \subset V$ of the sum of $S, T \in \operatorname{Hom}(V, U)$ is the sum of their restrictions. How about restriction of scalar multiple of T on W?

3. Show that if $S, T \in \operatorname{End} V$ and the subspace W is invariant under S and T, then $(ST)|_W = (S|_W)(T|_W)$.

4. Let $T \in \operatorname{End} V$ and $S \in \operatorname{End} W$ and $H \in \operatorname{Hom}(V, W)$ satisfy $SH = HT$. Prove that $\operatorname{Im} H$ is invariant under S and $\operatorname{Ker} H$ is invariant under T.

Solutions to Problems 1.7

1. Let $\{e_1, \ldots, e_m\}$ be a basis of W and extend it to $E = \{e_1, \ldots, e_m, e_{m+1}, \ldots, e_n\}$ a basis of V. Define $T \in \operatorname{Hom}(V, U)$ by $Te_i = T_1 e_i$, $i = 1, \ldots, m$ and $Te_j = w_j$ where $j = m+1, \ldots, n$ and $w_{m+1}, \ldots, w_n \in W$ are arbitrary. Clearly $T|_W = T_1$ but T is not unique.

2. $(S+T)|_W v = (S+T)v = Sv + Tv = S|_W v + T|_W v$ for all $v \in V$. So $(S+T)|_W = S|_W + T|_W$.

3. $(S|_W)(T|_W)v = S|_W(Tv) = STv$ since $Tv \in W$ and thus $(S|_W)(T|_W)v = ST|_W v$ for all v.

4. For any $v \in \operatorname{Ker} H$, $H(Tv) = SHv = 0$, i.e., $Tv \in \operatorname{Ker} H$ so that $\operatorname{Ker} H$ is an invariant under T. For any $Hv \in \operatorname{Im} H$ $(v \in V)$, $S(Hv) = HT(v) \in \operatorname{Im} H$.

1.8 Projections and direct sums

The vector space V is said to be a **direct sum** of the subspaces $W_1, \ldots, W_m$ if each $v \in V$ can be uniquely expressed as $v = \sum_{i=1}^n w_i$ where $w_i \in W_i$, $i = 1, \ldots, n$ and is denoted by $V = W_1 \oplus \cdots \oplus W_m$. In other words, $V = W_1 + \cdots + W_m$ and if $w_1 + \cdots + w_m = w_1' + \cdots + w_m'$ where $w_i \in W_i$, then $w_i = w_i'$ for all i.

Theorem 1.8.1. $V = W_1 \oplus \cdots \oplus W_m$ if and only if $V = W_1 + \cdots + W_m$ and $W_i \cap (W_1 + \cdots + \hat{W}_i + \cdots + W_m) = 0$ for all $i = 1, \ldots, m$. Here $\hat{W}_i$ denotes deletion. In particular, $V = W_1 \oplus W_2$ if and only if $W_1 + W_2 = V$ and $W_1 \cap W_2 = 0$.

Proof. Problem 7. □

In addition, if V is an inner product space, and $W_i \perp W_j$ if $i \neq j$, i.e., $w_i \perp w_j$ whenever $w_i \in W_i$, $w_j \in W_j$, we call V an **orthogonal sum** of $W_1, \ldots, W_m$, denoted by $V = W_1 \dot{+} \cdots \dot{+} W_m$.

Suppose $V = W_1 \oplus \cdots \oplus W_m$. Each $P_i \in \operatorname{End} V$ defined by $P_i v = w_i$, $i = 1, \ldots, m$ satisfies $P_i^2 = P_i$ and $\operatorname{Im} P_i = W_i$. In general $P \in \operatorname{End} V$ is called a **projection** if $P^2 = P$. Notice that P is a projection if and only if $I_V - P$ is a projection; in this case $\operatorname{Im} P = \operatorname{Ker}(I_V - P)$.

Theorem 1.8.2. Let V be a vector space and let $P \in \operatorname{End} V$ be a projection. Then the eigenvalues of P are either 0 or 1 and $\operatorname{rank} P = \operatorname{tr} P$.

Proof. Let $\operatorname{rank} P = k$. By Theorem 1.1.2, there are $v_1, \ldots, v_n \in V$ such that $Pv_1, \ldots, Pv_k$ is a basis of $\operatorname{Im} P$ and $\{v_{k+1}, \ldots, v_n\}$ is a basis of $\operatorname{Ker} P$. From $P^2 = P$, $E = \{Pv_1, \ldots, Pv_k, v_{k+1}, \ldots, v_n\}$ is linearly independent (check) and thus a basis of V. So $[P]_E^E = \operatorname{diag}(1, \ldots, 1, 0, \ldots, 0)$ in which there are k ones. So we have the desired results. □

The notions direct sum and projections are closely related.

Theorem 1.8.3. Let V be a vector space and let $P_1, \ldots, P_m$ be projections such that $P_1 + \cdots + P_m = I_V$. Then $V = \operatorname{Im} P_1 \oplus \cdots \oplus \operatorname{Im} P_m$. Conversely, if $V = W_1 \oplus \cdots \oplus W_m$, then there are unique projections $P_1, \ldots, P_m$ such that $P_1 + \cdots + P_m = I_V$ and $\operatorname{Im} P_i = W_i$, $i = 1, \ldots, m$.

Proof. From $P_1 + \cdots + P_m = I_V$, each $v \in V$ can be written as $v = P_1 v + \cdots + P_m v$ so that $V = \operatorname{Im} P_1 + \cdots + \operatorname{Im} P_m$. By Theorem 1.8.2

$$\dim V = \operatorname{tr} I_V = \sum_{i=1}^m \operatorname{tr} P_i = \sum_{i=1}^m \operatorname{rank} P_i = \sum_{i=1}^m \dim \operatorname{Im} P_i.$$

So $V = \operatorname{Im} P_1 \oplus \cdots \oplus \operatorname{Im} P_m$ (Problem 8.1).

Conversely if $V = W_1 \oplus \cdots \oplus W_m$, then for any $v \in V$, there is unique factorization $v = w_1 + \cdots + w_m$, $w_i \in W_i$, $i = 1, \ldots, m$. Define $P_i \in \operatorname{End} V$ by $P_i v = w_i$. It is easy to see each P_i is a projection and $\operatorname{Im} P_i = W_i$ and $P_1 + \cdots + P_m = I_V$. For the uniqueness, if there are projections $Q_1, \ldots, Q_m$ such that $Q_1 + \cdots + Q_m = I_V$ and $\operatorname{Im} Q_i = W_i$ for all i, then for each $v \in V$, from the unique factorization $Q_i v = P_i v$ so that $P_i = Q_i$ for all i. □

Corollary 1.8.4. If $P \in \operatorname{End} V$ is a projection, then $V = \operatorname{Im} P \oplus \operatorname{Ker} P$. Conversely if $V = W \oplus W'$, then there is a projection $P \in \operatorname{End} V$ such that $\operatorname{Im} P = W$ and $\operatorname{Ker} P = W'$.

We call $P \in \operatorname{End} V$ an **orthogonal projection** if $P^2 = P = P^*$. Notice that P is an orthogonal projection if and only if $I_V - P$ is an orthogonal projection. The notions orthogonal sum and orthogonal projections are closely related.

Theorem 1.8.5. Let V be an inner product space and let $P_1, \ldots, P_m$ be orthogonal projections such that $P_1 + \cdots + P_m = I_V$. Then $V = \operatorname{Im} P_1 \dot{+} \cdots \dot{+} \operatorname{Im} P_m$. Conversely, if $V = W_1 \dot{+} \cdots \dot{+} W_m$, then there are unique orthogonal projections $P_1, \ldots, P_m$ such that $P_1 + \cdots + P_m = I_V$ and $\operatorname{Im} P_i = W_i$, $i = 1, \ldots, m$.

Proof. From Theorem 1.8.3, if $P_1, \ldots, P_m$ are orthogonal projections such that $P_1 + \cdots + P_m = I_V$, then $V = \operatorname{Im} P_1 \oplus \cdots \oplus \operatorname{Im} P_m$. It suffices to show that the sum is indeed orthogonal. From the proof of Theorem 1.8.3, $P_i v = v_i \in \operatorname{Im} P_i$ for all i, where $v \in V$ is expressed as $v = \sum_{i=1}^m v_i$. Thus for all $u, v \in V$, if $i \neq j$, then

$$(P_i u, P_j v) = (P_i u, v_j) = (u, P_i^* v_j) = (u, P_i v_j) = (u, 0) = 0.$$

Conversely if $V = W_1 \dot{+} \cdots \dot{+} W_m$, then from Theorem 1.8.3 there are unique projections $P_1, \ldots, P_m$ such that $P_1 + \cdots + P_m = I_V$ and $\operatorname{Im} P_i = W_i$, for all i. It remains to show that each projection P_i is an orthogonal projection, i.e., $P_i = P_i^*$. For $u, v \in V$, write $u = \sum_{i=1}^m u_i$, $v = \sum_{i=1}^m v_i$ where $u_i, v_i \in W_i$ for all i. Then for all i

$$(P_i u, v) = (u_i, v) = (u_i, \sum_{i=1}^m v_i) = (u_i, v_i) = (\sum_{i=1}^m u_i, v_i) = (u, P_i v).$$

□

Corollary 1.8.6. Let V be an inner product space. If $P \in \operatorname{End} V$ is an orthogonal projection, then $V = \operatorname{Im} P \dot{+} \operatorname{Ker} P$. Conversely if $V = W \dot{+} W'$, then there is an orthogonal projection $P \in \operatorname{End} V$ such that $\operatorname{Im} P = W$ and $\operatorname{Ker} P = W'$.

Proof. Apply Theorem 1.8.5 on P and $I_V - P$. □

Problems

1. Show that if $V = W_1 + \cdots + W_m$, then $V = W_1 \oplus \cdots \oplus W_m$ if and only if $\dim V = \dim W_1 + \cdots + \dim W_m$.

2. Let $P_1, \ldots, P_m \in \operatorname{End} V$ be projections on V and $P_1 + \cdots + P_m = I_V$. Show that $P_i P_j = 0$ whenever $i \neq j$.

3. Let $P_1, \ldots, P_m \in \operatorname{End} V$ be projections on V. Show that $P_1 + \cdots + P_m$ is a projection if and only if $P_i P_j = 0$ whenever $i \neq j$.

4. Show that if $P \in \operatorname{End} V$ is an orthogonal projection, then $\|Pv\| \le \|v\|$ for all $v \in V$.

5. Prove that if $P_1, \ldots, P_m \in \operatorname{End} V$ are projections on V with $P_1 + \cdots + P_m = I_V$, then there is an inner product so that $P_1, \ldots, P_n$ are orthogonal projections.

6. Show that if $P \in \operatorname{End} V$ is an orthogonal projection on the inner product space and if W is P-invariant subspace of V, then $P|_W \in \operatorname{End} W$ is an orthogonal projection.

7. Prove Theorem 1.8.1.

Solutions to Problems 1.8

1.

2. By Theorem 1.8.3, $V = \operatorname{Im} P_1 \oplus \cdots \oplus \operatorname{Im} P_m$ so that for any $v \in V$, $P_i P_j v = P_i v_j = 0$ if $i \neq j$.

3. $P_1 + \cdots + P_m$ is a projection means

$$\sum_i P_i + \sum_{i \neq j} P_i P_j = \sum_i P_i^2 + \sum_{i \neq j} P_i P_j = (P_1 + \cdots + P_m)^2 = P_1 + \cdots + P_m.$$

So if $P_i P_j = 0$ for $i \neq j$, then $P_1 + \cdots + P_m$ is a projection. Conversely, if $P_1 + \cdots + P_m$ is a projection, so is $I_V - (P_1 + \cdots + P_m)$.

(Roy) Suppose that $P := P_1 + \cdots + P_m$ is a projection. Then the restriction of P on its image $P(V)$ is the identity operator and $P_i|_{P(V)}$ is still a projection for each $i = 1, \ldots, m$. According to Theorem 1.8.3, $P(V) = \oplus_{i=1}^m \operatorname{Im} P_i|_{P(V)} = \oplus_{i=1}^m \operatorname{Im} P_i$, which implies that $P_i P_j = 0$ whenever $i \neq j$.

4. $I - P$ is also an orthogonal projection. Write $v = Pv + (I - P)v$ so that $\|v\|^2 = \|Pv\|^2 + \|(I_P)v\|^2 \ge \|Pv\|^2$.

5. By Theorem 1.8.3, $V = \operatorname{Im} P_1 \oplus \cdots \oplus \operatorname{Im} P_m$. Let $\{e_{i1}, \ldots, e_{in_i}\}$ be a basis of $\operatorname{Im} P_i$ for all i. Then there is an inner product, by Theorem 1.6.3 so that the basis $\{e_{11}, \ldots, e_{1n_1}, \ldots, e_{m1}, \ldots, e_{mn_m}\}$ is orthonormal. Thus $P_1, \ldots, P_m$ are orthogonal projections by Theorem 1.8.5.

6. Notice that $P|_W^2 = P^2|_W = P|_W$ since $P^2 = P$ and from Theorem 1.7.1 $(P|_W)^{=} P^*|_W = P|_W$ since $P^* = P$.

7.

1.9 Dual spaces and Cartesian products

The space $V^* = \text{Hom}\,(V, \mathbb{C})$ is called the **dual space** of V. Elements in V^* are called (linear) **functionals**.

Theorem 1.9.1. Let $E = \{e_1, \ldots, e_n\}$ be a basis of V. Define $e_1^*, \ldots, e_n^* \in V^*$ by

$$e_i^*(e_j) = \delta_{ij}, \quad i, j = 1, \ldots, n.$$

Then $E^* = \{e_1^*, \ldots, e_n^*\}$ is a basis of V^*.

Proof. Notice that each $f \in V^*$ can be written as $f = \sum_{i=1}^n f(e_i)e_i^*$ since both sides take the same values on $e_1, \ldots, e_n$. Now if $\sum_{i=1}^n a_i e_i^* = 0$, then $0 = \sum_{i=1}^n a_i e_i^*(e_j) = a_j$, $j = 1, \ldots, n$. So $e_1^*, \ldots, e_n^*$ are linearly independent. □

The basis E^* is called a basis of V^* dual to E, or simply **dual basis**.

If $T \in \text{Hom}\,(V, W)$, then its **dual map** (or transpose) $T^* \in \text{Hom}\,(W^*, V^*)$ is defined by

$$T^*(\varphi) = \varphi \circ T, \qquad \varphi \in W^*.$$

The functional $T^*(\varphi)$ is in V^*, and is called the **pullback** of φ along T. Immediately the following identity holds for all $\varphi \in W^*$ and $v \in V$:

$$\langle T^*(\varphi),\, v\rangle = \langle \varphi,\, T(v)\rangle$$

where the bracket $\langle f, v\rangle := f(v)$ is the duality pairing of V with its dual space, and that on the right is the duality pairing of W with its dual. This identity characterizes the dual map T^*, and is formally similar to the definition of the adjoint (if V and W are inner product spaces). We remark that we use the same notation for adjoint and dual map of $T \in \text{Hom}\,(V, W)$. But it should be clear in the context since adjoint requires inner products but dual map does not.

Theorem 1.9.2. The map $* : \text{Hom}\,(V, W) \to \text{Hom}\,(W^*, V^*)$ defined by $*(T) = T^*$ is an isomorphism.

Proof. Problem 10. □

When V has an inner product, linear functionals have nice representation.

Theorem 1.9.3. (Riesz) Let V be an inner product space. For each $f \in V^*$, there is a unique $u \in V$ such that $f(v) = (v, u)$ for all $v \in V$. Hence the map $\xi : V \to V^*$ defined by $\xi(u) = (\cdot, u)$ is an isomorphism.

Proof. Let $E = \{e_1, \ldots, e_n\}$ be an orthonormal basis of V. Then for all $v \in V$, by Theorem 1.3.3

$$f(v) = f(\sum_{i=1}^n (v, e_i)e_i) = \sum_{i=1}^n f(e_i)(v, e_i) = (v, \sum_{i=1}^n \overline{f(e_i)} e_i)$$

so that $u := \sum_{i=1}^n \overline{f(e_i)} e_i$. Uniqueness follows from the positive definiteness of the inner product: If $(v, u') = (v, u)$ for all $v \in V$, then $(v, u - u') = 0$ for all v. Pick $v = u - u'$ to have $u = u'$.

The map $\xi : V \to V^*$ defined by $\xi(u) = (\cdot, u)$ is clearly a linear map bijective from the previous statement. □

Let $V_1, \ldots, V_m$ be m vector spaces over $\mathbb{C}$.

$$\times_i^m V_i = V_1 \times \cdots \times V_m = \{(v_1, \ldots, v_m) : v_i \in V_i,\ i = 1, \ldots, m\}$$

is called the **Cartesian product** of $V_1, \ldots, V_m$. It is a vector space under the natural addition and scalar multiplication.

Remark: If V is not finite-dimensional but has a basis $\{e_\alpha : \alpha \in A\}$ (axiom of choice is needed) where A is the (infinite) index set, then the same construction as in the finite-dimensional case (Theorem 1.9.1) yields linearly independent elements $\{e_\alpha^* : \alpha \in A\}$ in V^*, but they will not form a basis.

Example (for Alex's question): The space $\mathbb{R}^\infty$, whose elements are those sequences of real numbers which have only finitely many non-zero entries, has a basis $\{e_i : i = 1, 2, \ldots, \}$ where $e_i = (0, \ldots, 0, 1, 0, \ldots)$ in which the only nonzero entry 1 is at the ith position. The dual space of $\mathbb{R}^\infty$ is $\mathbb{R}^\aleph$, the space of all sequences of real numbers and such a sequence (a_n) is applied to an element $(x_n) \in \mathbb{R}^\infty$ to give $\sum_n a_n x_n$, which is a finite sum because there are only finitely many nonzero x_n. The dimension of $\mathbb{R}^\infty$ is countably infinite but $\mathbb{R}^\aleph$ does not have a countable basis.

Problems

1. Show that Theorem 1.9.3 remains true if the inner product is replaced by a nondegenerate bilinear form $B(\cdot, \cdot)$ on a real vector space. Nondegeneracy means that $B(u, v) = 0$ for all $v \in V$ implies that $u = 0$.

2. Let $\{e_1, \ldots, e_n\}$ be an orthonormal basis of the inner product space V. Show that $\{f_1, \ldots, f_n\}$ is a dual basis if and only if $f_j(v) = (v, e_i)$ for all $v \in V$ and $j = 1, \ldots, n$.

3. Let $\{f_1, \ldots, f_n\}$ be a basis of V^*. Show that if $v \in V$ such that $f_j(v) = 0$ for all $j = 1, \ldots, n$, then $v = 0$.

4. Let $\{f_1, \ldots, f\}$ be a basis of V^*. Prove that there is a basis $E = \{e_1, \ldots, e_n\}$ of V such that $f_i(e_j) = \delta_{ij}$, $i, j = 1, \ldots, n$.

5. Let $E = \{e_1, \ldots, e_n\}$ and $F = \{f_1, \ldots, f_n\}$ be two bases of V and let E^* and F^* be their dual bases. If $[I_V]_E^F = P$ and $[I_{V^*}]_{F^*}^{E^*} = Q$, prove that $Q = (P^{-1})^T$.

6. Show that if $S \in \mathrm{Hom}\,(W, U)$ and $T \in \mathrm{Hom}\,(V, W)$, then $(ST)^* = T^* S^*$.

7. Show that $\xi : V \to V^{**}$ defined by $\xi(v)(\varphi) = \varphi(v)$ for all $v \in V$ and $V \in V^*$, is an (canonical) isomorphism.

8. Suppose E and F are bases for V and W and let E^* and F^* be their dual bases. For any $T \in \operatorname{Hom}(V, W)$, what is the relation between $[T]_E^F$ and $[T^*]_{F^*}^{E^*}$?

9. Show that $\dim(V_1 \times \cdots \times V_m) = \dim V_1 + \cdots + \dim V_m$.

10. Prove Theorem 1.9.2.

Solutions to Problems 1.9

1. Let $B(\cdot, \cdot)$ be a nondegenerate bilinear form on a real vector space V. Define $\varphi : V \to V^*$ by $\varphi(v) = B(v, \cdot)$. Clearly φ is linear and we need to show that φ is surjective. Since $\dim V = \dim V^*$, it suffices to show that φ is injective (and thus bijective). Let $v \in \operatorname{Ker} \varphi$, i.e., $B(v, w) = 0$ for all $w \in V$. By the nondegeneracy of $B(\cdot, \cdot)$, $v = 0$.

2. Notice that $(e_i, e_j) = \delta_{ij}$. If F is a dual basis, then for each $v = \sum_{i=1}^n (v, e_i)e_i$, $f_j(v) = f_j(\sum_{i=1}^n (v, e_i)e_i) = \sum_{i=1}^n (v, e_i) f_j(e_i) = (v, e_j)$. On the other hand, if $f_j(v) = (v, e_j)$ for all v, then $f_j(e_i) = (e_i, e_j) = \delta_{ij}$, i.e., F is a dual basis.

3. The assumption implies that $f(v) = 0$ for all $f \in V^*$. If $v \neq 0$, extends it to basis $E = \{v, v_2, \ldots, v_n\}$ of V. Define $g \in V^*$ by $g(v) = 1$ and $g(v_i) = 0$ for all $i = 2, \ldots, n$.

4. Introduce an inner product on V so that by Theorem 1.9.3 we determine $u_1, \ldots, u_n$ via $f_j(v) = (v, u_j)$. By Theorem 1.6.3 and Theorem 1.6.1 there is a pd $T \in \operatorname{End} V$ such that $(Tu_i, u_j) = \delta_{ij}$. Set $e_i = Tu_i$ for all i. Uniqueness is clear.
   Another approach: Since $\{f_1, \ldots, f_n\}$ is a basis, each $f_i \neq 0$ so that $\dim \operatorname{Ker} f_i = n - 1$. So pick $v_i \in \cap_{j \neq i} \operatorname{Ker} f_j \neq 0$ (why?) such that $f_i(v_i) = 1$. Then $f_j(v_i) = 0$ for all $j \neq 0$.

5.

6. For any $\varphi \in U^*$, $(T^*S^*)\varphi = T^*(\varphi \circ S) = (\varphi \circ S) \circ T = \varphi \circ (ST) = (ST)^*\varphi$. So $(ST)^* = T^*S^*$.

7. Similar to Problem 10. Notice that $\xi(e_i) = e_i^{**}$ for all i. It is canonical because of $e \mapsto e^{**}$.

8. Let $A = [T]_E^F$, i.e., $Te_j = \sum_{i=1}^n a_{ij} f_i$. Write $v = \sum_i \alpha_i e_i$ so that

$$(T^* f_j^*)v = f_j^*(Tv) = f_j^*(\sum_i \alpha_i Te_i) = f_j^*(\sum_{i,k} \alpha_i a_{ki} f_k) = \sum_{i,k} \alpha_i a_{ki} f_j^*(f_k) = \sum_i \alpha_i a_{ji}.$$

On the other hand

$$(\sum_i a_{ji} e_i^*)(v) = \sum_i a_{ji} e_i^* (\sum_k \alpha_k e_k) = \sum_i \alpha_i a_{ji}.$$

So $T^* f_j^* = \sum_i a_{ji} e_i^*$, i.e., $[T^*]_{F^*}^{E^*} = A^T$, i.e., $([T]_E^F)^T = [T^*]_{F^*}^{E^*}$. This explains why T^* is also called the transpose of T.

(Roy) Suppose $E = \{e_1, \dots, e_n\}$ and $F = \{f_1, \dots, f_m\}$. Let $[T]_E^F = (a_{ij}) \in \mathbb{C}_{m\times n}$ and $[T^*]_{F^*}^{E^*} = (b_{pq}) \in \mathbb{C}_{n\times m}$. By definition, $Te_j = \sum_{i=1}^m a_{ij} f_i$ and $T^* f_q^* = \sum_{p=1}^n b_{pq} e_p^*$. Since $e_j^*(e_i) = \delta_{ij}$ and $f_j^*(f_i) = \delta_{ij}$, the definition $(T^* f_j^*) e_i = f_j^*(Te_i)$ implies that $b_{ij} = a_{ji}$. Thus $[T^*]_{F^*}^{E^*} = ([T]_E^F)^T$.

9. If $E_i = \{e_{i1}, \dots, e_{in_i}\}$ is a basis of V_i, $i = 1, \dots, m$, then

$$E = \{(e_{11}, \dots, e_{m1}), \dots, (e_{1n_1}, \dots, e_{mn_m})\}$$

is a basis of $V_1 \times \cdots \times V_m$ (check!).

10. $*$ is linear since for any scalars α, β, $(\alpha S + \beta T)^* \varphi = \varphi(\alpha S + \beta T) = \alpha\varphi S + \beta\varphi T = \alpha S^* \varphi + \beta T^* \varphi = (\alpha S^* + \beta T^*)\varphi$. It is injective since if $T^* = 0$, then $\varphi \circ T = T^* \varphi = 0$ for all $\varphi \in V^*$. By Problem 3, $T = 0$. Thus $*$ is an isomorphism.

1.10 Notations

Denote by S_m the symmetric group on $\{1, \dots, m\}$ which is the group of all bijections on $\{1, \dots, m\}$. Each $\sigma \in S_m$ is represented by

$$\sigma = \begin{pmatrix} 1, & \dots, & m \\ \sigma(1), & \dots, & \sigma(m) \end{pmatrix}.$$

The sign function ε on S_m is

$$\varepsilon(\sigma) = \begin{cases} 1 & \text{if } \sigma \text{ is even} \\ -1 & \text{if } \sigma \text{ is odd} \end{cases}$$

where $\sigma \in S_m$. Let α, β, γ denote finite sequences whose components are natural numbers

$$\begin{aligned}
\Gamma(n_1, \dots, n_m) &:= \{\alpha : \alpha = (\alpha(1), \dots, \alpha(m)), 1 \le \alpha(i) \le n_i, i = 1, \dots, m\} \\
\Gamma_{m,n} &:= \{\alpha : \alpha = (\alpha(1), \dots, \alpha(m)), 1 \le \alpha(i) \le n, i = 1, \dots, m\} \\
G_{m,n} &:= \{\alpha \in \Gamma_{m,n} : \alpha(1) \le \cdots \le \alpha(m)\} \\
D_{m,n} &:= \{\alpha \in \Gamma_{m,n} : \alpha(i) \ne \alpha(j) \text{ whenever } i \ne j\} \\
Q_{m,n} &:= \{\alpha \in \Gamma_{m,n} : \alpha(1) < \cdots < \alpha(m)\}.
\end{aligned}$$

For $\alpha \in Q_{m,n}$ and $\sigma \in S_m$, define

$$\alpha\sigma := (\alpha(\sigma(1)), \dots, \alpha\sigma(m))).$$

Then $\alpha\sigma \in D_{m,n}$ and

$$D_{m,n} = \{\alpha\sigma : \alpha \in Q_{m.n}, \sigma \in S_m\} = Q_{m,n}S_m. \tag{1.1}$$

Suppose $n > m$. For each $\omega \in Q_{m,n}$, denote by ω' the complementary sequence of ω, i.e., $\omega' \in Q_{n-m,n}$, the components of ω and ω' are $1, \dots, n$, and

$$\sigma = \begin{pmatrix} 1, & \dots, & m, & m+1, & \dots, & n \\ \omega(1), & \dots, & \omega(m), & \omega'(1), & \dots, & \omega'(n-m) \end{pmatrix} \in S_n.$$

It is known that

$$\varepsilon(\sigma) = (-1)^{s(\omega)+m(m+1)/2}, \tag{1.2}$$

where $s(\omega) := \omega(1) + \cdots + \omega(m)$. Similarly for $\omega \in Q_{m,n}$, $\theta \in S_m$ and $\pi \in S_{n-m}$,

$$\sigma = \begin{pmatrix} 1, & \dots, & m, & m+1, & \dots, & n \\ \omega\theta(1), & \dots, & \omega\theta(m), & \omega'\pi(1), & \dots, & \omega'\pi(n-m) \end{pmatrix} \in S_n$$

and

$$\varepsilon(\sigma) = \varepsilon(\theta)\varepsilon(\pi)(-1)^{s(\omega)+m(m+1)/2}, \tag{1.3}$$

Moreover

$$\begin{aligned} S_n \quad = \quad & \{\sigma = \begin{pmatrix} 1, & \dots, & m, & m+1, & \dots, & n \\ \omega\theta(1), & \dots, & \omega\theta(m), & \omega'\pi(1), & \dots, & \omega'\pi(n-m) \end{pmatrix} : \\ & \omega \in Q_{m,n}, \theta \in S_m, \pi \in S_{n-m}\}. \end{aligned} \tag{1.4}$$

Let $A \in \mathbb{C}_{n\times k}$. For any $1 \le m \le n$, $1 \le l \le k$, $\alpha \in Q_{m,n}$ and $\beta \in Q_{l,k}$. Let $A[\alpha|\beta]$ denote the submatrix of A by taking the rows $\alpha(1), \dots, \alpha(m)$ of A and the columns $\beta(1), \dots, \beta(l)$. So the (i,j) entry of $A[\alpha|\beta]$ is $a_{\alpha(i)\beta(j)}$. The submatrix of A complementary to $A[\alpha|\beta]$ is denoted by $A(\alpha|\beta) := A[\alpha'|\beta'] \in \mathbb{C}_{(n-m)\times(k-l)}$. Similarly we define $A(\alpha|\beta] := A[\alpha'|\beta] \in \mathbb{C}_{(n-m)\times k}$ and $A[\alpha|\beta) := A[\alpha|\beta'] \in \mathbb{C}_{n\times(k-l)}$.

Recall that if $A \in \mathbb{C}_{n\times n}$, the determinant function is given by

$$\det A = \sum_{\sigma\in S_n} \varepsilon(\sigma) \prod_{i=1}^n a_{i\sigma(i)}. \tag{1.5}$$

It is easy to deduce that

$$\det A = \sum_{\sigma\in S_n} \varepsilon(\pi)\varepsilon(\sigma) \prod_{i=1}^n a_{\pi(i),\sigma(i)}. \tag{1.6}$$

Let $A \in \mathbb{C}_{n\times k}$. For any $1 \le m \le \min\{n,k\}$, $\alpha \in Q_{m,n}$ and $\beta \in Q_{m,k}$,

$$\det A[\alpha|\beta] = \sum_{\sigma\in S_m} \varepsilon(\sigma) \prod_{i=1}^m a_{\alpha(i),\beta\sigma(i)}. \tag{1.7}$$

For any $\theta \in S_n$, we have

$$\det A[\alpha\theta|\beta] = \det A[\alpha|\beta\theta] = \varepsilon(\theta) \det A[\alpha|\beta]. \tag{1.8}$$

When $\alpha \in \Gamma_{m,n}$, $\beta \in \Gamma_{m,k}$, $A[\alpha|\beta] \in \mathbb{C}_{m\times m}$ whose (i,j) entry is $a_{\alpha(i)\beta(j)}$. However if $\alpha \in \Gamma_{m,n}$ and $\alpha \notin D_{m,n}$, then $A[\alpha|\beta]$ has two identical rows so that

$$\det A[\alpha|\beta] = 0, \quad \alpha \in \Gamma_{m,n} \setminus D_{m,n}. \tag{1.9}$$

Theorem 1.10.1. (Cauchy-Binet) Let $A \in \mathbb{C}_{r\times n}$, $B \in \mathbb{C}_{n\times l}$ and $C = AB$. Then for any $1 \le m \le \min\{r,n,l\}$ and $\alpha \in Q_{m,r}$ and $\beta \in Q_{m,l}$,

$$\det C[\alpha|\beta] = \sum_{\omega \in Q_{m,n}} \det A[\alpha|\omega] \det B[\omega|\beta].$$

Proof.

$$\begin{aligned}
\det C[\alpha|\beta] &= \sum_{\sigma\in S_m} \varepsilon(\sigma) \prod_{i=1}^m c_{\alpha(i),\beta\sigma(i)} \\
&= \sum_{\sigma\in S_m} \varepsilon(\sigma) \prod_{i=1}^m \sum_{j=1}^n a_{\alpha(i),j} b_{j,\beta\sigma(i)} \\
&= \sum_{\sigma\in S_m} \varepsilon(\sigma) \sum_{\tau\in\Gamma_{m,n}} \prod_{i=1}^m a_{\alpha(i),\gamma(i)} b_{\gamma(i),\beta\sigma(i)} \\
&= \sum_{\tau\in\Gamma_{m,n}} \prod_{i=1}^m a_{\alpha(i),\gamma(i)} \sum_{\sigma\in S_m} \varepsilon(\sigma) b_{\gamma(i),\beta\sigma(i)} \\
&= \sum_{\tau\in\Gamma_{m,n}} \sum_{i=1}^m a_{\alpha(i),\gamma(i)} \det B[\gamma|\beta] \\
&= \sum_{\omega\in D_{m,n}} \prod_{i=1}^m a_{\alpha(i),\gamma(i)} \det B[\gamma|\beta] \\
&= \sum_{\omega\in Q_{m,n}} \sum_{\sigma\in S_m} \prod_{i=1}^m a_{\alpha(i),\omega\sigma(i)} \det B[\omega\sigma|\beta] \\
&= \sum_{\omega\in Q_{m,n}} \Big(\sum_{\sigma\in S_m} \varepsilon(\sigma) \prod_{i=1}^m a_{\alpha(i),\omega\sigma(i)} \Big) \det B[\omega|\beta] \\
&= \sum_{\omega\in Q_{m,n}} \det A[\alpha|\omega] \det B[\omega|\beta].
\end{aligned}$$

□

Theorem 1.10.2. (Laplace) Let $A \in \mathbb{C}_{n\times n}$, $1 \le m \le n$ and $\alpha \in Q_{m,n}$. Then

$$\det A = \sum_{\omega\in Q_{m,n}} \det A[\alpha|\omega] (-1)^{s(\alpha)+s(\omega)} \det A(\alpha|\omega).$$

Proof. The right side of the formula is

$$\begin{aligned}
& (-1)^{s(\alpha)} \sum_{\omega \in Q_{m,n}} (-1)^{s(\omega)} \det A[\alpha|\omega] \det A[\alpha'|\omega'] \\
= & (-1)^{s(\alpha)} \sum_{\omega \in Q_{m,n}} \sum_{\theta \in S_m} \sum_{\pi \in S_{n-m}} \varepsilon(\theta)\varepsilon(\pi)(-1)^{s(\omega)} \prod_{i=1}^{m} a_{\alpha(i)\omega\theta(i)} \prod_{j=1}^{n-m} a_{\alpha'(j)\omega'\pi(j)} \\
= & (-1)^{s(\alpha)} \sum_{\sigma \in S_n} \varepsilon(\sigma)(-1)^{m(m+1)/2} \prod_{i=1}^{m} a_{\alpha(i)\sigma(i)} \prod_{i=m+1}^{n} a_{\alpha'(i-m)\sigma(i)} \\
= & \sum_{\sigma \in S_n} \varepsilon(\pi)\varepsilon(\sigma) \prod_{i=1}^{n} a_{\pi(i)\sigma(i)} \\
= & \det A.
\end{aligned}$$

□

Theorem 1.10.3. Let $U \in \mathrm{U}_n(\mathbb{C})$. Then for $\alpha, \beta \in Q_{m,n}$ with $1 \le m \le n$,

$$\det U \det \overline{U[\alpha|\beta]} = (-1)^{s(\alpha)+s(\beta)} \det U(\alpha|\beta). \tag{1.10}$$

In particular

$$|\det U[\alpha|\beta]| = |\det U(\alpha|\beta)|. \tag{1.11}$$

Proof. Because $\det I[\beta|\omega] = \delta_{\beta\omega}$ and $U^*[\beta|\gamma] = \overline{U[\gamma|\beta]^T}$, apply Cauchy-Binet theorem on $I_n = U^*U$ to yield

$$\delta_{\beta\omega} = \sum_{\tau \in Q_{m,n}} \det \overline{U[\gamma|\beta]} \det U[\gamma|\omega]. \tag{1.12}$$

Using Laplace theorem, we have

$$\delta_{\alpha\gamma} \det U = \sum_{\omega \in Q_{m,n}} \det U[\gamma|\omega](-1)^{s(\alpha)+s(\omega)} \det U(\alpha|\omega).$$

Multiplying both sides by $\overline{U[\gamma|\beta]}$ and sum over $\tau \in S_m$, the left side becomes

$$\sum_{\gamma \in Q_{m,n}} \delta_{\alpha\gamma} \det U \det \overline{U[\gamma|\beta]} = \det U \det \overline{U[\alpha|\beta]}$$

and the right side becomes

$$\begin{aligned}
& \sum_{\gamma \in Q_{m,n}} \sum_{\omega \in Q_{m,n}} \det \overline{U[\gamma|\beta]} \det U[\gamma|\omega](-1)^{s(\alpha)+s(\omega)} \det U(\alpha|\omega) \\
= & \sum_{\omega \in Q_{m,n}} \delta_{\beta\omega}(-1)^{s(\alpha)+s(\omega)} \det U(\alpha|\omega) \\
= & (-1)^{s(\alpha)+s(\beta)} \det U(\alpha|\beta)
\end{aligned}$$

Since $|\det U| = 1$, we have the desired result. □

The **permanent** function of A is

$$\operatorname{per} A = \sum_{\sigma \in S_n} \prod_{i=1}^{n} a_{i\sigma(i)}$$

which is also known as "positive determinant".

Problems

1. Prove equation (1.2).
2. Prove that $\prod_{i=1}^m \sum_{j=1}^n a_{ij} = \sum_{\gamma \in \Gamma_{m,n}} \prod_{i=1}^m a_{i\gamma(i)}$.
3. Prove a general Laplace expansion theorem: Let $A \in \mathbb{C}_{n\times n}$, $\alpha, \beta \in Q_{m,n}$, and $1 \le m \le n$.

$$\delta_{\alpha\beta} \det A = \sum_{\omega \in Q_{m,n}} \det A[\alpha|\omega](-1)^{s(\alpha)+s(\omega)} \det A(\beta|\omega).$$

4. Show that

$$\begin{aligned} |\Gamma(n_1, \ldots, n_m)| &= \prod_{i=1}^m n_i, \quad |\Gamma_{m,n}| = n^m, \quad |G_{m,n}| = \binom{n+m-1}{m}, \\ |D_{m,n}| &= \binom{n}{m} m!, \quad |Q_{m,n}| = \binom{n}{m}. \end{aligned}$$

Solutions to Problems 1.10

1.
2. (Roy) If $A = (a_{ij}) \in \mathbb{C}_{m\times n}$, the left side of $\prod_{i=1}^m \sum_{j=1}^n a_{ij} = \sum_{\gamma \in \Gamma_{m,n}} \prod_{i=1}^m a_{i\gamma(i)}$ is the product of row sums of A. Fix n and use induction on m. If $m = 1$, then both sides are equal to the sum of all entries of A. Suppose that the identity holds for $m-1$. We have

$$\begin{aligned} \sum_{\gamma \in \Gamma_{m,n}} \prod_{i=1}^m a_{i\gamma(i)} &= \left(\sum_{\gamma \in \Gamma_{m-1,n}} \prod_{i=1}^{m-1} a_{i\gamma(i)} \right) \left(\sum_{j=1}^n a_{mj} \right) \\ &= \left(\prod_{i=1}^{m-1} \sum_{j=1}^n a_{ij} \right) \left(\sum_{j=1}^n a_{mj} \right) \qquad \text{by induction hypothesis} \\ &= \prod_{i=1}^m \sum_{j=1}^n a_{ij} \end{aligned}$$

3.

4. The ith slot of each sequence in $\Gamma(n_1, \ldots, n_m)$ has n_i choices. So $|\Gamma(n_1, \ldots, n_m)| = \prod_{i=1}^m n_i$ and thus $|\Gamma_{m,n}| = n^m$ follows from $|\Gamma(n_1, \ldots, n_m)| = \prod_{i=1}^m n_i$ with $n_i = n$ for all i.

   $|G_{m,n}| = \binom{n+m-1}{m}$. Now $|Q_{m,n}|$ is the number of picking m distinct numbers from $1, \ldots, n$. So $|Q_{m,n}| = \binom{n}{m}$ is clear and $|D_{m,n}| = |Q_{m,n}|m! = \binom{n}{m}m!$

Chapter 2

Group representation theory

2.1 Symmetric groups

Let S_n be the symmetric group on $\{1, \ldots, n\}$. A matrix $P \in \mathbb{C}_{n\times n}$ is said to be a permutation matrix if

$$p_{i,\sigma(j)} = \delta_{i,\sigma(j)}, \quad i, j = 1, \ldots, n$$

where $\sigma \in S_n$ and we denote by $P(\sigma)$ for P because of its association with σ. Denote by P_n the set of all permutation matrices in $\mathbb{C}_{n\times n}$. Notice that $\varphi : S_n \to P_n$ $(\delta \mapsto P(\sigma))$ is an isomorphism.

Theorem 2.1.1. (Cayley) Each finite group G with n elements is isomorphic to a subgroup of S_n.

Proof. Let $\mathrm{Sym}\,(G)$ denotes the group of all bijections of G. For each $\sigma \in G$, define the left translation $l_\sigma : G \to G$ by $l_\sigma(x) = \sigma x$ for all $x \in G$. It is easy to show that $l_{\sigma\tau} = l_\sigma l_\tau$ for all $\sigma, \tau \in G$ so that $l : G \to \mathrm{Sym}\,(G)$ is a group homomorphism. The map l is injective since $l_\sigma = l_\tau$ implies that $\sigma x = \tau x$ for any $x \in G$ and thus $\sigma = \tau$. So G is isomorphic to some subgroup of $\mathrm{Sym}\,(G)$ and thus to some subgroup of S_n. □

The following is another proof in matrix terms. Let $G = \{\sigma_1, \ldots, \sigma_n\}$. Define the regular representation $Q : G \to \mathrm{GL}_n(\mathbb{C})$:

$$Q(\sigma)_{ij} := (\delta_{\sigma_i,\sigma\sigma_j}) \in \mathrm{GL}_n(\mathbb{C}).$$

It is easy to see that $Q(\sigma)$ is a permutation matrix and $Q : \sigma \mapsto Q(\sigma)$ is injective. So it suffices to show that Q is a homomorphism. Now

$$(Q(\sigma)Q(\pi))_{ij} = \sum_{k=1}^{n} Q(\sigma)_{ik}Q(\pi)_{kj} = \sum_{k=1}^{n} \delta_{\sigma_i,\sigma\sigma_k}\delta_{\sigma_k,\pi\sigma_j} = \delta_{\sigma_i,\sigma\pi\sigma_j} = Q(\sigma\pi)_{ij}.$$

So we view each finite group as a subgroup of S_n for some n.

The element $\sigma \in S_n$ is called a **cycle** of length k ($1 \le k \le n$), if there exist $1 \le i_1, \ldots, i_k \le n$ such that

$$\begin{aligned} \sigma(i_t) &= i_{t+1}, \quad t = 1, \ldots, k-1 \\ \sigma(i_k) &= i_1 \\ \sigma(i) &= i, \quad i \notin \{i_1, \ldots, i+k\} \end{aligned}$$

and we write $\sigma = (i_1, \ldots, i_k)$. Two cycles $(i_1, \ldots, i_k)$ and $(j_1, \ldots, j_m)$ are said to be **disjoint** if there is no intersection between the two set $\{i_1, \ldots, i_k\}$ and $\{j_1, \ldots, j_m\}$. From the definition, we may express the cycle σ as

$$(i_1, \sigma(i_1), \ldots, \sigma^{k-1}(i_1)).$$

For any $\sigma \in S_n$ there is $1 \le k \le n$ such that $\sigma^k(i) = i$ and let k be the smallest such integer. Then $(i\ \sigma(i)\ \cdots\ \sigma^{k-1}(i))$ is a cycle but is not necessarily equal to σ. The following theorem asserts that σ can be reconstructed from these cycles.

Theorem 2.1.2. Each element $\sigma \in S_n$ can be written as a product of disjoint cycles. The decomposition is unique up to permutation of the cycles.

Proof. Let $i \le n$ be a positive integer. Then there is a smallest positive integer r such that $\sigma^r(i) = i$. If $r = n$, then the cycle $(i\ \sigma(i)\ \cdots\ \sigma^{r-1}(i))$ is σ. If $r < n$, then there is positive integer $j \le n$ such that $j \notin \{i, \sigma(i), \ldots, \sigma^{r-1}(i)\}$. Then there is a smallest positive integer s such that $\sigma^s(j) = j$. Clearly the two cycles $(i\ \sigma(i)\ \cdots\ \sigma^{r-1}(i))$ and $(j\ \sigma(j)\ \cdots\ \sigma^{s-1}(j))$ are disjoint. Continue the process we have

$$\sigma = (i\ \sigma(i)\ \cdots\ \sigma^{r-1}(i))(j\ \sigma(j)\ \cdots\ \sigma^{s-1}(j))\cdots(k\ \sigma(k)\ \cdots\ \sigma^{t-1}(k))$$

□

For example,

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 2 & 5 & 1 & 7 & 3 & 6 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 5 & 3 \end{pmatrix} \begin{pmatrix} 4 & 7 \end{pmatrix} \begin{pmatrix} 6 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 5 & 3 \end{pmatrix} \begin{pmatrix} 4 & 7 \end{pmatrix}.$$

So $c(\sigma) = 3$ (which include the cycle of length 1)

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 5 & 4 & 3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 5 \end{pmatrix} \begin{pmatrix} 3 & 4 \end{pmatrix} = \begin{pmatrix} 3 & 4 \end{pmatrix} \begin{pmatrix} 1 & 2 & 5 \end{pmatrix} = \begin{pmatrix} 3 & 4 \end{pmatrix} \begin{pmatrix} 5 & 1 & 2 \end{pmatrix}.$$

So $c(\sigma) = 2$.

Cycles of length 2, i.e., $(i\,j)$, are called **transpositions**.

Theorem 2.1.3. Each element $\sigma \in S_n$ can be written (not unique) as a product of transposition.

Proof. Use Theorem 2.1.2

$$(i_1 \cdots i_k) = (i_1 i_2)(i_2 i_3) \cdots (i_{k-1} i_k) = (i_1 i_k)(i_1 i_{k-1}) \cdots (i_1 i_2). \tag{2.1}$$

□

Though the decomposition is not unique, the parity (even number or odd number) of transpositions in each $\sigma \in S_n$ is unique (see Merris p.56-57).

Problems

1. Prove that each element of S_n is a product of transpositions. (Hint: $(i_1\ i_2 \cdots i_k) = (i_1 i_2)(i_2 i_3) \cdots (i_{k-1} i_k) = (i_1 i_k)(i_1 i_{k-1}) \cdots (i_1 i_2)$).
2. Show that S_n is generated by the transposition $(12), (23), \ldots, (n-1, n)$ (Hint: express each tranposition as a product of $(12), (23), \ldots, (n-1, n)$).
3. Express the following permutations as a product of nonintersecting cycles:
   (a) $(a\, b)(a\, i_1 \ldots i_r\, b\, j_1 \ldots j_r)$.
   (b) $(a\, b)(a\, i_1 \ldots i_r\, b\, j_1 \ldots j_r)(a\, b)$.
   (c) $(a\, b)(a\, i_1 \ldots i_r)$.
   (d) $(a\, b)(a\, i_1 \ldots i_r)(a\, b)$.
4. Express $\sigma = (24)(1234)(34)(356)(56) \in S_7$ as a product of disjoint cycles and find $c(\sigma)$.

Solutions to Problems 2.1

1. Use Theorem 2.1.2 and the hint.
2. $(i, i+2) = (i+1, i+2)(i, i+1)(i+1, i+2)$ and each (ij) is obtained by conjugating $(i, i+1)$ by the product $(j, j+1) \cdots (i+1, i+2)$ where $i < j$. So S_n is generated by $(12), (23), \ldots, (n-1, n)$.

   (Zach) $(ij) = (1i)(1j)(1i)$. By Problem 1, every cycle is a product of transpositions. So S_n is generated by $(12), (13), \ldots, (1n)$. Now $(1k) = (1, k-1)(k-1, k)(1, k-1)$ for all $k > 1$.
3. (a) $(a\, b)(a\, i_1 \ldots i_r\, b\, j_1 \ldots j_r) = (a\, i_1 \ldots i_r)(b\, j_1 \ldots j_s)$.
   (b) $(a\, b)(a\, i_1 \ldots i_r\, b\, j_1 \ldots j_s)(a\, b) = (a\, i_1 \ldots i_r)(b\, j_1 \ldots j_s)(a\, b) = (a\, j_1 \ldots j_s\, b\, i_1 \ldots i_r)$.
   (c) $(a\, b)(a\, i_1 \ldots i_r) = (a\, i_1 \ldots i_r\, b)$.
   (d) $(a\, b)(a\, i_1 \ldots i_r)(a\, b) = (b\, i_1 \ldots i_r)$.
4. $\sigma = (24)(1234)(34)(356)(56) = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 4 & 3 & 5 & 2 & 1 & 6 & 7 \end{pmatrix} = (14235)(6)(7)$.

   So $c(\sigma) = 3$.

2.2 Representation theory

Let V be an n-dimensional complex vector space and $\mathrm{GL}(V)$ be the group of invertible operators $T \in \mathrm{End}\, V$. Clearly the group of $n \times n$ complex invertible matrices $\mathrm{GL}_n(\mathbb{C})$ is isomorphic to $\mathrm{GL}(V)$. Let G be a finite group. A (linear) **representation** of G of degree n is a homomorphism $T : G \to \mathrm{GL}(V)$, i.e.,

$$T(\sigma_1\sigma_2) = T(\sigma_1)T(\sigma_2), \quad \sigma_1, \sigma_2 \in G.$$

Then $T(e) = I$ where $e \in G$ is the identity element (reason: $T(e) = T(e^2) = T(e)T(e)$) and $T(\sigma^{-1}) = (T(\sigma))^{-1}$. The dimension of V is called the **degree** of the representation T. The representation T is called **faithful** (to the group) if it is injective, i.e., G is isomorphic to $T(G) < \mathrm{GL}(V)$.

Similarly we have the **matrix representation** $A : G \to \mathrm{GL}_n(\mathbb{C})$, i.e.,

$$A(\sigma_1\sigma_2) = A(\sigma_1)A(\sigma_2), \quad \sigma_1, \sigma_2 \in G.$$

So $A(e) = A$ and $A(\sigma^{-1}) = (A(\sigma))^{-1}$. Clearly if $P \in \mathrm{GL}_n(\mathbb{C})$, then $B(\sigma) := P^{-1}A(\sigma)P$, $\sigma \in G$, yields a matrix representation $B : G \to \mathrm{GL}_n(\mathbb{C})$. We call the representations A and B **isomorphic**.

Let E be a basis of V and let $T : G \to \mathrm{GL}(V)$ be a representation. Let $A(\sigma) := [T(\sigma)]_E^E \in \mathrm{GL}_n(\mathbb{C})$ where $n = \dim V$. Then $A : G \to \mathrm{GL}_n(\mathbb{C})$ is a matrix representation since

$$A(\sigma_1\sigma_2) = [T(\sigma_1\sigma_2)]_E^E = [T(\sigma_1)T(\sigma_2)]_E^E = [T(\sigma_1)]_E^E[T(\sigma_2)]_E^E = A(\sigma_1)A(\sigma_2).$$

Under another basis F, we have a matrix representation $B(\sigma) := [T(\sigma)]_F^F$. Then

$$B(\sigma) = [T(\sigma)]_F^F = [I]_E^F[T(\sigma)]_E^E[I]_F^E = P^{-1}A(\sigma)P, \quad \sigma \in G,$$

where $P := [I]_F^E$, i.e., A and B are isomorphic. In other words, the matrix representations with respect to different bases of a linear representation are isomorphic. Similarly the representations $S : G \to \mathrm{GL}(V)$ and $T : G :\to \mathrm{GL}(W)$ are said to be **isomorphic** if there is an invertible $L \in \mathrm{Hom}\,(V, W)$ such that $T(\sigma) = L^{-1}S(\sigma)L$ for all $\sigma \in G$. In other words, isomorphic representations have the same matrix representation by choosing appropriate bases accordingly.

Given any representation $T : G \to \mathrm{GL}(V)$ with $\dim V = n$, a choice of basis in V identifies with $\mathbb{C}^n$ (i.e., isomorphic to via the coordinate vector) and call this isomorphism L. Thus we have a "new" representation $S : G \to \mathrm{GL}_n(\mathbb{C})$ which is isomorphic to T. So any representation of G of degree n is isomorphic to a matrix representation on $\mathbb{C}^n$. The vector space representation does not lead to "new" representation, but it is a basis free approach; while at the same time the matrix representation may give us some concrete computation edge. So we often use both whenever it is convenient to us.

The **basic problem** of representation theory is to classify all representations of G, up to isomorphism. We will see that character theory is a key player.

Example 2.2.1. (a) The **principal representation**: $\sigma \mapsto 1$ for all $\sigma \in G$, is of degree 1.

(b) The **alternating representation** on S_n, $\sigma \mapsto \varepsilon(\sigma)$, is of degree 1.

(c) $\sigma \mapsto I_n$ is a representation of degree n.

(d) The **regular representation** $Q : G \to \mathrm{GL}_n(\mathbb{C})$, where $n = |G|$, is given by

$$Q(\sigma)_{ij} = \delta_{\sigma_i, \sigma\sigma_j}.$$

Here is another version: let V be an n-dimensional complex vector space with a (ordered) basis indexed by G, i.e., $E = \{f_{\sigma_1}, \ldots, f_{\sigma_n}\}$ where $G = \{\sigma_1, \ldots, \sigma_n\}$. For each $\sigma \in G$, define $T(\sigma)$ by $T(\sigma)f_\pi = f_{\sigma\pi}$. So the images $f_\sigma = T(\sigma)f_e$, where $\sigma \in G$, form a basis of V. Then $Q = [T(\sigma)]_F^F$ since the jth column of Q is e_i where $\sigma\sigma_j = \sigma_i$, i.e., $Q(\sigma)_{ij} = \delta_{\sigma_i, \sigma\sigma_j}$.

(e) S_3 has 6 elements: $e, (23), (12), (123), (132), (13)$. Define $T : S_3 \to \mathrm{GL}_2(\mathbb{C})$ by setting their images accordingly as

$$I_2, \begin{pmatrix} 1 & -1 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix}, \begin{pmatrix} -1 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ -1 & 1 \end{pmatrix}$$

which is a representation of degree 2.

The representation $A : G \to \mathrm{GL}_n(\mathbb{C})$ is said to be **reducible** if there is $P \in \mathrm{GL}_n(\mathbb{C})$ such that

$$P^{-1}A(\sigma)P = \begin{pmatrix} A_1(\sigma) & 0 \\ C(\sigma) & A_2(\sigma) \end{pmatrix}, \quad \text{for all } \sigma \in G, \tag{2.2}$$

where $A_1(\sigma) \in \mathrm{GL}_m(\mathbb{C})$ and $A_2(\sigma) \in \mathrm{GL}_{n-m}(\mathbb{C})$ with $1 \le m < n$. In addition if $C(\sigma) = 0$ for all $\sigma \in G$, then A is said to be **completely reducible**. Moreover, A is said to be **irreducible** if A is not reducible. It is easy to see that A is irreducible if and only if any representation isomorphic to A is irreducible. Similar notions are defined for representation $T : G \to \mathrm{GL}(V)$, for example, T is reducible if there is a basis E for V such that $[T(\sigma)]_E^E$ is of the form (2.2), etc.

Notice that if $A : G \to \mathrm{GL}_n(\mathbb{C})$ is reducible, i.e., (2.2) holds, then

$$\begin{aligned} \begin{pmatrix} A_1(\sigma\pi) & 0 \\ C(\sigma\pi) & A_2(\sigma\pi) \end{pmatrix} &= \begin{pmatrix} A_1(\sigma) & 0 \\ C(\sigma) & A_2(\sigma) \end{pmatrix} \begin{pmatrix} A_1(\pi) & 0 \\ C(\pi) & A_2(\pi) \end{pmatrix} \\ &= \begin{pmatrix} A_1(\sigma)A_1(\pi) & 0 \\ C(\sigma)A_1(\pi) + A_2(\sigma)C(\pi) & A_2(\sigma)A_2(\pi) \end{pmatrix}. \end{aligned} \tag{2.3}$$

So $A_1(\sigma\pi) = A_1(\sigma)A_1(\pi)$ and $A_2(\sigma\pi) = A_2(\sigma)A_1(\pi)$. So A_1 and A_2 are lower degree representations of G, known as **sub-representations**. If A is completely reducible, it is determined by A_1 and A_2.

From definition all linear representations are irreducible. So Example 2.2.1(a), (b) are irreducible; (c) is reducible if and only if $n > 1$; (d) is reducible if $n > 1$ (we will see, Problem 1 when $n = 3$); (e) Problem 2.

Theorem 2.2.2. The representation $T : G \to \mathrm{GL}(V)$ is reducible if and only if there is a nontrivial subspace W invariant under $T(\sigma)$ for all $\sigma \in G$, i.e., $T(\sigma)W \subset W$ (indeed $T(\sigma)W = W$) for all $\sigma \in G$.

Proof. If T is reducible, then there is a basis $E = \{e_1, \ldots, e_m, e_{m+1}, \ldots, e_n\}$ such that

$$[T(\sigma)]_E^E = \begin{pmatrix} A_1(\sigma) & 0 \\ C(\sigma) & A_2(\sigma) \end{pmatrix}, \quad \text{for all } \sigma \in G,$$

where $A_1(\sigma) \in \mathrm{GL}_m(\mathbb{C})$ and $A_2(\sigma) \in \mathrm{GL}_{n-m}(\mathbb{C})$. Clearly $W := \langle e_{m+1}, \ldots, e_n \rangle$ is the desired invariant subspace.

Conversely if W is nontrivial and invariant under $T(\sigma)$ for all $\sigma \in G$, then extend the basis F of W to a basis E for V in which F is the latter part of E. Then $[T(\sigma)]_E^E$ is the desired reducible representation. □

So $T|_W : G \to \mathrm{GL}(W)$ is a representation of G and is called a sub-representation of T, defined by restriction:

$$T|_W(\sigma) = T(\sigma)|_W.$$

The following theorem asserts that reducibility and completely reducibility coincide for finite group representations.

Theorem 2.2.3. (Maschke) Let A be a matrix representation of the finite group G. If A is reducible, then A is completely reducible.

Proof. Since A is reducible, we may assume that

$$A(\sigma) = \begin{pmatrix} A_1(\sigma) & 0 \\ C(\sigma) & A_2(\sigma) \end{pmatrix}, \quad \text{for all } \sigma \in G,$$

where $A_1(\sigma) \in \mathrm{GL}_m(\mathbb{C})$ and $A_2(\sigma) \in \mathrm{GL}_{n-m}(\mathbb{C})$. Let

$$P := \begin{pmatrix} I_m & 0 \\ Q & I_{n-m} \end{pmatrix}, \quad \text{so that } P^{-1} = \begin{pmatrix} I_m & 0 \\ -Q & I_{n-m} \end{pmatrix},$$

where Q is to be determined. Then

$$\begin{aligned} P^{-1}A(\sigma)P &= \begin{pmatrix} I_m & 0 \\ -Q & I_{n-m} \end{pmatrix} \begin{pmatrix} A_1(\sigma) & 0 \\ C(\sigma) & A_2(\sigma) \end{pmatrix} \begin{pmatrix} I_m & 0 \\ Q & I_{n-m} \end{pmatrix} \\ &= \begin{pmatrix} A_1(\sigma) & 0 \\ C(\sigma) - QA_1(\sigma) + A_2(\sigma)Q & A_2(\sigma) \end{pmatrix}. \end{aligned} \tag{2.4}$$

So we want to find Q such that $C(\sigma) - QA_1(\sigma) + A_2(\sigma)Q = 0$ for all $\sigma \in G$. From (2.3)

$$C(\sigma\pi) = C(\sigma)A_1(\pi) + A_2(\sigma)C(\pi)$$

so that

$$\begin{aligned} C(\sigma) &= C(\sigma\pi)A_1(\pi)^{-1} - A_2(\sigma)C(\pi)A_1(\pi)^{-1} \\ &= C(\sigma\pi)A_1(\sigma\pi)^{-1}A_1(\sigma) - A_2(\sigma)C(\pi)A_1(\pi)^{-1}. \end{aligned}$$

Sum over π to have

$$\begin{aligned} |G|C(\sigma) &= (\sum_{\pi\in G} C(\sigma\pi)A_1(\sigma\pi)^{-1})A_1(\sigma) - A_2(\sigma)(\sum_{\pi\in G} C(\pi)A_1(\pi)^{-1}) \\ &= (\sum_{\pi\in G} C(\pi)A_1(\pi)^{-1})A_1(\sigma) - A_2(\sigma)(\sum_{\pi\in G} C(\pi)A_1(\pi)^{-1}) \end{aligned}$$

Set $Q := \frac{1}{|G|}(\sum_{\pi\in G} C(\pi)A_1(\pi)^{-1})$. Then

$$C(\sigma) - QA_1(\sigma) + A_2(\sigma)Q = 0, \quad \text{for all } \sigma \in G.$$

Substitute into (2.4) to have the desired result.

□

Theorem 2.2.4. Every matrix representation A of the finite group G is a direct sum of irreducible representations, i.e., there is $P \in \mathrm{GL}_n(\mathbb{C})$ such that $P^{-1}A(\sigma)P = A_1(\sigma) \oplus A_2(\sigma) \oplus \cdots \oplus A_k(\sigma)$ for all $\sigma \in G$, where $A_1, \ldots, A_k$ are irreducible representations.

Proof. Use Theorem 2.2.3 with induction. □

So the study of representations of G, up to isomorphism, is reduced to the study of irreducible representations.

A representation $T : G \to \mathrm{GL}(V)$ is unitary if V has been equipped with an inner product such that $T(\sigma)$ are unitary for all $\sigma \in G$. It is **unitarisable** if it can be equipped with such an inner product. For example, the regular representation $Q : G \to \mathrm{GL}_n(\mathbb{C})$ ($n = |G|$) of a finite group is unitary (pick the standard inner product).

Theorem 2.2.5. Every representation $T : G \to \mathrm{GL}(V)$ is unitarisable.

Proof. Problem 3 "unitary trick". □

Problems

1. Prove that the degree 3 representation of S_3: $\sigma \mapsto (\delta_{i\sigma(j)}) \in \mathrm{GL}_3(\mathbb{C})$ is reducible (Hint: find a 1-dimensional invariant subspace).

2. Prove that the degree 2 representation in Example 2.2.1(e) is irreducible.

3. Let $T : G \to \mathrm{GL}_n(\mathbb{C})$ be a representation of a finite group G. Prove that there is an inner product such that $T(\sigma)$ is unitary for all $\sigma \in G$, i.e., for any $u, v \in V$, $\sigma \in G$ we have $(T(\sigma)u, T(\sigma)v) = (u, v)$. (Hint: Let $\langle\cdot,\cdot\rangle$ be any arbitrary inner product. Then consider it "average" $(u, v) := \frac{1}{|G|}\sum_{\sigma\in G}\langle T(\sigma)u, T(\sigma)v\rangle)$.

4. Prove that every matrix representation of a finite group is isomorphic to a unitary matrix representation.

5. Write out the regular representation of the cyclic group $C_3 = \{1, \omega, \omega^2\}$ where ω is the cubic primitive root of unity. Is it reducible?

6. Prove Maschke Theorem using Theorem 2.2.5 (Hint: If W is an invariant subspace for all $A(\sigma)$, so is $W^\perp$).

Solutions to Problems 2.2

1. The representation T is given as

$$e \mapsto I_3, (12) \mapsto \begin{pmatrix} & 1 & \\ 1 & & \\ & & 1 \end{pmatrix}, (13) \mapsto \begin{pmatrix} & & 1 \\ & 1 & \\ 1 & & \end{pmatrix},$$

$$(23) \mapsto \begin{pmatrix} 1 & & \\ & & 1 \\ & 1 & \end{pmatrix}, (123) \mapsto \begin{pmatrix} & & 1 \\ 1 & & \\ & 1 & \end{pmatrix}, (132) \mapsto \begin{pmatrix} & 1 & \\ & & 1 \\ 1 & & \end{pmatrix}$$

The subspace $W := \langle e_1 + e_2 + e_3 \rangle$ is invariant under all permutations in $\mathrm{GL}_3(\mathbb{C})$. By Theorem 2.2.2 $T : S_3 \to \mathrm{GL}_3(\mathbb{C})$ is reducible.

Remark: Using the same argument regular representation is always reducible if $|G| > 1$ since the subspace spanned by $e_1 + \cdots + e_n$, where $n := |G|$, is a nontrivial invariant subspace.

2. If the representation were reducible, there would exist a nontrivial invariant subspace, i.e., 1-dimensional W, by Theorem 2.2.2. Let $W = \langle c_1 e_1 + c_2 e_2 \rangle$ and $c_1 e_1 + c_2 e_2 \neq 0$. Then

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} (c_1 e_1 + c_2 e_2) = c_2 e_1 + c_1 e_2 \in W$$

i.e., $c_2 = \lambda c_1$ and $c_1 = \lambda c_2$. Hence $c_2 = \lambda^2 c_2$ and $c_1 = \lambda^2 = c_1$. So $\lambda^2 = 1$, i.e., $c_2 = \pm c_1$. From

$$\begin{pmatrix} 1 & -1 \\ 0 & -1 \end{pmatrix} (c_1 e_1 + c_2 e_2) = (c_1 - c_2) e_1 - c_2 e_2 \in W,$$

$c_1 - c_2 = \gamma c_1$ and $c_2 = -\gamma c_2$. If $c_2 \neq 0$, then $\gamma = -1$ but then $2c_1 = c_2$, which is impossible.

(Roy) The matrices corresponding to (23), (12), (13) are

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

with eigenvalues $1, -1$. If the representation were reducible, i.e., completely reducible, then all the matrices are simultaneously similar to diagonal matrices, in particular the above three matrices. Two would be the same, contradiction.

3. (a) It is straight forward computation to verify that $(\cdot,\cdot)$ is an inner product.
$\overline{(u,v)} = \frac{1}{|G|}\sum_{\sigma\in G}\overline{\langle T(\sigma)u, T(\sigma)v\rangle} = \frac{1}{|G|}\sum_{\sigma\in G}\langle T(\sigma)v, T(\sigma)u\rangle = (v,u)$.

(b)
$$\begin{aligned}
& (\alpha_1 u_1 + \alpha_2 u_2, v) \\
= & \frac{1}{|G|}\sum_{\sigma\in G}\langle T(\sigma)(\alpha_1 u_1 + \alpha_2 u_2), T(\sigma)v\rangle \\
= & \frac{1}{|G|}\sum_{\sigma\in G}\langle \alpha_1 T(\sigma)u_1 + \alpha_2 T(\sigma)u_2 T(\sigma)v\rangle \\
= & \alpha_1\frac{1}{|G|}\sum_{\sigma\in G}\langle T(\sigma)u_1, T(\sigma)v\rangle + \alpha_2\frac{1}{|G|}\sum_{\sigma\in G}\langle T(\sigma)u_2, T(\sigma)v\rangle \\
= & \alpha_1(u_1,v) + \alpha_2(u_2,v)
\end{aligned}$$

(c) $(v,v) = \frac{1}{|G|}\sum_{\sigma\in G}\langle T(\sigma)v, T(\sigma)v\rangle \geq 0$ since each summand is non-negative. Moreover $(v,v) = 0$ if and only if $T(\sigma)v = 0$ for all σ. Since $T(\sigma) \in \mathrm{GL}(V)$, $v = 0$.

For all $u, v \in V$

$$\begin{aligned}
(T(\pi)u, T(\pi)v) &:= \frac{1}{|G|}\sum_{\sigma\in G}\langle T(\sigma)T(\pi)u, T(\sigma)T(\pi)v\rangle \\
&= \frac{1}{|G|}\sum_{\sigma\in G}\langle T(\sigma\pi)u, T(\sigma\pi)v\rangle \\
&= \frac{1}{|G|}\sum_{\sigma\in G}\langle T(\sigma)u, T(\sigma)v\rangle = (u,v)
\end{aligned}$$

4. From Problem 3.

5. $e \mapsto I_3$, $\omega \mapsto \begin{pmatrix} & & 1\\ 1 & & \\ & 1 & \end{pmatrix}$, $\omega^2 \mapsto \begin{pmatrix} & 1 & \\ & & 1\\ 1 & & \end{pmatrix}$. It is reducible using the argument of Problem 1 (or observe that C_3 is abelian so all irreducible representations must be of degree 1 from Problem 3.1).

6. With respect to some inner product, each $A(\sigma)$ is unitary by Theorem 2.2.5, if $A : G \to \mathrm{GL}_n(V)$ is reducible, i.e., there is a nontrivial invariant subspace W. For each $v \in W^\perp$, we have $(A(\sigma)v, A(\sigma)w) = (v,w) = 0$ for all $w \in W$. But $A(\sigma)W = W$ so that $(A(\sigma)v, W) = \{0\}$, i.e., $A(\sigma)W^\perp \subset W^\perp$ for all $\sigma \in A$. So $W^\perp$ is invariant under $A(\sigma)$ for all $\sigma \in G$.

2.3 Irreducible representations

Theorem 2.3.1. (Schur's lemma) Let $T : G \to \mathrm{GL}(V)$ and $S : G \to \mathrm{GL}(W)$ be two irreducible representations of G. If there is nonzero $L \in \mathrm{Hom}\,(V, W)$ such that $L \circ T(\sigma) = S(\sigma) \circ L$ for all $\sigma \in G$, then $\dim V = \dim W$ and L is invertible; thus S and T are isomorphic.

Proof. Recall

$$\operatorname{Im} L = \{Lv : v \in V\} \subset W, \quad \operatorname{Ker} L = \{v \in V : Lv = 0\} \subset V.$$

From $L \neq 0$ we have $\operatorname{Im} L \neq 0$ and $\operatorname{Ker} L \neq V$. For any $v \in V$, from $S(\sigma)Lv = LT(\sigma)v \in \operatorname{Im} L$, $\operatorname{Im} L$ is $S(\sigma)$-invariant subspace of W for all $\sigma \in G$. By Theorem 2.2.2 $\operatorname{Im} L$ must be trivial, i.e., $\operatorname{Im} L = W$. Similarly $\operatorname{Ker} L$ is $T(\sigma)$-invariant subspace of V so that $\operatorname{Ker} L = 0$. Hence L is invertible; thus S and T are isomorphic. □

The following is simply a matrix version of Theorem 2.3.1.

Corollary 2.3.2. Let $A : G \to \mathrm{GL}_n(\mathbb{C})$ and $B : G \to \mathrm{GL}_m(\mathbb{C})$ be two irreducible representations of G. If there is nonzero $M \in \mathbb{C}_{m\times n}$ such that $MA(\sigma) = B(\sigma)M$ for all $\sigma \in G$, then $n = m$ and M is invertible; thus A and B are isomorphic.

The special case $A = B$ in Corollary 2.3.2 provides more information on M.

Corollary 2.3.3. Let $A : G \to \mathrm{GL}_n(\mathbb{C})$ be an irreducible representation of G. If there is $M \in \mathbb{C}_{n\times n}$ such that $MA(\sigma) = A(\sigma)M$ for all $\sigma \in G$, then $M = cI_n$.

Proof. Let c be an eigenvalue of M. So $M - cI_n$ is not invertible. Clearly $(M - cI_n)A(\sigma) = A(\sigma)(M - cI_n)$ for all $\sigma \in G$. Then apply Corollary 2.3.2. □

There are important orthogonal relations among the entries of an irreducible matrix representation and between matrix representations.

Theorem 2.3.4. Let $A(\sigma) = (a_{ij}(\sigma)) \in \mathrm{GL}_n(\mathbb{C})$ be a degree n irreducible matrix representation of G. Then

$$\sum_{\sigma\in G} a_{is}(\sigma^{-1})a_{tj}(\sigma) = \frac{|G|}{n}\delta_{ij}\delta_{st}. \tag{2.5}$$

If $B(\sigma) = (b_{ij}(\sigma))$ is an irreducible degree m matrix representation of G and not isomorphic to A, then

$$\sum_{\sigma\in G} a_{is}(\sigma^{-1})b_{tj}(\sigma) = 0. \tag{2.6}$$

Proof. For any $M \in \mathbb{C}_{n\times m}$, let

$$\varphi(M) := \sum_{\sigma\in G} A(\sigma^{-1})MB(\sigma).$$

For any $\pi \in G$,

$$A(\pi)\varphi(M) = \sum_{\sigma\in G} A(\pi\sigma^{-1})MB(\sigma) = \sum_{\tau\in G} A(\tau^{-1})MB(\tau\pi) = \varphi(M)B(\pi).$$

If A and B are not isomorphic, then from Corollary 2.3.2, $\varphi(M) = 0$; if $A = B$, then from Corollary 2.3.3 we have $\varphi(M) = c_M I_n$.

Pick $M = E_{st}$ which is the matrix with 1 as its (s, t)-entry and zeros elsewhere. Computing the (i, j)-entry of $\varphi(E_{st})$:

$$\begin{aligned}
\varphi(E_{st})_{ij} &= \sum_{\sigma\in G} (A(\sigma^{-1})E_{st}B(\sigma))_{ij} \\
&= \sum_{\sigma\in G}\sum_{p=1}^{n}\sum_{q=1}^{m} A(\sigma^{-1})_{ip}(E_{st})_{pq}B(\sigma)_{qj} \\
&= \sum_{\sigma\in G}\sum_{p=1}^{n}\sum_{q=1}^{m} a_{ip}(\sigma^{-1})\delta_{sp}\delta_{tq}b_{qj}(\sigma) \\
&= \sum_{\sigma\in G} a_{is}(\sigma^{-1})b_{tj}(\sigma).
\end{aligned}$$

If A and B are not isomorphic, then $\varphi(E_{st}) = 0$ so that we have (2.6). If $A = B$, then $\varphi(E_{st}) = c_{st}I_m$ (c_{st} are to be determined) so that $\varphi(E_{st})_{ij} = c_{st}\delta_{ij}$. Then

$$c_{st}\delta_{ij} = \sum_{\sigma\in G} a_{is}(\sigma^{-1})a_{tj}(\sigma). \tag{2.7}$$

Set $i = j$ in (2.7) and sum over i to have

$$nc_{st} = \sum_{\sigma\in G}\sum_{i=1}^{n} a_{ti}(\sigma)a_{is}(\sigma^{-1}) = \sum_{\sigma\in G}(A(\sigma)A(\sigma^{-1}))_{ts} = \sum_{\sigma\in G}(I_n)_{ts} = |G|\delta_{st}.$$

So $c_{st} = \frac{|G|}{n}\delta_{st}$ and substitute it into (2.5) to yield the desired result.

□

Theorem 2.3.5. Let $A(\sigma) = (a_{ij}(\sigma))$ and $B(\sigma) = (b_{ij}(\sigma))$ be two irreducible matrix representations of G of degree n and m respectively. Then for any $\pi \in G$,

$$\sum_{\sigma\in G} a_{is}(\sigma^{-1})b_{tj}(\sigma\pi) = \begin{cases} \frac{|G|}{n}a_{ij}(\pi)\delta_{st} & \text{if } A = B \\ 0 & \text{if } A \text{ and } B \text{ not isomorphic} \end{cases}$$

Proof. Multiply both sides of (2.5) by $a_{jk}(\pi)$ and sum over j:

$$\sum_{\sigma\in G} a_{is}(\sigma^{-1})\sum_{j=1}^{n} a_{tj}(\sigma)a_{jk}(\pi) = \frac{|G|}{n}\delta_{st}\sum_{j=1}^{n}\delta_{ij}a_{jk}(\pi) = \frac{|G|}{n}a_{ik}(\pi)\delta_{st},$$

i.e., $\sum_{\sigma\in G} a_{is}(\sigma^{-1})a_{tk}(\sigma\pi) = \frac{|G|}{n}a_{ik}(\pi)\delta_{st}$.

Similarly multiply both sides of (2.6) by $b_{jk}(\pi)$ and sum over j to have the second part of the theorem. □

Each entry of $A(\sigma)$ can be viewed as a function $G \to \mathbb{C}$. The totality of $A(\sigma)$ as σ runs over G is viewed as a $|G|$-dimensional complex vector.

Theorem 2.3.6. Let $A(\sigma) = (a_{ij}(\sigma))$ be an irreducible degree n matrix representation of G. Then the n^2 functions $a_{ij} : G \to \mathbb{C}$ are linearly independent.

Proof. Let $c_{ij} \in \mathbb{C}$ such that

$$\sum_{i,j}^{n} c_{ij} a_{ij}(\sigma) = 0, \quad \text{for all } \sigma \in G.$$

Multiply both side by $a_{pq}(\sigma^{-1})$ and sum over σ:

$$0 = \sum_{i,j}^{n} c_{ij} \sum_{\sigma \in G} a_{pq}(\sigma^{-1}) a_{ij}(\sigma) = \sum_{i,j}^{n} c_{ij} \delta_{pj} \delta_{qi} \frac{|G|}{n} = \frac{|G|}{n} c_{pq}.$$

□

Problems

1. Prove that the irreducible representations of G are of degree 1 if and only if G is abelian.

2. Let $A : G \to \mathrm{GL}_n(\mathbb{C})$ be an irreducible representation of a finite group G. Prove that $n^2 \le |G|$.

3. Let $A_1(\sigma) = (a^1_{ij}(\sigma)), \ldots, A_r(\sigma) = (a^r_{ij}(\sigma))$ be non-isomorphic representations of G with degrees $n_1, \ldots, n_r$ respectively. Prove that there are linearly independent $\sum_{t=1}^r n_t^2$ functions $a^t_{ij} : G \to \mathbb{C}$ such that $\sum_{t=1}^r n_t^2 \le |G|$.

Solutions to Problems 2.3

1. Suppose that G is abelian and let $A : G \to \mathrm{GL}_n(\mathbb{C})$ be irreducible. By Corollary 2.3.3, since $A(\sigma)A(\pi) = A(\pi)A(\sigma)$ for all π we have $A(\sigma) = c_\sigma I_n$ for all $\sigma \in G$; so $n = 1$ because A is irreducible.

   (Daniel): Indeed $\{A(\sigma) : \sigma \in G\}$ is a commuting family of matrices, they are simultaneously diagonalizable. So $n = 1$.

   (Daniel) Conversely suppose that all irreducible representations are of degree 1. The regular representation of G is injective, i.e., $Q(G) \cong G$ and is isomorphic to a diagonal representation T, i.e., $T(\sigma)$ is diagonal for all $\sigma \in G$. Clearly $T(\sigma)$ for all $\sigma \in G$ are commuting. We then see that $Q(G)$, and thus G is abelian.

2. By Theorem 2.3.6 and the remark before it.

3. (Roy) The second claim $\sum_{t=1}^r n_t^2 \le |G|$ follows from the first one, because $\dim \mathbb{C}[G] = |G|$. Suppose there are $c_{ij}^t \in \mathbb{C}$ such that

$$\sum_{t=1}^{r} \sum_{i,j} c_{ij}^t a_{ij}^t(\sigma) = 0, \qquad \text{for all } \sigma \in G.$$

Multiply (on the right) both sides by $a_{pq}^s(\sigma^{-1}\pi)$ and sum over σ, by Theorem 2.3.5

$$0 = \sum_{t=1}^{r} \sum_{i,j} c_{ij}^t \sum_{\sigma\in G} a_{ij}^t(\sigma) a_{pq}^s(\sigma^{-1}\pi) = \frac{|G|}{n_s} \sum_{i,j} c_{ij}^s a_{iq}^s(\pi)\delta_{jp} = \frac{|G|}{n_s} \sum_{i} c_{ip}^s a_{iq}^s(\pi)$$

for all $\pi \in G$. Since the above equation holds for all $s = 1, \dots, r$ and $1 \le p, q \le n_s$, we conclude by Theorem 2.3.6 again that $c_{ij}^s = 0$ for all $1 \le i, j \le n_s$. Thus all c_{ij}^s's are zero by induction hypothesis.

2.4 Characters

In this section, we assume that G is a finite group.

Let A be a (matrix or linear) representation of G. The function $\chi : G \to \mathbb{C}$ defined by

$$\chi(\sigma) := \operatorname{tr} A(\sigma)$$

is called a **character** of G. If A is irreducible, we call χ an **irreducible character**. Characters are **class functions** on G, i.e., $\chi(\tau)$ is a constant for $\tau \in [\sigma]$, where

$$[\sigma] = \{\pi\sigma\pi^{-1} : \pi \in G\}$$

denotes the conjugacy class in G containing $\sigma \in G$. So characters are independent of the choice of basis in the representation. Though non-similar matrices could have same trace, we will see in this section that characters completely determine the representation up to isomorphism, i.e., representations having the same characters if and only if they are isomorphic.

Clearly $\chi(e) = \operatorname{tr} A(e) = \operatorname{tr} I_n = n$ where $n = \chi(e)$ is the degree of the representation. When $\chi(e) = 1$, χ is said to be **linear**; otherwise it is nonlinear.

Example 2.4.1. $\chi \equiv 1$ is a character of G and is called the **principal** character. When $G = S_m$, the sign function ε is called the **alternating** character.

We denote by $I(G)$ the set of irreducible characters of G. We will see that $I(G)$ is a finite set.

Theorem 2.4.2. (Orthogonal relation of the first kind) Let $\chi, \mu \in I(G)$. Then

$$\frac{1}{|G|} \sum_{\sigma\in G} \chi(\sigma)\mu(\sigma^{-1}) = \begin{cases} 1 & \text{if } \chi = \mu \\ 0 & \text{if } \chi \ne \mu \end{cases} \tag{2.8}$$

Proof. Let A and B be the irreducible representations of $\chi, \mu \in I(G)$. Then the left side of (2.8) is

$$\frac{1}{|G|} \sum_{i,j} \sum_{\sigma \in G} a_{ii}(\sigma) b_{jj}(\sigma^{-1}) \tag{2.9}$$

If A and B are not isomorphic, from (2.5), (2.9) is zero. So if $\chi \neq \mu$, then A and B are not isomorphic so that $\frac{1}{|G|} \sum_{\sigma \in G} \chi(\sigma)\mu(\sigma^{-1}) = 0$. If $\chi = \mu$, then we may assume that $A = B$. Applying (2.5),

$$\frac{1}{|G|} \sum_{\sigma \in G} \chi(\sigma)\mu(\sigma^{-1}) = \frac{1}{|G|} \sum_{i,j} \frac{|G|}{n} \delta_{ij} = 1.$$

□

Immediately from Theorem 2.4.2 non-isomorphic irreducible representations have different characters.

Theorem 2.4.3. *Let $\chi, \mu \in I(G)$. For any $\pi \in G$,*

$$\frac{1}{|G|} \sum_{\sigma \in G} \chi(\sigma^{-1})\mu(\sigma\pi) = \begin{cases} \frac{\chi(\pi)}{\chi(e)} & \text{if } \chi = \mu \\ 0 & \text{if } \chi \neq \mu \end{cases}$$

Proof. Similar to the proof of Theorem 2.4.2 and use Theorem 2.3.5. □

Theorem 2.4.4. For any character χ of G, $\chi(\sigma^{-1}) = \overline{\chi(\sigma)}$ for all $\sigma \in G$.

Proof. From Problem 2.3, χ is obtained from a unitary representation, i.e., $A(\sigma^{-1}) = A(\sigma)^{-1} = A(\sigma)^*$. So

$$\chi(\sigma^{-1}) = \operatorname{tr} A(\sigma^{-1}) = \operatorname{tr}(A(\sigma)^*) = \overline{\operatorname{tr} A(\sigma)} = \overline{\chi(\sigma)}.$$

□

We remark that when χ is linear, $\chi(\sigma^{-1}) = \chi(\sigma)^{-1}$ because $\chi(\sigma) = A(\sigma)$. Moreover from Theorem 2.4.4 $|\chi(\sigma)| = 1$ for all $\sigma \in G$. But this is not true for nonlinear characters.

Let $\mathbb{C}[G]$ be the set of all functions $f : G \to \mathbb{C}$. Notice that $\mathbb{C}[G]$ is a vector space under natural operations and $\dim \mathbb{C}[G] = |G|$.

Theorem 2.4.5. On $\mathbb{C}[G]$, $(f, g) := \frac{1}{|G|} \sum_{\sigma \in G} f(\sigma)\overline{g(\sigma)}$ is an inner product.

Proof. Exercise. □

We sometimes denote it as $(\cdot, \cdot)_G$. With respect to this inner product, from Theorem 2.4.2 and Theorem 2.4.4 the irreducible characters of G are orthonormal in $\mathbb{C}[G]$ and thus are linearly independent. So we have finitely many irreducible characters of G. Denote them by

$$I(G) = \{\chi_1, \ldots, \chi_k\}.$$

From (2.4.2)

$$(\chi_i, \chi_j) = \delta_{ij}. \tag{2.10}$$

A function $f : G \to \mathbb{C}$ is said to be a **class function** if f is constant on each conjugacy class of G. Denote by $C(G) \subset \mathbb{C}[G]$ the set of all class functions. It is clear that $C(G)$ is a vector space. Since G is a disjoint union of its conjugacy classes, $\dim C(G)$ is the number of conjugacy classes in G.

Theorem 2.4.6. (Isomorphism criterion) Let A, B be two representations of G. Then A and B are isomorphic if and only if $\chi_A = \chi_B$.

Proof. It suffices to establish the sufficiency. By Theorem 2.2.3 we can decompose A and B into their irreducible direct sums:

$$A(\sigma) = A_1(\sigma) \oplus \cdots \oplus A_l(\sigma), \quad B(\sigma) = B_1(\sigma) \oplus \cdots \oplus B_m(\sigma)$$

Among $A_1, \dots, A_l$, $B_1, \dots, B_m$, denote by $D_1, \dots, D_t$ be the non-isomorphic ones and let χ_i be the characters of D_i, $i = 1, \dots, t$. Suppose there are r_i representations among $A_1, \dots, A_l$ and s_i representations among $B_1, \dots, B_m$ isomorphic to D_i, $i - 1, \dots, t$. Then

$$\chi_A(\sigma) = \operatorname{tr}(\sum_{i=1}^{l} A_i(\sigma)) = \sum_{i=1}^{t} r_i \operatorname{tr} D_i(\sigma) = \sum_{i=1}^{t} r_i \chi_i(\sigma) \tag{2.11}$$

Similarly

$$\chi_B(\sigma) = \sum_{i=1}^{t} s_i \chi_i(\sigma).$$

From $\chi_A = \chi_B$ we have

$$\sum_{i=1}^{t} (r_i - s_i) \chi_i(\sigma) = 0, \quad \sigma \in G.$$

Since $\chi_1, \dots, \chi_t$ are linearly independent, $r_i = t_i$ for all $i = 1, \dots, t$. So A and B are isomorphic. □

Theorem 2.4.7. (Character criterion) $\chi \in C(G)$ is character of G if and only if $\chi = \sum_{i=1}^{k} m_i \chi_i$, where $m_1, \dots, m_k$ are nonnegative integers.

Proof. From (2.11). □

The following theorem implies that the irreducibility of a representation can be deduced from its character.

Theorem 2.4.8. (Irreducibility criterion) A character χ of G is irreducible if and only if $(\chi, \chi) = 1$.

Proof. Suppose that $A, A_1, \ldots, A_k$ are the representations for χ and $\chi_1, \ldots, \chi_k$ respectively and $I(G) = \{\chi_1, \ldots, \chi_k\}$. So $A = \sum_{i=1}^k m_i A_i$. From Theorem 2.4.7, $\chi = \sum_{i=1}^k m_i \chi_i$. Since irreducible characters are orthonormal,

$$(\chi, \chi) = (\sum_{i=1}^k m_i \chi_i, \sum_{i=1}^k m_i \chi_i) = \sum_{i=1}^n m_i^2.$$

Now χ irreducible, i.e., A irreducible, if and only if only one $m_i = 1$ and other m_i are zeros. In other words, χ irreducible, if and only if $(\chi, \chi) = 1$. □

The regular representation $Q : G \to \mathrm{GL}_n(\mathbb{C})$, where $n = |G|$, contains all the irreducible representations of G as sub-repesentations. We now study its character structure.

Theorem 2.4.9. The multiplicity of χ_i in χ_Q is $\chi_i(e)$, i.e., the degree of χ_i.

$$\chi_Q = \sum_{i=1}^k \chi(e)\chi_i. \tag{2.12}$$

Proof. Recall that if $G = \{\sigma_1, \ldots, \sigma_n\}$, then the regular representation $Q : G \to \mathrm{GL}_n(\mathbb{C})$ is given by

$$Q(\sigma)_{ij} = \delta_{\sigma_i, \sigma\sigma_j}.$$

Let χ_Q be its character. Then

$$\chi_Q(\sigma) = \sum_{i=1}^n Q(\sigma)_{ii} = \sum_{i=1}^n \delta_{\sigma_i, \sigma\sigma_j} = \begin{cases} |G| & \text{if } \sigma = e \\ 0 & \text{if } \sigma \neq e \end{cases} \tag{2.13}$$

From Theorem 2.4.7, $\chi_Q = \sum_{i=1}^k m_i \chi_i$, where $I(G) = \{\chi_1, \ldots, \chi_k\}$. Then by (2.10) and (2.13)

$$m_i = (\chi_Q, \chi_i) = \frac{1}{|G|} \sum_{\sigma \in G} \chi_Q(\sigma)\overline{\chi_i(\sigma)} = \chi_i(e).$$

□

We have the following important relation between the order of G and the degrees of all irreducible characters.

Theorem 2.4.10. $|G| = \sum_{i=1}^k \chi_i(e)^2 = \sum_{\chi \in I(G)} \chi(e)^2$.

Proof. Use (2.13) and (2.12) to have $|G| = \chi_Q(e) = \sum_{i=1}^k \chi_i(e)^2$. □

Immediately we have an upper bound for the degree of any irreducible character χ which is not the principal character: , i.e. $\chi(e) \leq (|G| - 1)^{1/2}$.

Example 2.4.11. $|S_3| = 6$ so there are three irreducible characters: the principal character, the alternating character and an irreducible character of degree 2 (see Example 2.2.1(e)). It is because $6 = 1 + 1 + 2^2$.

Theorem 2.4.12. 1. (Completeness of characters) The irreducible characters form an orthonormal basis of $C(G)$.

2. The number of irreducible characters $|I(G)|$ of G is the number of conjugacy classes of G.

Proof. It suffices to establish the first statement.

Since the space $C(G)$ of class functions on G is the number of conjugacy classes in G and $I(G) = \{\chi_1, \ldots, \chi_k\}$ is a linearly independent set in $C(G)$, it suffices to show that every class function is a linear combination of $\chi_1, \ldots, \chi_k$. Let A_t be the corresponding irreducible representation of χ_t for all $t = 1, \ldots, k$, and denote by $A_t(\sigma) = (a^t_{ij})$. Let $\chi_i(e) = n_i$ for all i. So from Theorem 2.4.10 $|G| = \sum_{t=1}^k n_i^2$. From Problem 3.3 the $|G|$ functions a^t_{ij} are linearly independent and $\dim \mathbb{C}[G] = |G|$. So

$$\{a^t_{ij} : i, j = 1, \ldots, n, t = 1, \ldots, k\}$$

is a basis of $\mathbb{C}[G]$. Let $f \in C(G) \subset \mathbb{C}[G]$. Then for some c^t_{ij}, $i, j = 1, \ldots, n$ and $t = 1, \ldots, k$,

$$f(\sigma) = \sum_{t=1}^{k} \sum_{i,j=1}^{n_t} c^t_{ij} a^t_{ij}(\sigma), \qquad \text{for all } \sigma \in G.$$

Since f is a class function,

$$\begin{aligned}
f(\sigma) &= f(\pi^{-1}\sigma\pi) = \frac{1}{|G|} \sum_{\pi \in G} f(\pi^{-1}\sigma\pi) \\
&= \frac{1}{|G|} \sum_{\pi \in G} \sum_{t=1}^{k} \sum_{i,j=1}^{n_t} c^t_{ij} a^t_{ij}(\pi^{-1}\sigma\pi) \\
&= \frac{1}{|G|} \sum_{\pi \in G} \sum_{t=1}^{k} \sum_{i,j=1}^{n_t} c^t_{ij} \sum_{p,q=1}^{n_t} a^t_{ip}(\pi^{-1}) a^t_{pq}(\sigma) a^t_{qj}(\pi) \\
&= \sum_{t=1}^{k} \sum_{i,j,p,q=1}^{n_t} c^t_{ij} a^t_{pq}(\sigma) \left(\frac{1}{|G|} \sum_{\pi \in G} a^t_{ip}(\pi^{-1}) a^t_{qj}(\pi) \right) \\
&= \sum_{t=1}^{k} \sum_{i,j,p,q=1}^{n_t} c^t_{ij} a^t_{pq}(\sigma) \frac{1}{n_t} \delta_{ij} \delta_{pq} \qquad \text{by (2.5)} \\
&= \sum_{t=1}^{k} \sum_{i,p=1}^{n_t} c^t_{ii} a^t_{pp}(\sigma) \frac{1}{n_t} \\
&= \sum_{t=1}^{k} \left(\frac{1}{n_t} \sum_{i=1}^{n_t} c^t_{ii} \right) \chi_t(\sigma)
\end{aligned}$$

a linear combination of $\chi_1, \ldots, \chi_k$. ☐

Example 2.4.13. Since $|S_4| = 24$ and there are 5 conjugacy classes, there are 5 irreducible characters of G: principal character, alternating character, two irreducible characters of degree 3 and one irreducible character of degree 2 as $24 = 1 + 1 + 2^2 + 3^2 + 3^2$ (the principal and alternating characters are the only linear characters).

Let $[\sigma]$ be the conjugacy class in G containing σ.

Theorem 2.4.14. (Orthogonal relation of the second kind)

$$\frac{|[\sigma]|}{|G|} \sum_{\chi \in I(G)} \chi(\sigma)\overline{\chi(\pi)} = \begin{cases} 1 & \text{if } [\sigma] = [\pi] \\ 0 & \text{if } [\sigma] \neq [\pi] \end{cases} \tag{2.14}$$

Proof. Let $I(G) = \{\chi_1, \ldots, \chi_k\}$. From Theorem 2.4.12, k is also the number of conjugacy classes. Let $\sigma_1, \ldots, \sigma_k \in G$ be from the distinct classes. Let $U = (u_{ij}) \in \mathbb{C}_{k\times k}$ be defined by

$$u_{ij} = \left(\frac{|[\sigma_j]|}{|G|}\right)^{1/2} \chi_i(\sigma_j), \quad i, j = 1, \ldots k.$$

Then

$$(UU^*)_{ij} = \sum_{t=1}^{k} u_{it}\overline{u}_{jt} = \frac{1}{|G|}\sum_{t=1}^{k} \chi_i(\sigma_t)\overline{\chi_j(\sigma_i)}|[\sigma_t]| = \frac{1}{|G|}\sum_{\sigma\in G} \chi_i(\sigma)\overline{\chi_j(\sigma)} = \delta_{ij},$$

i.e., $UU^* = I$ and thus $U^*U = I$. So

$$\delta_{ij} = (U^*U)_{ij} = \sum_{i=1}^{k} \overline{u}_{ti}u_{tj} = \frac{(|[\sigma_i]||[\sigma_j]|)^{1/2}}{|G|}\sum_{t=1}^{k} \chi_t(\sigma_i)\overline{\chi_t(\sigma_j)} \tag{2.15}$$

and (2.15) is another form of (2.14). ☐

Theorem 2.4.15.

$$\frac{1}{|G|}\sum_{\chi\in I(G)} \chi(e)\chi(\sigma) = \begin{cases} 1 & \text{if } \sigma = e \\ 0 & \text{if } \sigma \neq e \end{cases} \tag{2.16}$$

Proof. Take $\sigma = e$ in (2.14). ☐

Theorem 2.4.16. If $\chi \in I(G)$, then $\chi(e)$ divides $|G|$.

Proof. Omitted ☐

The above theorem has some interesting implications. For example, if $|G|$ is a prime, the all characters of G are linear and thus G must be abelian from Problem 3.1.

Problems

1. Let χ, μ be characters of G. Prove
   (a) $\chi(\sigma\pi) = \chi(\pi\sigma)$, for all $\sigma, \pi \in G$.
   (b) $|\chi(\sigma)| \leq \chi(e)$, for all $\sigma \in G$.
   (c) (χ, μ) is a nonnegative integer.

2. Let $\chi \in I(G)$. Prove that
   (a)
$$\sum_{\sigma \in G} \chi(\sigma) = \begin{cases} |G| & \text{if } \chi \equiv 1 \\ 0 & \text{otherwise} \end{cases}$$
   (b) $\sum_{\sigma \in G} |\chi(\sigma)|^2 = |G|$.

3. Prove that $\sum_{\chi \in I(G)} |\chi(\sigma)|^2 = \frac{|G|}{|[\sigma]|}$ for any $\sigma \in G$.

4. Prove that $\chi \equiv 1$ and $\chi = \varepsilon$ are the only linear characters of S_n.

5. What is $|I(S_3)|$? Find $I(S_3)$.

6. Let $A : G \to \mathrm{GL}_n(\mathbb{C})$ be a representation of G. Show that $C(\sigma) := A(\sigma^{-1})^T$ and $D(\sigma) := \overline{A(\sigma)}$ are representations and they are isomorphic. Is $E(\sigma) := A(\sigma)^*$ a representation of G?

7. Show that if χ is a character, so is $\overline{\chi}$. Is it true that $\overline{\chi} \in I(G)$ if $\chi \in I(G)$?

8. Find all the irreducible characters of the alternating group A_3 (A_3 is isomorphic to the cyclic group C_3).

Solutions to Problems 2.4

1. Let $I(G) = \{\chi_1, \ldots, \chi_k\}$. Let χ, μ be characters of G with representations A and B.
   (a) $\chi(\sigma\pi) = \operatorname{tr} A(\sigma\pi) = \operatorname{tr} A(\sigma)A(\pi) = \operatorname{tr} A(\pi)A(\sigma) = \operatorname{tr} A(\pi\sigma) = \chi(\pi\sigma)$.
   (b) Since $A(\sigma)$ can be viewed as a unitary matrix for all $\sigma \in G$, and diagonal entries of a unitary matrix have moduli no greater than 1, we have $|\chi(\sigma)| = |\operatorname{tr} A_{ii}(\sigma)| \leq \operatorname{tr} |A_{ii}(\sigma)| \leq \chi(e)$.
   (c) By Theorem 2.4.12, $\chi = \sum_{i=1}^k n_i\chi_i$ and $\mu = \sum_{i=1}^k m_i\chi_i$ where n_i and m_i are nonnegative integers for all i. So $(\chi, \mu) = \sum_{i=1}^k n_i m_i$ is a nonnegative integer.

2. (a) Use Theorem 2.4.2 with $\mu \equiv 1$.
   (b) Use Theorem 2.4.2 with $\chi = \mu$ and $\pi = e$.

3. From Theorem 2.4.14 with $\pi = \sigma$.

4. The transpositions form a conjugacy class $[(12)]$ in S_n (check!). Notice that if χ is linear, then it is a representation so that $\chi^2(12) = \chi((12)^2) = 1$. Hence $\chi(\tau) = \pm 1$ for all transpositions $\tau \in S_n$. Now S_n is generated by the transpositions and $\chi(\sigma\pi) = \chi(\sigma)\chi(\pi)$, $\chi \equiv 1$ and $\chi = \varepsilon$ are the only linear characters of S_n.

   (Alex) Let χ be a linear character of S_n, i.e., $\chi : S_n \to \mathbb{C}^*$ where $\mathbb{C}^* = \mathbb{C} \setminus \{0\}$ is a homomorphism. Since $\operatorname{Im}\chi$ is abelian, the commutator subgroup $[S_n, S_n] = \{xyx^{-1}y^{-1} : x, y \in S_n\}$ of S_n is contained in $\operatorname{Ker}\chi$. But the commutator group $[S_n, S_n]$ is the alternating group A_n. Since A_n is the largest proper normal subgroup of S_n, $\operatorname{Ker}\chi$ is either S_n or A_n. When $\operatorname{Ker}\chi = S_n$, $\chi \equiv 1$; when $\operatorname{Ker}\chi = A_n$, then $\chi(12) = -1$ since $(12)^2 = 1$ implies that $\chi^2(12) = 1$.

5. The conjugacy class of S_3 are:

   $$[e] = \{e\}, \quad [(12)] = \{(12), (23), (13)\}, \quad [(123)] = \{(123), (132)\}$$

   So $|I(S_3)| = 3$ by Theorem 2.4.12. Notice that $\chi_1 \equiv 1$ and $\chi_2 = \varepsilon$ are the linear characters of S_3. Let ξ_3 be the last irreducible character. So from Theorem 2.4.10, $6 = \chi_1^2(e) + \chi_3^2(e) + \chi_3^2(e) = 1 + 1 + \chi_3^2(e)$ so that χ_3 is of degree 2. It is the character corresponding to Problem 2.2.

6. $C(\sigma\pi) = A((\sigma\pi)^{-1})^T = (A(\pi^{-1})A(\sigma^{-1}))^T = A((\sigma)^{-1})^T A((\pi)^{-1})^T = C(\sigma)C(\pi)$ for all $\sigma, \pi \in G$. $D(\sigma\pi) = \overline{A(\sigma\pi)} = \overline{A(\sigma)A(\pi)} = \overline{A(\sigma)}\,\overline{A(\pi)} = D(\sigma)D(\pi)$. By Theorem 2.4.6 C and D are isomorphic since $\chi_D = \operatorname{tr} D(\sigma) = \operatorname{tr} C(\sigma) = \chi_C$. However $E(\sigma\pi) = A(\sigma\pi)^* = (A(\sigma)A(\pi))^* = A(\pi)^* A(\sigma)^* \neq E(\sigma)E(\pi)$.

7. From Problem 6: D is a representation. $\overline{\chi}$ is irreducible if and only if χ is irreducible. Just use $(\chi, \chi) = 1$, i.e., Theorem 2.4.8.

8. $|A_3| = |S_3|/2 = 3$. The conjugacy classes of A_3 are

   $$[e] = \{e\}, \quad [(123)] = \{(123)\}, \quad [(132)] = \{(132)\}.$$

   Let χ_1 be the principal character. By Theorem 2.4.10 there are only two irreducible characters left, both of degree 1, χ_2, χ_3, say. So $\chi_2(e) = \chi_3(e) = 1$ and $|\chi_2(\sigma)| = |\chi_3(\sigma)| = 1$ for all $\sigma \in A_3$. Since χ_1, χ_2, χ_3 are orthonormal, $\chi_2(123) = \xi$, $\chi_2(132) = \xi^2$, $\chi_2(123) = \xi^2$, and $\chi_2(132) = \xi$, where ξ is the primitive cubic root of unity. Notice that $\overline{\chi}_2 = \chi_3$.

Chapter 3

Multilinear maps and tensor spaces

3.1 Multilinear maps and tensor maps

Let $V_1, \ldots, V_m, W$ be $m+1$ vector spaces over a field $\mathbb{F}$. A map $\varphi : V_1 \times \cdots \times V_m \to W$ is called m-**multilinear** or simply multilinear if

$$\varphi(v_1, \ldots, \alpha v_i + \alpha' v_i', \ldots, v_m) = \alpha\varphi(v_1, \ldots, v_i, \ldots, v_m) + \alpha'\varphi(v_1, \ldots, v_i', \ldots, v_m),$$

for all $i = 1, \ldots, m$, i.e., T is linear while restricted on the ith coordinate for all i. Indeed the condition can be weakened to

$$\varphi(v_1, \ldots, v_i + \alpha v_i', \ldots, v_m) = \varphi(v_1, \ldots, v_i, \ldots, v_m) + \alpha\varphi(v_1, \ldots, v_i', \ldots, v_m)$$

for all i. A linear map $T : V \to W$ can be viewed as a 1-multilinear map. There is some significiant difference between a linear map and a multilinear map.

Consider the linear map $T \in \operatorname{Hom}(V_1 \times V_2, W)$ and a multilinear map $\varphi : V_1 \times V_2 \to W$. Since T is linear,

$$T(v_1+v_1', v_2+v_2') = T(v_1, v_2)+T(v_1', v_2') = T(v_1, 0)+T(0, v_2)+T(v_1', 0)+T(0, v_2').$$

Since φ is multilinear,

$$\varphi(v_1+v_1', v_2+v_2') = \varphi(v_1, v_2+v_2')+\varphi(v_1', v_2+v_2') = \varphi(v_1, v_2)+\varphi(v_1, v_2')+\varphi(v_1', v_2)+\varphi(v_1', v_2').$$

In particular $\varphi(v_1, 0) = \varphi(0, v_2) = 0$, but $T(v_1, 0)$ and $T(0, v_2)$ are not necessarily zero.

Example 3.1.1. The following maps are multilinear.

(a) $f : \mathbb{C} \times \mathbb{C} \to \mathbb{C}$ defined by $f(x, y) = xy$.

(b) $\varphi : V^* \times V \to \mathbb{C}$ defined by $\varphi(f, v) = f(v)$.

(c) $\varphi : \mathbb{C}^m \times \mathbb{C}^n \to \mathbb{C}$ defined by $\varphi(x, y) = x^T Ay$, where $A \in \mathbb{C}_{m\times n}$ is given.

(d) $\otimes : \mathbb{C}^m \times \mathbb{C}^n \to \mathbb{C}_{m\times n}$ defined by $\otimes(x, y) = xy^T$.

(e) $\det : \mathbb{C}^n \times \cdots \times \mathbb{C}^n \to \mathbb{C}$ defined by $\det(x_1, \ldots, x_n) = \det A$, where $A = [x_1 \ \cdots \ x_n]$.

(f) $f : V_1 \times \cdots \times V_m \to \mathbb{C}$ defined by $f(v_1, \ldots, v_m) = \prod_{i=1}^m f_i(v_i)$, where $f_i \in V_i^*$, $i = 1, \ldots, m$, are given. We will write $f = \prod_{i=1}^m f_i$. So $\prod_{i=1}^m f_i(v_1, \ldots, v_m) = \prod_{i=1}^m f_i(v_i)$.

(g) $g : V_1^* \times \cdots \times V_m^* \to \mathbb{C}$ defined by $g(f_1, \ldots, f_m) = \prod_{i=1}^m f_i(v_i)$ where $v_i \in V$, $i = 1, \ldots, m$, are given.

(h) Let $\varphi : V_1 \times \cdots \times V_m \to W$ and $\psi : V_1 \times \cdots \times V_m \to W$. Then $\alpha\varphi + \beta\psi$ is also multilinear. So the set of multilinear maps $M(V_1, \ldots, V_m; W)$ is a vector space.

Recall the notation in Chapter 1:

$$\Gamma := \Gamma(n_1, \ldots, n_m) := \{\gamma : \gamma = (\gamma(1), \ldots, \gamma(m)), 1 \le \gamma(i) \le n_i, i = 1, \ldots, m\}$$

with $|\Gamma| = \prod_{i=1}^m n_i$. We will use Γ frequently. Notice that we can order Γ according to the lexicographic order. Moreover

$$\prod_{i=1}^m \sum_{j=1}^{n_i} a_{ij} = \sum_{\gamma\in\Gamma} \prod_{i=1}^m a_{i\gamma(i)}.$$

Let $E_i = \{e_{i1}, \ldots, e_{in_i}\}$ be a basis of V_i, $i = 1, \ldots, m$. So each $v_i \in V_i$ can be written as $v_i = \sum_{j=1}^{n_i} a_{ij} e_{ij}$, $i = 1, \ldots, m$. Let $\psi : V_1 \times \cdots \times V_m \to W$ be multilinear. From definition,

$$\begin{aligned}
\psi(v_1, \ldots, v_m) &= \psi(\sum_{j_1=1}^{n_1} a_{1j_1} e_{1j_1}, \ldots, \sum_{j_m=1}^{n_m} a_{mj_m} e_{mj_m}) \\
&= \sum_{j_1=1}^{n_1} \cdots \sum_{j_m=1}^{n_m} a_{1j_1} \cdots a_{mj_m} \psi(e_{1j_1} \cdots e_{mj_m}) \\
&= \sum_{\gamma\in\Gamma} a_{1\gamma(1)} \cdots a_{m\gamma(m)} \psi(e_{1\gamma(1)}, \ldots, e_{m\gamma(m)}) \\
&= \sum_{\gamma\in\Gamma} a_\gamma \psi(e_\gamma) \qquad (3.1)
\end{aligned}$$

where

$$\begin{aligned}
a_\gamma &:= \prod_{i=1}^m a_{i\gamma(i)} \in \mathbb{C}, \\
e_\gamma &:= (e_{1\gamma(1)}, \ldots, e_{m\gamma(m)}) \in \times_{i=1}^m V_i, \qquad (3.2)
\end{aligned}$$

for all $\gamma \in \Gamma(n_1, \ldots, n_m)$. $\psi(e_\gamma)$ in (3.1) completely determine the multilinear ψ.

Theorem 3.1.2. (Multilinear extension) Let $E_i = \{e_{i1}, \ldots, e_{in_i}\}$ be a basis of V_i, $i = 1, \ldots, m$. There exists a unique multilinear map $\varphi : V_1 \times \cdots \times V_m \to W$ such that $\varphi(e_\gamma) = w_\gamma$ for all $\gamma \in \Gamma(n_1, \ldots, n_m)$, where e_γ is given in (3.2) and $w_\gamma \in W$ are given.

Proof. Since we want $\varphi(e_\gamma) = w_\gamma$ for all $\gamma \in \Gamma$, from (3.1) we need to define

$$\varphi(v_1, \ldots, v_m) = \sum_{\gamma \in \Gamma} a_\gamma w_\gamma$$

where a_γ, w_γ are given in (3.2) and $v_i = \sum_{j=1}^{n_i} a_{ij} e_{ij}$. Let $v_i' = \sum_{j=1}^{n_i} a_{ij}' e_{ij}$ where $i = 1, \ldots, m$. From the definition of φ,

$$\begin{aligned} & \varphi(v_1, \ldots, v_i + cv_i', \ldots, v_m) \\ = & \sum_{\gamma \in \Gamma} a_{1\gamma(1)} \cdots (a_{i\gamma(i)} + ca_{i\gamma(i)}') \cdots a_{m\gamma(m)} w_\gamma \\ = & \sum_{\gamma \in \Gamma} a_{1\gamma(1)} \cdots a_{i\gamma(i)} \cdots a_{m\gamma(m)} w_\gamma + c \sum_{\gamma \in \Gamma} a_{1\gamma(1)} \cdots a_{i\gamma(i)}' \cdots a_{m\gamma(m)} w_\gamma \\ = & \varphi(v_1, \ldots, v_i, \ldots, v_m) + c\varphi(v_1, \ldots, v_i', \ldots, v_m) \end{aligned}$$

i.e., φ is multilinear. We are going to show that $\varphi(e_\alpha) = w_\alpha$ for all $\alpha \in \Gamma$. For any $\alpha \in \Gamma$, write

$$e_{i\alpha(i)} = \sum_{j=1}^{n_j} \delta_{\alpha(i)j} e_{ij}.$$

From the definition of φ and $e_\alpha = (e_{1\alpha(1)}, \ldots, e_{m\alpha(m)})$, we have

$$\varphi(e_\alpha) = \varphi(e_{1\alpha(1)}, \ldots, e_{m\alpha(m)}) = \sum_{\gamma \in \Gamma} \delta_{\alpha(1)\gamma(1)} \cdots \delta_{\alpha(m)\gamma(m)} w_\gamma = \sum_{\gamma \in \Gamma} \delta_{\alpha\gamma} w_\gamma = w_\alpha.$$

Thus we just established the existence.

Suppose that there is another multilinear map $\psi : V_1 \times \cdots \times V_m \to W$ such that $\psi(e_\gamma) = w_\gamma$, $\gamma \in \Gamma$. Then from (3.1)

$$\psi(v_1, \ldots, v_m) = \sum_{\gamma \in \Gamma} a_\gamma \psi(e_\gamma) = \sum_{\gamma \in \Gamma} a_\gamma w_\gamma = \varphi(v_1, \ldots, v_m).$$

So $\psi = \varphi$. □

Let us point some more differences between linear and mutlilinear maps. When $T : V \to W$ is linear, T is completely determined by the $n := \dim V$ values $T(e_1), \ldots, T(e_n)$ where $E = \{e_1, \ldots, e_n\}$ is a basis of V. But for a multilinear map φ, we need $|\Gamma| = \prod_{i=1}^m \dim V_i$ values. It is much more than $\dim(V_1 \times \cdots \times V_m) = \sum_{i=1}^m \dim V_i$.

Recall Example 3.1.1(d) with $m = n = 2$, i.e., $\otimes : \mathbb{C}^2 \times \mathbb{C}^2 \to \mathbb{C}_{2\times 2}$ defined by $\otimes(x, y) = xy^T$. Since $\operatorname{rank}(xy^T) \le \min\{\operatorname{rank} x, \operatorname{rank} y^T\} \le 1$, we have $\det(\otimes(x, y)) = 0$. But if $x_1 := (1, 0)^T$ and $x_2 := (0, 1)^T$, then

$$\det(\otimes(x_1, x_1) + \otimes(x_2, x_2)) = \det I_2 = 1.$$

So $\otimes(x_1, x_1) + \otimes(x_2, x_2) \notin \operatorname{Im} \otimes$. So $\operatorname{Im} \otimes$ is not a subspace.

In general the image $\operatorname{Im} \varphi = \{\varphi(v_1, \dots, v_m) : v_i \in V_i, i = 1, \dots, m\}$ is not necessarily a subspace of W. The **rank** of φ,

$$\operatorname{rank} \varphi = \dim \langle \operatorname{Im} \varphi \rangle.$$

Clearly $\operatorname{rank} \varphi \leq \prod_{i=1}^m \dim V_i$. The multilinear map φ is called a **tensor map** if $\operatorname{rank} \varphi = \prod_{i=1}^m \dim V_i$. In other words, a tensor map is a multilinear map with **maximal image span**. Example 3.1.1 (a) is a trivial tensor map.

Theorem 3.1.3. The multilinear map $\varphi : V_1 \times \cdots \times V_m \to P$ is a tensor map if and only if the set $\{\varphi(e_\gamma) : \gamma \in \Gamma\}$ is linear independent, where e_γ is given in (3.2).

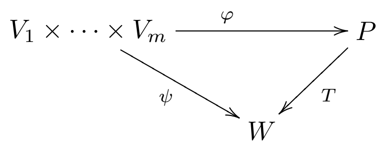
Proof. From (3.1) $\langle \varphi(e_\gamma) : \gamma \in \Gamma \rangle = \langle \operatorname{Im} \varphi \rangle$ and $|\Gamma| = \prod_{i=1}^m \dim V_i$. □

Theorem 3.1.4. Tensor map of $V_1, \dots, V_m$ exists, i.e., there are W and $\varphi : V_1 \times \cdots \times V_m \to W$ such that φ is a tensor map.

Proof. By Theorem 3.1.2, simply pick W so that $\dim W = \prod_{i=1}^m \dim V_i$ and let $\{w_\gamma : \gamma \in \Gamma\}$ be a basis so that w_γ $(\gamma \in \Gamma)$ determine the multilinear φ which is obviously a tensor map. □

Clearly tensor maps on $V_1 \times \cdots \times V_m$ are not unique.

The study of multilinear map is reduced to the study of some linear map (not unique) via a tensor map. A multilinear map $\varphi : V_1 \times \cdots \times V_m \to P$ is said to have **universal factorization property** if for any multilinear map $\psi : V_1 \times \cdots \times V_m \to W$, there is $T \in \operatorname{Hom}(P, W)$ such that $\psi = T \circ \varphi$.



Theorem 3.1.5. The multilinear map $\varphi : V_1 \times \cdots \times V_m \to P$ is a tensor map if and only if φ has universal factorization property.

Proof. Suppose that φ is a tensor map. Then $\{\varphi(e_\gamma) : \gamma \in \Gamma\}$ is a basis of $\langle \operatorname{Im} \varphi \rangle$. From Theorem 1.1.1, there is a unique $T_1 \in \operatorname{Hom}(\langle \operatorname{Im} \varphi \rangle, W)$ such that $T_1 \varphi(e_\gamma) = \psi(e_\gamma)$, $\gamma \in \Gamma$. Since $\langle \operatorname{Im} \varphi \rangle \subset P$, there is $T \in \operatorname{Hom}(P, W)$ such that $T|_{\langle \operatorname{Im} \varphi \rangle} = T_1$. So $T\varphi(e_\gamma) = \psi(e_\gamma)$, $\gamma \in \Gamma$. Since $T\varphi$ and ψ are multilinear maps on $V_1 \times \cdots \times V_m$ (Problem 2), from Theorem 3.1.2, $T\varphi = \psi$.

Conversely, suppose that φ has universal factorization property. In particular consider a tensor map ψ on $V_1 \times \cdots \times V_m$, i.e., $\dim \langle \operatorname{Im} \psi \rangle = \prod_{i=1}^m \dim V_i$. Then $T\varphi = \psi$ for some linear T. Thus $T(\langle \operatorname{Im} \varphi \rangle) = \langle \operatorname{Im} \psi \rangle$. Hence

$$\dim \langle \operatorname{Im} \psi \rangle \leq \dim \langle \operatorname{Im} \varphi \rangle.$$

So $\operatorname{rank} \varphi = \prod_{i=1}^m \dim V_i$ and φ is a tensor map. □

Problems

1. Suppose that $\varphi : W_1 \times \cdots \times W_m \to W$ is multilinear and $T_i : V_i \to W_i$ is linear, $i = 1, \ldots, m$. Define $\psi : V_1 \times \cdots \times V_m \to W$ by $\psi(v_1, \ldots, v_m) = \varphi(T_1 v_1, \ldots, T_m v_m)$. Show that ψ is multilinear.
2. Prove that if $\varphi : V_1 \times \cdots \times V_m \to W$ is multilinear and $T : V \to W$ is linear, then $T \circ \varphi$ is multilinear.
3. Show that when $n > 1$, the determinant function $\det : \mathbb{C}^n \times \cdots \times \mathbb{C}^n \to \mathbb{C}$ is not a tensor map.
4. Suppose that the multilinear map $\varphi : V_1 \times \cdots \times V_m \to P$ has universal factorization property. Show that the linear map T is unique if and only if $\langle \operatorname{Im} \varphi \rangle = P$.

Solutions to Problems 3.1

1.

$$\begin{aligned} & \psi(v_1, \ldots, v_i + cv_i', \ldots, v_m) \\ = \; & \varphi(T_1 v_1, \ldots T_i(v_i + cv_i'), \ldots, T_m v_m) \\ = \; & \varphi(T_1 v_1, \ldots, T_i(v_i) + cT_i(v_i'), \ldots, T_m v_m) \\ = \; & \varphi(T_1 v_1, \ldots, T_i(v_i), \ldots, T_m v_m) + c\varphi(T_1 v_1, \ldots, T_i(v_i'), \ldots, T_m v_m) \\ = \; & \psi(v_1, \ldots, v_i, \ldots, v_m) + c\psi(v_1, \ldots, v_i', \ldots, v_m). \end{aligned}$$

2.

$$\begin{aligned} & T\varphi(v_1, \ldots, v_i + cv_i', \ldots, v_m) \\ = \; & T(\varphi(v_1, \ldots, v_i, \ldots, v_m) + c\varphi(v_1, \ldots, v_i', \ldots, v_m)) \\ = \; & T\varphi(v_1, \ldots, v_i, \ldots, v_m) + cT\varphi(v_1, \ldots, v_i', \ldots, v_m) \end{aligned}$$

3. $\dim(\mathbb{C}^n \times \cdots \times \mathbb{C}^n) = n^n$ but $\dim\langle \operatorname{Im} \det \rangle = \dim \mathbb{C} = 1$. So $\det : \mathbb{C}^n \times \cdots \times \mathbb{C}^n \to \mathbb{C}$ is not a tensor map.
4. Suppose that $T_1 \circ \varphi = T_2 \circ \varphi$ where $T_1, T_2 : P \to W$. Then $T_1 z = T_2 z$ for all $z \in \operatorname{Im} \varphi$. Since $\langle \operatorname{Im} \varphi \rangle = P$, we have $T_1 z = T_2 z$ for all $z \in P$. So $T_1 = T_2$.

3.2 Tensor products and unique factorization

Let P be a vector space. If there is a tensor map $\otimes : V_1 \times \cdots \times V_m \to P$ such that $\langle \operatorname{Im} \otimes \rangle = P$, then P is said to be the **tensor product** of $V_1, \cdots, V_m$ and is written as $\otimes_{i=1}^m V_i = V_1 \otimes \cdots \otimes V_m$. Clearly

$$\dim \otimes_{i=1}^m V_i = \dim\langle \operatorname{Im} \otimes \rangle = \prod_{i=1}^n \dim V_i.$$

The elements of $\otimes_{i=1}^m V_i$ are called **tensors**. The tensors of the form

$$\otimes(v_1, \ldots, v_m) = v_1 \otimes \cdots \otimes v_m$$

are called **decomposable tensors**, i.e., tensor in $\Im \otimes$ are decomposable. The decomposable tensors span $\otimes_{i=1}^m V_i$ so we can find a basis of decomposable tensors for $\otimes_{i=1}^m V_i$. The tensors which are not decomposable, i.e., tensors in $\langle \operatorname{Im} \otimes \rangle \setminus \operatorname{Im} \otimes$ are called **indecomposable** tensors.

From Theorem 3.1.4:

Theorem 3.2.1. Let P be a vector space with $\dim P = \prod_{i=1}^n \dim V_i$. Then there exists $\otimes : V_1 \times \cdots \times V_m \to P$ so that P is the tensor space.

Theorem 3.2.2. Tensor spaces of $V_1, \ldots, V_m$ are isomorphic, i.e., if $P = \otimes_{i=1}^m V_i$ and $Q = \boxtimes_{i=1}^m V_i$ are tensor spaces of $V_1, \ldots, V_m$, then there exists an invertible $T \in \operatorname{Hom}(P, Q)$ such that $T \circ \otimes = \boxtimes$.

Proof. Since P and Q are tensor spaces of $V_1, \ldots, V_m$, $\dim P = \dim Q$,

$$\{\otimes e_\gamma : \gamma \in \Gamma\}, \quad \{\boxtimes e_\gamma : \gamma \in \Gamma\}$$

are bases of P and Q. By Theorem 3.1.5 (universal factorization property) there is $T \in \operatorname{Hom}(P, Q)$ such that $T \otimes (e_\gamma) = \boxtimes e_\gamma$, $\gamma \in \Gamma$. Thus T is invertible and $T \otimes = \boxtimes$. ☐

So all tensor products are isomorphic and thus we use $V_1 \otimes \cdots \otimes V_m$ to denote any one of those.

Theorem 3.2.3. Let $\psi : V_1 \times \cdots \times V_m \to W$ be a multilinear map. Then there exists a unique linear $T : \otimes_{i=1}^m V_i \to W$ such that $\psi = T \circ \otimes$, i.e.,

$$\psi(v_1, \ldots, v_m) = T \otimes (v_1, \ldots, v_m) = T v^\otimes.$$

Proof. From Problem 3.1 #.4. ☐

In other words, the tensor map on $\otimes_{i=1}^m V_i$ has **unique factorization property.**

Corollary 3.2.4. Let $\varphi : V_1 \times \cdots \times V_m \to P$ be a multilinear map. Then φ is a tensor map and $\langle \operatorname{Im} \varphi \rangle = P$ if and only if φ has a unique universal factorization property.

Problems

1. Show that if some $v_i = 0$, then $v_1 \otimes \cdots \otimes v_m = 0$.

2. Let $z \in U \otimes V$ so that z can be represented as $z = \sum_{i=1}^k u_i \otimes v_i$. Prove that if k is the smallest natural number among all such representations, then $\{u_1, \ldots, u_k\}$ and $\{v_1, \ldots, v_k\}$ are linear independent sets respectively.

3. Suppose that $e_1, e_2 \in V$ are linearly independent. Prove that $e_1 \otimes e_2 + e_2 \otimes e_1 \in V \otimes V$ is indecomposable.

4. Let $P = \otimes_{i=1}^m V_i$ with tensor map φ and let $T \in \operatorname{Hom}(P, Q)$ be invertible linear map. Prove that $\psi = T \circ \varphi$ is also a tensor map and Q is the tensor space with tensor map ψ.

Solutions to Problems 3.2

1. The tensor map is a multilinear map. So $v_1 \otimes \cdots \otimes v_m = 0$ if $v_i = 0$ for some i.

2. (Daniel and Roy) Without loss of generality, we suppose on the contrary that $\{u_1, \ldots, u_k\}$ is not linearly independent and $u_k = \sum_{i=1}^{k-1} a_i u_i$ for some $a_i \in \mathbb{C}$. Then

$$z = \sum_{i=1}^{k} u_i \otimes v_i = \sum_{i=1}^{k-1} u_i \otimes v_i + \sum_{i=1}^{k-1} (a_i u_i) \otimes v_k = \sum_{i=1}^{k-1} u_i \otimes (v_i + a_i v_k),$$

which implies that the rank of z is no greater than $k-1$, a contradiction.

3. Extend e_1, e_2 to a basis $\{e_1, \ldots, e_n\}$ of V. If $e_1 \otimes e_2 + e_2 \otimes e_1$ were decomposable, then we would have

$$e_1 \otimes e_2 + e_2 \otimes e_1 = (\sum_{i=1}^{n} \alpha_i e_i) \otimes (\sum_{i=1}^{n} \beta_i e_i) = \sum_{i,j} \alpha_i \beta_j e_i \otimes e_j.$$

Since $e_i \otimes e_j$, $i, j = 1, \ldots, n$, are linearly independent, $\alpha_1 \beta_2 = \alpha_2 \beta_1 = 1$ and $\alpha_1 \beta_1 = \alpha_2 \beta_2 = 0$ which is impossible.

4. (Roy) Since φ is multilinear and T is linear, for each $i = 1, \ldots, m$, we have

$$\begin{aligned} & \psi(v_1, \ldots, v_i + a v_i', \ldots, v_m) \\ = \; & T(\varphi(v_1, \ldots, v_i, \ldots, v_m) + a\varphi(v_1, \ldots, v_i', \ldots, v_m)) \\ = \; & T(\varphi(v_1, \ldots, v_i, \ldots, v_m)) + aT(\varphi(v_1, \ldots, v_i', \ldots, v_m)) \\ = \; & \psi(v_1, \ldots, v_i, \ldots, v_m) + a\psi(v_1, \ldots, v_i', \ldots, v_m) \end{aligned}$$

Thus ψ is multilinear. Notice that $\operatorname{rank} \psi = \operatorname{rank} \varphi = \prod_{i=1}^m \dim V_i$, since T is invertible. Therefore, ψ is a tensor map and Q is the corresponding tensor space.

(Daniel) By Theorem 3.1.5, in order to show that ψ is a tensor map, we need to show that ψ has universal factorization property. Let $f : \times_{i=1}^m V_i \to R$ be any multilinear map. Since $\varphi : \times_{i=1}^m V_i \to P$ is a tensor map and $P = \otimes_{i=1}^m V_i$, there is $\bar{f} : P \to R$ such that $\bar{f} \circ \varphi = f$. Set $\tilde{f} = \bar{f} \circ T^{-1} : Q \to R$. Then

$$f = \bar{f} \circ \varphi = \bar{f} \circ T^{-1} \circ T \circ \varphi = \tilde{f} \circ \psi.$$

3.3 Basic properties of tensors and induced inner products

Theorem 3.3.1. If $u_1 \otimes \cdots \otimes u_m + \cdots + v_1 \otimes \cdots \otimes v_m = 0$, then $\varphi(u_1, \ldots, u_m) + \cdots + \varphi(v_1, \ldots, v_m) = 0$ for any multilinear $\varphi : V_1 \times \cdots \times V_m \to W$.

Proof. From Theorem 3.2.3, $\varphi = T\otimes$ for some $T \in \text{Hom}\,(\otimes_{i=1}^m V_i, W)$. So

$$\varphi(u_1, \cdots, u_m) = T \otimes (u_1, \cdots, u_m) = T(u_1 \otimes \cdots \otimes u_m)$$

and similarly $\varphi(v_1, \ldots, v_m) = T(v_1 \otimes \cdots \otimes v_m)$. Hence

$$\begin{aligned} & \varphi(u_1, \ldots, u_m) + \cdots + \varphi(v_1, \ldots, v_m) \\ = \ & T(u_1 \otimes \cdots \otimes u_m) + \cdots + T(v_1 \otimes \cdots \otimes v_m) \\ = \ & T(u_1 \otimes \cdots \otimes u_m + \cdots + v_1 \otimes \cdots \otimes v_m) \\ = \ & 0. \end{aligned}$$

□

Theorem 3.3.2. Let $v_i \in V_i$, $i = 1, \ldots, m$. Then $v_1 \otimes \cdots \otimes v_m = 0$ if and only if some $v_i = 0$.

Proof. One implication is trivial. Suppose $v_1 \otimes \cdots \otimes v_m = 0$ but $v_i \neq 0$ for all i. Then there were $f_i \in V^*$ such that $f_i(v_i) = 1$ for all i. Let

$$\varphi = \prod_{i=1}^m f_i : V_1 \times \cdots \times V_m \to \mathbb{C}$$

which is multilinear. Then

$$\varphi(v_1, \ldots, v_m) = \prod_{i=1}^m f_i(v_i) = 1.$$

By Theorem 3.3.1 $\varphi(v_1, \ldots, v_m) = 0$, a contradiction. □

Theorem 3.3.3. $u_1 \otimes \cdots \otimes u_m = v_1 \otimes \cdots \otimes v_m \neq 0$ if and only if $v_i = c_i u_i \neq 0$ for all $i = 1, \ldots, m$ and $\prod_{i=1}^m c_i = 1$.

Proof. One implication is trivial. Suppose $u_1 \otimes \cdots \otimes u_m = v_1 \otimes \cdots \otimes v_m \neq 0$. From Theorem 3.3.2 all u_i, v_i are nonzero. From Theorem 3.3.1 for all multilinear φ,

$$\varphi(u_1, \ldots, u_m) = \varphi(v_1, \ldots, v_m).$$

Suppose that u_k and v_k are not linearly dependent for some k. Then there is $V_k \in V^*$ such that $f_k(v_k) = 1$ and $f_k(u_k) = 0$. For other $i \neq k$, choose $f_i \in V_i^*$ such that $f(v_i) = 1$. Set $\varphi := \prod_{i=1}^m f_i$. Then

$$\varphi(v_1, \ldots, v_m) = \prod_{i=1}^m f_i(v_i) = 1$$

and

$$\varphi(u_1, \ldots, u_m) = \prod_{i=1}^{m} f_i(u_i) = 0$$

contradicting $\varphi(u_1, \ldots, u_m) = \varphi(v_1, \ldots, v_m)$. Hence $v_k = cu_k$ for all k. From

$$0 \neq u_1 \otimes \cdots \otimes u_m = v_1 \otimes \cdots \otimes v_m = (\prod_{i=1}^{m} c_i) u_1 \otimes \cdots \otimes u_m$$

we have $\prod_{i=1}^{m} c_i = 1$. □

Since the decomposable elements span the tensor space $\otimes_{i=1}^{m} V_i$, each $z \in \otimes_{i=1}^{m} V_i$ is a linear combination of decomposable tensors. Let k be the smallest number of decomposable tensors in all such linear combinations for z. We call k the **smallest length** or the **rank** of z. The smallest length of $z \neq 0$ is one if and only if z is decomposable.

Theorem 3.3.4. Let $z \in U \otimes V$ so that z can be represented as $z = \sum_{i=1}^{k} u_i \otimes v_i$. Then k the smallest length of z if and only if $\{u_1, \ldots, u_k\}$ and $\{v_1, \ldots, v_k\}$ are linear independent sets, respectively.

Proof. One implication follows from Problem 3.2 #2.

For the sufficiency, let $z = \sum_{j=1}^{r} x_j \otimes y_j$ and we are going to show that $k \leq r$. Since $v_1, \ldots, v_k$ are linearly independent, there is $g \in V^*$ such that $g(v_l) = 1$ (l is arbitrary fixed), $g(v_j) = 0$ when $j \neq l$. Let $f \in U^*$ be arbitrary. Then $fg : U \times V \to \mathbb{C}$ is bilinear. Since

$$\sum_{i=1}^{k} u_i \otimes v_i = \sum_{j=1}^{r} x_j \otimes y_j$$

from Theorem 3.3.1 with $\varphi = fg$, we have

$$f(u_l) = \sum_{i=1}^{k} f(u_i) g(v_i) = \sum_{j=1}^{r} f(x_j) g(y_j) = f(\sum_{j=1}^{r} g(y_j) x_j).$$

Since $f \in V^*$ is arbitrary, $u_l = \sum_{j=1}^{r} g(y_j) x_j$, i.e., $u_l \in \langle x_1, \ldots, x_r \rangle$ for all $l = 1, \ldots, k$. Since $u_1, \ldots, u_k$ are linearly independent, we have $k \leq r$. □

We now consider an induced inner product of $\otimes_{i=1}^{m} V_i$. Suppose that $(\cdot, \cdot)_i$ is an inner product on V_i and $E_i = \{e_{i1}, \ldots, e_{in_i}\}$ is an orthonormal basis of V_i for all $i = 1, \ldots, m$. We know that

$$E := \{e_\gamma^\otimes := e_{1\gamma(1)} \otimes \cdots \otimes e_{m\gamma(m)} : \gamma \in \Gamma\}$$

is a basis of $\otimes_{i=1}^{m} V_i$, where $\Gamma = \Gamma(n_1, \ldots, n_m)$. We would like to have an inner product $(\cdot, \cdot)$ on $\otimes_{i=1}^{m} V_i$ such that E is an orthonormal basis, i.e.,

$$(e_\alpha^\otimes, e_\beta^\otimes) = \delta_{\alpha,\beta}.$$

Such inner product is unique from Theorem 1.6.3. From Problem 1.6 #1

$$(u, v) := \sum_{\gamma\in\Gamma} a_\gamma \bar{b}_\gamma \tag{3.3}$$

where $u = \sum_{\gamma\in\Gamma} a_\gamma e^\otimes_\gamma, v = \sum_{\gamma\in\Gamma} b_\gamma e^\otimes_\gamma \in \otimes^m_{i=1} V_i$. It appears depending on the choice of basis but it does not.

Theorem 3.3.5. Let V_i be inner product space with orthonormal basis $E_i = \{e_{i1}, \dots, e_{in_i}\}$, $i = 1, \dots, m$. The inner product obtained from (3.3) satisfies

$$(u_1 \otimes \cdots \otimes u_m, v_1 \otimes \cdots \otimes v_m) = \prod_{i=1}^m (u_i, v_i)_i \tag{3.4}$$

$u_i, v_i \in V_i$ for all i.

Proof. Since $u_i, v_i \in V_i$ we have

$$u_i = \sum_{j=1}^{n_i} a_{ij} e_{ij}, \quad v_i = \sum_{j=1}^{n_i} b_{ij} e_{ij}, \quad i = 1, \dots, m.$$

Since $\otimes$ is multilinear, from (3.1)

$$\begin{aligned} u^\otimes &= \otimes(u_1, \dots, u_m) = \sum_{\gamma\in\Gamma} \prod_{i=1}^m a_{i\gamma(i)} e^\otimes_\gamma \\ v^\otimes &= \otimes(v_1, \dots, v_m) = \sum_{\gamma\in\Gamma} \prod_{i=1}^m b_{i\gamma(i)} e^\otimes_\gamma. \end{aligned}$$

From (3.3) and $\prod_{i=1}^m \sum_{j=1}^{n_i} c_{ij} = \sum_{\gamma\in\Gamma} \prod_{i=1}^m c_{i\gamma(i)}$,

$$\begin{aligned} (u^\otimes, v^\otimes) &= \sum_{\gamma\in\Gamma} \prod_{i=1}^m a_{i\gamma(i)} \bar{b}_{i\gamma(i)} \\ &= \prod_{i=1}^m \sum_{j=1}^{n_i} a_{ij} \bar{b}_{ij} \\ &= \prod_{i=1}^m (\sum_{j=1}^{n_i} a_{ij} e_{ij}, \sum_{j=1}^{n_i} b_{ij} e_{ij})_i \\ &= \prod_{i=1}^m (u_i, v_i)_i. \end{aligned}$$

□

There are many bilinear maps from $(\otimes^k_{i=1} V_i) \times (\otimes^m_{i=k+1} V_i)$ to $\otimes^m_{i=1} V_i$ since

$$\dim \otimes^m_{i=1} V_i = \prod_{i=1}^m n_i = \prod_{i=1}^k n_i \prod_{i=k+1}^m n_i = \dim(\otimes^k_{i=1} V_i) \dim(\otimes^m_{i=k+1} V_i). \tag{3.5}$$

What we like is one that maps $(v_1 \otimes \cdots \otimes v_k, v_{k+1} \otimes \cdots \otimes v_m)$ to $v_1 \otimes \cdots \otimes v_m$.

Theorem 3.3.6. There is a unique multilinear map $\boxtimes : (\otimes_{i=1}^k V_i)\times(\otimes_{i=k+1}^m V_i) \to \otimes_{i=1}^m V_i$ such that

$$\boxtimes (v_1 \otimes \cdots \otimes v_k, v_{k+1} \otimes \cdots \otimes v_m) = v_1 \otimes \cdots \otimes v_m, \tag{3.6}$$

and

$$(V_1 \otimes \cdots \otimes V_k) \boxtimes (V_{k+1} \cdots \otimes V_m) = V_1 \otimes \cdots \otimes V_m. \tag{3.7}$$

Proof. By (3.5) the tensor map $\boxtimes$ that satisfies (3.6) exists and is unique. From $\langle \operatorname{Im} \boxtimes \rangle = \langle v_1 \otimes \cdots \otimes v_m : v_i \in V_i \rangle = \otimes_{i=1}^m V_i$, (3.7) follows. See Problem 3.2 #4 for details. $\square$

We also write $\otimes$ for $\boxtimes$ in Theorem 3.3.6. So (3.7) can be written as

$$(V_1 \otimes \cdots \otimes V_k) \otimes (V_{k+1} \cdots \otimes V_m) = V_1 \otimes \cdots \otimes V_m$$

and (3.6) can be written as

$$(v_1 \otimes \cdots \otimes v_k) \otimes (v_{k+1} \otimes \cdots \otimes v_m) = v_1 \otimes \cdots \otimes v_m.$$

Problems

1. Suppose that $v_1, \dots, v_k \in V$ are linearly independent and $u_1, \dots, u_k \in U$. Prove that $\sum_{i=1}^k u_i \otimes v_i = 0$ if and only if $u_1 = \cdots = u_k = 0$.

2. Let $v_1, \dots, v_k \in V$ and $A \in \mathbb{C}_{k\times k}$. Suppose $AA^T = I_k$ and $u_j = \sum_{i=1}^k a_{ij} v_i$, $j = 1, \dots, k$. Prove that $\sum_{i=1}^k u_i \otimes u_i = \sum_{i=1}^k v_i \otimes v_i$.

3. Define $\otimes : \mathbb{C}^k \times \mathbb{C}^n \to \mathbb{C}_{k\times n}$ by $x \otimes y = xy^T$. Let $\mathbb{C}^k$ and $\mathbb{C}^n$ be equipped with the standard inner products. Prove that for any $A, B \in \mathbb{C}_{k\times n} = \mathbb{C}^k \otimes \mathbb{C}^n$, the induced inner product is given by $(A, B) = \operatorname{tr} B^* A$.

4. Let $E_i = \{e_{i1}, \dots, e_{in_i}\}$ be a basis of V_i, $i = 1, \dots, m$. Define

$$\varphi : (V_1 \otimes \cdots \otimes V_k) \times (V_{k+1} \otimes \cdots \otimes V_m) \to V_1 \otimes \cdots \otimes V_m$$

by $\varphi(e_{1i_1} \otimes \cdots \otimes e_{ki_k}, e_{k+1,i_{k+1}} \otimes \cdots \otimes e_{mi_m}) = e_{1i} \otimes \cdots \otimes e_{mi_m}$ (with bilinear extension). Show that φ is the tensor map satisfying

$$\varphi(v_1 \otimes \cdots \otimes v_k, v_{k+1} \otimes \cdots \otimes v_m) = v_1 \otimes \cdots \otimes v_m.$$

5. Let $z = \sum_{i=1}^k u_i \otimes v_i \otimes w_i \in U \otimes V \otimes W$. Prove that if $\{u_1, \dots, u_k\}$ and $\{v_1, \dots, v_k\}$ are linearly independent and $w_i \neq 0$ for all k, then k is the smallest length of z.

Solutions to Problem 3.3

1. By Theorem 3.3.1 $\sum_{i=1}^k u_i \otimes v_i = 0$ implies that $\sum_{i=1}^k \varphi(u_i, v_i) = 0$ for any multilinear map φ. Since $v_1, \dots, v_k$ are linearly independent, we can choose $f \in V^*$ such that $f(v_i) = 1$ for all i. If some $u_i \neq 0$, then there would be $g \in U^*$ such that $g(u_i) = 1$ and $g(u_j) = 0$ for $j \neq i$. The map $fg : U \times V \to \mathbb{C}$ is multilinear so that $\sum_{i=1}^k fg(u_i, v_i) = 1$, a contradiction.

2.

$$\begin{aligned}
\sum_{j=1}^k u_j \otimes u_j &= \sum_{j=1}^k \left(\left(\sum_{i=1}^k a_{ij} v_i \right) \otimes \left(\sum_{i=1}^k a_{ij} v_i \right) \right) \\
&= \sum_{j=1}^k \left(\sum_{\ell=1}^k \sum_{i=1}^k a_{ij} a_{\ell j} v_i \otimes v_\ell \right) \\
&= \sum_{\ell=1}^k \sum_{i=1}^k \left(\sum_{j=1}^k a_{ij} a_{\ell j} \right) v_i \otimes v_\ell \\
&= \sum_{\ell=1}^k \sum_{i=1}^k \delta_{i\ell} v_i \otimes v_\ell = \sum_{j=1}^k v_j \otimes v_j
\end{aligned}$$

3. (Roy) Suppose $A = x \otimes y$ and $B = z \otimes w$. Then by Theorem 3.3.5, we have

$$(A, B) = (x, z) \cdot (y, w) = \left(\sum_{i=1}^k x_i \overline{z_i} \right) \left(\sum_{j=1}^n y_j \overline{w_j} \right).$$

On the other hand,

$$\operatorname{tr} B^* A = \sum_{i=1}^k \sum_{j=1}^n a_{ij} \overline{b_{ij}} = \sum_{i=1}^k \sum_{j=1}^n x_i y_j \overline{z_i w_j} = \left(\sum_{i=1}^k x_i \overline{z_i} \right) \left(\sum_{j=1}^n y_j \overline{w_j} \right).$$

Thus the induced inner product is given by $(A, B) = \operatorname{tr} B^* A$.

4.

5.

3.4 Induced maps

In this section we study $\operatorname{Hom}(\otimes_{i=1}^m V_i, \otimes_{i=1}^m W_i)$. Let $T_i \in \operatorname{Hom}(V_i, W_i)$. Define the multilinear map from $\times_{i=1}^m V_i$ to $\otimes_{i=1}^m W_i$ by

$$\varphi(v_1, \dots, v_m) := T_1 v_1 \otimes \cdots \otimes T_m v_m.$$

By Theorem 3.2.3, there is a unique $T \in \operatorname{Hom}(\otimes_{i=1}^m V_i, \otimes_{i=1}^m W_i)$ such that

$$T(v_1 \otimes \cdots \otimes v_m) = T_1 v_1 \otimes \cdots \otimes T_m v_m.$$

Denote such T by $T_1 \otimes \cdots \otimes T_m$ and call it the **induced map** of $T_1, \ldots, T_m$, i.e.,

$$(\otimes_{i=1}^m T_i)(v_1 \otimes \cdots \otimes v_m) = T_1 v_1 \otimes \cdots \otimes T_m v_m.$$

We will see that in Section 3.7 that $T_1 \otimes \cdots \otimes T_m$ is a tensor of $\otimes_{i=1}^m \text{Hom}\,(V_i, W_i)$. The main focus in this section is to study $T_1 \otimes \cdots \otimes T_m \in \text{Hom}\,(\otimes_{i=1}^m V_i, \otimes_{i=1}^m W_i)$ as a linear map.

Theorem 3.4.1. Let $S_i \in \text{Hom}\,(W_i, U_i)$, $T_i \in \text{Hom}\,(V_i, W_i)$, $i = 1, \ldots, m$. Then

$$(\otimes_{i=1}^m S_i)(\otimes_{i=1}^m T_i) = \otimes_{i=1}^m (S_i T_i).$$

Proof.

$$\begin{aligned} & (\otimes_{i=1}^m S_i)(\otimes_{i=1}^m T_i)(v_1 \otimes \cdots \otimes v_m) \\ = \;& (\otimes_{i=1}^m S_i)(T_1 v_1 \otimes \cdots \otimes T_m v_m) \\ = \;& S_i T_i v_1 \otimes \cdots \otimes S_m T_m v_m \\ = \;& \otimes_{i=1}^m (S_i T_i)(v_1 \otimes \cdots \otimes v_m). \end{aligned}$$

Since $\otimes_{i=1}^m V_i$ is spanned by the decomposable elements, we have the desired result. □

Theorem 3.4.2. Let $T_i \in \text{Hom}\,(V_i, W_i)$, $i = 1, \ldots, m$. Then

$$\text{rank}\, \otimes_{i=1}^m T_i = \prod_{i=1}^m \text{rank}\, T_i.$$

Proof. Let $\text{rank}\, T_i = k_i$ for all i. So there is a basis $\{e_{i1}, \ldots, e_{ik_i}, e_{ik_i+1}, \ldots, e_{in_i}\}$ for V_i such that $Te_{i1}, \ldots, Te_{ik_i}$ are linearly independent in W_i and $Te_{ik_i+1} = \cdots Te_{in_i} = 0$, $i = 1, \ldots, m$. Notice that $\{e_\gamma^\otimes : \gamma \in \Gamma(n_1, \ldots, n_m)\}$ is a basis for $\otimes_{i=1}^m V_i$. Moreover

$$(\otimes_{i=1}^m T_i) e_\gamma^\otimes = Te_{1\gamma(1)} \otimes \cdots \otimes T_m e_{m\gamma(m)}$$

so that if $\gamma \notin \Gamma(k_1, \ldots, k_m)$, then $(\otimes_{i=1}^m T_i) e_\gamma^\otimes = 0$ (because some $\gamma(i) > k_i$). Since $Te_{i1}, \ldots, Te_{ik_i}$ are linearly independent in W_i for all i, the vectors

$$(\otimes_{i=1}^m T_i) e_\gamma^\otimes, \qquad \gamma \in \Gamma(k_1, \ldots, k_m)$$

are linearly independent in $\otimes_{i=1}^m W_i$ (Why?) Hence

$$\text{rank}\, \otimes_{i=1}^m T_i = |\Gamma(k_1, \ldots, k_m)| = \prod_{i=1}^m k_i = \prod_{i=1}^m \text{rank}\, T_i.$$

□

Theorem 3.4.3. Let $T_i \in \text{Hom}\,(V_i, W_i)$, where V_i, W_i are inner product spaces, $i = 1, \ldots, m$. Then

$$(\otimes_{i=1}^m T_i)^* = \otimes_{i=1}^m T_i^*.$$

Proof. We use the same notation $(\cdot, \cdot)$ for all the inner products on V_i, W_i.

$$\begin{aligned}
(\otimes_{i=1}^m T_i v^{\otimes}, w^{\otimes}) &= (T_1 v_1 \otimes \cdots \otimes T_m v_m, w_1 \otimes \cdots \otimes w_m) \\
&= \prod_{i=1}^m (T_i v_i, w_i) \\
&= \prod_{i=1}^m (v_i, T_i^* w_i) \\
&= (v_1 \otimes \cdots \otimes v_m, T_1^* w_1 \otimes \cdots \otimes T_m^* w_m) \\
&= (v^{\otimes}, \otimes_{i=1}^m T_i^* w^{\otimes})
\end{aligned}$$

Since $\otimes_{i=1}^m V_i$ is spanned by decomposable tensors, we have the desired result. □

Problems

1. Prove that

   (a) $T_1 \otimes \cdots \otimes T_m = 0$ if and only if some $T_i = 0$,

   (b) $T_1 \otimes \cdots \otimes T_m$ is invertible if and only if T_i are all invertible.

2. Let $S_i, T_i \in \text{Hom}\,(V_i, W_i)$, $i = 1, \ldots, m$. Prove that $\otimes_{i=1}^m T_i = \otimes_{i=1}^m S_i \neq 0$ if and only if $T_i = c_i S_i \neq 0$, $i = 1, \ldots, m$ and $\prod_{i=1}^m c_i = 1$.

3. Let $T_i \in \text{End}\, V_i$, $i = 1, \ldots, m$. Prove that be $\otimes_{i=1}^m T_i$ is invertible if and only if T_i $(i = 1, \ldots, m)$ are invertible. In this case $(\otimes_{i=1}^m T_i)^{-1} = \otimes_{i=1}^m T_i^{-1}$.

4. Let $T_i \in \text{Hom}\,(V_i, W_i)$, $i = 1, \ldots, m$. Define

   $$\varphi : \text{Hom}\,(V_i, W_i) \times \cdots \times \text{Hom}\,(V_m, W_m) \to \text{Hom}\,(\otimes_{i=1}^m V_i, \otimes_{i=1}^m W_i)$$

   by $\varphi(T_1, \ldots, T_m) = T_1 \otimes \cdots \otimes T_m$. Prove that φ is multilinear.

Solutions to Problems 3.4

1. (a) Directly from Theorem 3.4.2: $\otimes_{i=1}^m T_i = 0 \Leftrightarrow 0 = \text{rank} \otimes_{i=1}^m T_i = \prod_{i=1}^m \text{rank}\, T_i$, i.e., some $T_i = 0$. The converse is trivial. (b) From Theorem 3.4.2.
   Another approach: use Theorem 3.7.2 and Theorem 3.3.2.

2. (Roy) If $T_i = c_i S_i$ with $\prod_{i=1}^m c_i = 1$, then clearly

$$\begin{aligned} T_1 \otimes \cdots \otimes T_m(v_1 \otimes \cdots \otimes v_m) &= T_1 v_1 \otimes \cdots \otimes T_m v_m \\ &= c_1 S_1 v_1 \otimes \cdots \otimes c_m S_m v_m \\ &= (\prod_{i=1}^m c_i) S_1 v_1 \otimes \cdots \otimes S_m v_m \\ &= S_1 \otimes \cdots \otimes S_m(v_1 \otimes \cdots \otimes v_m) \end{aligned}$$

So $\otimes_{i=1}^m T_i = \otimes_{i=1}^m S_i$. Conversely if $\otimes_{i=1}^m T_i = \otimes_{i=1}^m S_i \neq 0$, then

$$T_1 v_1 \otimes \cdots \otimes T_m v_m = S_1 v_1 \otimes \cdots \otimes S_m v_m \neq 0$$

for all $v_1 \otimes \cdots \otimes v_m \notin \operatorname{Ker} \otimes_{i=1}^m T_i$. In particular, if $v_i \notin \operatorname{Ker} T_i$ for all i, then by Theorem 3.3.2, $v_1 \otimes \cdots \otimes v_m \notin \operatorname{Ker} \otimes_{i=1}^m T_i$. By Theorem 3.3.3 $T_i v_i = c_i S_i v_i$ for all $v_i \notin \operatorname{Ker} T_i$ with $\prod_{i=1}^m c_i = 1$. By symmetry $S_i v_i = c_i' T_i v_i$ for all $v_i \notin \operatorname{Ker} S_i$ with $\prod_{i=1}^m c_i' = 1$. So $T_i = c_i S_i$ in the complement of Clearly $S_i v_i = \frac{1}{c_i} T_i v_i$ So T_i and $c_i S_i$

Another approach: apply Theorem 3.7.2 and Theorem 3.3.3.

3. The first part follows from Theorem 3.4.2. By Theorem 3.4.1

$$(\otimes_{i=1}^m T_i^{-1})(\otimes_{i=1}^m T_i) = \otimes_{i=1}^m (T_i^{-1} T_i) = I.$$

So $(\otimes_{i=1}^m T_i)^{-1} = (\otimes_{i=1}^m T_i^{-1})$.

4.

$$\begin{aligned} & \varphi(T_1, \ldots, T_i + cT_i', \ldots, T_m)(v_1 \otimes \cdots \otimes v_m) \\ = \; & T_1 \otimes \cdots \otimes T_i + cT_i' \otimes \cdots \otimes T_m(v_1 \otimes \cdots \otimes v_m) \\ = \; & T_1 v_1 \otimes \cdots \otimes (T_i + cT_i') v_i \otimes \cdots \otimes T_m v_m \\ = \; & T_1 v_1 \otimes \cdots \otimes (T_i v_i + cT_i' v_i) \otimes \cdots \otimes T_m v_m \\ = \; & T_1 v_1 \otimes \cdots \otimes T_m v_m + cT_1 v_1 \otimes \cdots \otimes T_i' v_i \otimes \cdots \otimes T_m v_m \\ = \; & \varphi((T_1, \ldots, T_i, \ldots, T_m) + c(T_1, \ldots, T_i', \ldots, T_m))(v_1 \otimes \cdots \otimes v_m) \end{aligned}$$

So $\varphi(T_1, \ldots, T_i + cT_i', \ldots, T_m) = \varphi((T_1, \ldots, T_i, \ldots, T_m) + c(T_1, \ldots, T_i', \ldots, T_m))$.

3.5 Matrix representations of induced maps and Kronecker product

Let $E_i = \{e_{i1}, \ldots, e_{in_i}\}$ be a basis of V_i and $F_i = \{f_{i1}, \ldots, f_{ik_i}\}$ be a basis of W_i, $i = 1, \ldots, m$. Then

$$E_\otimes = \{e_\beta^\otimes : \beta \in \Gamma(n_1, \ldots, n_m)\}, \quad F_\otimes = \{f_\alpha^\otimes : \alpha \in \Gamma(k_1, \ldots, k_m)\}$$

are bases for $\otimes_{i=1}^m V_i$ and $\otimes_{i=1}^m W_i$ respectively, where $\Gamma(n_1, \ldots, n_m)$ and $\Gamma(k_1, \ldots, k_m)$ are in lexicographic ordering.

Theorem 3.5.1. Let $T_i \in \operatorname{Hom}(V_i, W_i)$ such that $A_i = [T_i]_{E_i}^{F_i} = (a_{st}^i) \in \mathbb{C}_{k_i \times n_i}$ $i = 1, \dots, m$. That is

$$T_i e_{it} = \sum_{s=1}^{k_i} a_{st}^i f_{is}, \quad t = 1, \dots, n_i, \quad i = 1, \dots, m.$$

Then the matrix representation of $\otimes_{i=1}^m T_i$ with respect to $E_\otimes$ and $F_\otimes$ is given by

$$[\otimes_{i=1}^m T_i]_{E_\otimes}^{F_\otimes} = \left(\prod_{i=1}^m a_{\alpha(i)\beta(i)}^i \right)_{\substack{\alpha \in \Gamma(k_1,\dots,k_m) \\ \beta \in \Gamma(n_1,\dots,n_m)}}$$

In other words, the (α, β) entry of $[\otimes_{i=1}^m T_i]_{E_\otimes}^{F_\otimes}$ is $\prod_{i=1}^m a_{\alpha(i)\beta(i)}^i$.

Proof. For any $\beta \in \Gamma(n_1, \dots, n_m)$, we have

$$\begin{aligned}
\otimes_{i=1}^m T_i e_\beta^\otimes &= T_1 e_{1\beta(1)} \otimes \cdots \otimes T_m e_{m\beta(m)} \\
&= \sum_{s=1}^{k_1} a_{s\beta(1)}^1 f_{1s} \otimes \cdots \otimes \sum_{s=1}^{k_m} a_{s\beta(m)}^m f_{ms} \\
&= \sum_{\alpha \in \Gamma(k_1,\dots,k_m)} a_{\alpha(1)\beta(1)}^1 f_{1\alpha(1)} \otimes \cdots \otimes a_{\alpha(m)\beta(m)}^m f_{m\alpha(m)} \\
&= \sum_{\alpha \in \Gamma(k_1,\dots,k_m)} \left(\prod_{i=1}^m a_{\alpha(i)\beta(i)}^i \right) f_\alpha^\otimes.
\end{aligned}$$

□

Let $A_i = (a_{st}^i) \in \mathbb{C}_{k_i \times n_i}$, $i = 1, \dots, m$. The **Kronecker product** of $A_1, \dots, A_m$, denoted by $A_1 \otimes \cdots \otimes A_m \in \mathbb{C}_{(\prod_{i=1}^m k_i) \times (\prod_{i=1}^m n_i)}$, is defined by

$$(\otimes_{i=1}^m A_i)_{\alpha,\beta} = \prod_{i=1}^m a_{\alpha(i)\beta(i)}^i,$$

where $\alpha \in \Gamma(k_1, \dots, k_m), \beta \in \Gamma(n_1, \dots, n_m)$, both arranged in lexicographic ordering respectively. We immediately have the following

Theorem 3.5.2. Let $A_i = [T_i]_{E_i}^{F_i}$ $i = 1, \dots, m$. Then

$$[\otimes_{i=1}^m T_i]_{E_\otimes}^{F_\otimes} = \otimes_{i=1}^m [T_i]_{E_i}^{F_i} = \otimes_{i=1}^m A_i.$$

Example 3.5.3. Let $A \in \mathbb{C}_{m \times n}$ and $B \in \mathbb{C}_{p \times q}$. Then

$$A \otimes B = \begin{pmatrix} a_{11}B & \dots & a_{1n}B \\ \cdots & \cdots & \cdots \\ a_{m1}B & \cdots & a_{mn}B \end{pmatrix} = (a_{ij}B) \in \mathbb{C}_{mp \times nq}$$

In general $A \otimes B \neq B \otimes A$.

Notice that

$$\Gamma(m,p) = \{(1,1),\ldots,(1,p),(2,1),\ldots,(2,p),\cdots,(m,1),\ldots,(m,p)\}$$

and

$$\Gamma(n,q) = \{(1,1),\ldots,(1,q),(2,1),\ldots,(2,q),\cdots,(n,1),\ldots,(n,q)\}.$$

Consider $\alpha = (1,i) \in \Gamma(m,p)$, $\beta = (1,j) \in \Gamma(n,q)$. Then

$$(A \otimes B)_{(1,i),(1,j)} = a_{\alpha(1)\beta(1)} b_{\alpha(2)\beta(2)} = a_{11} b_{ij}.$$

That is, the $(A \otimes B)[1,\ldots,p|1,\ldots,q] = a_{11}B$. Similar for other blocks.

Problems

1. Let $A_i, B_i \in \mathbb{C}_{k_i,n_i}$, $i = 1,\ldots,m$. Prove the following:
   (i) $\otimes_{i=1}^m A_i = 0$ if and only if some $A_i = 0$.
   (ii) $\otimes_{i=1}^m A_i = \otimes_{i=1}^m B_i \neq 0$ if and only if $A_i = c_i B_i \neq 0$ for all i and $\prod_{i=1}^m c_i = 1$.
   (iii) $\operatorname{rank}(\otimes_{i=1}^m A_i) = \prod_{i=1}^m \operatorname{rank} A_i$.
   (iv) $(\otimes_{i=1}^m A_i)^* = \otimes_{i=1}^m A_i^*$.

2. From the definition of Kronecker product prove that $(\otimes_{i=1}^m A_i)(\otimes_{i=1}^m B_i) = \otimes_{i=1}^m A_i B_i$. (Hint:

$$\sum_{\gamma \in \Gamma(t_1,\ldots,t_m)} (\prod_{i=1}^m a^i_{\alpha(i)\gamma(i)})(\prod_{i=1}^m b^i_{\gamma(i)\beta(i)}) = \prod_{i=1}^m (\sum_{j=1}^{t_i} a^i_{\alpha(i)j} b^i_{j\beta(i)})$$

   where $A_k = (a^k_{ij})$).

3. Show that if $A_i \in \mathbb{C}_{n_i \times n_i}$, $i = 1,\ldots,m$, are triangular (diagonal, resp.) matrices, then $A_1 \otimes \cdots \otimes A_m$ is also triangular (diagonal, resp.). In particular the (α,α)-entry of $A_1 \otimes \cdots \otimes A_m$ is $\prod_{i=1}^m a^i_{\alpha(i)\alpha(i)}$.

4. Show that if $A_i, B_i \in \mathbb{C}_{n_i \times n_i}$ are similar for all $i = 1,\ldots,m$, then $\otimes_{i=1}^m A_i$ and $\otimes_{i=1}^m B_i$ are similar.

5. Which one of $x \otimes y = xy^T$ and $x \boxtimes y = yx^T$, $x \in \mathbb{C}^m$ and $y \in \mathbb{C}^n$ is the Kronecker product? Explain.

6. (Merris) Show that for any $A, B \in \mathbb{C}_{n\times n}$, (1) If $AB = BA$, then $e^{A+B} = e^A e^B$ and (2) $e^A \otimes e^B = e^{A\otimes I + \otimes B}$, where $e^A := \sum_{k=1}^\infty A^k/k$.

Solutions to Problems 3.5

1. (i) From Problem 3.4 #1 and Theorem 3.5.2

   (Roy) If $A_i = 0$, then $\otimes_{i=1}^m A_i = 0$ because every entry of $\otimes_{i=1}^m A_i$ has a factor of some entry of A_i.

   ($\Rightarrow$). Suppose on the contrary that $A_i \neq 0$ for all i. Then each A_i contains some nonzero entry $a^i_{s_i t_i}$. Let $\alpha \in \Gamma(k_1, \ldots, k_m)$ and $\beta \in \Gamma(n_1, \ldots, n_m)$ be such that $\alpha(i) = s_i$ and $\beta(i) = t_i$ for all i. It follows that the (α, β) entry of $\otimes_{i=1}^m A_i$ is nonzero, which means $\otimes_{i=1}^m A_i \neq 0$.

   (ii) From Problem 3.4 #2 and Theorem 3.5.2

   (iii) From Theorem 3.4.2 and Theorem 3.5.2

   (iv) From Theorem 3.4.3 and Theorem 3.5.2

2. Let $A = (a^i_{st})$ and $B = (b^i_{kl})$. The (α, β)-entry of $(\otimes_{i=1}^m A_i)(\otimes_{i=1}^m B_i)$ is

$$\begin{aligned} & \sum_{\gamma \in \Gamma(t_1, \ldots, t_m)} (\prod_{i=1}^m a^i_{\alpha(i)\gamma(i)})(\prod_{i=1}^m b^i_{\gamma(i)\beta(i)}) \\ = & \sum_{\gamma \in \Gamma(t_1, \ldots, t_m)} \prod_{i=1}^m (a^i_{\alpha(i)\gamma(i)}) b^i_{\gamma(i)\beta(i)}) \\ = & \prod_{i=1}^m (\sum_{j=1}^{t_i} a^i_{\alpha(i)j} b^i_{j\beta(i)}) \end{aligned}$$

   which is the (α, β)-entry of $\otimes_{i=1}^m A_i B_i$.

3. Recall that

$$(\otimes_{i=1}^m A_i)_{\alpha,\beta} = \prod_{i=1}^m a^i_{\alpha(i)\beta(i)},$$

   where $\alpha \in \Gamma(k_1, \ldots, k_m), \beta \in \Gamma(n_1, \ldots, n_m)$, both arranged in lexicographic ordering respectively. When $\alpha < \beta$ in lexicographic ordering, $\alpha(i) < \beta(i)$ for some i (remark: converse is not true) so that $a^i_{\alpha(i)\beta(i)} = 0$, i.e., $(\otimes_{i=1}^m A_i)_{\alpha,\beta} = 0$. In other words, $(\otimes_{i=1}^m A_i)$ is upper triangular.

4. Suppose $B_i = P_i A_i P_i^{-1}$ for all i. Then from Problem 2

$$\begin{aligned} \otimes_{i=1}^m B_i &= \otimes_{i=1}^m (P_i A_i P_i^{-1}) = (\otimes_{i=1}^m P_i)(\otimes_{i=1}^m A_i)(\otimes_{i=1}^m P_i^{-1}) \\ &= (\otimes_{i=1}^m P_i)(\otimes_{i=1}^m A_i)(\otimes_{i=1}^m P_i)^{-1}. \end{aligned}$$

5. (Roy) None of them is the Kronecker product since $x \otimes y \in \mathbb{C}_{mn \times 1}$.

6. (Roy) Let $E = \{e_1, e_2, \ldots, e_n\}$ be a basis of $\mathbb{C}^n$ so that $[A]_E^E$ is upper triangular. Since $[I]_E^E$ is also upper triangular, both $A \otimes I_n$ and $I_n \otimes A$ are upper triangular with respect to the basis $E \otimes E$ in the lexicographic order by Theorem 3.6.1(a). Therefore, we may assume, without loss of generality, that A is upper triangular with diagonal entries $\lambda_1, \ldots, \lambda_n$ in order. The

diagonal entries of $A \otimes I_n$ are $\lambda_1, \ldots, \lambda_1, \lambda_2, \ldots, \lambda_2, \ldots, \lambda_n, \ldots, \lambda_n$ in order, and the diagonal entries of $I_n \otimes A$ are $\lambda_1, \ldots, \lambda_n, \lambda_1, \ldots, \lambda_n, \ldots, \lambda_1, \ldots, \lambda_n$ in order. Thus the eigenvalues of $A \otimes I_n - I_n \otimes A$ are $\lambda_i - \lambda_j$, $i, j = 1, \ldots, n$.

3.6 Induced operators

We now study some basic properties of the induced map $T_1 \otimes \cdots \otimes T_m$.

Theorem 3.6.1. Let $T_i \in \operatorname{End} V_i$ and let $A_i = [T_i]_{E_i}^{E_i}$, where E_i is a basis of V_i, $i = 1, \ldots, m$.

(a) Suppose that A_i is upper triangular with $\operatorname{diag} A_i = (\lambda_{i1}, \ldots, \lambda_{in_i})$. Then $[\otimes_{i=1}^m T_i]_{E_\otimes}^{F_\otimes} = \otimes_{i=1}^m A_i$ is also upper triangular with (γ, γ)-diagonal entry $\prod_{i=1}^m \lambda_{i\gamma(i)}$, $\gamma \in \Gamma$.

(b) If λ_{ij}, $j = 1, \ldots, n_i$, are the eigenvalues of T_i, then the eigenvalues of $\otimes_{i=1}^m T_i$ are $\prod_{i=1}^m \lambda_{i\gamma(i)}$, $\gamma \in \Gamma$.

(c) If s_{ij}, $j = 1, \ldots, n_i$, are the singular values of T_i, then the singular values of $\otimes_{i=1}^m T_i$ are $\prod_{i=1}^m s_{i\gamma(i)}$, $\gamma \in \Gamma$.

(d) $\operatorname{tr} \otimes_{i=1}^m T_i = \prod_{i=1}^m \operatorname{tr} T_i$.

(e) $\det \otimes_{i=1}^m T_i = \prod_{i=1}^m (\det T_i)^{n/n_i}$, where $n := \prod_{i=1}^m n_i$.

Proof. (a) Use Problem 3.4 #3 for matrix approach.

The following is linear operator approach. Since A_i is upper triangular, For any $\gamma \in \Gamma(n_1, \ldots, n_m)$, we have

$$T_i e_{i\gamma(i)} = \lambda_{i\gamma(i)} e_{i\gamma(i)} + v_i,$$

where $v_i \in \langle e_{ij} : j < \gamma(i) \rangle$. Then

$$\begin{aligned} (\otimes_{i=1}^m T_i) e_\gamma^\otimes &= T_1 e_{1\gamma(1)} \otimes \cdots \otimes T_m e_{m\gamma(m)} \\ &= (\lambda_{1\gamma(1)} e_{1\gamma(1)} + v_1) \otimes \cdots \otimes (\lambda_{m\gamma(m)} e_{m\gamma(m)} + v_m) \\ &= \prod_{i=1}^m \lambda_{i\gamma(i)} e_\gamma^\otimes + w \end{aligned}$$

where w is a linear combination of $u_1 \otimes \cdots \otimes u_m$. Each u_i is either $e_{i\gamma(i)}$ or v_i and at least one is v_i. But v_i is a linear combination of e_{ij}, $j < \gamma(i)$. So $u^\otimes$ and thus w is a linear combination of $e_\alpha^\otimes$ such that α is before γ in lexicographic order. So $[\otimes_{i=1}^m T_i]_{E_\otimes}^{F_\otimes}$ is upper triangular.

(b) By Schur's triangularization theorem we may assume that the matrix representation of A_i is upper triangular for all i (cf. Theorem 3.4.1 and Problem 3.4 #3). Then use (a).

(c) The singular values of A are the eigenvalue nonnegative square roots of A^*A. Apply (b) on A^*A.

(d) $\operatorname{tr}(\otimes_{i=1}^m T_i) = \sum_{\gamma\in\Gamma}\prod_{i=1}^m \lambda_{i\gamma(i)} = \prod_{i=1}^m \sum_{j=1}^{n_i} \lambda_{ij} = \prod_{i=1}^m \operatorname{tr} T_i$.
(e) Since determinant is equal to the product of eigenvalues,

$$\det \otimes_{i=1}^m T_i = \prod_{\gamma\in\Gamma}\prod_{i=1}^m \lambda_{i\gamma(i)} = \prod_{i=1}^m\prod_{\gamma\in\Gamma} \lambda_{i\gamma(i)}$$

Since $\gamma(i)$ is chosen from $1,\dots,n_i$ and such $\gamma(i)$ appears in $|\Gamma|/n_i$ times in those $n := |\Gamma|$ elements in Γ. So

$$\prod_{\gamma\in\Gamma}\lambda_{i\gamma(i)} = \prod_{j=1}^{n_i}\lambda_{ij}^{|\Gamma|/n_i} = \prod_{j=1}^{n_i}\lambda_{ij}^{n/n_i}$$

and hence

$$\det(\otimes_{i=1}^m T_i) = \prod_{i=1}^m\left(\prod_{j=1}^{n_i}\lambda_{ij}\right)^{n/n_i} = \prod_{i=1}^m (\det T_i)^{n/n_i}.$$

□

Theorem 3.6.2. Let $T_i \in \operatorname{End} V_i$ where V_i are inner product spaces, $i = 1,\dots,m$. If all T_i are (a) normal, (b) Hermitian, (c) psd, (d) pd, and (e) unitary respectively, so is $\otimes_{i=1}^m T_i$.

Proof. Recall from Theorem 3.4.1 $(\otimes_{i=1}^m S_i)(\otimes_{i=1}^m T_i) = \otimes_{i=1}^m (S_iT_i)$ and Theorem 3.4.3 $\otimes_{i=1}^m T_i^* = (\otimes_{i=1}^m T_i)^*$.
(a) If all T_i are normal, then

$$(\otimes_{i=1}^m T_i)^*(\otimes_{i=1}^m T_i) = (\otimes_{i=1}^m T_i^*)(\otimes_{i=1}^m T_i) = \otimes_{i=1}^m T_i^*T_i = \otimes_{i=1}^m T_iT_i^* = (\otimes_{i=1}^m T_i)(\otimes_{i=1}^m T_i)^*.$$

So $\otimes_{i=1}^m T_i$ is normal.
(b) If all T_i are Hermitian, then

$$(\otimes_{i=1}^m T_i)^* = \otimes_{i=1}^m T_i^* = \otimes_{i=1}^m T_i$$

i.e., $\otimes_{i=1}^m T_i$ is Hermitian.
(c) and (d) If all T_i are psd (pd), then $\otimes_{i=1}^m T_i$ is Hermitian and by Theorem 3.6.1 all eigenvalues of $\otimes_{i=1}^m T_i$ are nonnegative (positive). So $\otimes_{i=1}^m T_i$ is psd (pd).
(e) If all T_i are unitary, then

$$(\otimes_{i=1}^m T_i)^*(\otimes_{i=1}^m T_i) = \otimes_{i=1}^m T_i^*T_i = \otimes_{i=1}^m I_{V_i} = I_{\otimes_{i=1}^m V_i}.$$

□

We now study the converse of Theorem 3.6.2.

Theorem 3.6.3. Let $T_i \in \operatorname{End} V_i$ where V_i are inner product spaces, $i = 1,\dots,m$.

(a) If $\otimes_{i=1}^m T_i \neq 0$ is normal, then all $T_i \neq 0$ are normal.

(b) If $\otimes_{i=1}^m T_i \neq 0$ is Hermitian, then $T_i = c_i H_i$ where H_i are Hermitan and $\prod_{i=1}^m c_i \in \mathbb{R}$.

Proof. (a) From Problem 3.4 #1, $T_i \neq 0$. There is an orthonormal basis $E_i = \{e_{i1}, \dots, e_{in_i}\}$ such that $[T_i]_{E_i}^{E_i}$ is upper triangular. Let $E_\otimes = \{e_\gamma^\otimes : \gamma \in \Gamma\}$ which is an orthonormal basis for $\otimes_{i=1}^m V_i$. From Theorem 3.6.1, $[\otimes_{i=1}^m T_i]_{E_\otimes}^{E_\otimes}$ is upper triangular. On the other hand $[\otimes_{i=1}^m T_i]_{E_\otimes}^{E_\otimes}$ is a normal matrix. By Lemma 1.5.3 $[\otimes_{i=1}^m T_i]_{E_\otimes}^{E_\otimes}$ must be diagonal. now

$$[\otimes_{i=1}^m T_i]_{E_\otimes}^{E_\otimes} = [T_1]_{E_1}^{E_1} \otimes \cdots \otimes [T_m]_{E_m}^{E_m} \neq 0.$$

From Problem 3.5 #3, each $[T_i]_{E_i}^{E_i}$ is diagonal. Hence T_i is normal for all i.

(b) From (a) each $T_i \neq 0$ and is normal. So some eigenvalue of T_i is nonzero, i.e., there exists some $\alpha \in \Gamma$ such that the eigenvalues of T_i are $\lambda_{i\alpha(i)} \neq 0$ for all i. Consider any fixed T_k. Since the eigenvalues of the Hermitian $\otimes_{i=1}^m T_i$ are real, from Theorem 3.6.1(b)

$$(\prod_{i=1, i\neq k}^m \lambda_{i\alpha(i)})\lambda_{kj} \in \mathbb{R}, \quad j = 1, \dots, n_k$$

since they are (not all) eigenvalues of $\otimes_{i=1}^m T_i \neq 0$ and are nonzero. Let

$$a_k := \prod_{i\neq k}^m \lambda_{i\alpha(i)} \neq 0$$

since $\lambda_{k\alpha(k)} \neq 0$. Then $H_k := a_k T_k$ is normal with real eigenvalues, i.e., H_k is Hermitian. Set $c_k := a_k^{-1}$. Then

$$0 \neq \otimes_{i=1}^m T_i = (\prod_{k=1}^m c_k)(\otimes_{k=1}^m H_k).$$

Since $\otimes_{i=1}^m T_i$ and $\otimes_{k=1}^m H_k$ are Hermitian, we have $\prod_{k=1}^m c_k \in \mathbb{R}$.

□

Problems

1. Prove that if $\otimes_{i=1}^m T_i \neq 0$ is pd (psd), then $T_i = c_i H_i$ and H_i is pd (psd) and $\prod_{i=1}^m c_i > 0$.

2. Prove that $\otimes_{i=1}^m T_i \neq 0$ is unitary if and only if $T_i = c_i U_i$ and U_i is unitary and $|\prod_{i=1}^m c_i| = 1$.

3. Suppose $S_i, T_i \in \operatorname{End} V_i$, $i = 1, \ldots, m$, are psd. Show that $\otimes_{i=1}^m (S_i + T_i) \geq \otimes_{i=1}^m S_i + \otimes_{i=1}^m T_i$.

4. Show that if the eigenvalues of $A \in \mathbb{C}_{n\times n}$ are $\lambda_1, \ldots, \lambda_n$, then the eigenvalues of $A \otimes I_n - I_n \otimes A$ are $\lambda_i - \lambda_j$, $i, j = 1, \ldots, n$.

5. Let $T_i \in \operatorname{End} V$ and $v_i \in V$, $i = 1, \ldots, m$ where $m \leq \dim V$. Prove that $(\otimes_{i=1}^m T_i v^\otimes, v^\otimes) = 0$ for all $(v_i, v_j) = \delta_{ij}$ if and only if some $T_i = 0$.

Solutions to Problems 3.6

1. By Theorem 3.6.2 and Theorem 3.6.3(b).

2. Suppose that $\otimes_{i=1}^m T_i \neq 0$ is unitary. By Theorem 3.6.2, all T_i are normal and by Problem 3.4 #3, all T_i are invertible. The converse is easy by Theorem 3.4.1 and Theorem 3.4.3.

3.

4. By Schur's triangularization theorem, there is unitary matrix U such that $U^*AU = T$ is upper triangular with $\operatorname{diag} T = (\lambda_1, \ldots, \lambda_n)$. The matrix $U \otimes I$ is unitary by Problem 2. So

$$\begin{aligned} & (U^* \otimes I)(A \otimes I_n - I_n \otimes A)(U \otimes I) \\ = \; & (U \otimes I)^*(A \otimes I_n - I_n \otimes A)(U \otimes I) \\ = \; & T \otimes I - I \otimes T \end{aligned}$$

whose spectrum is the spectrum of $A \otimes I_n - I_n \otimes A$ counting multiplicities. From Theorem 3.6.1(a), the eigenvalues of $T \otimes I - I \otimes T$ are $\lambda_i - \lambda_j$ for all i, j. In particular, $A \otimes I_n - I_n \otimes A$ is not invertible.

5.

3.7 Tensor product of linear maps

Given $T_i \in \operatorname{Hom}(V_i, W_i)$, $i = 1, \ldots, m$, we studied basic properties of the induced map $\otimes_{i=1}^m T_i \in \operatorname{Hom}(\otimes_{i=1}^m V_i, \otimes_{i=1}^m W_i)$. Not every member of $\operatorname{Hom}(\otimes_{i=1}^m V_i, \otimes_{i=1}^m W_i)$ is of the form $\otimes_{i=1}^m T_i$ (see Problem 1), but the span of all the induced maps is $\operatorname{Hom}(\otimes_{i=1}^m V_i, \otimes_{i=1}^m W_i)$.

Theorem 3.7.1. $\operatorname{Hom}(\otimes_{i=1}^m V_i, \otimes_{i=1}^m W_i) = \langle \otimes_{i=1}^m T_i : T_i \in \operatorname{Hom}(V_i, W_i) \rangle$.

Proof. Let $E_i = \{e_{i1}, \ldots, e_{in_i}\}$ be a basis of V_i and $F_i = \{f_{i1}, \ldots, f_{ik_i}\}$ be a basis of W_i, $i = 1, \ldots, m$. Then

$$E_\otimes = \{e_\beta^\otimes : \beta \in \Gamma(n_1, \ldots, n_m)\}, \quad F_\otimes = \{f_\alpha^\otimes : \alpha \in \Gamma(k_1, \ldots, k_m)\}$$

are bases for $\otimes_{i=1}^m V_i$ and $\otimes_{i=1}^m W_i$ respectively. Set

$$\Gamma_1 := \Gamma(n_1, \ldots, n_m), \quad \Gamma_2 := \Gamma(k_1, \ldots, k_m)$$

and both are in lexicographic ordering. Let $S \in \operatorname{Hom}(\otimes_{i=1}^m V_i, \otimes_{i=1}^m W_i)$ be arbitrary. For any $\beta \in \Gamma_1$,

$$Se_\beta^\otimes = \sum_{\alpha\in\Gamma_2} c_{\alpha\beta} f_\alpha^\otimes.$$

Define $T_{jq}^i \in \operatorname{Hom}(V_i, W_i)$ such that

$$T_{jq}^i e_{ip} = \delta_{pq} f_{ij},$$

where $p, q = 1, \ldots, n_i$, $j = 1, \ldots k_i$, $i = 1, \ldots, m$. For any $\alpha \in \Gamma_2$, $\gamma \in \Gamma_1$, set

$$T_{\alpha\gamma}^\otimes := T_{\alpha(1)\gamma(1)}^1 \otimes \cdots \otimes T_{\alpha(m)\gamma(m)}^m.$$

Then

$$\begin{aligned} T_{\alpha\gamma}^\otimes e_\beta^\otimes &= T_{\alpha(1)\gamma(1)}^1 e_{1\beta(1)} \otimes \cdots \otimes T_{\alpha(m)\gamma(m)}^m e_{m\beta(m)} \\ &= \delta_{\beta(1)\gamma(1)} f_{1\alpha(1)} \otimes \cdots \otimes \delta_{\beta(m)\gamma(m)} f_{m\alpha(m)} \\ &= \delta_{\beta\gamma} f_\alpha^\otimes. \end{aligned}$$

So

$$\left(\sum_{\alpha\in\Gamma_2}\sum_{\gamma\in\Gamma_1} c_{\alpha\gamma} T_{\alpha\gamma}^\otimes\right) e_\beta^\otimes = \sum_{\alpha\in\Gamma_2}\sum_{\gamma\in\Gamma_1} c_{\alpha\gamma}\delta_{\beta\gamma} f_\alpha^\otimes = \sum_{\alpha\in\Gamma_2} c_{\alpha\beta} f_\alpha^\otimes = Se_\beta^\otimes.$$

Since $\{e_\beta^\otimes : \beta \in \Gamma_1\}$ is a basis of $\otimes_{i=1}^m V_i$ we have

$$S = \sum_{\alpha\in\Gamma_2}\sum_{\gamma\in\Gamma_1} c_{\alpha\gamma} T_{\alpha\gamma}^\otimes.$$

In other words, S is a linear combination of induced maps $T_{\alpha\gamma}^\otimes$ ($\alpha \in \Gamma_2, \gamma \in \Gamma_1$). □

We now see that $T_1 \otimes \cdots \otimes T_m$ can be viewed as a (decomposable) tensor when each $T_i \in \operatorname{Hom}(V_i, W_i)$ is viewed as a vector, i.e., we have the following result.

Theorem 3.7.2. There exist a unique tensor map $\boxtimes$ such that $\operatorname{Hom}(\otimes_{i=1}^m V_i, \otimes_{i=1}^m W_i)$ is the tensor space of $\operatorname{Hom}(V_1, W_1), \ldots, \operatorname{Hom}(V_m, W_m)$, i.e.,

$$\operatorname{Hom}(\otimes_{i=1}^m V_i, \otimes_{i=1}^m W_i) = \operatorname{Hom}(V_1, W_1) \boxtimes \cdots \boxtimes \operatorname{Hom}(V_m, W_m).$$

In other words, $T_1 \otimes \cdots \otimes T_m = \boxtimes(T_1, \ldots, T_m)$.

Proof. From Problem 3.4 #4, the map $\boxtimes(T_1, \dots, T_m) = T_1 \otimes \cdots \otimes T_m \in \operatorname{Hom}(\otimes_{i=1}^m V_i, \otimes_{i=1}^m W_i)$ is multilinear and is unique.

From Theorem 3.7.1, $\operatorname{Hom}(\otimes_{i=1}^m V_i, \otimes_{i=1}^m W_i)$ is spanned by the induced maps $\otimes_{i=1}^m T_i$, $T_i \in \operatorname{Hom}(V_i, W_i)$, i.e.,

$$\langle \operatorname{Im} \boxtimes \rangle = \langle \otimes_{i=1}^m T_i : T_i \in \operatorname{Hom}(V_i, W_i) \rangle = \operatorname{Hom}(\otimes_{i=1}^m V_i, \otimes_{i=1}^m W_i).$$

Moreover

$$\begin{aligned} \dim \langle \operatorname{Im} \boxtimes \rangle &= \dim \operatorname{Hom}(\otimes_{i=1}^m V_i, \otimes_{i=1}^m W_i) \\ &= (\dim \otimes_{i=1}^m V_i)(\dim \otimes_{i=1}^m W_i) \\ &= \prod_{i=1}^m \dim V_i \dim W_i \\ &= \prod_{i=1}^m \dim \operatorname{Hom}(V_i, W_i). \end{aligned}$$

So $\boxtimes$ is a tensor map and

$$\operatorname{Hom}(V_1, W_1) \boxtimes \cdots \boxtimes \operatorname{Hom}(V_m, W_m) = \operatorname{Hom}(\otimes_{i=1}^m V_i, \otimes_{i=1}^m W_i),$$

i.e., $\operatorname{Hom}(\otimes_{i=1}^m V_i, \otimes_{i=1}^m W_i)$ is a tensor product of $\operatorname{Hom}(V_1, W_1), \dots, \operatorname{Hom}(V_m, W_m)$. □

Similarly we can view $\mathbb{C}_{\prod_{i=1}^m k_i \times \prod_{i=1}^m n_m}$ as the tensor product of $\mathbb{C}_{k_1 \times n_1}, \dots, \mathbb{C}_{k_m \times n_m}$ such that the Kronecker products $A_1 \otimes \cdots \otimes A_m$ are the decomposable tensors.

Problems

1. Define the bilinear map $\varphi : V \times V \to V \otimes V$ by $\varphi(v_1, v_2) = v_2 \otimes v_1$ so that there is $T \in \operatorname{End}(V \otimes V)$ such that $T\otimes = \varphi$, i.e., $T(v_1, v_2) = v_2 \otimes v_1$. Prove that when $\dim V = 1$, T is not an induced map, i.e., there are no $T_1, T_2 \in \operatorname{End}(V)$ such that $T = T_1 \otimes T_2$.

2. Let $\{e_1, \dots, e_n\}$ be a basis for V. Define $T_{ij} \in \operatorname{End} V$ such that $T_{ij} e_k = \delta_{kj} e_i$. Let $T = \sum_{i=1}^n \sum_{j=1}^n T_{ij} \otimes T_{ji}$. Show that $T(v_1 \otimes v_2) = v_2 \otimes v_1$ for all $v_1, v_2 \in V$ (cf. Theorem 3.7.1 and notice that the T in Problem 1 is a linear combination of induced maps).

3. Prove that in the proof of Theorem 3.7.1

   (a) (Daniel) $\{T^i_{jp} : j = 1, \dots, k_i, q = 1, \dots, n_i\}$ is a basis of $\operatorname{Hom}(V_i, W_i)$.

   (b) $\{T^{\otimes}_{\alpha\gamma} : \alpha \in \Gamma_2, \gamma \in \Gamma_1\}$ is a basis of $\operatorname{Hom}(\otimes_{i=1}^m V_i, \otimes_{i=1}^m W_i)$.

4. Let V_i be an inner product space, $i = 1, \dots, m$. Equip $\otimes_{i=1}^m V_i$ with the inner product. Let $T \in \operatorname{End}(\otimes_{i=1}^m V_i)$. Prove that $(Tv^{\otimes}, v^{\otimes}) = 0$ for all $v^{\otimes} \in \otimes_{i=1}^m V_i$ if and only if $T = 0$. (Hint: Show that $(Tu^{\otimes}, v^{\otimes}) = 0$).

Solutions to Problems 3.7

1.

2.

3.

4.

3.8 Some models of tensor products

Let $M(V, \ldots, V; \mathbb{C})$ denote the space of all m-multilinear maps $f : \times^m V \to \mathbb{C}$. Let $E = \{e_1, \ldots, e_n\}$ be a basis of V and let $E^* = \{f_1, \ldots, f_n\}$ be the dual basis of V^*, i.e.,

$$f_i(e_j) = \delta_{ij}, \quad i, j = 1, \ldots, n.$$

Theorem 3.8.1. 1. $\{\prod_{i=1}^m f_{\alpha(i)} : \alpha \in \Gamma_{m,n}\}$ is a basis of $M(V, \ldots, V; \mathbb{C})$.

2. $M(V, \ldots, V; \mathbb{C}) = \otimes^m V^*$, i.e., there is tensor map $\otimes : V^* \times \cdots \times V^* \to M(V, \ldots, V; \mathbb{C})$ and $\dim M(V, \ldots, V; \mathbb{C}) = n^m$.

3. $M(V^*, \ldots, V^*; \mathbb{C}) = \otimes^m V$, i.e., there is tensor map $\otimes : V \times \cdots \times V \to M(V^*, \ldots, V^*; \mathbb{C})$ and $\dim M(V^*, \ldots, V^*; \mathbb{C}) = n^m$.

Proof. (1) We first show that $S := \{\prod_{i=1}^m f_{\alpha(i)} : \alpha \in \Gamma_{m,n}\}$ spans $M(V, \ldots, V; \mathbb{C})$. First observe that for each $e_\beta = (e_{\beta(1)}, \ldots, e_{\beta(m)}) \in V \times \cdots \times V$ and $\beta \in \Gamma_{m,n}$,

$$(\prod_{i=1}^m f_{\alpha(i)}) e_\beta = \prod_{i=1}^m f_{\alpha(i)}(e_{\beta(i)}) = \delta_{\alpha\beta}. \tag{3.8}$$

Let $f \in M(V, \ldots, V; \mathbb{C})$. Then we claim

$$f = \sum_{\alpha \in \Gamma_{m,n}} f(e_\alpha) \prod_{i=1}^m f_{\alpha(i)}$$

where $e_\alpha = (e_{\alpha(1)}, \ldots, e_{\alpha(m)})$. It is because from (3.8)

$$\left(\sum_{\alpha \in \Gamma_{m,n}} f(e_\alpha) \prod_{i=1}^m f_{\alpha(i)} \right) e_\beta = \sum_{\alpha \in \Gamma_{m,n}} f(e_\alpha) \delta_{\alpha\beta} = f(e_\beta), \quad \beta \in \Gamma_{m,n}.$$

We now show that S is a linearly independent set. Set

$$\sum_{\alpha \in \Gamma_{m,n}} c_\alpha \prod_{i=1}^m f_{\alpha(i)} = 0.$$

Then

$$0 = \left(\sum_{\alpha \in \Gamma_{m,n}} c_\alpha \prod_{i=1}^m f_{\alpha(i)} \right) (e_\beta) = c_\beta, \quad \beta \in \Gamma_{m,n}.$$

(2) It is easy to see that the map $\otimes : \times^m V^* \to M(V, \ldots, V; \mathbb{C})$ defined by

$$\otimes(g_1, \ldots, g_m) = \prod_{i=1}^m g_i$$

is multilinear. From (1) $\{\prod_{i=1}^m f_{\alpha(i)} : \alpha \in \Gamma_{m,n}\}$ is a basis of $M(V, \ldots, V; \mathbb{C})$. So

$$\dim M(V, \ldots, V; \mathbb{C}) = |\Gamma_{m,n}| = n^m = (\dim V)^m = (\dim V^*)^m.$$

So $\otimes$ is a tensor map and $\langle \operatorname{Im} \otimes \rangle = M(V, \ldots, V; \mathbb{C})$, i.e., $M(V, \ldots, V; \mathbb{C}) = \otimes^m V^*$.

(3) Similar to (2). Indeed we can define $\otimes : \times^m V \to M(V^*, \ldots, V^*; \mathbb{C})$ by $\otimes(v_1, \ldots, v_m) = \prod_{i=1}^m v_i$ (a notation) where

$$(\prod_{i=1}^m v_i)(g_1, \ldots, g_m) := \prod_{i=1}^m g_i(v_i)$$

□

So $M(V, \ldots, V; \mathbb{C})$ is a model for $\otimes^m V^*$ and $M(V^*, \ldots, V^*; \mathbb{C})$ is a model for $\otimes^m V$.

We have another model for $\otimes^n V$, namely $M(V, \ldots, V; \mathbb{C})^*$, i.e., the dual space of $M(V, \ldots, V; \mathbb{C})$.

Theorem 3.8.2. $M(V, \ldots, V; \mathbb{C})^*$ is a model for $\otimes^m V$, i.e., there is tensor map $\otimes : V \times \cdots \times V \to M(V, \ldots, V; \mathbb{C})^*$ and $\dim M(V, \ldots, V; \mathbb{C})^* = n^m$.

Proof. For $v_1, \ldots, v_m \in V$ define $\prod_{i=1}^m v_i \in M(V, \ldots, V; \mathbb{C})^*$ by

$$(\prod_{i=1}^m v_i) f = f(v_1, \ldots, v_m), \quad f \in M(V, \ldots, V; \mathbb{C}).$$

From Theorem 3.8.1(1), $\{\prod_{i=1}^m f_{\alpha(i)} : \alpha \in \Gamma_{m,n}\}$ is a basis of $M(V, \ldots, V; \mathbb{C})$. Now from (3.8) $\{\prod_{i=1}^m e_{\alpha(i)} : \alpha \in \Gamma_{m,n}\}$ is the dual basis of $\{\prod_{i=1}^m f_{\alpha(i)} : \alpha \in \Gamma_{m,n}\}$ and thus is a basis of $M(V, \ldots, V; \mathbb{C})^*$. Then define $\otimes : V \times \cdots, \times V \to M(V, \ldots, V; \mathbb{C})^*$ by $\otimes(v_1, \ldots, v_m) = \prod_{i=1}^m v_i$. □

Elements in $M(V, \ldots, V; \mathbb{C})$ are called **contra-variant tensors**; elements in $M(V^*, \ldots, V^*; \mathbb{C})$ are called **covariant tensors**. They are useful tools in differential geometry.

The tensor space

$$V_q^p = \overbrace{V \otimes \cdots \otimes V}^{p} \otimes \overbrace{V^* \otimes \cdots \otimes V^*}^{q}$$

is called a **tensor space of type** (p, q) (with covariant type of degree p and with contra-variant type of degree q). Analogous to the previous treatment, under some tensor map, $M(V^*, \ldots, V^*, V, \ldots, V; \mathbb{C})$ (p copies of V and q copies of V^*) is a model of V_q^p:

Let $E = \{e_1, \ldots, e_n\}$ be a basis of V and let $E^* = \{f_1, \ldots, f_n\}$ be the dual basis of V^*. Then

$$\{\prod_{i=1}^{p} e_{\alpha(i)} \prod_{j=1}^{q} f_{\beta(j)} : \alpha \in \Gamma_{p,n}, \beta \in \Gamma_{q,n}\}$$

is a basis for

$$M(V^*, \ldots, V^*, V, \ldots, V; \mathbb{C}) \quad (p \text{ copy of } V, q \text{ copy of } V^*)$$

Define $\otimes : V^* \times \cdots \times V^* \times V \times \cdots \times V \to M(V^*, \ldots, V^*, V, \ldots, V; \mathbb{C})$ by

$$\otimes(e_{\alpha(1)}, \ldots, e_{\alpha(p)}, f_{\beta(1)}, \ldots, f_{\beta(q)}) = \prod_{i=1}^{p} e_{\alpha(i)} \prod_{j=1}^{q} f_{\beta(j)}.$$

$$\begin{aligned} & \{e_{\alpha(1)} \otimes \cdots \otimes e_{\alpha(p)} \otimes f_{\beta(1)} \otimes \cdots \otimes f_{\beta(q)} : \alpha \in \Gamma_{p,n}, \beta \in \Gamma_{q,n}\} \\ = \; & \{e_\alpha^\otimes \otimes f_\beta^\otimes : \alpha \in \Gamma_{p,n}, \beta \in \Gamma_{q,n}\} \end{aligned}$$

is a basis of V_q^p.

Problems

1. Define a simple tensor map $\boxtimes : \times^m V^* \to (\otimes^m V)^*$ such that $\boxtimes^m V^* = (\otimes V)^*$.

2. Let $M(V_1, \ldots, V_m; W)$ be the set of all multilinear maps from $V_1 \times \cdots \times V_m$ to W. Prove that $\dim M(V_1, \ldots, W_m; W) = \dim W \cdot \prod_{i=1}^m \dim V_i$.

Solutions to Problems 3.8

1. Let $f_1, \ldots, f_m \in V^*$. Define $\boxtimes(f_1, \ldots, f_m) = f \in (\otimes^m V)^*$ by $f(v_1 \otimes \cdots \otimes v_m) = \prod_{i=1}^m f_i(v_i)$. Clearly $\boxtimes$ is multilinear.

2.

Chapter 4

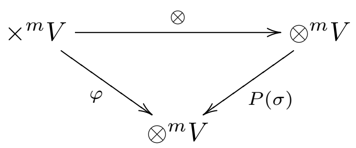
Symmetry classes of tensors

4.1 Permutation operators

Let V be an n-dimensional inner product space. Let S_m be the symmetric group of degree m on the set $\{1, \ldots, m\}$. Each $\sigma \in S_m$ yields a multilinear map

$$\varphi(v_1, \ldots, v_m) = v_{\sigma^{-1}(1)} \otimes \cdots \otimes v_{\sigma^{-1}(m)}, \quad v_1, \ldots, v_m \in V.$$

By the unique factorization property,



there is unique $P(\sigma) \in \operatorname{End}(\otimes^m V)$ such that $P\otimes = \varphi$, i.e.,

$$P(\sigma)(v_1 \otimes v_2 \otimes \cdots \otimes v_m) := v_{\sigma^{-1}(1)} \otimes v_{\sigma^{-1}(2)} \otimes \cdots \otimes v_{\sigma^{-1}(m)},$$

for all $v_1, \ldots, v_m \in V$. The operator $P(\sigma)$ called the **permutation operator** associated with σ on $\otimes^m V$.

Theorem 4.1.1. For each $\sigma \in S_m$, $([P(\sigma)]_{E_\otimes}^{E_\otimes})_{\alpha,\beta} = \delta_{\alpha,\beta\sigma^{-1}}$, where $\alpha, \beta \in \Gamma_{m,n}$.

Proof. Let $E = \{e_1, \ldots, e_m\}$ be a basis of V and let $E_\otimes = \{e_\gamma : \gamma \in \Gamma_{m,n}\}$ which is a basis of $\otimes^m V$. Then

$$\begin{aligned} P(\sigma)e^\otimes_\beta &= P(\sigma)e_{\beta(1)} \otimes \cdots \otimes e_{\beta(m)} \\ &= e_{\beta(\sigma^{-1}(1))} \otimes \cdots \otimes e_{\beta(\sigma^{-1}(m))} \\ &= e^\otimes_{\beta\sigma^{-1}} \\ &= \sum_{\alpha \in \Gamma_{m,n}} \delta_{\alpha,\beta\sigma^{-1}} e^\otimes_\alpha, \qquad \beta \in \Gamma_{m,n}. \end{aligned}$$

So the matrix representation is

$$([P(\sigma)]_{E_\otimes}^{E_\otimes})_{\alpha,\beta} = (\delta_{\alpha,\beta\sigma^{-1}})_{\substack{\alpha\in\Gamma_{m,n}\\ \beta\in\Gamma_{m,n}}}$$

□

From the proof we see that

$$P(\sigma)e_\beta^\otimes = e_{\beta\sigma^{-1}}^\otimes \tag{4.1}$$

The permutation operator yields a representation of S_m, i.e., $P : S_m \to \mathrm{GL}(\otimes_{i=1}^m V_i)$.

Theorem 4.1.2. Let $\sigma, \pi \in S_m$. Then

(a) $P(\sigma\pi) = P(\sigma)P(\pi)$.

(b) $P(e) = I_{\otimes^m V}$.

(c) $P(\sigma)$ is invertible and $P(\sigma)^{-1} = P(\sigma^{-1})$.

(d) $P(\sigma)$ is unitary, i.e., $P(\sigma)^{-1} = P(\sigma)^*$ with respect to the induced inner product of $\otimes^m V$.

(e) If $\dim V \geq 2$, then $P(\sigma) = P(\pi)$ implies $\sigma = \pi$, i.e., P is a faithful representation.

(f) $\operatorname{tr} P(\sigma) = n^{c(\sigma)}$, where $n = \dim V$ and $c(\sigma)$ is the number of factors in the disjoint cycle factorization of σ.

Proof. (a) From (4.1)

$$P(\sigma)P(\pi)v^\otimes = P(\sigma)v_{\pi^{-1}}^\otimes = v_{\pi^{-1}\sigma^{-1}}^\otimes = v_{(\sigma\pi)^{-1}}^\otimes = P(\sigma\pi)v^\otimes$$

(b) Clear or from Theorem 4.1.1.

(c) From (a) and (b).

(d) With respect to the inner product

$$(P(\sigma)u^\otimes, v^\otimes) = \prod_{i=1}^m (v_{\sigma^{-1}(i)}, v_i) = \prod_{i=1}^m (u_i, v_{\sigma(i)}) = (u^\otimes, v_\sigma^\otimes) = (u^\otimes, P(\sigma^{-1})v^\otimes).$$

Since $\otimes^m V$ is spanned by the decomposable tensors, $P(\sigma)^{-1} = P(\sigma^{-1}) = P(\sigma)^*$, i.e., $P(\sigma)$ is unitary for all $\sigma \in S_m$.

(e) Since $\dim V \geq 2$, there are linearly independent vectors $e_1, e_2 \in V$. Set

$$v_k := e_1 + ke_2, \quad k = 1, \ldots, m.$$

When $i \neq j$, v_i and v_j are linearly independent. If $P(\sigma) = P(\pi)$, then $v_{\sigma^{-1}}^\otimes = v_{\pi^{-1}}^\otimes \neq 0$. By Theorem 3.3.3, $v_{\sigma^{-1}(i)}$ and $v_{\pi^{-1}(i)}$ are linearly dependent for all i. So $\sigma^{-1}(i) = \pi^{-1}(i)$ for all i, i.e., $\sigma = \pi$.

(f) From Theorem 4.1.1

$$\operatorname{tr} P(\sigma) = \sum_{\Gamma_{m,n}} \delta_{\alpha,\alpha\sigma^{-1}} = \sum_{\alpha\in\Gamma_{m,n}} \delta_{\alpha\sigma,\alpha} = |\{\alpha \in \Gamma_{m,n} : \alpha\sigma = \alpha\}|.$$

In other word, $\operatorname{tr} P(\sigma)$ is the number of elements $\alpha \in \Gamma_{m,n}$ that is fixed by σ ($\alpha \mapsto \alpha\sigma$). Decompose σ into the product of $c(\sigma)$ disjoint cycles. It is easy to see that for each cycle all the α-components must be the same and they can be chosen from $1, \dots, n$. So $\operatorname{tr} P(\sigma) = n^{c(\sigma)}$.

□

We remark that when S_m is replaced by any subgroup G of S_m, it yields a representation $P : G \to \mathrm{GL}(\otimes^m V)$.

Problems

1. Suppose $m > 1$. Prove that unless $\sigma = e$ or $\dim V = 1$, $P(\sigma)$ is not an induced operator, i.e., there are no $T_i \in \operatorname{End} V$ such that $P(\sigma) = \otimes_{i=1}^m T_i$.
2. Suppose that $\tau \in S_m$ is a transposition. Prove that if $P(\tau)$ is a sum of k induced operators, then $k \ge n^2$, where $n = \dim V$.
3. If $\dim V > 1$, then $\sigma \mapsto P(\sigma)$ is a reducible representation of S_m.

Solutions to Problems 4.1

1. Suppose that there were $T_i \in \operatorname{End} V$ such that $P(\sigma) = \otimes_{i=1}^m T_i$. For all $0 \ne v \in V$ we have

$$T_1 v \otimes \cdots \otimes T_m v = (\otimes_{i=1}^m T_i)(\otimes_{i=1}^m v) = P(\sigma)(\otimes_{i=1}^m v) = v \otimes \cdots \otimes v.$$

Thus by Theorem 3.3.2 and Theorem 3.3.3 we have $T_i v = c_i v$, $\prod_{i=1}^m c_i = 1$ for all $i = 1, \dots, m$. Notice that c_i apparently (but really not) depends on v. choose an orthonormal basis $E = \{e_1, \dots e_n\}$ of V (after equip V with an inner product). Each T_i is unitary since $P(\sigma)$ is unitary (Problem 3.6 #2). Now $T_i e_j = c_i^{(j)} e_j$, for all i, j with $|c_i^{(j)}| = 1$, and $T(e_s + e_t) = c(e_s + e_t)$ with $|c| = 1$ where c depends on T_i. Since $n = \dim V > 1$,

$$c(e_s + e_t) = T_i(e_s + e_t) = c_i^{(s)} e_s + c_i^{(t)} e_t,$$

for any $s \ne j$. So $c_s = c_t$ for all $s \ne t$, i.e., $T_i = cI$. Hence $\otimes_{i=1}^m T_i$ is a scalar multiple of the identity, contradiction.

2.

3. Each $P(\sigma)$ is a permutation matrix in $\mathbb{C}_{n^m \times n^m}$. If $\dim V > 1$, then $n^m > 1$. Notice that the span of $\sum_{\gamma\in\Gamma_{m,n}} e_\gamma^{\otimes}$ is an invariant subspace under $P(\sigma)$ for all $\gamma \in \Gamma_{m,n}$. Hence by Theorem 2.2.2, $P : G \to \mathrm{GL}(\otimes^m V)$ is reducible.

4.2 Symmetric multilinear maps and symmetrizers

A multilinear map $\psi : \times^m V \to W$ is said to be **completely symmetric** if

$$\psi(v_{\sigma(1)}, \ldots, v_{\sigma(m)}) = \psi(v_1, \ldots, v_m), \quad \text{for all } \sigma \in S_m. \tag{4.2}$$

Each multilinear map $\varphi : \times^m V \to W$ induces a completely symmetric multilinear map:

$$1_\varphi(v_1, \ldots, v_m) = \frac{1}{m!} \sum_{\sigma \in S_m} \varphi(v_{\sigma(1)}, \ldots, v_{\sigma(m)}).$$

It is because

$$\begin{aligned} & 1_\varphi(v_{\pi(1)}, \ldots, v_{\pi(m)}) \\ = & \frac{1}{m!} \sum_{\sigma \in S_m} \varphi(v_{\pi\sigma(1)}, \ldots, v_{\pi\sigma(m)}) \\ = & \frac{1}{m!} \sum_{\tau \in S_m} \varphi(v_{\tau(1)}, \ldots, v_{\tau(m)}) \quad (\tau = \pi\sigma) \\ = & 1_\varphi(v_1, \ldots, v_m), \end{aligned}$$

i.e., 1_φ is a completely symmetric multilinear map according to (4.2).

It is easy to see that (4.2) is equivalent to the following

$$\frac{1}{m!} \sum_{\sigma \in S_m} \psi(v_{\sigma(1)}, \ldots, v_{\sigma(m)}) = \psi(v_1, \ldots, v_m) \quad \text{for all } v_1, \ldots, v_m \in V. \tag{4.3}$$

Similarly, ψ is said to be **skew symmetric** if

$$\psi(v_{\sigma(1)}, \ldots, v_{\sigma(m)}) = \varepsilon(\sigma)\psi(v_1, \ldots, v_m) \quad \text{for all } \sigma \text{ and } v_1, \ldots, v_m \in V. \tag{4.4}$$

Notice that (4.4) is equivalent to the following

$$\frac{1}{m!} \sum_{\sigma \in S_m} \varepsilon(\sigma^{-1})\psi(v_{\sigma(1)}, \ldots, v_{\sigma(m)}) = \psi(v_1, \ldots, v_m). \tag{4.5}$$

Each multilinear $\varphi : \times^m V \to W$ yields a skew symmetric multilinear $\varepsilon_\varphi : \times^m V \to W$:

$$\varepsilon_\varphi(v_1, \ldots, v_m) = \frac{1}{m!} \sum_{\sigma \in S_m} \varepsilon(\sigma^{-1})\varphi(v_{\sigma(1)}, \ldots, v_{\sigma(m)}).$$

It is because

$$\begin{aligned}
& \varepsilon_\varphi(v_{\pi(1)}, \dots, v_{\pi(m)}) \\
= \ & \frac{1}{m!} \sum_{\sigma \in S_m} \varepsilon(\sigma^{-1}) \varphi(v_{\pi\sigma(1)}, \dots, v_{\pi\sigma(m)}) \\
= \ & \frac{1}{m!} \sum_{\tau \in S_m} \varepsilon(\tau^{-1}\pi) \varphi(v_{\tau(1)}, \dots, v_{\tau(m)}) \quad (\tau = \pi\sigma) \\
= \ & \frac{1}{m!} \sum_{\tau \in S_m} \varepsilon(\tau^{-1}) \varepsilon(\pi) \varphi(v_{\tau(1)}, \dots, v_{\tau(m)}) \quad (\text{by } \varepsilon(\tau^{-1}\pi) = \varepsilon(\tau^{-1})\varepsilon(\pi)) \\
= \ & \varepsilon(\pi) \varepsilon_\varphi(v_1, \dots, v_m).
\end{aligned}$$

When $m = 2$, φ can be decomposed as

$$\begin{aligned}
\varphi(v_1, v_2) &= \frac{1}{2}[\varphi(v_1, v_2) + \varphi(v_2, v_1)] + \frac{1}{2}[\varphi(v_1, v_2) - \varphi(v_2, v_1)] \\
&= 1_\varphi(v_1, v_2) + \varepsilon_\varphi(v_1, v_2), \qquad (4.6)
\end{aligned}$$

i.e., $\varphi = 1_\varphi + \varepsilon_\varphi$.

When $m > 2$, the decomposition of φ may have more than two summands. The reason for $m = 2$ case having two summands is that

$$I(S_2) = \{1, \varepsilon\}$$

is the set of irreducible characters of S_2.

For general subgroup G of S_m, the set of irreducible characters $I(G)$ of G will come into the picture.

We say that the multilinear map $\psi : \times^m V \to W$ is **symmetric with respect to G and χ** if

$$\frac{\chi(e)}{|G|} \sum_{\sigma \in G} \chi(\sigma^{-1}) \psi(v_{\sigma(1)}, \dots, v_{\sigma(m)}) = \psi(v_1, \dots, v_m) \qquad (4.7)$$

for all $v_1, \dots, v_m \in V$. If χ is linear, then $\psi : \times^m V \to W$ is symmetric with respect to G and χ if and only if $\psi(v_{\sigma(1)}, \dots, v_{\sigma(m)}) = \chi(\sigma)\psi(v_1, \dots, v_m)$ for all $\sigma \in G$ (Problem 3). When χ is not linear, they are not equivalent.

Motivated by 1_φ and ε_φ, each multilinear map $\varphi : \times^m V \to W$ induces a mulitlinear map $\chi_\varphi : \times^m V \to W$ symmetric with respect to G and χ:

$$\chi_\varphi(v_1, \dots, v_m) := \frac{\chi(e)}{|G|} \sum_{\sigma \in G} \chi(\sigma^{-1}) \varphi(v_{\sigma(1)}, \dots, v_{\sigma(m)})$$

for all $v_1, \dots, v_m \in V$. The map χ_φ is called the **component of φ with respect to G and χ**.

Theorem 4.2.1. Let $\varphi : \times^m V \to W$ be multilinear. Then χ_φ is symmetric with respect to G and χ, i.e., $\chi_{\chi_\varphi} = \chi_\varphi$.

Proof. We now show that χ_φ satisfies (4.7) by using the irreducible character orthogonal relation of the first kind:

$$\begin{aligned}
& \chi_{\chi_\varphi}(v_1,\ldots,v_m) \\
= & \frac{\chi(e)}{|G|}\sum_{\pi\in G}\chi(\pi^{-1})\chi_\varphi(v_{\pi(1)},\ldots,v_{\pi(m)}) \\
= & \frac{\chi(e)}{|G|}\sum_{\pi\in G}\chi(\pi^{-1})\cdot\frac{\chi(e)}{|G|}\sum_{\sigma\in G}\chi(\sigma^{-1})\varphi(v_{\pi\sigma(1)},\ldots,v_{\pi\sigma(m)}) \\
= & \frac{\chi(e)^2}{|G|^2}\sum_{\tau\in G}\left(\sum_{\pi\in G}\chi(\pi^{-1})\chi(\tau^{-1}\pi)\right)\varphi(v_{\tau(1)},\ldots,v_{\tau(m)}) \quad (\tau=\pi\sigma) \\
= & \frac{\chi(e)^2}{|G|^2}\sum_{\tau\in G}\left(\frac{|G|}{\chi(e)}\chi(\tau^{-1})\right)\varphi(v_{\tau(1)},\ldots,v_{\tau(m)}) \quad (\text{by Theorem 2.4.2}) \\
= & \chi_\varphi(v_1,\ldots,v_m). \qquad (4.8)
\end{aligned}$$

□

Clearly if $\psi : \times^m V \to W$ is symmetric with respect to G and χ, then $\chi_\psi = \psi$, and vice versa.

Recall $I(G)$ denotes the set of irreducible characters of G. Each $\varphi : \times^m V \to W$ induces χ_φ for all $\chi \in I(G)$. The following is an extension of (4.6).

Theorem 4.2.2. Let $\varphi : \times^m V \to W$. Then $\varphi = \sum_{\chi\in I(G)} \chi_\varphi$. In particular, for $\otimes : \times^m V \to \otimes^m V$

$$\otimes = \sum_{\chi\in I(G)} \chi_\otimes \qquad (4.9)$$

where $\chi_\otimes : \times^m V \to \otimes^m V$.

Proof.

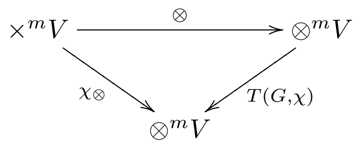
$$\begin{aligned}
& \sum_{\chi\in I(G)} \chi_\varphi(v_1,\ldots,v_m) \\
= & \sum_{\chi\in I(G)}\frac{\chi(e)}{|G|}\sum_{\sigma\in G}\chi(\sigma^{-1})\varphi(v_{\sigma(1)},\ldots,v_{\sigma(m)}) \\
= & \sum_{\sigma\in G}\left(\frac{1}{|G|}\sum_{\chi\in I(G)}\chi(e)\chi(\sigma^{-1})\right)\varphi(v_{\sigma(1)},\ldots,v_{\sigma(m)}) \\
= & \sum_{\sigma\in G}\delta_{e,\sigma^{-1}}\varphi(v_{\sigma(1)},\ldots,v_{\sigma(m)}) \quad (\text{by Corollary 2.4.15}) \\
= & \varphi(v_1,\ldots,v_m). \qquad (4.10)
\end{aligned}$$

Since $\otimes : \times^m V \to \otimes^m V$ is multilinear, $\otimes = \sum_{\chi\in I(G)} \chi_\otimes$ follows immediately. □

Since $\chi_\otimes : \times^m V \to \otimes^m V$ is multilinear, by the unique factorization property of $\otimes^m V$ there is a unique $T(G,\chi) \in \operatorname{End}(\otimes^m V)$ such that

$$\chi_\otimes = T(G,\chi)\otimes, \tag{4.11}$$

i.e.,



and thus

$$\chi_\otimes(v_1,\dots,v_m) = T(G,\chi)v^\otimes,$$

To determine $T(G,\chi)$ consider

$$\begin{aligned} T(G,\chi)v^\otimes &= \chi_\otimes(v_1,\dots,v_m) \\ &= \frac{\chi(e)}{|G|}\sum_{\sigma\in G}\chi(\sigma^{-1})\otimes(v_{\sigma(1)},\dots,v_{\sigma(m)}) \\ &= \frac{\chi(e)}{|G|}\sum_{\sigma\in G}\chi(\sigma^{-1})v_\sigma^\otimes \\ &= \frac{\chi(e)}{|G|}\sum_{\sigma\in G}\chi(\sigma^{-1})P(\sigma^{-1})v^\otimes \\ &= \frac{\chi(e)}{|G|}\sum_{\sigma\in G}\chi(\sigma)P(\sigma)v^\otimes. \end{aligned}$$

So

$$T(G,\chi) = \frac{\chi(e)}{|G|}\sum_{\sigma\in G}\chi(\sigma)P(\sigma) \in \operatorname{End}(\otimes^m V) \tag{4.12}$$

and is called the **symmetrizer** associated with G and χ since from (4.11), it turns $\otimes$ into $\chi_\otimes$ which is symmetric to G and χ. We now see some basic properties of $T(G,\chi)$.

Theorem 4.2.3. Let $\chi,\mu \in I(G)$ where $G < S_m$. Then

(a) $T(G,\chi)$ is an orthogonal projection with respect to the induced inner product on $\otimes^m V$, i.e., $T(G,\chi)^2 = T(G,\chi)$ and $T(G,\chi)^* = T(G,\chi)$.

(b) $\sum_{\chi\in I(G)} T(G,\chi) = I_{\otimes^m V}$.

(c) If $\chi \neq \mu$, then $T(G,\chi)T(G,\mu) = 0$.

Proof. (a) From Theorem 4.1.2 $P(\sigma\pi) = P(\sigma)P(\pi)$ so that

$$\begin{aligned}
T(G,\chi)^2 &= \left(\frac{\chi(e)}{|G|}\sum_{\sigma\in G}\chi(\sigma)P(\sigma)\right)\left(\frac{\chi(e)}{|G|}\sum_{\pi\in G}\chi(\pi)P(\pi)\right)\\
&= \frac{\chi^2(e)}{|G|^2}\sum_{\sigma\in G}\sum_{\pi\in G}\chi(\sigma)\chi(\pi)P(\sigma)P(\pi)\\
&= \frac{\chi^2(e)}{|G|^2}\sum_{\tau\in G}\left(\sum_{\sigma\in G}\chi(\sigma)\chi(\sigma^{-1}\tau)\right)P(\tau) \qquad (\tau=\sigma\pi)\\
&= \frac{\chi^2(e)}{|G|^2}\sum_{\tau\in G}\left(\frac{|G|}{\chi(e)}\chi(\tau)\right)P(\tau) \quad \text{(by Theorem 2.4.2)}\\
&= T(G,\chi).
\end{aligned}$$

So $T(G,\chi)$ is a projection. By Theorem 2.4.4 $\overline{\chi(\sigma)} = \chi(\sigma^{-1})$ so that

$$\begin{aligned}
& T(G,\chi)^*\\
=\ & \frac{\chi(e)}{|G|}\sum_{\sigma\in G}\overline{\chi(\sigma)}P(\sigma)^*\\
=\ & \frac{\chi(e)}{|G|}\sum_{\sigma\in G}\chi(\sigma^{-1})P(\sigma^{-1}) \quad (P(\sigma)^* = P(\sigma)^{-1} \text{ by Theorem 4.1.2}(d))\\
=\ & T(G,\chi).
\end{aligned}$$

So $T(G,\chi)$ is an orthogonal projection.

(b) By Theorem 2.4.15

$$\begin{aligned}
\sum_{\chi\in I(G)} T(G,\chi) &= \sum_{\chi\in I(G)}\frac{\chi(e)}{|G|}\sum_{\sigma\in G}\chi(\sigma)P(\sigma)\\
&= \sum_{\sigma\in G}\left(\frac{1}{|G|}\sum_{\chi\in I(G)}\chi(e)\chi(\sigma)\right)P(\sigma)\\
&= P(e)\\
&= I_{\otimes^m V}.
\end{aligned}$$

(c) By Theorem 2.4.3, $\sum_{\sigma\in G}\chi(\sigma)\mu(\sigma^{-1}\tau) = 0$ if $\chi\neq\mu$. So

$$\begin{aligned}
& T(G,\chi)T(G,\mu)\\
=\ & \frac{\chi(e)\mu(e)}{|G|^2}\sum_{\sigma\in G}\sum_{\pi\in G}\chi(\sigma)\mu(\pi)P(\sigma\pi)\\
=\ & \frac{\chi(e)\mu(e)}{|G|^2}\sum_{\tau\in G}\left(\sum_{\sigma\in G}\chi(\sigma)\mu(\sigma^{-1}\tau)\right)P(\tau)\\
=\ & 0.
\end{aligned}$$

□

The range of $T(G,\chi)$

$$V_\chi^m(G) := T(G,\chi)(\otimes^m V)$$

is called the **symmetry class of tensors** over V associated with G and χ. Theorem 4.2.2 implies that

$$\otimes^m V = \sum_{\chi \in I(G)} V_\chi(G).$$

The following asserts that the sum is indeed an orthogonal sum with respect to the induced inner product on $\otimes^n V$.

Theorem 4.2.4. The tensor space $\otimes^m V$ is an orthogonal sum of $V_\chi(G)$, $\chi \in I(G)$, i.e.,

$$\otimes^m V = \perp_{\chi \in I(G)} V_\chi(G).$$

In other words

$$\otimes^m V = V_{\chi_1}(G) \perp \cdots \perp V_{\chi_k}(G),$$

where $I(G) = \{\chi_1, \ldots, \chi_k\}$ is the set of irreducible characters of G.

Proof. From Theorem 4.2.3(a) and (c), $T(G,\chi)$ is an orthogonal projection. □

The elements in $V_\chi^m(G)$ of the form

$$v_1 * \cdots * v_m := T(G,\chi)(v_1 \otimes \cdots \otimes v_m)$$

are called **decomposable symmetrized tensors**. Such notation does not reflect its dependence on G and χ. From Theorem 4.2.4, $v_1 * \cdots * v_m$ is the piece of $v_1 \otimes \cdots \otimes v_m$ in $V_\chi(G)$. So $V_\chi(G)$ is spanned by the decomposable symmetrized tensors $v_1 * \cdots * v_m$.

As we saw before, when $G = S_2$ we have $I(G) = \{1, \varepsilon\}$. Then

$$v_1 \bullet v_2 = T(S_2, 1)(v_1 \otimes v_2) = \frac{1}{2}(v_1 \otimes v_2 + v_2 \otimes v_1)$$

and

$$v_1 \wedge v_2 = T(S_2, \varepsilon)(v_1 \otimes v_2) = \frac{1}{2}(v_1 \otimes v_2 - v_2 \otimes v_1).$$

Clearly

$$v_1 \otimes v_2 = v_1 \bullet v_2 + v_1 \wedge v_2.$$

Moreover

$$\begin{aligned}
& (v_1 \bullet v_2, v_1 \wedge v_2) \\
= \;& \frac{1}{4}(v_1 \otimes v_2 + v_2 \otimes v_1, v_1 \otimes v_2 - v_2 \otimes v_1) \\
= \;& \frac{1}{4}((v_1 \otimes v_2, v_1 \otimes v_2) - (v_1 \otimes v_2, v_2 \otimes v_1) + (v_2 \otimes v_1, v_1 \otimes v_2) - (v_2 \otimes v_1, v_2 \otimes v_1)) \\
= \;& \frac{1}{4}((v_1, v_1)(v_2, v_2) - (v_1, v_2)(v_2, v_1) + (v_2, v_1)(v_1, v_2) - (v_2, v_2)(v_1, v_1)) \\
= \;& 0,
\end{aligned}$$

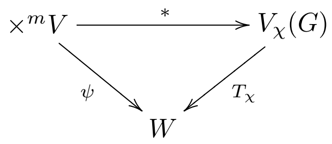
verifying that

$$V \otimes V = V_1(S_2) \perp V_\varepsilon(S_2)$$

as indicated in Theorem 4.2.4.

The following is the unique factorization property of symmetrizer map $* : \times^m V \to V_\chi(G)$.

Theorem 4.2.5. Let $\psi : \times^m V \to W$ be a multilinear map symmetric with respect to G and χ. There exists a unique $T_\chi \in \operatorname{Hom}(V_\chi(G), W)$ such that $\psi = T_\chi *$,

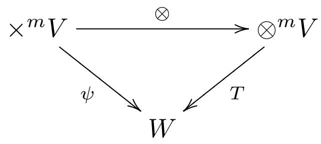


i.e,

$$\psi(v_1, \ldots, v_m) = T_\chi v^*.$$

Proof. Uniqueness follows immediately since the decomposable symmetrized tensors v^* span $V_\chi(G)$.

For the existence, from the factorization property of $\otimes$ (Theorem 3.2.3),



there is $T \in \operatorname{Hom}(\otimes^m V, W)$ such that $\psi = T\otimes$. Then from the symmetry of ψ with respect to G and χ,

$$\begin{aligned}
\psi(v_1, \ldots, v_m) &= \chi_\psi(v_1, \ldots, v_m) \\
&= \frac{\chi(e)}{|G|} \sum_{\sigma \in G} \chi(\sigma^{-1}) \psi(v_{\sigma(1)}, \ldots, v_{\sigma(m)}) \\
&= \frac{\chi(e)}{|G|} \sum_{\sigma \in G} \chi(\sigma^{-1}) T v_\sigma^\otimes \quad (\psi = T\otimes) \\
&= T \left(\frac{\chi(e)}{|G|} \sum_{\sigma \in G} \chi(\sigma^{-1}) P(\sigma^{-1}) \right) v^\otimes \\
&= T \circ T(G, \chi) v^\otimes \\
&= T v^*
\end{aligned}$$

Set $T_\chi := T|_{V_\chi(G)} \in \operatorname{Hom}(V_\chi(G), W)$. So

$$\psi(v_1, \ldots, v_m) = T_\chi v^*.$$

□

Problems

1. Prove that if $\pi \in G$, then $P(\pi)T(G,\chi) = T(G,\chi)P(\pi)$. In addition, if $\chi(e) = 1$, then $P(\pi)T(G,\chi) = \chi(\pi^{-1})T(G,\chi)$. (Hint: If χ is linear, then $\chi(\sigma\pi) = \chi(\sigma)\chi(\pi)$)
2. Prove $P(\sigma^{-1})v^* = v^*_\sigma$. In addition, if $\chi(e) = 1$, then $v^*_\sigma = \chi(\sigma)v^*$.
3. Show that if χ is linear, then $\psi : \times^m V \to W$ is symmetric with respect to G and χ if and only if $\psi(v_{\sigma(1)}, \dots, v_{\sigma(m)}) = \chi(\sigma)\psi(v_1, \dots, v_m)$ for all $\sigma \in G$.
4. Let $f_1, \dots, f_m \in V^*$. Prove that $\psi : \times^m V \to \mathbb{C}$ defined by $\psi = \sum_{\pi\in G} \chi(\pi) \prod_{t=1}^m f_{\pi(t)}$ is symmetric with respect to G and χ.
5. Prove that the multilinear map $\psi : \times^m V \to \mathbb{C}$ is skew symmetric with respect to S_m if and only if $v_i = v_j$ implies that $\psi(v_1, \dots, v_m) = 0$ whenever $i \neq j$. (Hint: σ is a product of transpositions).

Solutions to Problems 4.2

1. Essentially from the proof of Theorem 4.2.3

$$\begin{aligned}
P(\pi)T(G,\chi) &= P(\pi)\left(\frac{\chi(e)}{|G|}\sum_{\sigma\in G}\chi(\sigma)P(\sigma)\right) \\
&= \frac{\chi(e)}{|G|}\sum_{\sigma\in G}\chi(\sigma)P(\pi\sigma) \qquad (4.13) \\
&= \frac{\chi(e)}{|G|}\sum_{\tau\in G}\chi(\pi^{-1}\tau\pi)P(\tau\pi) \quad (\tau\pi = \pi\sigma) \\
&= \left(\frac{\chi(e)}{|G|}\sum_{\tau\in G}\chi(\tau)P(\tau)\right)P(\pi) \quad (\chi(\pi^{-1}\tau\pi) = \chi(\tau)) \\
&= T(G,\chi)P(\pi).
\end{aligned}$$

In addition, if χ is linear, then from (4.13)

$$\begin{aligned}
& P(\pi)T(G,\chi) \\
= & \frac{\chi(e)}{|G|}\sum_{\sigma\in G}\chi(\sigma)P(\pi\sigma) \\
= & \frac{\chi(e)}{|G|}\sum_{\tau\in G}\chi(\pi^{-1}\tau)P(\tau) \\
= & \chi(\pi^{-1})\frac{\chi(e)}{|G|}\sum_{\tau\in G}\chi(\tau)P(\tau) \quad (\chi(\sigma\pi) = \chi(\sigma)\chi(\pi) \text{ since } \chi \text{ linear }) \\
= & \chi(\pi^{-1})T(G,\chi).
\end{aligned}$$

2. By Problem 1,

$$P(\sigma^{-1})v^* = P(\sigma^{-1})T(G,\chi)v^{\otimes} = T(G,\chi)P(\sigma^{-1})v^{\otimes} = T(G,\chi)v_\sigma^{\otimes} = v_\sigma^*.$$

In addition, if χ is linear, then again by Problem 1,

$$v_\sigma^* = P(\sigma^{-1})v^* = P(\sigma^{-1})T(G,\chi)v^{\otimes} = \chi(\sigma)T(G,\chi)v^{\otimes} = \chi(\sigma)v^*.$$

3. Suppose χ is linear so that it is a representation.

If $\psi(v_{\sigma(1)},\dots,v_{\sigma(m)}) = \chi(\sigma)\psi(v_1,\dots,v_m)$ for all $\sigma \in G$, then

$$\begin{aligned} & \frac{\chi(e)}{|G|}\sum_{\sigma\in G}\chi(\sigma^{-1})\psi(v_{\sigma(1)},\dots,v_{\sigma(m)}) \\ = & \frac{\chi(e)}{|G|}\sum_{\sigma\in G}\chi(\sigma^{-1})\chi(\sigma)\psi(v_1,\dots,v_m) \\ = & \chi(e)^2\psi(v_1,\dots,v_m) \quad (\chi(\sigma^{-1})\chi(\sigma) = \chi(\sigma^{-1}\sigma) = \chi(e)) \\ = & \psi(v_1,\dots,v_m) \quad (\text{ since } \chi(e) = 1) \end{aligned}$$

(Answer to Alex's question: Indeed for general irreducible character χ, if $\psi(v_{\sigma(1)},\dots,v_{\sigma(m)}) = \frac{\chi(\sigma)}{\chi(e)}\psi(v_1,\dots,v_m)$ for all $\sigma \in G$, then ψ is symmetric with respect to G and χ since $\sum_{\sigma\in G}\chi(\sigma^{-1})\chi(\sigma) = |G|$ by Theorem 2.4.2)

If ψ is symmetric with respect to G and χ, then $\psi = T_\chi *$ for some $T_\chi \in \operatorname{Hom}(V_\chi(G), W)$ by Theorem 4.2.5. Thus

$$\begin{aligned} \psi(v_{\sigma(1)},\dots,v_{\sigma(m)}) &= T_\chi * (v_{\sigma(1)},\dots,v_{\sigma(m)}) \\ &= T_\chi v_\sigma^* \\ &= T_\chi(\chi(\sigma)v^*) \quad (\text{by Problem 2}) \\ &= \chi(\sigma)T_\chi(v^*) \\ &= \chi(\sigma)\varphi(v_1,\dots,v_m) \end{aligned}$$

4. (Roy)

$$\begin{aligned}
& \frac{\chi(e)}{|G|}\sum_{\sigma\in G}\chi(\sigma^{-1})\psi(v_{\sigma(1)},\ldots,v_{\sigma(m)}) \\
= & \frac{\chi(e)}{|G|}\sum_{\sigma\in G}\chi(\sigma^{-1})\left(\sum_{\pi\in G}\chi(\pi)\prod_{t=1}^{m}f_{\pi(t)}(v_{\sigma(t)})\right) \\
= & \frac{\chi(e)}{|G|}\sum_{\sigma,\tau\in G}\chi(\sigma^{-1})\chi(\tau\sigma)\prod_{t=1}^{m}f_{\tau\sigma(t)}(v_{\sigma(t)}) \quad (\pi=\tau\sigma) \\
= & \frac{\chi(e)}{|G|}\sum_{\tau\in G}\left(\sum_{\sigma\in G}\chi(\sigma^{-1})\chi(\sigma\tau)\right)\prod_{t=1}^{m}f_{\tau(t)}(v_t) \quad (\chi(\sigma\tau)=\chi(\tau\sigma)) \\
= & \frac{\chi(e)}{|G|}\sum_{\tau\in G}\left(\frac{|G|}{\chi(e)}\chi(\tau)\right)\prod_{t=1}^{m}f_{\tau(t)}(v_t) \quad (\text{Theorem } 2.4.3) \\
= & \sum_{\tau\in G}\chi(\tau)\prod_{t=1}^{m}f_{\tau(t)}(v_t) \\
= & \psi(v_1,\ldots,v_m)
\end{aligned}$$

Therefore ψ is symmetric with respect to G and χ.

5. From Problem 3, ψ is skew symmetric if and only if $\psi(v_{\sigma(1)},\ldots,v_{\sigma(m)}) = \varepsilon(\sigma)\psi(v_1,\ldots,v_m)$. Thus, if $\sigma = (ij)$, then ψ skew symmetric implies that $\psi(v_1,\ldots,v_i,\ldots,v_j,\ldots,v_m) = -\psi(v_1,\ldots,v_j,\ldots,v_i,\ldots,v_m)$. So if $v_i = v_j$, then $\psi(v_1,\ldots,v_m) = 0$. Conversely suppose that $v_i = v_j$ implies that $\psi(v_1,\ldots,v_m) = 0$ whenever $i \neq j$. Then

$$\begin{aligned}
0 &= \psi(v_1,\ldots,v_i+v_j,\ldots,v_i+v_j,\ldots,v_m) \\
&= \psi(v_1,\ldots,v_i,\ldots,v_i,\ldots,v_m)+\psi(v_1,\ldots,v_j,\ldots,v_j,\ldots,v_m) \\
&\quad +\psi(v_1,\ldots,v_i,\ldots,v_j,\ldots,v_m)+\psi(v_1,\ldots,v_j,\ldots,v_i,\ldots,v_m)
\end{aligned}$$

So $\psi(v_1,\ldots,v_i,\ldots,v_j,\ldots,v_m) = -\psi(v_1,\ldots,v_j,\ldots,v_i,\ldots,v_m)$. Each $\sigma \in S_m$ is a product of transpositions, i.e., $\sigma = \sigma_1\sigma_2\cdots\sigma_k$. So

$$\psi(v_{\sigma(1)},\ldots,v_{\sigma(m)}) = -\psi(v_{\sigma_2\cdots\sigma_k(1)},\ldots,v_{\sigma_2\cdots\sigma_k(m)}) = \cdots = \varepsilon(\sigma)\psi(v_1,\ldots,v_m).$$

4.3 Basic properties of symmetrized tensor

The symmetrized tensors, i.e., the elements in $V_\chi(G)$ is a subspace of $\otimes^m V$ but the decomposable symmetrized tensors in $V_\chi(G)$ may not be decomposable elements in $\otimes^m V$. When $G = \{e\}$, χ must be trivial and $V\chi(e) = \otimes^m V$ so $\otimes^m V$ can be viewed a special kind of $V_\chi(G)$. So basic properties of $V_\chi(G)$ apply to $\otimes^m V$ as well.

Theorem 4.3.1. If $u^* + \cdots + v^* = 0$, then for any multilinear map $\psi : \times^m V \to W$ symmetric with respect to G and χ,

$$\psi(u_1, \ldots, u_m) + \cdots + \psi(v_1, \ldots, v_m) = 0.$$

Proof. From Theorem 4.2.5, there is $T_\chi \in \operatorname{Hom}(V_\chi(G), W)$ such that

$$\psi(u_1, \ldots, u_m) = T_\chi u^*, \ldots, \psi(v_1, \ldots, v_m) = T_\chi v^*.$$

So

$$\psi(u_1, \ldots, u_m) + \cdots + \psi(v_1, \ldots, v_m) = T_\chi(u^* + \cdots + v^*) = 0.$$

□

Notice that when $G = \{e\}$, the symmetric multilinear function φ with respect to $\{e\}$ and $\chi \equiv 1$ simply means that φ is simply multilinear. So Theorem 4.3.1 is a generalization of Theorem 3.3.1.

Theorem 4.3.2. If $v_1 * \cdots * v_m = 0$, then $v_1, \ldots, v_m$ are linearly dependent.

Proof. Suppose on the contrary that $v_1 * \cdots * v_m = 0$ and $v_1, \ldots, v_m$ were linearly independent. Extend it to a basis $\{v_1, \ldots, v_m, v_{m+1}, \ldots, v_n\}$ of V. So

$$\{v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(m)} : \sigma \in G\} \subset \{v_{\gamma(1)} \otimes \cdots \otimes v_{\gamma(m)} : \gamma \in \Gamma_{m,n}\}$$

must be linearly independent. Now

$$0 = v^* = \frac{\chi(e)}{|G|} \sum_{\sigma \in G} \chi(\sigma^{-1}) v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(m)}.$$

So all the coefficients are zero, contradicting $\chi(e) = n \neq 0$. □

Necessary and sufficient condition for $v_1 * \cdots * v_m = 0$ is still unknown. See [16, 17, 6] for some partial results.

The conclusion in Theorem 4.3.2 is weaker than Theorem 3.3.2.

The necessary and sufficient condition for $u^* = v^* (\neq 0)$ is not known for general G and χ. When $G = S_m$ (and its irreducible characters are called **immanent**), necessary and sufficient conditions are given [25]. A necessary condition for equality of immanantal decomposable tensors, unconstrained by the families of vectors $(u_1, \ldots, u_m)$ and $(v_1, \ldots, v_m)$, was given in [23].

The following is a necessary condition.

Theorem 4.3.3. If $u^* = v^* \neq 0$, then $\langle u_1, \ldots, u_m \rangle = \langle v_1, \ldots, v_m \rangle$.

Proof. Set $W := \langle v_1, \ldots, v_m \rangle$. Let $\{e_1, \ldots, e_k\}$ $(k \leq m)$ be a basis of W. Suppose on the contrary, $u_i \notin W$ for some i. Then extend $\{e_1, \ldots, e_k\}$ to a basis $E = \{e_1, \ldots, e_k, e_{k+1} = u_i, e_{k+1}, \ldots, e_n\}$ of V and let $\{f_1, \ldots, f_n\}$ be the dual basis of V^*, i.e., $f_i(e_j) = \delta_{ij}$. So

$$f_r(v_j) = 0, \quad \text{if } k < r \leq n \text{ and } j = 1, \ldots, m \tag{4.14}$$
$$f_s(u_i) = 0, \quad \text{if } 1 \leq s \leq k \tag{4.15}$$

Now for each $\alpha \in \Gamma_{m,n}$, $\prod_{t=1}^{m} f_{\alpha(t)} : \times^m V \to \mathbb{C}$ is multilinear. So by the unique factorization property of $\otimes$, there is $T_\alpha \in (\otimes^m V)^*$ such that

$$\prod_{t=1}^{m} f_{\alpha(t)} = T_\alpha \otimes .$$

In particular

$$T_\alpha e_\gamma^\otimes = T_\alpha \otimes e_\gamma = (\prod_{t=1}^{m} f_{\alpha(t)}) e_\gamma = \prod_{t=1}^{m} f_{\alpha(t)}(e_{\gamma(t)}) = \delta_{\alpha\gamma}.$$

In other words, $\{T_\alpha : \alpha \in \Gamma_{m,n}\}$ is a basis of $(\otimes^m V)^*$ dual to $E_\otimes$. From $v^* \neq 0$, we claim that there is $\beta \in \Gamma_{m,n}$ such that $T_\beta v^* \neq 0$. Reason: if $T_\beta v^* = 0$ for all β, then write (since $v^* \in \otimes^m V$)

$$v^* = \sum_{\gamma \in \Gamma_{m,n}} a_\gamma e_\gamma^\otimes$$

so that for all $\beta \in \Gamma_{m,n}$

$$0 = T_\beta v^* = \sum_{\gamma \in \Gamma_{m,n}} a_\gamma T_\beta e_\gamma^\otimes = \sum_{\gamma \in \Gamma_{m,n}} a_\gamma \delta_{\beta\gamma} = a_\beta,$$

contradicting that $v^* \neq 0$. Now

$$0 \neq T_\beta v^* = \frac{\chi(e)}{|G|} \sum_{\sigma \in G} \chi(\sigma^{-1}) T_\beta v_\sigma^\otimes = \frac{\chi(e)}{|G|} \sum_{\sigma \in G} \chi(\sigma^{-1}) \prod_{t=1}^{m} f_{\beta(t)}(v_{\sigma(t)}). \quad (4.16)$$

So $\beta \in \Gamma_{m,k}$ (otherwise some $\beta(t) > k$ so that $f_{\beta(t)}(v_{\sigma(t)}) = 0$ from (4.14)). From (4.15),

$$0 \neq T_\beta v^* = T_\beta u^* = \frac{\chi(e)}{|G|} \sum_{\sigma \in G} \chi(\sigma^{-1}) \prod_{t=1}^{m} f_{\beta(t)}(u_{\sigma(t)}) = 0,$$

a contradiction. So $u_1, \ldots, u_m \in \langle v_1, \ldots, v_m \rangle$. By symmetry, $\langle u_1, \ldots, u_m \rangle = \langle v_1, \ldots, v_m \rangle$. □

The conclusion in Theorem 4.3.3 is weaker than Theorem 3.3.3.

The inner product on V induces an inner product on $\otimes^m V$. Such induced inner product descends on $V_\chi(G)$ via restriction. We now see the explicit form of the induced inner product on $V_\chi(G)$.

Theorem 4.3.4. Let V be an inner product space. Then the induced inner product on $V_\chi(G)$ is given by

$$(u^*, v^*) = (u^\otimes, v^*) = (u^*, v^\otimes) = \frac{\chi(e)}{|G|} \sum_{\sigma \in G} \chi(\sigma) \prod_{t=1}^{m} (u_t, v_{\sigma(t)}).$$

Proof. Since $T(G,\chi)$ is an orthogonal projection Theorem 4.2.3(a) and $\chi(\sigma^{-1}) = \overline{\chi(\sigma)}$,

$$\begin{aligned}
(u^*, v^*) &= (T(G,\chi)u^{\otimes}, T(G,\chi)v^{\otimes}) \\
&= (u^{\otimes}, T(G,\chi)v^{\otimes}) = (u^{\otimes}, v^*) = (u^*, v^{\otimes}) \\
&= \left(\frac{\chi(e)}{|G|} \sum_{\sigma\in G} \chi(\sigma) u^{\otimes}_{\sigma^{-1}}, v^{\otimes} \right) \\
&= \frac{\chi(e)}{|G|} \sum_{\sigma\in G} \chi(\sigma)(u^{\otimes}_{\sigma^{-1}}, v^{\otimes}) \\
&= \frac{\chi(e)}{|G|} \sum_{\sigma\in G} \chi(\sigma) \prod_{t=1}^{m} (u_{\sigma^{-1}(t)}, v_t) \quad \text{(induced inner product)} \\
&= \frac{\chi(e)}{|G|} \sum_{\sigma\in G} \chi(\sigma) \prod_{t=1}^{m} (u_t, v_{\sigma(t)}).
\end{aligned}$$

□

Let

$$G_\alpha := \{\sigma \in G : \alpha\sigma = \alpha\} < G$$

be the **stabilizer** of $\alpha \in \Gamma_{m,n}$.

Theorem 4.3.5. (a) Let $E = \{e_1, \ldots, e_n\}$ be an orthonormal basis of V. Then

$$(e^*_\alpha, e^*_\beta) = \frac{\chi(e)}{|G|} \sum_{\sigma\in G} \chi(\sigma)\delta_{\alpha,\beta\sigma}, \quad \alpha, \beta \in \Gamma_{m,n}.$$

(b)

$$\|e^*_\alpha\|^2 = \frac{\chi(e)}{|G|} \sum_{\sigma\in G_\alpha} \chi(\sigma), \quad \alpha \in \Gamma_{m,n}.$$

Proof. (a) From Theorem 4.3.4

$$(e^*_\alpha, e^*_\beta) = \frac{\chi(e)}{|G|} \sum_{\sigma\in G} \chi(\sigma) \prod_{i=1}^{m} (e_{\alpha(t)}, e_{\beta\sigma(t)}) = \frac{\chi(e)}{|G|} \sum_{\sigma\in G} \chi(\sigma)\delta_{\alpha,\beta\sigma}.$$

(b) It follows from (a) by setting $\beta = \alpha$:

$$\|e^*_\alpha\|^2 = \frac{\chi(e)}{|G|} \sum_{\sigma\in G} \chi(\sigma)\delta_{\alpha,\alpha\sigma} = \frac{\chi(e)}{|G|} \sum_{\sigma\in G_\alpha} \chi(\sigma).$$

□

Theorem 4.3.6. Let V be a vector space with $\dim V = n$. Then

$$\dim V_\chi(G) = \frac{\chi(e)}{|G|} \sum_{\sigma\in G} \chi(\sigma) n^{c(\sigma)}.$$

Proof. Since $T(G,\chi)$ is an orthogonal projection by Theorem 4.2.3(a),

$$\begin{aligned}\dim V_\chi(G) &= \operatorname{rank} T(G,\chi) = \operatorname{tr} T(G,\chi)\\ &= \frac{\chi(e)}{|G|}\sum_{\sigma\in G}\chi(\sigma)\operatorname{tr} P(\sigma)\\ &= \frac{\chi(e)}{|G|}\sum_{\sigma\in G}\chi(\sigma)n^{c(\sigma)}. \quad \text{(by Theorem 4.1.2(f))}\end{aligned}$$

□

Since $\dim \otimes^m V = n^m$ we have

$$n^m = \dim \otimes^m V = \sum_{\chi\in I(G)} V_\chi(G) = \sum_{\chi\in I(G)} \frac{\chi(e)}{|G|}\sum_{\sigma\in G}\chi(\sigma)n^{c(\sigma)}$$

which also follows from Theorem 2.4.15.

Problems

1. Let $E = \{e_1,\ldots,e_n\}$ be an orthonormal basis of V. Prove that if $\alpha \in Q_{m,n}$, then $\|e^*_\alpha\|^2 = \frac{\chi(e)^2}{|G|}$.
2. Let $E = \{e_1,\ldots,e_n\}$ be an orthonormal basis of V and $\alpha \in \Gamma_{m,n}$, $\tau\in G$. Prove that $(e^*_\alpha, e^*_{\alpha\tau}) = \frac{\chi(e)}{|G|}\sum_{\sigma\in G_\alpha}\chi(\tau^{-1}\sigma)$.
3. Suppose $m \le n$. Prove that $V_\chi(G) \neq 0$.
4. Suppose χ is an irreducible character of G and $\alpha \in \Gamma_{m,n}$. Prove that $\frac{1}{|G_\alpha|}\sum_{\sigma\in G_\alpha}\chi(\sigma)$ is a nonnegative integer (Hint: consider the restriction of χ onto G_α).

Solutions to Problem 4.3

1. By Theorem 4.3.5(b) with $G_\alpha = e$.
2. By Theorem 4.3.5(a),

$$(e^*_\alpha, e^*_{\alpha\tau}) = \frac{\chi(e)}{|G|}\sum_{\sigma\in G}\chi(\sigma)\delta_{\alpha,\alpha\tau\sigma} = \frac{\chi(e)}{|G|}\sum_{\sigma\in G}\chi(\tau^{-1}\sigma)\delta_{\alpha,\alpha\sigma} = \frac{\chi(e)}{|G|}\sum_{\sigma\in G_\alpha}\chi(\tau^{-1}\sigma).$$

3.
4. which is a nonnegative integer.

4.4 Bases of symmetry classes of tensors

In the last section we have a formula for the dimension of $V_\chi(G)$ and the expression

$$\dim V_\chi(G) = \frac{\chi(e)}{|G|} \sum_{\sigma \in G} \chi(\sigma) n^{c(\sigma)}$$

is not simple at all. In this section we want to construct a basis of $V_\chi(G)$. It involves two steps: (1) express $V_\chi(G)$ as a direct sum of orbital subspaces (Theorem 4.4.5 and (2) find bases for orbital subspaces. To this end we first decompose $\Gamma_{m,n}$ via G.

Let $G < S_m$. Recall

$$\Gamma_{m,n} = \{\alpha = (\alpha(1), \ldots, \alpha(m)),\ 1 \le \alpha(j) \le n,\ j = 1, \ldots, m\}.$$

Each $\sigma \in G$ acts on $\Gamma_{m,n}$, i.e., $\bar\sigma : \Gamma_{m,n} \to \Gamma_{m,n}$ defined by

$$\bar\sigma(\alpha) = \alpha\sigma^{-1}, \quad \alpha \in \Gamma_{m,n}.$$

Since

$$\bar\sigma\bar\tau(\alpha) = \bar\sigma(\alpha\tau^{-1}) = \alpha\tau^{-1}\sigma^{-1} = \overline{\sigma\tau}(\alpha), \quad \alpha \in \Gamma_{m,n}$$

the action $\sigma \mapsto \bar\sigma$ is a a homomorphism (isomorphism if $n > 1$) from G into the manifestation of S_{n^m} as a group of permutations of $\Gamma_{m,n}$, i.e.,

$$\bar G = \{\bar\sigma : \sigma \in G\}$$

is a group of permutations of $\Gamma_{m,n}$.

Two sequences α and β in $\Gamma_{m,n}$ are said to be equivalent modulo G (more precisely, modulo $\bar G$), denoted by $\alpha \equiv \beta\,(\mathrm{mod}\ G)$, if there exists $\sigma \in G$ such that $\beta = \overline{\sigma^{-1}}(\alpha) = \alpha\sigma$, where $\alpha\sigma := (\alpha(\sigma(1)), \ldots, \alpha(\sigma(m)))$. This equivalence relation partitions $\Gamma_{m,n}$ into equivalence classes.

For each $\alpha \in \Gamma_{m,n}$ the equivalence class

$$\Gamma_\alpha = \{\alpha\sigma : \sigma \in G\}$$

is called the **orbit** containing α. So we have the following disjoint union

$$\Gamma_{m,n} = \cup_{\alpha \in \Delta} \Gamma_\alpha.$$

Let Δ be a system of representatives for the orbits such that each sequence in Δ is first in its orbit relative to the lexicographic order. Now

$$H < G \Rightarrow \Delta_G \subset \Delta_H$$

since $\sigma \in \Delta_G$ implies that $\sigma \le \sigma\pi$ for all $\pi \in G$ and thus for all $\pi \in H$.

Consider the example $\Gamma_{3,2}$.

1. If $G = S_3$, then

   $$\Gamma_{3,2} = \{(1,1,1), (1,1,2), (1,2,1), (1,2,2), (2,1,1), (2,1,2), (2,2,1), (2,2,2)\}$$

   has 4 orbits:

(a) $\{(1,1,1)\}$.
(b) $\{(1,1,2),(1,2,1),(2,1,1)\}$.
(c) $\{(1,2,2),(2,1,2),(2,2,1)\}$.
(d) $\{(2,2,2)\}$.

Then $\Delta = \{(1,1,1),(1,1,2),(1,2,2),(2,2,2)\}$.

2. If $G = \{e,(12)\}$, then $\Gamma_{3,2}$ has 6 orbits:

$$\Delta = \{(1,1,1),(1,1,2),(1,2,1),(1,2,2),(2,2,1),(2,2,2)\}.$$

We use the notations $\alpha \le \beta$ and $\alpha < \beta$ with respect to the lexicographical order. The following theorem is clear.

Theorem 4.4.1. If $\alpha \in \Gamma_{m,n}$, then $\alpha \in \Delta$ if and only if $\alpha \le \alpha\sigma$ for all $\sigma \in G$.

Theorem 4.4.2. (a) $G_{m,n} \subset \Delta$ for all $G < S_m$.

(b) If $G = S_m$, then $\Delta = G_{m,n}$.

(c) If $m \le n$ and $\Delta = G_{m,n}$, then $G = S_m$.

Proof. (a) If $\alpha \in G_{m,n}$, then clearly $\alpha \le \alpha\sigma$ for all $\sigma \in G$. So $\alpha \in \Delta$.

(b) If $\alpha \in \Delta$, then there is $\sigma \in S_m$ such that $\alpha\sigma \in G_{m,n} \subset \Delta$. So $\alpha = \alpha\sigma \in G_{m,n}$. Then apply (a).

(c) Suppose $G \ne S_m$. Then there is a transposition $\tau \notin G$. Since $m \le n$, $\alpha = (1,\dots,m) \in Q_{m,n} \subset G_{m,n}$. Now $\alpha\tau$ and α are not in the same orbit (otherwise there is $\sigma \in G$ such that $\alpha\tau = \alpha\sigma$; but then $\tau = \sigma \in G$, a contradiction). In other words, $\alpha\tau\pi \ne \alpha$ for any $\pi \in G$. Moreover, since $\tau\pi \ne e$, we have $\alpha\tau\pi \notin G_{m,n}$ for any $\pi \in G$. Thus the representative (in Δ) of the orbit $\Gamma_{\alpha\tau}$ is not contained in $G_{m,n}$, contradicting $\Delta = G_{m,n}$. $\square$

??? The conclusion in (c) is not valid if $m > n$, for example, $\Gamma_{3,2}$ with $G = \{e,(12)\}$.

We develop some combinatorial machinery in order to discuss the structure of a basis for $V_\chi(G)$.

Recall that the stabilizer $G_\alpha < G$ is a subgroup of G. For any $\tau \in G$, recall the right coset of G_α in G

$$G_\alpha\tau := \{\sigma\tau : \sigma \in G_\alpha\}$$

containing τ. Since G is a disjoint union of right cosets, i.e., there are $\tau_1,\dots,\tau_k \in G$ such that

$$G = G_\alpha\tau_1 \cup \dots \cup G_\alpha\tau_k.$$

Moreover $|G_\alpha\tau_i| = |G_\alpha|$ so that $k|G_\alpha| = |G|$. Clearly $\Gamma_\alpha = \{\alpha\tau_1,\dots,\alpha\tau_k\}$ ($\tau_1,\dots,\tau_k$ depend on α).

The following is a combinatorial formula involving Δ and $\Gamma_{m,n}$ that we will need in the construction of a basis for $V_\chi(G)$.

Theorem 4.4.3. Let $\varphi : \Gamma_{m,n} \to W$ be a map. Then

$$\sum_{\gamma \in \Gamma_{m,n}} \varphi(\gamma) = \sum_{\alpha \in \Delta} \frac{1}{|G_\alpha|} \sum_{\sigma \in G} \varphi(\alpha\sigma).$$

Proof. Recall $\Gamma_{m,n} = \cup_{\alpha \in \Delta} \Gamma_\alpha$ so that

$$\sum_{\gamma \in \Gamma_{m,n}} \varphi(\gamma) = \sum_{\alpha \in \Delta} \sum_{\gamma \in \Gamma_\alpha} \varphi(\gamma).$$

Write $\Gamma_\alpha = \{\alpha\tau_1, \ldots, \alpha\tau_k\}$. Notice that

$$|\{\sigma \in G : \alpha\sigma = \alpha\tau_j\}| = |\{\sigma \in G : \alpha\sigma\tau_j^{-1} = \alpha\}| = |\{\pi \in G : \alpha\pi = \alpha\}| = |G_\alpha|$$

$j = 1, \ldots, k$, i.e., when σ runs through G, $\alpha\sigma$ meets $\alpha\tau_j$ with $|G_\alpha|$ occurrences. So

$$\sum_{\gamma \in \Gamma_\alpha} \varphi(\gamma) = \sum_{j=1}^{k} \varphi(\alpha\tau_j) = \frac{1}{|G_\alpha|} \sum_{\sigma \in G} \varphi(\alpha\sigma).$$

□

Let $E = \{e_1, \ldots, e_n\}$ be a basis of V. Since $E_\otimes = \{e_\gamma^\otimes : \gamma \in \Gamma_{m,n}\}$ is a basis of $\otimes^m V$ and $\Gamma_{m,n} = \cup_{\alpha \in \Delta} \Gamma_\alpha$,

$$V_\chi(G) = \langle e_\gamma^* : \gamma \in \Gamma_{m,n} \rangle = \sum_{\alpha \in \Delta} \langle e_\gamma^* : \gamma \in \Gamma_\alpha \rangle = \sum_{\alpha \in \Delta} \langle e_{\alpha\sigma}^* : \sigma \in G \rangle. \tag{4.17}$$

The subspace

$$O_\alpha := \langle e_{\alpha\sigma}^* : \sigma \in G \rangle \subset V_\chi(G)$$

is called the **orbital subspace** associated with $\alpha \in \Gamma_{m,n}$. The following gives some relation between orbital spaces.

Theorem 4.4.4. Let $E = \{e_1, \ldots, e_n\}$ be an orthonormal basis of the inner product space V. Then

(a) $(e_\alpha^*, e_\beta^*) = 0$ if $\alpha \not\equiv \beta \,(\text{mod } G)$.

(b) $\|e_\alpha^*\| = \|e_\beta^*\|$ if $\alpha \equiv \beta \,(\text{mod } G)$.

Proof. (a) Notice that $(e_\alpha^*, e_\beta^*) = 0$ if $\alpha \not\equiv \beta \,(\text{mod } G)$ by Theorem 4.3.5(a).

(b) By Theorem 4.3.5(b) it suffices to show that if $\alpha = \beta\theta$ for some $\theta \in G$, then $\sum_{\sigma \in G_\alpha} \chi(\sigma) = \sum_{\tau \in G_\beta} \chi(\tau)$. Indeed $G_\alpha = \theta^{-1} G_\beta \theta$ by Problem 2. So

$$\sum_{\sigma \in G_\beta} \chi(\tau) = \sum_{\tau \in G_\beta} \chi(\theta^{-1}\tau\theta) = \sum_{\sigma \in \theta^{-1} G_\beta \theta} \chi(\sigma) = \sum_{\sigma \in G_\alpha} \chi(\sigma).$$

□

So for any $\alpha \in \Gamma_{m,n}$, $e^*_\alpha = 0$ if and only if $e^*_{\alpha\sigma} = 0$ for all $\sigma \in G$. In other words, if one element in the orbital space O_α is zero, then all elements in O_α are zero. By Theorem 4.3.5 $e^*_\alpha \neq 0$ if and only if $\sum_{\sigma\in G_\alpha} \chi(\sigma) \neq 0$. So we define

$$\bar{\Delta} := \{\alpha \in \Delta : \sum_{\sigma\in G_\alpha} \chi(\sigma) \neq 0\} \subset \Delta \tag{4.18}$$

where $G_\alpha = \{\sigma \in G : \alpha\sigma = \alpha\}$ is the stabilizer (a subgroup of G) of α. Notice that

$$\frac{1}{|G_\alpha|} \sum_{\sigma\in G_\alpha} \chi(\sigma) = (\chi|_{G_\alpha}, 1), \tag{4.19}$$

the inner product (see Chapter 2) of the principal character and the character $\chi|_{G_\alpha}$ (not necessarily irreducible), the restriction of χ onto G_α (it is a character of G_α because the restriction of the representation $A : G \to \mathrm{GL}_n(\mathbb{C})$ to G_α is a representation). Notice that $(\chi|_{G_\alpha}, 1)$ is a nonnegative integer since it denotes the multiplicity of the principal character in $\chi|_{G_\alpha}$ by Theorem 2.4.7 and Theorem 2.4.12.

Theorem 4.4.5. Let $E = \{e_1, \ldots, e_n\}$ be a basis of V. Then

$$V_\chi(G) = \oplus_{\alpha\in\bar{\Delta}} \langle e^*_{\alpha\sigma} : \sigma \in G\rangle. \tag{4.20}$$

Proof. Equip V with an inner product so that E is an orthonormal basis (Theorem 1.6.3). When $\alpha \in \Delta \setminus \bar{\Delta}$, $e^*_\alpha = e^*_{\alpha\sigma} = 0$ for all $\sigma \in G$. So (4.17) becomes (4.20). By Theorem 4.4.4, the orbital subspaces are orthogonal, i.e., (4.20) is an orthogonal sum and thus is a direct sum. □

Notice that Δ depends on $\Gamma_{m,n}$ and G alone but $\bar{\Delta}$ depends on χ as well and is a very important set. We now see some basic properties of $\bar{\Delta}$ (Compare Theorem 4.4.2).

Theorem 4.4.6. (a) $Q_{m,n} \subset \bar{\Delta}$ for all $G < S_m$ ($Q_{m,n} = \emptyset$ if $m > n$).

(b) If $G = S_m$ and $\chi = \varepsilon$, then $\bar{\Delta} = Q_{m,n}$.

(c) If $1 < m \leq n$ and $\bar{\Delta} = Q_{m,n}$, then $G = S_m$ and $\chi = \varepsilon$.

Proof. (a) By Theorem 4.4.2(a) $Q_{m,n} \subset G_{m,n} \subset \Delta$. If $\alpha \in Q_{m,n}$, then $G_\alpha = \{e\}$ so that $\sum_{\sigma\in G_\alpha} \chi(\sigma) = \chi(e) \neq 0$. Thus $\alpha \in \bar{\Delta}$ by (4.20).

(b) Since $G = S_m$, by Theorem 4.4.2(b) $\Delta = G_{m,n}$. If $\gamma \in \Delta \setminus Q_{m,n}$, then there are $i \neq j$ such that $\gamma(i) = \gamma(j)$. Define the transposition $\tau = (ij)$. Then $\tau \in G_\gamma$. Since $\varepsilon(\tau) = -1$,

$$\sum_{\sigma\in G_\gamma} \varepsilon(\sigma) = \sum_{\sigma\in G_\gamma} \varepsilon(\tau\sigma) = -\sum_{\sigma\in G_\gamma} \varepsilon(\sigma).$$

So $\sum_{\sigma\in G_\gamma} \varepsilon(\sigma) = 0$, i.e., $\gamma \notin \bar{\Delta}$. So $\bar{\Delta} \subset Q_{m,n}$ and hence $\bar{\Delta} = Q_{m,n}$ by (a).

(c) From $1 < m \leq n$, $\alpha = (1, \ldots, i-1, i, i, i+2, \ldots, m) \in G_{m,n}$. Since $\alpha \notin Q_{m,n} = \bar{\Delta}$, $\sum_{\sigma\in G_\alpha} \chi(\sigma) = 0$. Notice that $(i, i+1) \in G$ for all $i = 1, \ldots, m-1$

(if the transposition $(i,i+1)\notin G$, then we would have $G_\alpha=\{e\}$, contradicting $\sum_{\sigma\in G_\alpha}\chi(\sigma)=0$). Then $G=S_m$ by Problem 2.1 #2.

Let $\tau:=(i,i+1)\in G$. Then $G_\alpha=\{e,\tau\}$ so that

$$\chi(e)+\chi(\tau)=\sum_{\sigma\in G_\alpha}\chi(\sigma)=0.$$

So $\chi(\tau)=-\chi(e)=-k$ where k is the degree of χ. Then from Problem 1, $A(\tau)=-I_k$ if A is the (unitary) representation for χ. By Problem 2.1 #2 for any $\sigma\in G$

$$\sigma=\tau_1\cdots\tau_r$$

where each τ_j is a transposition of the form $(i,i+1)$. So

$$A(\sigma)=A(\tau_1\cdots\tau_r)=A(\tau_1)\cdots A(\tau_r)=(-1)^r I_k.$$

So

$$\chi(\sigma)=\operatorname{tr}A(\sigma)=(-1)^r k=\varepsilon(\tau_1)\cdots\varepsilon(\tau_r)k=k\varepsilon(\sigma)$$

for all $\sigma\in G$, i.e., $\chi=k\varepsilon$. Since χ is irreducible, we must have $k=1$. □

Let

$$\Omega:=\{\alpha\in\Gamma_{m,n}:\sum_{\alpha\in G_\alpha}\chi(\alpha)\neq 0\}.$$

Note that $\bar\Delta=\Delta\cap\Omega$ and

$$\Omega=\cup_{\alpha\in\bar\Delta}\Gamma_\alpha=\cup_{\alpha\in\bar\Delta}\{\alpha\sigma:\sigma\in G\}. \tag{4.21}$$

So for $\alpha\in\Gamma_{m,n}$, $e^*_\alpha\neq 0$ if and only if $\alpha\in\Omega$. So the set $\{e^*_\alpha:\alpha\in\Omega\}$ consists of the nonzero elements of $\{e^*_\alpha:\alpha\in\Gamma_{m,n}\}$ (a spanning set of $V_\chi(G)$).

Moreover from (4.20)

$$V_\chi(G)=\oplus_{\alpha\in\bar\Delta}O_\alpha, \tag{4.22}$$

where $O_\alpha=\langle e^*_{\alpha\sigma}:\sigma\in G\rangle$. In order to find a basis for $V_\chi(G)$, it suffices to find bases of the orbital subspaces O_α, $\alpha\in\bar\Delta$. Define

$$s_\alpha:=\dim O_\alpha=\dim\langle e^*_{\alpha\sigma}:\sigma\in G\rangle.$$

Notice that s_α depends on χ.

Theorem 4.4.7. (Freese)

$$s_\alpha=\frac{\chi(e)}{|G_\alpha|}\sum_{\sigma\in G_\alpha}\chi(\sigma)=\chi(e)(\chi|_{G_\alpha},1).$$

Proof. Since G is a disjoint union of right cosets of G_α in G, i.e., there are $\tau_1,\dots,\tau_k\in G$ such that

$$G=G_\alpha\tau_1\cup\cdots\cup G_\alpha\tau_k$$

and $\Gamma_\alpha = \{\alpha\tau_1, \dots, \alpha\tau_k\}$. Moreover $|G_\alpha\tau_i| = |G_\alpha|$ so that $k|G_\alpha| = |G|$. Notice that $O_\alpha = T(G,\chi)W_\alpha$, where

$$W_\alpha := \langle e^{\otimes}_{\alpha\sigma} : \sigma \in G\rangle = \langle e^{\otimes}_{\beta} : \beta \in \Gamma_\alpha\rangle.$$

Then $E_\alpha := \{e^{\otimes}_{\alpha\tau_1}, \dots, e^{\otimes}_{\alpha\tau_k}\}$ is a basis of W_α but $\{e^*_{\alpha\tau_1}, \dots, e^*_{\alpha\tau_k}\}$ may not be a basis for O_α.

Since W_α is invariant under $P(\sigma)$ and thus invariant under $T(G,\chi)$, the restriction $T_\alpha(G,\chi) = T(G,\chi)|_{W_\alpha}$ of $T(G,\chi)$ is a linear map on W_α and is a projection (Problem 1.8 #8).

Let $C = (c_{ij}) := [T_\alpha(G,\chi)]^{E_\alpha}_{E_\alpha} \in \mathbb{C}_{k\times k}$ (C depends on α and $s_\alpha := \dim O_\alpha \le k$). From

$$T_\alpha(G,\chi)W_\alpha = T(G,\chi)\langle e^{\otimes}_{\alpha\sigma} : \sigma \in G\rangle = \langle e^*_{\alpha\sigma} : \sigma \in G\rangle,$$

we have (by Theorem 1.8.2)

$$s_\alpha = \dim\langle e^*_{\alpha\sigma} : \sigma \in G\rangle = \operatorname{rank} T_\alpha(G,\chi) = \operatorname{tr} T_\alpha(G,\chi) = \operatorname{tr} C.$$

We now compute C:

$$\begin{aligned}
T_\alpha(G,\chi)e^{\otimes}_{\alpha\tau_j} &= e^*_{\alpha\tau_j} = T(G,\chi)e^{\otimes}_{\alpha\tau_j} \\
&= \frac{\chi(e)}{|G|}\sum_{\sigma\in G}\chi(\sigma)e^{\otimes}_{\alpha\tau_j\sigma^{-1}} \\
&= \frac{\chi(e)}{|G|}\sum_{\pi\in G}\chi(\pi^{-1}\tau_j)e^{\otimes}_{\alpha\pi} \quad (\pi = \tau_j\sigma^{-1}) \\
&= \frac{\chi(e)}{|G|}\sum_{i=1}^{k}\sum_{\pi\in G_\alpha\tau_i}\chi(\pi^{-1}\tau_j)e^{\otimes}_{\alpha\pi} \quad (G = \cup_{i=1}^k G_\alpha\tau_i) \\
&= \sum_{i=1}^{k}\frac{\chi(e)}{|G|}\sum_{\theta\in G_\alpha}\chi(\tau_i^{-1}\theta^{-1}\tau_j)e^{\otimes}_{\alpha\tau_i} \quad (\pi = \theta\tau_i)
\end{aligned}$$

for $j = 1, \dots, k$. So

$$c_{ij} = \frac{\chi(e)}{|G|}\sum_{\sigma\in G_\alpha}\chi(\tau_i^{-1}\sigma\tau_j), \quad i,j = 1,\dots,k.$$

Since $\chi(\tau_i^{-1}\sigma\tau_i) = \chi(\sigma)$ for all i,

$$s_\alpha = \operatorname{tr} C = \sum_{i=1}^{k} c_{ii} = \sum_{i=1}^{k}\frac{\chi(e)}{|G|}\sum_{\sigma\in G_\alpha}\chi(\sigma) = \frac{\chi(e)}{|G_\alpha|}\sum_{\sigma\in G_\alpha}\chi(\sigma).$$

□

Recall that $(\chi|_{G_\alpha}, 1)$ is a nonnegative integer when $\alpha \in \Gamma_{m,n}$. So $s_\alpha = \chi(e)(\chi|_{G_\alpha}, 1) > 1$ for each $\alpha \in \bar{\Delta}$ if χ is not linear. When χ is linear, $s_\alpha = 1$ (Problem 6).

Let $E = \{e_1, \ldots, e_n\}$ be a basis of V. We now construct a basis of $V_\chi(G)$. For each $\alpha \in \bar{\Delta}$, we find a basis of the orbital subspace O_α: choose a lexicographically ordered set $\{\alpha_1, \ldots, \alpha_{s_\alpha}\}$ from $\{\alpha\sigma : \sigma \in G\}$ such that $\{e^*_{\alpha_1}, \ldots, e^*_{\alpha_{s_\alpha}}\}$ is a basis of O_α. Execute this procedure for each $\gamma \in \bar{\Delta}$. If $\{\alpha, \beta, \cdots\}$ is the lexicographically ordered set $\bar{\Delta}$, take

$$\hat{\Delta} = \{\alpha_1, \ldots, \alpha_{s_\alpha}; \beta_1, \ldots, \beta_{s_\beta}; \ldots\}$$

to be ordered as indicated (often we pick $\alpha_1 = \alpha$). Then $\{e^*_\alpha : \alpha \in \hat{\Delta}\}$ is a basis of $V_\chi(G)$ (Clearly $\bar{\Delta} = \{\alpha_1, \beta_1, \ldots,\}$ is in lexicographic order but note that the elements of $\hat{\Delta}$ need not be in lexicographical order, for example, it is possible that $\alpha_2 > \beta_1$). Such order (not unique) in $\hat{\Delta}$ is called an **orbital order**. Clearly

$$\bar{\Delta} \subseteq \hat{\Delta} \subseteq \Omega.$$

Although $\hat{\Delta}$ is not unique, it does not depend on the choice of basis E since Δ, $\bar{\Delta}$ and C do not depend on E, i.e., if $F = \{f_1, \ldots, f_n\}$ is another basis of V, then $\{f^*_\alpha : \alpha \in \hat{\Delta}\}$ is still a basis of $V_\chi(G)$.

Example: Consider $m = n = 3$ and $G = S_3$. Recall that there are three irreducible characters of S_3, namely, the principal character 1, the alternating character ε and the character with degree 2:

$$\chi(e) = 2, \ \chi(23) = \chi(12) = \chi(13) = 0, \ \chi(123) = \chi(132) = -1.$$

We now construct Δ, $\bar{\Delta}$ and $\hat{\Delta}$ for this degree two χ. Since $G = S_m$, $\Delta = G_{m,n}$ for $\Gamma_{3,3}$ by Theorem 4.4.2 and the orbits (equivalence classes) are

$$\begin{aligned}
&\{(1,1,1)\};\\
&\{(1,1,2),(1,2,1),(2,1,1)\};\\
&\{(1,1,3),(1,3,1),(3,1,1)\};\\
&\{(1,2,2),(2,1,2),(2,1,1)\};\\
&\{(1,2,3),(1,3,2),(2,1,3),(2,3,1),(3,1,2),(3,2,1)\};\\
&\{(1,3,3),(3,1,3),(3,3,1)\};\\
&\{(2,2,2)\};\\
&\{(2,2,3),(2,3,2),(3,2,2)\};\\
&\{(2,3,3),(3,2,3),(3,3,2)\};\\
&\{(3,3,3)\}.
\end{aligned}$$

Then by direct computation

$$\bar{\Delta} = \{(1,1,2),(1,1,3),(1,2,2),(1,2,3),(1,3,3),(2,2,3),(2,3,3)\}.$$

For example, if $\alpha = (1,1,1)$, then $G_\alpha = S_3$ and

$$\sum_{\sigma \in G_\alpha} \chi(\sigma) = \sum_{\sigma \in G} \chi(\sigma) = \chi(e) + \chi(123) + \chi(132) = 2 - 1 - 1 = 0.$$

If $\alpha = (1,1,2)$, $G_\alpha = \{e, (12)\}$ so that

$$\sum_{\sigma \in G_\alpha} \chi(\sigma) = \chi(e) + \chi(12) = 2 + 0 = 2.$$

The orbit Γ_ω $(\omega := (1,2,3))$ $|\Gamma_\omega| = 6$ and other orbits are of size 3.

First we consider the "small" orbits. For $\alpha = (1,1,2)$, we have $G_\alpha = \{e, (12)\}$ so that

$$s_\alpha = \frac{\chi(e)}{|G_\alpha|} \sum_{\sigma \in G_\alpha} \chi(\sigma) = \chi(e) + \chi((1\,2)) = 2.$$

With $\tau_1 = e$, $\tau_2 = (23)$, $\tau_3 = (13)$,

$$\Gamma_\alpha = \{(1,1,2), (1,2,1), (2,1,1)\} = \{\alpha\tau_1, \alpha\tau_2, \alpha\tau_3\}.$$

From

$$c_{ij} = \frac{\chi(e)}{|G_\alpha|} \sum_{\sigma \in G} \chi(\tau_i^{-1}\sigma\tau_j) = \frac{1}{3}(\chi(\tau_i^{-1}\tau_j) + \chi(\tau_i^{-1}(12)\tau_j))$$

we have

$$C = \begin{pmatrix} \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \end{pmatrix}$$

Any two columns of C are linearly independent so that we can take $\alpha_1 = \alpha = (1,1,2)$, $\alpha_2 = (1,2,1)$. The computation for the other five "small orbits" is similar, i.e.,

$$(1,1,3),\ (1,2,2),\ (1,3,3,),\ (2,2,3),\ (2,3,3)$$

Now we consider the larger orbit. For $\omega = (1,2,3)$, we have $G_\omega = \{e\}$ so that

$$s_\omega = \frac{\chi(e)}{|G_\omega|} \sum_{\sigma \in G_\omega} \chi(\sigma) = 2\chi(e) = 4$$

With $\tau_1 = e$, $\tau_2 = (23)$, $\tau_3 = (12)$, $\tau_4 = (123)$, $\tau_5 = (132)$, $\tau_6 = (13)$,

$$\Gamma_\omega = \{(1,2,3), (1,3,2), (2,1,3), (2,3,1), (3,1,2), (3,2,1)\} = \{\alpha\tau_1, \ldots, \alpha\tau_6\}.$$

From

$$c_{ij} = \frac{\chi(e)}{|G_\alpha|} \sum_{\sigma \in G} \chi(\tau_i^{-1}\sigma\tau_j) = \frac{1}{3}\chi(\tau_i^{-1}\tau_j))$$

we have

$$C = \begin{pmatrix} \frac{2}{3} & 0 & 0 & -\frac{1}{3} & -\frac{1}{3} & 0 \\ 0 & \frac{2}{3} & -\frac{1}{3} & 0 & 0 & -\frac{1}{3} \\ 0 & -\frac{1}{3} & \frac{2}{3} & 0 & 0 & -\frac{1}{3} \\ -\frac{1}{3} & 0 & 0 & \frac{2}{3} & -\frac{1}{3} & 0 \\ -\frac{1}{3} & 0 & 0 & -\frac{1}{3} & \frac{2}{3} & 0 \\ 0 & -\frac{1}{3} & -\frac{1}{3} & 0 & 0 & \frac{2}{3} \end{pmatrix}$$

Notice that $\operatorname{rank} C = 4$ (for example the columns $1, 2, 3, 4$ or columns $1, 2, 3, 5$ are linearly independent. So we can take $\omega_1 = \omega = (1, 2, 3)$, $\omega_2 = (1, 3, 2)$, $\omega_3 = (2, 1, 3)$ and $\omega_4 = (2, 3, 1)$. Hence

$$\begin{aligned} \hat{\Delta} = \; & \{(1,1,2),(1,2,1); \\ & (1,1,3),(1,3,1); \\ & (1,2,2),(2,1,2); \\ & (1,2,3),(1,3,2),(2,1,3),(2,3,1); \\ & (1,3,3),(3,1,3); \\ & (2,2,3),(2,3,2); \\ & (2,3,3),(3,2,3)\} \end{aligned}$$

in which the order is not lexicographic.

It is known (Problem 6) that $\bar{\Delta} = \hat{\Delta}$ if and only if χ is linear. In such cases, $\{e^*_\alpha : \alpha \in \bar{\Delta}\}$ is an orthogonal basis of $V_\chi(G)$ if $E = \{e_1, \dots, e_n\}$ is an orthogonal basis of V. If $\chi(e) > 1$, can we choose $\hat{\Delta}$ such that $\{e^*_\alpha : \alpha \in \bar{\Delta}\}$ is an orthogonal basis for $V_\chi(G)$? When $G = D_m$ (the diheral group) such $\hat{\Delta}$ exists for every $\chi \in I(G)$ if and only if m is a power of 2 (Wang and Gong (1991) [34, 35], Holmes and Tam (1992) [4]). On the other hand, every doubly transitive subgroup $G < S_m$ has an irreducible character for which no such $\hat{\Delta}$ exists (Holmes (1995) [3])

We give several common examples of symmetry classes of tensors and induced operators.

Example 4.4.8. Assume $1 \le m \le n$, $G = S_m$, and $\chi = \varepsilon$. Then $V_\chi(G)$ is the mth exterior space $\wedge^m V$, $\bar{\Delta} = \hat{\Delta} = Q_{m,n}$, the set of strictly increasing sequences in $\Gamma_{m,n}$, $\Delta = G_{m,n}$, the set of nondecreasing sequences in $\Gamma_{m,n}$. See Problem 6 and Theorem 4.4.6(b).

Example 4.4.9. Assume $G = S_m$ and $\chi \equiv 1$, the principal character. Then $V_1(G)$ is the mth completely symmetric space $\bullet^m V$, $\bar{\Delta} = \hat{\Delta} = \Delta = G_{m,n}$. See Problem 7.

Example 4.4.10. Assume $G = \{e\} < S_m$ ($\chi \equiv 1$). Then $V_\chi(G) = \otimes^m V$, $\bar{\Delta} = \hat{\Delta} = \Delta = \Gamma_{m,n}$. See Problem 7.

We know that from Theorem 4.3.6

$$\dim V_\chi(G) = |\hat{\Delta}| = \frac{\chi(e)}{|G|} \sum_{\sigma \in G} \chi(\sigma) n^{c(\sigma)}.$$

The following gives another expression for $|\hat{\Delta}|$ in terms of s_α.

Theorem 4.4.11.

$$\begin{aligned}\dim V_\chi(G) &= |\hat{\Delta}| = \sum_{\alpha\in\bar{\Delta}} s_\alpha = \sum_{\alpha\in\Delta} s_\alpha\\ &= \sum_{\alpha\in\Delta} \frac{\chi(e)}{|G_\alpha|}\sum_{\sigma\in G_\alpha}\chi(\sigma)\\ &= \frac{\chi(e)}{|G|}\sum_{\alpha\in\Gamma_{m,n}}\sum_{\sigma\in G_\alpha}\chi(\sigma).\end{aligned}$$

Proof. It suffices to establish the last equality. By the proofs of Theorem 4.4.3 and Theorem 4.4.4

$$\begin{aligned}\sum_{\alpha\in\Gamma_{m,n}}\sum_{\pi\in G_\alpha}\chi(\pi) &= \sum_{\alpha\in\Delta}\frac{1}{|G_\alpha|}\sum_{\sigma\in G}\sum_{\pi\in G_{\alpha\sigma}}\chi(\pi)\\ &= \sum_{\alpha\in\Delta}\frac{1}{|G_\alpha|}\sum_{\sigma\in G}\sum_{\pi\in G_\alpha}\chi(\pi) \qquad (G_{\alpha\sigma} = \sigma^{-1}G_\alpha\sigma)\\ &= |G|\sum_{\alpha\in\bar{\Delta}}\frac{1}{|G_\alpha|}\sum_{\pi\in G_\alpha}\chi(\pi).\end{aligned}$$

□

Problems

1. If $U \in \mathbb{C}_{n\times n}$ is unitary and $\operatorname{tr} U = -n$, prove that $U = -I_n$.
2. Let $\theta \in G < S_m$, $\beta \in \Gamma_{m,n}$. Prove that $G_{\beta\theta} = \theta^{-1}G_\beta\theta$.
3. Prove that $\sum_{\alpha\in\Delta}\frac{|G|}{|G_\alpha|} = |\Gamma_{m,n}|$ and $|\Delta| = \sum_{\alpha\in\Gamma_{m,n}}\frac{|G_\alpha|}{|G|}$.
4. Let $\{e_1,\ldots,e_n\}$ be a basis of V. Prove that $\{e^*_\alpha : \alpha \in \bar{\Delta}\}$ is linearly independent.
5. Write $s_\alpha(\chi) = s_\alpha = \sum_{\alpha\in\Delta}\frac{\chi(e)}{|G_\alpha|}\sum_{\sigma\in G_\alpha}\chi(\sigma)$ to indicate its dependence on χ, an irreducible character of $G < S_m$ and $\alpha \in \Gamma_{m,n}$.
   (a) $s_\alpha(\chi)$ divides $\chi(e)$ for all $\chi \in I(G)$.
   (b) $\sum_{\chi\in I(G)} s_\alpha(\chi) = \frac{|G|}{|G_\alpha|}$.
   (c) Unless $G_\alpha(\chi) = G$ and $\chi \equiv 1$, we have $s_\alpha < \frac{|G|}{|G_\alpha|}$.
   (d) For each $\tau \in S_n$ and $\pi \in G$, $s_{\tau\alpha\pi}(\chi) = s_\alpha(\chi)$. Here $\tau\alpha = (\tau(\alpha(1)),\ldots,\tau(\alpha(m))) \in \Gamma_{m,n}$.

(e) if $\alpha \in \Omega$, then $\chi(e) \le s_\alpha(\chi) \le \chi(e)^2$. (Hint: Problem 2.4 #1).

6. Prove that $\bar{\Delta} = \hat{\Delta}$ if and only if $\chi(e) = 1$.

7. Prove that $\Delta = \bar{\Delta}$ (or $\Delta = \hat{\Delta}$) if and only if $\chi \equiv 1$.

8. Let $\alpha \in \Omega$. Prove that if $\chi(e) = 1$, then for any $\sigma \in G_\alpha$, $\chi(\sigma) = 1$, i.e., $\chi|_{G_\alpha} \equiv 1$.

9. Prove the identities: $\binom{n}{m} = \frac{1}{m!}\sum_{\sigma\in S_m} \varepsilon(\sigma) n^{c(\sigma)}$, $\binom{n+m-1}{m} = \frac{1}{m!}\sum_{\sigma\in S_m} n^{c(\sigma)}$.

10. Suppose $V_\chi(G) \ne 0$. Prove that if $G \ne S_m$ or $\chi \ne \varepsilon$, then there is $\alpha \in \bar{\Delta}$ such that $\alpha(i) = \alpha(j)$ for some $i \ne j$.

11. Let $C \in \mathbb{C}_{k\times k}$ be the matrix in the proof of Theorem 4.4.7, i.e., $c_{ij} = \frac{\chi(e)}{|G|}\sum_{\sigma\in G_\alpha} \chi(\tau_i^{-1}\sigma\tau_j)$. Give a direct proof of $C^2 = C$.

Solution to Problems 4.4

1. Since the diagonal entries u_{ii} satisfies $|u_{ii}| \le 1$ since $U \in \mathbb{C}_{n\times n}$ is unitary. So $-n = \operatorname{tr} U$ implies that $n = |\operatorname{tr} U| = \sum_{i=1}^n u_{ii} \le \sum_{i=1}^n |u_{ii}| = n$. By the equality case of the triangle inequality, $u_{ii} = -n$ and hence $U = -I_n$ since U has orthonormal rows (columns).

2. $\gamma \in G_{\beta\theta} \Leftrightarrow \beta\theta\gamma = \beta\theta \Leftrightarrow \beta\theta\gamma\theta^{-1} = \beta \Leftrightarrow \theta\gamma\theta^{-1} \in G_\beta$, i.e., $\gamma \in \theta^{-1}G_\beta\theta$.

3.

4. By Theorem 4.4.5 $V_\chi(G) = \oplus_{\alpha\in\bar{\Delta}}\langle e^*_{\alpha\sigma} : \sigma \in G\rangle$. So $\{e^*_\alpha : \alpha \in \bar{\Delta}\}$ is linearly independent.

5.

6. From Theorem 4.4.7 and the definition of $\hat{\Delta}$, $\bar{\Delta} = \hat{\Delta} \Leftrightarrow 1 = s_\alpha(= \chi(e)(\chi|_{G_\alpha}, 1))$ for all $\alpha \in \bar{\Delta}$. Then notice that $(\chi|_{G_\alpha}, 1)$ is a positive integer for each $\alpha \in \bar{\Delta}$.

7.

8. Suppose $\chi \equiv 1$. For any $\alpha \in \Delta$, $\sum_{\sigma\in G_\alpha} \chi(\sigma) = |G_\alpha| \ne$ so that $\alpha \in \bar{\Delta}$. Hence $\bar{\Delta} = \Delta$. Then use Problem 6 to have $\bar{\Delta} = \hat{\Delta}$ since $\chi(e) = 1$.

   Conversely suppose $\bar{\Delta} = \Delta$, i.e., $\frac{1}{|G_\alpha|}\sum_{\sigma\in G_\alpha} \chi(\sigma)$ is a positive integer for all $\alpha \in \Delta$. So
$$|\sum_{\sigma\in G_\alpha} \chi(\sigma)| \le \sum_{\sigma\in G_\alpha} |\chi(\sigma)| \le |G_\alpha|.$$
   But for all $\sigma \in G$, $|\chi(\sigma)| \le \chi(e) = 1$ by Problem 2.4 #1. So by the triangle inequality, $\chi(\sigma) = 1$ for all $\sigma \in G_\alpha$, i.e., $\chi|_{G_\alpha} \equiv 1$. Pick $\alpha = (1, \ldots, 1) \in \Delta = \bar{\Delta}$ to have the desired result since $G_\alpha = G$.

9. Since the alternating character ε and the principal character 1 of S_m are linear, By Problem 6, $\bar{\Delta} = \hat{\Delta}$ for $\chi = \varepsilon, 1$. So on one hand

$$\dim V_\varepsilon(S_m) = |\bar{\Delta}| = |Q_{m,n}| = \binom{n}{m},$$

$$\dim V_1(S_m) = |\Delta| = |G_{m,n}| = \binom{n+m-1}{m} \quad \text{(by Problem 7)}.$$

On the other hand, by Theorem 4.3.6

$$\dim V_\varepsilon(S_m) = \frac{1}{m!} \sum_{\sigma \in S_m} \varepsilon(\sigma) n^{c(\sigma)}$$

and

$$\dim V_1(S_m) = \frac{1}{m!} \sum_{\sigma \in S_m} n^{c(\sigma)}.$$

10. Suppose $V_\chi(G) \neq 0$ so that $\bar{\Delta}$ is nonempty. Assume that all $\alpha \in \bar{\Delta}$ have distinct entries. Then $m \le n$ evidently. Suppose $G = S_m$ and we need to show that $\chi \neq \varepsilon$. Since $G = S_m$, $\Delta = G_{m,n}$ (Theorem 4.4.2(b)) we have $\bar{\Delta} \subset Q_{m,n}$. By Theorem 4.4.6(a) $\bar{\Delta} = Q_{m,n}$ and then by Theorem 4.4.6(c) $\chi = \varepsilon$.

11. Remark: In the proof of Theorem 4.4.7 we know that $C = (c_{ij}) := [T_\alpha(G,\chi)]_{E_\alpha}^{E_\alpha} \in \mathbb{C}_{k\times k}$ is a projection matrix since $T_\alpha(G,\chi) \in \operatorname{End}(W_\alpha)$ is a projection.

$$\begin{aligned}
(C^2)_{ij} &= \sum_{s=1}^k c_{is}c_{sj} \\
&= \frac{\chi(e)^2}{|G|^2} \sum_{s=1}^k \sum_{\sigma\in G_\alpha} \chi(\tau_i^{-1}\sigma\tau_s) \sum_{\pi\in G_\alpha} \chi(\tau_s^{-1}\pi\tau_j) \\
&= \frac{\chi(e)^2}{|G|^2} \sum_{s=1}^k \sum_{\sigma\in G_\alpha} \chi(\tau_i^{-1}\sigma\tau_s) \sum_{\mu\in G_\alpha} \chi(\tau_s^{-1}\sigma^{-1}\mu\tau_j) \quad (\pi = \sigma^{-1}\mu) \\
&= \frac{\chi(e)^2}{|G|^2} \sum_{s=1}^k \sum_{\sigma,\mu\in G_\alpha} \chi(\tau_i^{-1}\sigma\tau_s)\chi(\tau_s^{-1}\pi\tau_j) \\
&= \frac{\chi(e)^2}{|G|^2} \sum_{\mu\in G_\alpha} \sum_{\gamma\in G} \chi(\tau_i^{-1}\gamma)\chi(\gamma^{-1}\mu\tau_j) \quad (\gamma = \sigma\tau_s) \\
&= \frac{\chi(e)^2}{|G|^2} \sum_{\mu\in G_\alpha} \left(\sum_{\theta\in G} \chi(\theta)\chi(\theta^{-1}\tau_i^{-1}\mu\tau_j) \right) \\
&= \frac{\chi(e)^2}{|G|} \sum_{\mu\in G_\alpha} \frac{\chi(\tau_i^{-1}\mu\tau_j)}{\chi(e)} \quad \text{by Theorem 2.4.3} \\
&= c_{ij}.
\end{aligned}$$

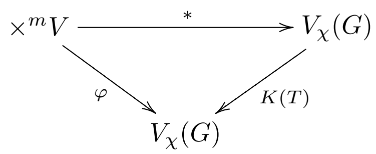
So C is a projection matrix.

4.5 Linear operators on $V_\chi(G)$

Let $T \in \operatorname{End} V$. Similar to the induced operator $\otimes^m T \in \operatorname{End}(\otimes^m V)$, we want to study the induced operator on $V_\chi(G)$ induced by T. The map $\varphi : \times^m V \to V_\chi(G)$ defined by

$$\varphi(v_1, \ldots, v_m) = Tv_1 * \cdots * Tv_m$$

is multilinear and symmetric with respect to G and χ. By Theorem 4.2.5, there is a unique $K(T) \in \operatorname{End}(V_\chi(G))$



satisfying

$$K(T)v_1 * \cdots * v_m = Tv_1 * \cdots * Tv_m.$$

Such $K(T)$ is called the induced operator of T on $V_\chi(G)$. Notice that $K(T)$ depends on G, χ, m, n. We now discuss some basic properties of $K(T)$.

Theorem 4.5.1. Let $T \in \operatorname{End}(V)$. Then $V_\chi(G)$ is an invariant subspace of $\otimes^m T$ and

$$K(T) = \otimes^m T|_{V_\chi(G)}.$$

So $K(T)v^* = (\otimes^m T)v^*$ for all $v^* \in V_\chi(G)$.

Proof. Notice that

$$(\otimes^m T)P(\sigma) = P(\sigma)(\otimes^m T)$$

so that $(\otimes^m T)T(G,\chi) = T(G,\chi)(\otimes^m T)$ (Problem 1). So $V_\chi(G)$ is an invariant subspace of $(\otimes^m T)$. Thus $(\otimes^m T)$ induces a linear operator on $V_\chi(G)$ which is $K(T)$, because

$$\begin{aligned} K(T)v^* &= Tv_1 * \cdots * Tv_m = T(G,\chi)Tv_1 \otimes \cdots \otimes Tv_m \\ &= T(G,\chi)(\otimes^m T)v^\otimes = (\otimes^m T)T(G,\chi)v^\otimes \\ &= (\otimes^m T)v^* \end{aligned}$$

and the decomposable v^* span $V_\chi(G)$, i.e., $K(T) = \otimes^m T|_{V_\chi(G)}$. □

Theorem 4.5.2. Let $S, T \in \operatorname{End}(V)$. Then

(a) $K(I_V) = I_{V_\chi(G)}$.

(b) $K(ST) = K(S)K(T)$. So $T \mapsto K(T)$ is a representation of $\operatorname{GL}(V)$ in $\operatorname{GL}(V_\chi(G))$.

(c) $\operatorname{rank} K(T) = |\hat{\Delta} \cap \Gamma_{m,r}|$, where $r = \operatorname{rank} T$.

Proof. (a) Clear.

(b)

$$\begin{aligned} K(ST) &= \otimes^m (ST)|_{V_\chi(G)} \\ &= (\otimes^m S \otimes^m T)|_{V_\chi(G)} \quad \text{by Theorem 3.4.1} \\ &= \otimes^m S|_{V_\chi(G)} \otimes^m T|_{V_\chi(G)} \quad \text{by Theorem 4.5.1} \\ &= K(S)K(T). \end{aligned}$$

(c) Since $r = \operatorname{rank} T$, there is a basis $\{v_1, \dots, v_n\}$ of V such that $Tv_1, \dots, Tv_r$ are linearly independent and $Tv_{r+1} = \cdots = Tv_n = 0$. Let $e_i = Tv_i$, $i = 1, \dots, r$, and extend them to a basis $E = \{e_1, \dots, e_r, e_{r+1}, \dots, e_n\}$ of V. Now

$$E_* = \{e^*_\alpha : \alpha \in \hat{\Delta}\} \quad \text{and} \quad \{v^*_\alpha : \alpha \in \hat{\Delta}\}$$

are bases of $V_\chi(G)$. If $\alpha \in \Gamma_{m,r}$, then

$$K(T)v^*_\alpha = Tv_{\sigma(1)} * \cdots * Tv_{\alpha(m)} = e^*_\alpha.$$

So

$$\{e^*_\alpha = K(T)v^*_\alpha : \alpha \in \hat{\Delta} \cap \Gamma_{m,r}\}$$

is a subset of the basis E_* of $V_\chi(G)$ so that it is a linearly independent set. When $\alpha \notin \hat{\Delta} \cap \Gamma_{m,r}$, there is some i such that $\alpha(i) > r$. Then $Tv_{\alpha(i)} = 0$ so that $K(T)v^*_\alpha = 0$. So $\operatorname{rank} K(T) = |\hat{\Delta} \cap \Gamma_{m,r}|$. □

Theorem 4.5.3. Let V be an inner product space and $T \in \operatorname{End} V$. Equip $V_\chi(G)$ with the induced inner product.

1. $K(T)^* = K(T^*)$.

2. If T is normal, Hermitian, positive definite, positive semidefinite, unitary, so is $K(T)$.

Proof. (a) From Theorem 4.5.1 $V_\chi(G)$ is invariant under $\otimes^m T$ and $\otimes^m T^*$. So

$$\begin{aligned} K(T^*) &= (\otimes^m T^*)|_{V_\chi(G)} \\ &= (\otimes^m T)^*|_{V_\chi(G)} \quad \text{by Theorem 3.4.3} \\ &= (\otimes^m T|_{V_\chi(G)})^* \quad \text{Theorem 1.7.1(a)} \\ &= K(T)^*. \end{aligned}$$

(b) Since $\otimes^m T$ has the corresponding property by Theorem 3.6.2, by Theorem 1.7.1(b) $K(T) = \otimes^n T|_{V_\chi(G)}$ has the corresponding property. □

Theorem 4.5.4. Let $E = \{e_1, \dots, e_n\}$ be a basis of V and $T \in \operatorname{End} V$. Suppose that $A = [T]^E_E$ is upper triangular with $\operatorname{diag} A = (\lambda_1, \dots, \lambda_n)$. Then $B = [K(T)]^{E_*}_{E_*}$ is upper triangular with diagonal entries $\prod_{t=1}^m \lambda_{\beta(t)}$, $\beta \in \hat{\Delta}$, where $E_* = \{e^*_\alpha : \alpha \in \hat{\Delta}\}$ (in the orbital order).

Proof. Since $A = [T]_E^E$ is upper triangular,

$$Te_j = \lambda_j e_j + \sum_{i<j} a_{ij} e_i, \quad j = 1, \dots, n,$$

where $\lambda_1, \dots, \lambda_n$ are the eigenvalues of T. For each $\beta \in \hat{\Delta}$, there is $\pi \in G$ such that $\beta\pi = \beta_1 \in \bar{\Delta}$. Then

$$\begin{aligned} K(T)e^*_\beta &= Te_{\beta(1)} * \cdots * Te_{\beta(m)} \\ &= (\lambda_{\beta(1)} e_{\beta(1)} + \sum_{i<\beta(1)} a_{i\beta(1)} e_i) * \cdots * (\lambda_{\beta(m)} e_{\beta(m)} + \sum_{i<\beta(m)} a_{i\beta(m)} e_i) \\ &= (\prod_{t=1}^m \lambda_{\beta(t)}) e^*_\beta + \sum_{\alpha \in S_\beta} c_\alpha e^*_\alpha. \end{aligned}$$

where

$$S_\beta := \{\alpha \in \Gamma_{m,n} : \alpha(i) \le \beta(i),\ i = 1, \dots, m, \text{ with at least one strict inequality}\}.$$

We remark that some $\alpha \in S_\beta$ may not be in $\hat{\Delta}$. Fix $\alpha \in S_\beta$ and we may assume that $e^*_\alpha \neq 0$ (note: $\alpha \in \Omega$ from (4.21) but may not be in $\hat{\Delta}$). Clearly $\alpha_1 = \alpha\theta \in \bar{\Delta}$ for some $\theta \in G$. From the definition of S_β,

$$\alpha_1 \le \alpha\pi < \beta\pi = \beta_1.$$

In other words α_1 comes strictly before β_1 in lexicographic order.

Case 1. If $\alpha \in \hat{\Delta}$, then $\alpha < \beta$ in the orbital order, according to the construction of $\hat{\Delta}$.

Case 2. If $\alpha \notin \hat{\Delta}$, then $e^*_\alpha \in O_\alpha = \langle e^*_{\alpha\sigma} : \sigma \in G\}$ is a linear combination of those e^*_γ in the orbital subspace O_α, where $\gamma \in \Gamma_\alpha \cap \hat{\Delta}$. Each $\gamma < \beta$ in the orbital order by Case 1.

So $[K(T)]_{E_*}^{E_*}$ is upper triangular, where $E_* = \{e_\alpha : \alpha \in \hat{\Delta}\}$ in the orbital order. □

In Theorem 4.5.5, the eigenvalues of $K(T)$ are evidently $\prod_{i=1}^m \lambda_{\omega(i)}$, $\omega \in \hat{\Delta}$, counting multiplicity. However the order of these eigenvalues apparently depends on the order of $\lambda_1, \dots, \lambda_n$; but it is merely a deception.

Theorem 4.5.5. Let $\lambda_1, \dots, \lambda_n$ and $s_1, \dots, s_n$ be the eigenvalues and the singular values of $T \in \operatorname{End} V$, respectively. Then

1. The eigenvalues of $K(T)$ are $\prod_{i=1}^m \lambda_{\omega(i)}$, $\omega \in \hat{\Delta}$, counting multiplicities. Moreover for any $\tau \in S_n$,

$$\prod_{i=1}^m \lambda_{\tau\omega(i)}, \quad \omega \in \hat{\Delta}$$

are the eigenvalues of $K(T)$, counting multiplicity. Hence $f(\lambda_1,\dots,\lambda_n) :=$ $\operatorname{tr} K(T)$ is a symmetric function of $\lambda_1,\dots,\lambda_n$, i.e.,

$$f(\lambda_1,\dots,\lambda_n) = f(\lambda_{\tau(1)},\dots,\lambda_{\tau(n)}), \quad \tau \in S_n.$$

2. The singular values of $K(T)$ are $\prod_{i=1}^m s_{\omega(i)}$, $\omega \in \hat{\Delta}$, counting multiplicities. Moreover for any $\tau \in S_n$,

$$\prod_{i=1}^m s_{\tau\omega(i)}, \quad \omega \in \hat{\Delta}$$

are the singular values of $K(T)$, counting multiplicities.

Proof. (a) The first part follows from Theorem 4.5.4. Schur's triangular theorem or Jordan form allows any prescribed ordering of $\lambda_1,\dots,\lambda_n$. In other words, for any $\tau \in S_n$,

$$\prod_{i=1}^m \lambda_{\tau\omega(i)}, \quad \omega \in \hat{\Delta}$$

are the eigenvalues of $K(T)$, counting multiplicities.

(b) Apply (a) on A^*A. □

In order to compute $\det K(T)$ we rewrite $\prod_{i=1}^m \lambda_{\omega(i)}$ as $\prod_{t=1}^n \lambda_t^{m_t(\omega)}$, where $m_t(\omega)$ denotes the multiplicity of the integer t, i.e., the number of times t appearing in ω. Clearly for all $\omega \in \Gamma_{m,n}$,

$$\prod_{i=1}^m \lambda_{\omega(i)} = \prod_{t=1}^n \lambda_t^{m_t(\omega)}.$$

Theorem 4.5.6. Let $T \in \operatorname{End} V$. Then

$$\det K(T) = (\det T)^{\frac{m}{n}|\hat{\Delta}|} \tag{4.23}$$

Proof. Since the determinant is the product of eigenvalues, by Theorem 4.5.5

$$\det K(T) = \prod_{\omega\in\hat{\Delta}} \prod_{t=1}^n \lambda_t^{m_t(\omega)} = \prod_{t=1}^n \prod_{\omega\in\hat{\Delta}} \lambda_t^{m_t(\omega)} = \prod_{t=1}^n \lambda_t^{q_t},$$

where $q_t := \sum_{\omega\in\hat{\Delta}} m_t(\omega)$ and $m_t(\omega)$ is the multiplicity of t in ω. In other words, q_t is the total number of times the integer t appearing among all the sequences in $\hat{\Delta}$. By Theorem 4.5.5(a)

$$\prod_{t=1}^n \lambda_t^{q_t} = \prod_{t=1}^n \lambda_{\tau(t)}^{q_t} = \prod_{t=1}^n \lambda_t^{q_{\tau^{-1}(t)}}$$

which is true for all scalars $\lambda_1,\dots,\lambda_n$ (view them as indeterminants). So $q_t = q$, a constant, for all $t = 1,\dots,n$. Since $\sum_{t=1}^n m_t(\omega) = m$,

$$nq = \sum_{t=1}^n q_t = \sum_{t=1}^n \sum_{\omega\in\hat\Delta} m_t(\omega) = \sum_{\omega\in\hat\Delta}\sum_{t=1}^n m_t(\omega) = m|\hat\Delta|.$$

So

$$q = \frac{m}{n}|\hat\Delta|.$$

Then

$$\det K(T) = \prod_{t=1}^n \lambda_t^{\frac{m}{n}|\hat\Delta|} = \left(\prod_{t=1}^n \lambda_t\right)^{\frac{m}{n}|\hat\Delta|} = (\det T)^{\frac{m}{n}|\hat\Delta|}.$$

□

The following is an alternate approach for $q_t = q$ for all t:

From Problem 4.4 #2 ($G_{\alpha\sigma} = \sigma^{-1}G_\alpha\sigma$) and #5 for any $\alpha \in \Gamma_{m,n}$, $\sigma \in G$ we have

$$|G_{\alpha\sigma}| = |G_\alpha|, \quad s_{\alpha\sigma} = s_\sigma, \quad m_t(\alpha\sigma) = m_t(\alpha).$$

For any arbitrary $\tau \in S_n$,

$$G_{\tau\alpha} = G_\alpha, \quad s_{\tau\alpha} = s_\alpha, \quad m_t(\tau\alpha) = m_{\tau^{-1}(t)}(\alpha).$$

Moreover, as α runs over $\Gamma_{m,n}$, $\tau\alpha$ runs over $\Gamma_{m,n}$ as well. Then

$$\begin{aligned}
q_t &= \sum_{\omega\in\hat\Delta} m_t(\omega) = \sum_{\alpha\in\bar\Delta} s_\alpha m_t(\alpha) = \sum_{\alpha\in\Delta} s_\alpha m_t(\alpha) \quad (s_\alpha = 0 \text{ if } \alpha\notin\bar\Delta)\\
&= \frac{1}{|G|}\sum_{\alpha\in\Delta}\frac{1}{|G_\alpha|}\sum_{\sigma\in G}|G_\alpha| s_\alpha m_t(\alpha)\\
&= \frac{1}{|G|}\sum_{\alpha\in\Delta}\frac{1}{|G_\alpha|}\sum_{\sigma\in G}|G_{\alpha\sigma}| s_{\alpha\sigma} m_t(\alpha\sigma) \quad (|G_{\alpha\sigma}| = |G_\alpha|, s_{\alpha\sigma} = s_\sigma, m_t(\alpha\sigma) = m_t(\alpha))\\
&= \frac{1}{|G|}\sum_{\gamma\in\Gamma_{m,n}} |G_\gamma| s_\gamma m_t(\gamma) \quad (\text{Theorem 4.4.3}, \varphi(\gamma) = |G_\gamma| s_\gamma m_t(\gamma))
\end{aligned}$$

Then for any $\tau \in S_n$

$$\begin{aligned}
q_t &= \frac{1}{|G|}\sum_{\alpha\in\Gamma_{m,n}} |G_{\tau\alpha}| s_{\tau\alpha} m_t(\tau\alpha) \quad (\tau\Gamma_{m,n} = \Gamma_{m,n})\\
&= \frac{1}{|G|}\sum_{\alpha\in\Gamma_{m,n}} |G_\alpha| s_\alpha m_{\tau^{-1}(t)}(\alpha) \quad (G_{\tau\alpha} = G_\alpha, s_{\tau\alpha} = s_\alpha, m_t(\tau\alpha) = m_{\tau^{-1}(t)}(\alpha))\\
&= q_{\tau^{-1}(t)}.
\end{aligned}$$

Theorem 4.5.7. Let $E = \{e_1, \dots, e_n\}$ be an orthonormal basis of the inner product space V and equip $V_\chi(G)$ with the induced inner product. Let $T \in \operatorname{End} V$ and $A = (a_{ij}) = [T]_E^E$. Then for $\alpha, \beta \in \Gamma_{m,n}$,

$$(K(T)e_\beta^*, e_\alpha^*) = \frac{\chi(e)}{|G|}\sum_{\sigma\in G}\chi(\sigma)\prod_{t=1}^m a_{\alpha\sigma(t)\beta(t)}. \tag{4.24}$$

Proof. Notice that $K(T)e^*_\beta = Te_{\beta(1)} * \cdots * Te_{\beta(1)}$. Since E is an orthonormal basis,

$$(Te_{\beta(t)}, e_{\alpha\sigma(t)}) = a_{\alpha(\sigma(t))\beta(t)}.$$

Then use Theorem 4.3.4 with $u_t = Te_{\beta(t)}$ and $v_t = e_{\alpha(t)}$:

$$(K(T)e^*_\beta, e^*_\alpha) = \frac{\chi(e)}{|G|}\sum_{\sigma\in G}\chi(\sigma)\prod_{t=1}^m (Te_{\beta(t)}, e_{\alpha\sigma(t)}) = \frac{\chi(e)}{|G|}\sum_{\sigma\in G}\chi(\sigma)\prod_{t=1}^m a_{\alpha\sigma(t)\beta(t)}.$$

□

We now discuss the matrix representation of $K(T)$. Let $E = \{e_1, \ldots, e_n\}$ be an orthonormal basis of the inner product space V and equip $V_\chi(G)$ with the induced inner product. From the construction of the basis $E_* = \{e^*_\alpha : \alpha \in \hat{\Delta}\}$, different orbit subspaces are orthogonal (Theorem 4.3.5). However the vectors within the same orbit subspace may not be orthogonal. So the right side of (4.24) is not necessarily the corresponding element of $K(T)$ with respect to the basis E_* of $V_\chi(G)$.

When $\chi(e) = 1$, ($\bar{\Delta} = \hat{\Delta}$ by Problem 4.4 #6) $E_* = \{e^*_\alpha : \alpha \in \bar{\Delta}\}$ is an orthogonal basis of $V_\chi(G)$ (but **not** necessarily orthonormal). Upon normalization,

$$E'_* = \left\{\sqrt{\frac{|G|}{|G_\alpha|}}e^*_\alpha : \alpha \in \bar{\Delta}\right\}$$

is an orthonormal basis of $V_\chi(G)$ (Problem 6). Thus if $\chi(e) = 1$, then the matrix representation of $K(T)$ with respect to E'_* follows from the following theorem.

Theorem 4.5.8. Suppose $\chi(e) = 1$ and $E = \{e_1, \ldots, e_n\}$ is a basis of V (so $E'_* = \{\sqrt{\frac{|G|}{G_\alpha|}}e^*_\alpha : \alpha \in \bar{\Delta}\}$ is a basis of $V_\chi(G)$ according to the lexicographic ordering). If $T \in \operatorname{End} V$ and $A = [T]^E_E$, then

$$[K(T)]^{E'_*}_{E'_*} = \left(\frac{1}{\sqrt{|G_\alpha||G_\beta|}}\sum_{\sigma\in G}\chi(\sigma)\prod_{t=1}^m a_{\alpha\sigma(t)\beta(t)}\right)_{\alpha,\beta\in\bar{\Delta}} \tag{4.25}$$

Proof. Equip V with an inner product so that E is an orthonormal basis (Theorem 1.6.3). Then E'_* is an orthonormal basis of $V_\chi(G)$ with respect to the induced inner product. Since $\chi(e) = 1$, use (4.24) to have

$$(K(T)\sqrt{\frac{|G|}{|G_\beta|}}e^*_\beta, \sqrt{\frac{|G|}{|G_\alpha|}}e^*_\alpha) = \frac{1}{\sqrt{|G_\alpha||G_\beta|}}\sum_{\sigma\in G}\chi(\sigma)\prod_{t=1}^m a_{\alpha\sigma(t)\beta(t)}$$

□

When χ is **linear**, similar to the Kronecker product of matrices, we define the induced matrix $K(A)$ using (4.25). Let $\chi(e) = 1$ and let $A \in \mathbb{C}_{n\times n}$. The matrix $K(A) \in \mathbb{C}_{|\bar{\Delta}|\times|\bar{\Delta}|}$ defined by

$$K(A)_{\alpha,\beta} = \frac{1}{\sqrt{|G_\alpha||G_\beta|}} \sum_{\sigma\in G} \chi(\sigma) \prod_{t=1}^{m} a_{\alpha\sigma(t),\beta(t)}, \quad \alpha,\beta \in \bar{\Delta} \tag{4.26}$$

is called the induced matrix. If T and A are as in Theorem 4.5.8, then

$$[K(T)]_{E'_*}^{E'_*} = K(A).$$

Since the induced matrix $K(A)$ is the matrix representation of $K(T)$, we have the following properties.

Theorem 4.5.9. Let $A \in \mathbb{C}_{n\times n}$ with eigenvalues $\lambda_1, \ldots, \lambda_n$ and singular values $s_1 \geq \cdots \geq s_n \geq 0$. Suppose χ is linear and let $K(A)$ be the induced matrix of A. Then

(a) $K(I) = I_{|\bar{\Delta}|}$, where $|\bar{\Delta}| = \frac{1}{|G|}\sum_{\sigma\in G}\chi(\sigma)n^{c(\sigma)}$.

(b) $K(AB) = K(A)K(B)$, where $A, B \in \mathbb{C}_{n\times n}$.

(c) $K(A^{-1}) = K(A)^{-1}$ if $A \in \mathbb{C}_{n\times n}$ is invertible.

(d) $\operatorname{rank} K(A) = |\hat{\Delta} \cap \Gamma_{m,r}|$, where $r = \operatorname{rank} A$.

(e) If A is upper triangular, then $K(A)$ is also upper triangular.

(f) The eigenvalues of $K(A)$ are $\prod_{i=1}^m \lambda_{\omega(i)}$, $\omega \in \bar{\Delta}$. The singular values of $K(A)$ are $\prod_{i=1}^m s_{\omega(i)}$, $\omega \in \bar{\Delta}$.

(g) $\operatorname{tr} K(A) = \sum_{\alpha\in\bar{\Delta}} \prod_{t=1}^m \lambda_{\alpha(t)} = \sum_{\alpha\in\bar{\Delta}} \frac{1}{|G_\alpha|}\sum_{\sigma\in G}\chi(\sigma)\prod_{t=1}^m a_{\alpha\sigma(t),\alpha(t)}$.

(h) $\det K(A) = (\det A)^{\frac{m}{n}|\hat{\Delta}|}$.

(i) $K(A^*) = K(A)^*$.

(j) If A is normal, Hermitian, psd, pd, or unitary, so is $K(A)$.

For $\chi(e) > 1$, see [18] for the construction of orthonormal basis of $V_\chi(G)$, the matrix representation of $K(T)$, where $T \in \operatorname{End} V$, and the induced matrix $K(A)$.

We now provide an example of $K(T)$ with nonlinear irreducible character.

Assume that $G = S_3$ and χ is the degree 2 irreducible character. We have

$$\chi(e) = 2,\ \chi(12) = \chi(23) = \chi(13) = 0,\ \chi(123) = \chi(132) = -1.$$

Assume $\dim V = n = 2$. Then by direct computation,

$$\begin{aligned}
\Gamma_{3,2} &= \{(1,1,1),(1,1,2),(1,2,1),(1,2,2),(2,1,1),(2,1,2),(2,2,1),(2,2,2)\}\\
\Delta &= \{(1,1,1),(1,1,2),(1,2,2),(2,2,2)\}\\
\bar{\Delta} &= \{(1,1,2),(1,2,2)\}\\
\hat{\Delta} &= \{(1,1,2),(1,2,1),(1,2,2),(2,1,2)\}.
\end{aligned}$$

Let $E = \{e_1, e_2\}$ be a basis of V. Then

$$E_* = \{e^*_\alpha : \alpha \in \hat{\Delta}\} = \{e^*_{(1,1,2)}, e^*_{(1,2,1)}, e^*_{(1,2,2)}, e^*_{(2,1,2)}\}$$

is a basis of $V_\chi(G)$. Observe that E_* is not an orthogonal basis even if E is an orthonormal basis, since

$$(e^*_{(1,1,2)}, e^*_{(1,2,1)}) = (e^*_{(1,2,2)}, e^*_{(2,1,2)}) = -\frac{1}{3}.$$

For example, by Theorem 4.3.5

$$(e^*_{(1,1,2)}, e^*_{(1,2,1)}) = \frac{\chi(e)}{|S_3|} \sum_{\sigma \in S_3} \chi(\sigma)\delta_{(1,1,2),(1,2,1)\sigma} = \frac{2}{6}(\chi(23) + \chi(123)) = -\frac{1}{3}.$$

Moreover (check by working out the C matrices!)

$$e^*_{(2,1,1)} = -e^*_{(1,1,2)} - e^*_{(1,2,1)}, \quad e^*_{(2,2,1)} = -e^*_{(1,2,2)} - e^*_{(2,1,2)}. \tag{4.27}$$

Let $T \in \operatorname{End} V$ be defined by

$$[T]^E_E = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

By direct computation,

$$[K(T)]^{E_*}_{E_*} = \begin{pmatrix} a^2d - abc & 0 & abd - b^2c & 0 \\ 0 & a^2d - abc & abd - b^2c & b^2c - abd \\ acd - bc^2 & 0 & ad^2 - bcd & 0 \\ acd - bc^2 & bc^2 - acd & 0 & ad^2 - bcd \end{pmatrix}$$

For example

$$\begin{aligned}
K(T)e^*_{(1,1,2)} &= Te_1 * Te_1 * Te_2 \\
&= (ae_1 + ce_2) * (ae_1 + ce_2) * (be_1 + de_2) \\
&= a^2be^*_{(1,1,1)} + a^2de^*_{(1,1,2)} + acbe^*_{(1,2,1)} + acde^*_{(1,2,2)} \\
&\quad + cabe^*_{(2,1,1)} + cade^*_{(2,1,2)} + c^2be^*_{(2,2,1)} + c^2be^*_{(2,2,2)} \\
&= a^2de^*_{(1,1,2)} + acbe^*_{(1,2,1)} + acde^*_{(1,2,2)} \\
&\quad + cabe^*_{(2,1,1)} + cade^*_{(2,1,2)} + c^2be^*_{(2,2,1)} \qquad \text{by (4.26)} \\
&= (a^2d - abc)e^*_{(1,1,2)} + 0e^*_{(1,2,1)} + (acd - bc^2)e^*_{(1,2,2)} + (acd - bc^2)e^*_{(2,1,2)}.
\end{aligned}$$

The computations for $K(T)e^*_{(1,2,1)}, K(T)e^*_{(1,2,2)}, K(T)e^*_{(2,1,2)}$ are similar.

Furthermore, one can define the derivation operator $D_K(T)$ of T by

$$D_K(T) = \frac{d}{dt} K(I + tT)\Big|_{t=0},$$

which acts on $V_\chi(G)$ in the following way:

$$D_K(T)v_1 * \cdots * v_m = \sum_{j=1}^m v_1 * \cdots * v_{j-1} * Tv_j * v_{j+1} * \cdots * v_m.$$

Clearly $T \mapsto D_K(T)$ is linear. With respect to the above example

$$[D_K(T)]_{E_*}^{E_*} = \begin{pmatrix} 2a+d & 0 & b & 0 \\ 0 & 2a+d & b & -b \\ c & 0 & a+2d & 0 \\ c & -c & 0 & a+2d \end{pmatrix}.$$

Research problem: Given the Jordan structure of $T \in \mathbb{C}_{n\times n}$ ($T \in \operatorname{End} V$), what is the Jordan structure of $C_m(T)$ ($G = S_m$ and $\chi = \varepsilon$) where $1 \le m \le n$? The answer is given by Atkin [1]; also see Littlewood [13] for some simplified treatment. Atkin also studied the problem for Kronecker product. In general the corresponding problem for $K(T)$ is not known.

Problems

1. Prove that $(\otimes^m T)P(\sigma) = P(\sigma)(\otimes^m T)$ and $(\otimes^m T)T(G,\chi) = T(G,\chi)(\otimes^m T)$.

2. If $V_\chi(G) \neq 0$, prove that $K(T)$ is invertible if and only if T is invertible. Moreover, $K(T^{-1}) = K(T)^{-1}$.

3. Let $S, T \in \operatorname{End} V$ are psd. Show that $K(S+T) \ge K(S) + K(T)$, i.e., $K(S+T) - K(S) + K(T)$ is psd. Hence $S \ge T$ implies $K(S) \ge K(T)$. (Hint: Use Problem 3.6 #2)

4. Prove that if $G = S_m$, $\chi = \varepsilon$, $\operatorname{rank} T < m$, then $K(T) = 0$.

5. Suppose $T \in \operatorname{End} V$ is positive semidefinte with eigenvalues $\lambda_1 \ge \cdots \ge \lambda_n$. Prove that $K(T)$ on $V_1(G)$ has largest eigenvalue λ_1^m and smallest eigenvalues λ_n^m.

6. Suppose that $\chi(e) = 1$ and $E = \{e_1, \ldots, e_n\}$ is an orthonormal basis of the inner product V. Prove that $E'_* = \{\sqrt{\frac{|G|}{|G_\alpha|}} e^*_\alpha : \alpha \in \bar{\Delta}\}$ is an orthonormal basis for $V_\chi(G)$.

7. Suppose that $\chi(e) = 1$ and $E = \{e_1, \ldots, e_n\}$ is a basis of V. Let $T \in \operatorname{End} V$ and $[T]_E^E = A = (a_{ij})$. Prove that

$$([K(T)]_{E_*}^{E_*})_{\alpha,\beta} = \frac{1}{|G_\alpha|} \sum_{\sigma \in G} \chi(\sigma) \prod_{i=1}^m a_{\alpha\sigma(t)\beta(t)}, \quad \alpha, \beta \in \bar{\Delta}.$$

8. Let $\alpha \in \Gamma_{m,n}$ and χ be linear. Prove that if $\sum_{\sigma \in G_\alpha} \chi(\sigma) = 0$, i.e., $\alpha \notin \Omega$, then $\sum_{\sigma \in G} \chi(\sigma)\varphi(\alpha\sigma) = 0$, where $\varphi : \Gamma_{m,n} \to W$ is any map.

9. Suppose that χ is linear. If $A, B \in \mathbb{C}_{n\times n}$, deduce directly from (4.26) that $K(AB) = K(A)K(B)$ (Hint: Compare Theorem 1.10.1).

10. Show that $\operatorname{tr} K(A) = \chi(e)\sum_{\alpha\in\Delta}(\chi|_{G_\alpha}, 1)\prod_{t=1}^m \lambda_{\alpha(t)}$.

Solution to Problems 4.5

1. $(\otimes^m T)P(\sigma)v^\otimes = (\otimes^m T)v^\otimes_{\sigma^{-1}} = Tv_{\sigma^{-1}(1)}\otimes\cdots\otimes Tv_{\sigma^{-1}(m)} = P(\sigma)(\otimes^m T)v^\otimes$. Then $(\otimes^m T)T(G,\chi) = T(G,\chi)(\otimes^m T)$ follows since $T(G,\chi) = \frac{\chi(e)}{|G|}\sum_{\sigma\in G}\chi(\sigma)P(\sigma)$ is a linear combination of $P(\sigma)$, $\sigma \in G$.

2. Suppose $V_\chi(G) \neq 0$, i.e., $|\hat\Delta| \neq 0$. Then use Theorem 4.5.6 since a matrix is invertible if and only if its determinant is nonzero.

3. By Problem 3.6 #4, $\otimes^m(S+T) \geq \otimes^m S + \otimes^m T$ So

$$\begin{aligned} K(S+T) &= \otimes^m(S+T)|_{V_\chi(G)} \geq (\otimes^m S + \otimes^m T)|_{V_\chi(G)} \\ &= \otimes^m S|_{V_\chi(G)} + \otimes^m T|_{V_\chi(G)} = K(S) + K(T). \end{aligned}$$

So if $S \geq T$, then $S = T + R$ where $R \geq 0$. Then $K(R) \geq 0$ by Theorem 4.5.3 and hence

$$K(S) = K(T+R) \geq K(T) + K(R) \geq K(T).$$

4. Let $r = \operatorname{rank} T$.

$$\begin{aligned} \operatorname{rank} K(T) &= |\Gamma_{m,r}\cap\hat\Delta| \quad \text{(Theorem 4.5.2)} \\ &= |\Gamma_{m,r}\cap Q_{m,n}| \quad \text{(Theorem 4.4.6 since } G = S_m, \chi = \varepsilon) \\ &= 0 \end{aligned}$$

since all elements of $\Gamma_{m,r}$ has some repeated entries as $r < m$ so that $\Gamma_{m,r}\cap Q_{m,n} = \emptyset$.

5. By Theorem 4.5.5, since $\hat\Delta = \Gamma_{m,n}$, the eigenvalues are $\prod_{i=1}^m \lambda_{\omega(i)}$, $\omega \in \Gamma_{m,n}$. Since $\lambda_1 \geq \cdots \geq \lambda_n \geq 0$, the largest eigenvalue of $K(T)$ is λ_1^m and the smallest eigenvalue is λ_n^m.

6. By Theorem 4.3.5(a), E'_* is an orthogonal basis. It remains to show that each vector in E'_* has unit length. By Theorem 4.3.5(b)

$$\left\|\sqrt{\frac{|G|}{|G_\alpha|}}e^*_\alpha\right\|^2 = \frac{|G|}{|G_\alpha|}\|e^*_\alpha\|^2 = \frac{\chi(e)}{|G_\alpha|}\sum_{\sigma\in G_\alpha}\chi(\sigma) = \frac{1}{|G_\alpha|}\sum_{\sigma\in G_\alpha}\chi(\sigma) = (\chi|_{G_\alpha}, 1)$$

which is a positive integer since $\alpha \in \Delta$. Notice that

$$\frac{1}{|G_\alpha|}\sum_{\sigma\in G_\alpha}\chi(\sigma) \leq \frac{1}{|G_\alpha|}\sum_{\sigma\in G_\alpha}|\chi(\sigma)| \leq 1$$

since each $|\chi(\sigma)| \leq 1$ (χ is linear). So $(\chi|_{G_\alpha}, 1) = 1$.

7.

8.

9.

10. Notice that $\prod_{t=1}^m \lambda_{\alpha(t)} = \prod_{t=1}^m \lambda_{\alpha\sigma(t)}$ for all $\sigma \in G$, $s_\alpha = \chi(e)(\chi|_{G_\alpha}, 1)$ and $G_{\alpha\theta} = \theta^{-1}G_\alpha\theta$, for all $\alpha \in \Gamma_{m,n}$ and $\theta \in G$. So

$$\begin{aligned} \operatorname{tr} K(T) &= \sum_{\alpha\in\hat{\Delta}} \prod_{t=1}^m \lambda_{\alpha(t)} = \chi(e) \sum_{\alpha\in\bar{\Delta}} (\chi|_{G_\alpha}, 1) \prod_{t=1}^m \lambda_{\alpha(t)} \\ &= \chi(e) \sum_{\alpha\in\Delta} (\chi|_{G_\alpha}, 1) \prod_{t=1}^m \lambda_{\alpha(t)}. \end{aligned}$$

4.6 Generalized matrix functions

Let $G < S_m$, $\chi \in I(G)$ and let $A \in \mathbb{C}_{m\times m}$. The function

$$d_G^\chi(A) = \sum_{\sigma\in G} \chi(\sigma) \prod_{t=1}^m a_{t\sigma(t)} \tag{4.28}$$

is called the **generalized matrix function** association with G and χ.

Example:

(a) When $G = S_m$ and $\chi = \varepsilon$,

$$d_G^\chi(A) = \det A := \sum_{\sigma\in S_m} \varepsilon(\sigma) \prod_{t=1}^m a_{t\sigma(t)}.$$

(b) When $G = S_m$ and $\chi \equiv 1$,

$$d_G^\chi(A) = \operatorname{per} A := \sum_{\sigma\in S_m} \prod_{t=1}^m a_{t\sigma(t)}$$

is called the **permanent** of A.

(c) When $G = \{e\} < S_m$, then

$$d_G^\chi(A) = h(A) := \prod_{t=1}^m a_{tt}$$

which is the product of the diagonal entries.

Theorem 4.6.1. Let $A \in \mathbb{C}_{m\times m}$. Then

(a) $d_G^\chi(I_m) = \chi(e)$.

(b) If A is upper triangular, then

$$d_G^\chi(A) = \chi(e)\det A = \chi(e)\operatorname{per} A = \chi(e)h(A).$$

(c) $d_G^\chi(A^T) = d_G^{\bar\chi}(A)$ ($\bar\chi(\sigma) := \overline{\chi(\sigma)}$ for all $\sigma \in G$).

(d) $d_G^\chi(A^*) = \overline{d_G^\chi(A)}$. So if A is Hermitian, then $d_G^\chi(A)$ is real.

Proof. (a) and (b) are trivial.

(c)

$$d_G^\chi(A^T) = \sum_{\sigma\in G}\chi(\sigma)\prod_{t=1}^m a_{\sigma(t)t} = \sum_{\sigma\in G}\overline{\chi(\sigma)}\prod_{t=1}^m a_{t\sigma(t)} = d_G^{\bar\chi}(A).$$

(d) $d_G^\chi(A^*) = d_G^\chi((\bar A)^T) = d_G^{\bar\chi}(\bar A) = \overline{d_G^\chi(A)}$. □

Theorem 4.3.4 and Theorem 4.5.7 can be restated respectively using generalized matrix function as in the next two theorems.

Theorem 4.6.2. Let $v_1, \dots, v_m, u_1, \dots, u_m \in V$ and let $A \in \mathbb{C}_{m\times m}$ where $a_{ij} := (u_i, v_j)$. Then for each $\alpha, \beta \in \Gamma_{m,n}$,

$$\begin{aligned}(u^*, v^*) &= \frac{\chi(e)}{|G|}\sum_{\sigma\in G}\chi(\sigma)\prod_{t=1}^m (u_t, v_{\sigma(t)}) \quad \text{(Theorem 4.3.4)}\\ &= \frac{\chi(e)}{|G|}d_G^\chi(A) = \frac{\chi(e)}{|G|}d_G^\chi((u_i, v_j))\end{aligned}$$

Theorem 4.6.3. Let $E = \{e_1, \dots, e_n\}$ be an orthonormal basis of V and let $T \in \operatorname{End} V$. Let $A = (a_{ij}) = [T]_E^E$. Then for each $\alpha, \beta \in \Gamma_{m,n}$,

$$\begin{aligned}(K(T)e_\beta^*, e_\alpha^*) &= \frac{\chi(e)}{|G|}\sum_{\sigma\in G}\overline{\chi(\sigma)}\prod_{t=1}^m a_{\sigma(t)\beta\sigma(t)} \quad \text{(Theorem 4.5.7)}\\ &= \frac{\chi(e)}{|G|}d_G^{\bar\chi}(A[\alpha|\beta])\\ &= \frac{\chi(e)}{|G|}d_G^\chi(A^T[\beta|\alpha]) \quad \text{(Theorem 4.6.1(c))}\end{aligned}$$

where $A[\alpha|\beta] = (a_{\alpha(i)\beta(j)}) \in \mathbb{C}_{m\times m}$. In particular when $m = n$,

$$(K(T)e_1 * \cdots * e_n, e_1 * \cdots * e_n) = \frac{\chi(e)}{|G|}d_G^{\bar\chi}(A) = \frac{\chi(e)}{|G|}d_G^\chi(A^T).$$

We will use the above two theorems to establish some inequalities and identities. Notice that χ is irreducible if and only if $\bar\chi$ is irreducible. Clearly $\chi(e) = \bar\chi(e)$. So statements about $d_G^\chi(A)$ apply to $d_G^{\bar\chi}(A)$.

Theorem 4.6.4. Let $A, B \in \mathbb{C}_{m\times n}$. Then

$$|d_G^\chi(AB^*)|^2 \le d_G^\chi(AA^*)\, d_G^\chi(BB^*). \tag{4.29}$$

Proof. Let $u_i = A_{(i)}$ and $v_j = B_{(j)}$, where $A_{(i)}$ denotes the ith row of A. Then

$$\begin{aligned}(u^*, v^*) &= \frac{\chi(e)}{|G|} d_G^\chi((u_i, v_j)) \quad \text{(Theorem 4.6.2)}\\ &= \frac{\chi(e)}{|G|} d_G^\chi(AB^*) \quad ((u,v) := v^*u)\end{aligned}$$

Then apply Cauchy-Schwarz inequality

$$|(u^*, v^*)| \le (u^*, u^*)(v^*, v^*)$$

to have (4.29). □

Theorem 4.6.5. Let $A, B \in \mathbb{C}_{m\times m}$. Then

$$|d_G^\chi(A)|^2 \le \chi(e)\, d_G^\chi(AA^*). \tag{4.30}$$

Proof. In (4.29) pick $B = I_m$ and apply Theorem 4.6.1 to have (4.30). □

Theorem 4.6.6. (Schur) If $A \in \mathbb{C}_{m\times m}$ is psd, then

$$\chi(e) \det A \le d_G^\chi(A). \tag{4.31}$$

Proof. Since A is psd, there is an upper triangular matrix L such that $A = LL^*$ (Problem 1.5 #9). By Theorem 4.6.1(b),

$$|d_G^\chi(L)|^2 = \chi(e)^2 |\det L|^2 = \chi(e)^2 \det A.$$

Applying (4.30),

$$|d_G^\chi(L)|^2 \le \chi(e) d_G^\chi(LL^*) = \chi(e) d_G^\chi(A).$$

□

From Theorem 4.6.6, if $A \ge 0$, then

$$d_G^\chi(A) \ge 0, \quad \det A \le \operatorname{per} A, \quad \det A \le h(A).$$

Theorem 4.6.7. (Fischer) If $A \in \mathbb{C}_{m\times m}$ is psd, then

$$\det A \le \det A[1, \dots p|1, \dots, p] \det A[p+1, \dots, m|p+1, \dots, m]. \tag{4.32}$$

Proof. This is indeed a special case of Theorem 4.6.4. Let $G < S_m$ be the group consisting of $\sigma\pi$, where σ is a permutation on $\{1, \dots, p\}$, and π is a permutation on $\{p+1, \dots, m\}$. Let $\chi(\sigma\pi) = \varepsilon(\sigma)\varepsilon(\pi)$. Then χ is a linear irreducible character of G (check!) and $d_G^\chi(A)$ is the right side of (4.32).

□

See Horn and Johnson's Matrix Analysis p.478 for a matrix proof of Theorem 4.6.7.

Theorem 4.6.8. Let $A \in \mathbb{C}_{m\times m}$. Then

$$h(A) = \frac{1}{|G|}\sum_{\chi\in I(G)} \chi(e) d_G^\chi(A).$$

In particular, if $A \geq 0$, then

$$h(A) \geq \frac{\chi(e)}{|G|} d_G^\chi(A), \quad h(A) \geq \frac{1}{m!}\operatorname{per} A$$

Proof. Since $A \in \mathbb{C}_{m\times m}$, there are $u_1, \dots, u_m; v_1, \dots, v_m \in V$ where $\dim V = m$ such that $a_{ij} = (u_i, v_j)$ (Problem 1.6 #3). Then

$$\begin{aligned}
\frac{1}{|G|}\sum_{\chi\in I(G)} \chi(e) d_G^\chi(A) &= \sum_{\chi\in I(G)} (u^*, v^*) \quad \text{(by Theorem 4.6.2)} \\
&= \sum_{\chi\in I(G)} (T(G,\chi)u^\otimes, T(G,\chi)v^\otimes) \\
&= (\sum_{\chi\in I(G)} T(G,\chi)u^\otimes, v^\otimes) \quad \text{(by Theorem 4.2.3)} \\
&= \prod_{i=1}^m (u_i, v_i) \quad \text{(by Theorem 4.2.3(b))} \\
&= h(A).
\end{aligned}$$

When $A \geq 0$, by Theorem 4.6.6 each summand $d_G^\chi(A)$ in $\frac{1}{|G|}\sum_{\chi\in I(G)} \chi(e) d_G^\chi(A)$ is nonnegative so we have

$$h(A) \geq \frac{\chi(e)}{|G|} d_G^\chi(A)$$

for all $\chi \in I(G)$. In particular with $G = S_m$ and $\chi \equiv 1$, we have $h(A) \geq \frac{1}{m!}\operatorname{per} A$. □

We can express $u^* = v^*$ in terms of the generalized matrix function when $u_1, \dots, u_m$ are linearly independent (so $m \leq \dim V$).

Theorem 4.6.9. Let $u_1, \dots, u_m \in V$ be linearly independent. Then $u^* = v^*$ if and only if $v_i = \sum_{j=1}^m a_{ij}u_j$, $j = 1, \dots, m$ and $d_G^\chi(AA^*) = d_G^\chi(A) = \chi(e)$.

Proof. Since $u_1, \dots, u_m$ are linearly independent, by Theorem 4.3.2, $u^* \neq 0$.

($\Rightarrow$) Suppose $u^* = v^*(\neq 0)$. From Theorem 4.3.3,

$$\langle u_1, \dots, u_m\rangle = \langle v_1, \dots, v_m\rangle.$$

So there is $A \in \mathbb{C}_{m\times m}$ such that $v_i = \sum_{j=1}^m a_{ij}u_j$, $i = 1, \ldots, m$. Equip an inner product on V so that $\{u_1, \ldots, u_m\}$ is an orthonormal set. Then

$$(u_i, u_j) = \delta_{ij}, \quad (v_i, u_j) = a_{ij}, \quad (v_i, v_j) = (AA^*)_{ij}.$$

Apply Theorem 4.6.2 we have

$$\begin{aligned}
(v^*, u^*) &= \frac{\chi(e)}{|G|} d_G^\chi((v_i, u_j)) = \frac{\chi(e)}{|G|} d_G^\chi(A) \\
(u^*, u^*) &= \frac{\chi(e)}{|G|} d_G^\chi(I_m) = \frac{\chi^2(e)}{|G|} \\
(v^*, v^*) &= \frac{\chi(e)}{|G|} d_G^\chi((v_i, v_j)) = \frac{\chi(e)}{|G|} d_G^\chi(AA^*)
\end{aligned}$$

Then from $u^* = v^*$ we have $d_G^\chi(AA^*) = d_G^\chi(A) = \chi(e)$.

($\Leftarrow$) From the assumption and the above argument,

$$(v^*, u^*) = (u^*, u^*) = (v^*, v^*) \tag{4.33}$$

so that

$$|(v^*, u^*)|^2 = (u^*, u^*)(v^*, v^*)$$

with $u^* \neq 0$. By the equality case of Cauchy-Schwarz inequality, $v^* = cu^*$ for some constant. Then substitute into (4.33) to have $c = 1$. □

Theorem 4.6.10. For all psd $A, B \in \mathbb{C}_{m\times m}$,

$$d_G^\chi(A + B) \geq d_G^\chi(A) + d_G^\chi(B).$$

In particular, if $A \geq B$, then $d_G^\chi(A) \geq d_G^\chi(B)$.

Proof. Let E be an orthonormal basis of V with $\dim V = m$. There are psd $S, T \in \operatorname{End} V$ such that $[S]_E^E = A$ and $[T]_E^E = B$. Notice that $K(S + T) \geq K(S) + K(T)$ (Problem 4.5 #3) and $[T + S]_E^E = A + B$. Apply Theorem 4.6.3 (switching χ to $\bar{\chi}$) to have the desired result.

Suppose $A \geq B$, i.e., $A = B + C$ where $C \geq 0$. The

$$d_G^\chi(A) \geq d_G^\chi(B + C) \geq d_G^\chi(B) + d_G^\chi(C) \geq d_G^\chi(B)$$

since $d_G^\chi(C) \geq 0$ by Theorem 4.6.5. □

We now discuss the generalized Cauchy-Binet Theorem (compare Theorem 1.10.1).

Theorem 4.6.11. (Generalized Cauchy-Binet) Let $A, B \in \mathbb{C}_{n\times n}$. For any $\alpha, \beta \in \Omega = \{\omega : \Gamma_{m,n} : \sum_{\sigma\in G_\omega} \chi(\sigma) \neq 0\}$

$$d_G^\chi((AB)[\alpha|\beta]) = \frac{\chi(e)}{|G|} \sum_{\omega\in\Omega} d_G^\chi(A[\alpha|\omega]) d_G^\chi(B[\omega|\beta])$$

Proof. Let $E = \{e_1, \dots, e_n\}$ be an orthonormal basis of V. There are $S, T \in \operatorname{End} V$ such that $[S]_E^E = A$ and $[T]_E^E = B$. Notice that $E_\otimes = \{e^\otimes_\omega : \omega \in \Gamma_{m,n}\}$ is an orthonormal basis of $\otimes^m V$ with respect to the induced inner product. We have

$$\begin{aligned}
&\frac{\chi(e)}{|G|} d_G^{\bar\chi}((AB)[\alpha|\beta]\\
=\ & (K(ST)e^*_\beta, e^*_\alpha) \quad \text{(by Theorem 4.6.3)}\\
=\ & (K(S)K(T)e^*_\beta, e^*_\alpha) \quad (K(ST) = K(S)K(T))\\
=\ & (K(T)e^*_\beta, K(S^*)e^*_\alpha) \quad (K(S)^* = K(S^*))\\
=\ & \sum_{\omega\in\Gamma_{m,n}} (K(T)e^*_\beta, e^\otimes_\omega)(e^\otimes_\omega, K(S^*)e^*_\alpha) \quad \text{(by Theorem 1.3.3)}\\
=\ & \sum_{\omega\in\Gamma_{m,n}} (K(T)e^*_\beta, e^*_\omega)(e^*_\omega, K(S^*)e^*_\alpha) \quad \text{(by Theorem 4.2.4 } \otimes^m V = \perp_{\chi\in I(G)} V_\chi(G))\\
=\ & \sum_{\omega\in\Omega} (K(S)e^*_\omega, e^*_\alpha)(K(T)e^*_\beta, e^*_\omega) \quad (e^*_\omega \neq 0 \Leftrightarrow \omega \in \Omega,\ \text{if } \omega \in \Gamma_{m,n})\\
=\ & \frac{\chi(e)^2}{|G|^2} \sum_{\omega\in\Omega} d_G^{\bar\chi}(A[\alpha|\omega]) d_G^{\bar\chi}(B[\omega|\beta]). \quad \text{(by Theorem 4.6.3)}
\end{aligned}$$

Then switch back to χ. □

When $G = S_m$ and $\chi = \varepsilon$ it reduces to Theorem 1.10.1.

Theorem 4.6.12. Let $H < G < S_m$ and $\chi \in I(G)$. Suppose $\chi|_H \in I(H)$. Then for all psd $A \in \mathbb{C}_{m\times m}$,

$$\frac{1}{|H|} d_H^\chi(A) \geq \frac{1}{|G|} d_G^\chi(A).$$

Proof.

$$\begin{aligned}
T(G,\chi)T(H,\chi) &= \frac{\chi(e)^2}{|G||H|} \sum_{\sigma\in G}\sum_{\pi\in H} \chi(\sigma)\chi(\pi)P(\sigma\pi)\\
&= \frac{\chi(e)^2}{|G||H|} \sum_{\tau\in G}\sum_{\pi\in H} \chi(\tau\pi^{-1})\chi(\pi)P(\tau) \quad (\tau = \sigma\pi)\\
&= \frac{\chi(e)}{|G|} \sum_{\tau\in G} \chi(\tau)P(\tau) \quad \text{(Theorem 2.4.3 since } \chi(\tau\pi^{-1}) = \chi(\pi^{-1}\tau))\\
&= T(G,\chi).
\end{aligned}$$

With respect to the induced inner product on $\otimes^m V$ where V is an m-dimensional inner product space, $T(G,\chi)$ is an orthogonal projection by Theorem 4.2.5(a). So for any $z \in \otimes^m V$

$$\|z\| \geq \|T(G,\chi)z\|.$$

Since $T(G,\chi)T(H,\chi) = T(G,\chi)$, apply the above inequality on $T(H,\chi)z$,

$$\|T(H,\chi)z\| \geq \|T(G,\chi)T(H,\chi)z\| = \|T(G,\chi)z\|.$$

Since $A \in \mathbb{C}_{m\times m}$ is psd and $\dim V = m$, there are $v_1,\dots,v_m \in V$ such that $a_{ij} = (v_i, v_j)$ (Problem 1.5 #10). Set $z = v^{\otimes}$ in the above inequality to have

$$(T(H,\chi)v^{\otimes}, T(H,\chi)v^{\otimes}) \geq (T(G,\chi)v^{\otimes}, T(G,\chi)v^{\otimes}).$$

Then use Theorem 4.6.2 to have

$$\frac{1}{|H|}d_H^\chi(A) \geq \frac{1}{|G|}d_G^\chi(A).$$

□

The following famous conjecture is still open.

Permanent-on-top conjecture: If $A \in \mathbb{C}_{m\times m}$ is psd, then

$$d_G^\chi(A) \leq \chi(e)\text{per}\, A.$$

Problems

1. Show that if $A, B \in \mathbb{C}_{m\times m}$ are psd, then $\det(A+B) \geq \det A + \det B$ and $\text{per}\,(A+B) \geq \text{per}\, A + \text{per}\, B$.
2. (Hadamard) Prove that if $A \in \mathbb{C}_{m\times m}$, then $|\det A|^2 \leq \prod_{i=1}^m \sum_{j=1}^m |a_{ij}|^2$. Hint: $\det(AA^*) \leq h(AA^*)$.
3. Prove that when χ is linear, Cauchy-Binet formula takes the form: for any $\alpha, \beta \in \bar{\Delta}$,
$$d_G^\chi((AB)[\alpha|\beta]) = \sum_{\omega\in\bar{\Delta}} \frac{1}{|G_\omega|} d_G^\chi(A[\alpha|\omega]) d_G^\chi(B[\omega|\beta]).$$
4. Suppose that $H < G$ and assume that $\chi|_H$ is also irreducible, where χ is an irreducible character of G. Show that $V_\chi(G) \subset V_\chi(H)$.
5. If $A \in \mathbb{C}_{m\times m}$ is psd, then
$$\text{per}\, A \geq \text{per}\, A[1,\dots,p|1,\dots,p]\, \text{per}\, A[p+1,\dots,m|p+1,\dots,m].$$

Solution to Problems 4.5

1. Follows from Theorem 4.6.10 with $V = \mathbb{C}^m$, $G = S_m$, $\chi = \varepsilon$; $V = \mathbb{C}^m$, $G = S_m$, $\chi \equiv 1$.

2. By Theorem 4.6.6, $|\det A|^2 = \det(AA^*) \leq h(AA^*)$ since AA^* is psd. But $h(AA^*) = \prod_{i=1}^m \sum_{j=1}^m |a_{ij}|^2$.

3. From Theorem 4.6.11, since $\chi(e) = 1$,

$$\begin{aligned} d_G^\chi((AB)[\alpha|\beta]) &= \frac{1}{|G|}\sum_{\omega\in\Omega} d_G^\chi(A[\alpha|\omega])d_G^\chi(B[\omega|\beta]) \\ &= \frac{1}{|G|}\sum_{\omega\in\bar{\Delta}}\sum_{\sigma\in G} d_G^\chi(A[\alpha|\omega\sigma])d_G^\chi(B[\omega\sigma|\beta]) \quad (\Omega = \bar{\Delta}G) \end{aligned}$$

   Let $A = (a_{ij}) \in \mathbb{C}_{m\times m}$ and let $u_1, \ldots, u_m; v_1, \ldots, v_m \in V$ such that $a_{ij} = (u_i, v_j)$. If $\sigma \notin G_\omega$, then $\omega\sigma \notin \bar{\Delta}$ so that $v^*_{\omega\sigma} = v_{\omega\sigma(1)} * \cdots * v_{\omega\sigma(m)} = 0$. By Theorem 4.6.2, if $\sigma \notin G_\omega$, then

$$\frac{1}{|G|}d_G^\chi(A[\alpha|\omega\sigma]) = (u^*_{\alpha(i)}, v^*_{\omega\sigma(j)}) = 0.$$

   So we have the desired result.

4.

5.

Chapter 5

Exterior spaces and completely symmetric spaces

5.1 Exterior spaces

Let $\dim V = n$. When $G = S_m$ with $m \le n$, and $\chi = \varepsilon$, $V_\varepsilon(S_m)$, denoted by $\wedge^m V$ is called the mth **exterior space** or **Grassmannian space**. Denote by

$$v^\wedge = v_1 \wedge \cdots \wedge v_m$$

the decomposable element v^*. The induced operator $K(T)$ is denoted by $C_m(T)$. Since the alternating character is linear, the induced matrix $K(A)$ is defined by (4.25) and is denoted by $C_m(A)$. It is called the mth compound of A.

Theorem 5.1.1. Let $\dim V = n \ge m$. For $\wedge^m V$,

(a) $\Delta = G_{m,n}$, $\bar{\Delta} = \hat{\Delta} = Q_{m,n}$, $\dim \wedge^m V = \binom{n}{m}$.

(b) Let $E = \{e_1, \ldots, e_n\}$ be a basis of V. Then $E_\wedge := \{e_\alpha^\wedge : \alpha \in Q_{m,n}\}$ is a basis of $\wedge^m V$. If E is an orthonormal basis of V, then $E'_\wedge := \sqrt{m!}E^\wedge$ is an orthonormal basis of $\wedge^m V$.

(c) For each $\sigma \in S_m$, $v_\sigma^\wedge = \varepsilon(\sigma) v^\wedge$. When $i \ne j$ and $v_i = v_j$, $v^\wedge = 0$.

(d) $v_1 \wedge \cdots \wedge v_m = 0$ if and only if $v_1, \ldots, v_m$ are linearly dependent.

(e) If $u_i = \sum_{j=1}^m a_{ij} v_j$, $i = 1, \ldots, m$, then $u^\wedge = \det(a_{ij}) v^\wedge$.

(f) If $u^\wedge = v^\wedge \ne 0$, then $u_i = \sum_{j=1}^m a_{ij} v_j$ $i = 1, \ldots, m$ and $\det(a_{ij}) = 1$.

(g) If $v_i = \sum_{j=1}^n a_{ij} e_j$, $i = 1, \ldots, m$, then $v^\wedge = \sum_{\alpha \in Q_{m,n}} \det A[1, \ldots, m|\alpha] e_\alpha^\wedge$.

Proof. The proofs of (a), (b), (c) are obtained in Chapter 4. Notice that $\varepsilon(e) = 1$ and $\alpha \in Q_{m,n}$, $G_\alpha = \{e\}$.

(d) By Theorem 4.3.2 it suffices to show the sufficiency. Let $v_1, \dots, v_m$ be linearly dependent. Then for some j, $v_j = \sum_{t \neq j} c_t v_t$. So by (c)

$$v_1 \wedge \cdots \wedge v_m = \sum_{t \neq j} c_t v_1 \wedge \cdots \wedge v_{j-1} \wedge v_t \wedge v_{j+1} \wedge \cdots \wedge v_m = 0$$

because each summand has repeated v's.

(e) Direct computation yields

$$\begin{aligned}
u^\wedge &= (\sum_{j=1}^m a_{1j} v_j) \wedge \cdots \wedge (\sum_{j=1}^m a_{mj} v_j) \\
&= \sum_{\alpha \in \Gamma_{m,m}} (\prod_{i=1}^m a_{i\alpha(i)}) v_\alpha^\wedge \\
&= \sum_{\alpha \in D_{m,m}} (\prod_{i=1}^m a_{i\alpha(i)}) v_\alpha^\wedge \quad ((d) \text{ and } D_{m,m} = \{(1, \dots, m)\sigma : \sigma \in S_m\}) \\
&= \sum_{\sigma \in S_m} \left(\prod_{i=1}^m a_{i\sigma(i)} \right) v_\sigma^\wedge \\
&= \left(\sum_{\sigma \in S_m} \varepsilon(\sigma) \prod_{i=1}^m a_{i\sigma(i)} \right) v^\wedge \\
&= \det(a_{ij}) v^\wedge \quad \text{(Problem 1)}
\end{aligned}$$

(f) By Theorem 4.3.3

$$\langle u_1, \dots, u_m \rangle = \langle v_1, \dots, v_m \rangle.$$

Then $u_i = \sum_{j=1}^m a_{ij} v_j$, $i = 1, \dots, m$. Use (e) to have $\det(a_{ij}) = 1$ (a special case of Theorem 4.6.9)

(g) Similar to (e)

$$\begin{aligned}
v^\wedge &= (\sum_{j=1}^n a_{1j}e_j)\wedge\cdots\wedge(\sum_{j=1}^n a_{mj}e_j)\\
&= \sum_{\alpha\in\Gamma_{m,n}}\prod_{i=1}^m a_{i\omega(i)}e_\omega^\wedge\\
&= \sum_{\alpha\in D_{m,n}}\prod_{i=1}^m a_{i\omega(i)}e_\omega^\wedge \quad ((d) \text{ and } D_{m,n}=\{\alpha\sigma:\alpha\in Q_{m,n},\sigma\in S_m\})\\
&= \sum_{\alpha\in Q_{m,n}}\sum_{\sigma\in S_m}(\prod_{i=1}^m a_{i\alpha\sigma(i)})e_{\alpha\sigma}^\wedge\\
&= \sum_{\alpha\in Q_{m,n}}\left(\sum_{\sigma\in S_m}\varepsilon(\sigma)\prod_{i=1}^m a_{i\alpha\sigma(i)}\right)e_\alpha^\wedge\\
&= \sum_{\alpha\in Q_{m,n}}\det A[1,\ldots,m|\alpha]e_\alpha^\wedge
\end{aligned}$$

□

We now study the induced matrix $C_m(A)$. From (4.26),

$$K(A)_{\alpha,\beta}=\frac{1}{\sqrt{|G_\alpha||G_\beta|}}\sum_{\sigma\in G}\chi(\sigma)\prod_{t=1}^m a_{\alpha\sigma(t),\beta(t)}=\frac{1}{\sqrt{|G_\alpha||G_\beta|}}d_G^{\bar\chi}(A[\alpha|\beta]),\quad (5.1)$$

for all $\alpha,\beta\in\bar\Delta$. Hence

$$C_m(A)_{\alpha,\beta}=\det A[\alpha|\beta],$$

for all $\alpha,\beta\in Q_{m,n}$.

For example, if $n=3$ and $k=2$, then

$$C_2(A)=\begin{pmatrix}\det A[1,2|1,2] & \det A[1,2|1,3] & \det A[1,2|2,3]\\ \det A[1,3|1,2] & \det A[1,3|1,3] & \det A[1,3|2,3]\\ \det A[2,3|1,2] & \det A[2,3|1,3] & \det A[2,3|2,3]\end{pmatrix}.$$

In general $C_1(A)=A$ and $C_n(A)=\det A$.

From Theorem 4.5.8 we have the following result.

Theorem 5.1.2. Let $E=\{e_1,\ldots,e_n\}$ be a basis of V and $E_\wedge := \{e_\alpha^\wedge : \alpha\in Q_{m,n}\}$ and $E'_\wedge=\sqrt{m!}E_\wedge$, a basis of $\wedge^m V$ (in lexicographic order). Let $T\in \operatorname{End} V$ and $[T]_E^E=A$. For any $\alpha,\beta\in Q_{m,n}$,

$$([C_m(T))]_{E_\wedge}^{E_\wedge})_{\alpha,\beta}=([C_m(T))]_{F^\wedge}^{F^\wedge})_{\alpha,\beta}=C_m(A)_{\alpha,\beta}=\det A[\alpha|\beta].$$

The induced matrix $C_m(A)$ is called the mth compound of A. Clearly

$$C_m(AB)=C_m(A)C_m(B).$$

Indeed it is the Cauchy-Binet determinant identity (Theorem 1.10.1) since

$$C_m(AB)_{\alpha,\beta} = \sum_{\omega\in Q_{m,n}} C_m(A)_{\alpha,\omega} C_m(B)_{\omega,\beta}, \quad \alpha,\beta\in Q_{m,n}.$$

By Theorem 5.1.2

$$\det(AB)[\alpha|\beta] = \sum_{\omega\in Q_{m,n}} \det A[\alpha|\omega] \det B[\omega|\beta].$$

We now list some basic properties of $C_m(A)$.

Theorem 5.1.3. Let $m \le n$ and $A \in \mathbb{C}_{n\times n}$ with eigenvalues $\lambda_1,\ldots,\lambda_n$ and singular values $s_1,\ldots,s_n$. Then

(a) $C_m(A^*) = C_m(A)^*$, $C_m(A^{-1}) = C_m(A)^{-1}$ if A is invertible.

(b) If $\operatorname{rank} A = r \ge m$, then $\operatorname{rank} C_m(A) = \binom{r}{m}$. If $r < m$, then $C_m(A) = 0$.

(c) If A is upper triangular, so is $C_m(A)$.

(d) $\prod_{t=1}^m \lambda_{\sigma(t)}$ are the eigenvalues and $\prod_{t=1}^m s_{\sigma(t)}$ $(\sigma\in Q_{m,n})$ are the singular values of $C_m(A)$.

(e) $\operatorname{tr} C_m(A) = \sum_{\alpha\in Q_{m,n}} \prod_{t=1}^m \lambda_{\sigma(t)} = \sum_{\alpha\in Q_{m,n}} \det A[\alpha|\alpha]$.

(f) (Sylvester-Franke) $\det C_m(A) = (\det A)^{\binom{n-1}{m-1}}$.

(g) If A is normal, Hermitian, pd, psd or unitary, so is $C_m(A)$.

Compound matrix is a very powerful tool and the following results are two applications.

Theorem 5.1.4. Let $A\in\mathbb{C}_{n\times n}$ with eigenvalues $|\lambda_1|\ge\cdots\ge|\lambda_n|$. Denote by $R_i = \sum_{j=1}^n |a_{ij}|$, $i=1,\ldots,n$, and let $R_{[1]}\ge\cdots\ge R_{[n]}$ be an rearrangement of $R_1,\ldots,R_n$ in nonincreasing order. Then

$$\prod_{t=1}^m |\lambda_i| \le \prod_{t=1}^m R_{[i]}, \quad 1\le m\le n. \tag{5.2}$$

Proof. Let $Ax = \lambda_1 x$ where x is an eigenvector of A. Then

$$\sum_{j=1}^n a_{ij}x_j = \lambda_1 x_i, \quad i=1,\ldots,n.$$

Let $1\le s\le n$ be an integer such that

$$|x_s| = \max_{i=1} |x_i| > 0.$$

Then

$$|\lambda_1| = \left|\sum_{j=1}^n a_{sj}\frac{x_j}{x_s}\right| \le \sum_{j=1}^n \left|a_{sj}\frac{x_j}{x_s}\right| \le \sum_{j=1}^n |a_{sj}| = R_s \le R_{[1]}.$$

Use Theorem 5.1.3(d) and apply the inequality on $C_m(A)$ to have

$$\prod_{i=1}^m |\lambda_i(A)| = |\lambda_1(C_m(A))| \le R_{[1]}(C_m(A)).$$

For any $\alpha \in Q_{m,n}$, use Theorem 5.1.2

$$\begin{aligned} R_\alpha(C_m(A)) &= \sum_{\beta\in Q_{m,n}} |\det A[\alpha|\beta]| \\ &= \sum_{\beta\in Q_{m,n}} \left|\sum_{\sigma\in S_m}\varepsilon(\sigma)\prod_{i=1}^m a_{\alpha(i)\beta\sigma(i)}\right| \\ &\le \sum_{\omega\in D_{m,n}}\prod_{i=1}^m |a_{\alpha(i)\omega(i)}| \quad (D_{m,n} = \{\alpha\sigma : \alpha\in Q_{m,n}, \sigma\in S_m\}) \\ &\le \sum_{\gamma\in\Gamma_{m,n}}\prod_{i=1}^m |a_{\alpha(i)\gamma(i)}| \\ &= \prod_{i=1}^m\sum_{j=1}^n |a_{\alpha(i)j}| = \prod_{i=1}^m R_{\alpha(i)} \le \prod_{i=1}^m R_{[i]} \end{aligned}$$

□

We can use (5.8) to have

$$|\lambda_n| \ge \frac{|\det A|}{\prod_{i=1}^{n-1} R_{[i]}} \tag{5.3}$$

since

$$\det A = \prod_{i=1}^n \lambda_i.$$

Theorem 5.1.5. Suppose that $A \in \mathbb{C}_{n\times n}$ is invertible with singular values $s_1 \ge \cdots \ge s_n > 0$. Let $A = QR$ be the unique QR decomposition where positive diagonal entries of R. Let $\operatorname{diag} R = (a_1, \dots, a_n)$. Let $a_{[1]} \ge \cdots \ge a_{[n]}$ be the rearrangement of $a_1, \dots, a_n$. Then

$$\prod_{t=1}^m a_{[t]} \le \prod_{t=1}^m s_t, \quad m = 1, \dots, n-1, \quad \prod_{t=1}^n a_{[t]} = \prod_{t=1}^n s_t. \tag{5.4}$$

Proof. We first establish $a_{[1]} \le s_1$. Let $A = QR$ be the QR decomposition of A. Recall that s_1 is the square root of the largest eigenvalue of the psd

$$A^*A = R^*Q^*QR = R^*R.$$

It is known that any diagonal entry of a Hermitian B is less than or equal to the largest eigenvalue of B (by using Spectral Theorem for Hermitian matrices). Suppose $a_{[1]} = a_r$ for some $1 \le r \le n$. Then

$$a_{[1]}^2 = |a_r|^2 \le (R^*R)_{rr} \le s_1^2. \tag{5.5}$$

Hence $a_{[1]} \le s_1$. and use Theorem 5.1.3 (c) and (g),

$$C_m(A) = C_m(Q)C_m(R)$$

is the QR decomposition of $C_m(A)$. Now the diagonal entries of $C_m(R)$ are $\prod_{i=1}^m a_{\alpha(i)}$ $\alpha \in Q_{m,n}$ and the largest is $\prod_{i=1}^m a_{[i]}$ By Theorem 5.1.3(d), the largest singular value of $C_m(A)$ is $\prod_{i=1}^m s_i$. So by (5.5) we have

$$\prod_{i=1}^m a_{[i]} \le \prod_{i=1}^m s_i, \quad i = 1, \ldots, n-1.$$

The equality is obtained via the determinant equation

$$\det A = \det Q \det R$$

so that

$$\prod_{i=1}^n s_i = |\det A| = |\det R| = \prod_{i=1}^n a_{[i]}.$$

□

Remark: The converse of Theorem 5.1.5 is true, i.e., given positive numbers $a_1, \ldots, a_n$ and $s_1 \ge \cdots \ge s_n > 0$, if (5.4) is true, then there exists $A \in \mathbb{C}_{n\times n}$ with singular values s's and $\operatorname{diag} R = (a_1, \ldots, a_n)$ where $A = QR$ is the QR decomposition of A. This result is known as Kostant's convexity theorem [11]. Indeed Kostant's result is in a broader context known as real semisimple Lie groups.

For $A \in \mathbb{C}_{n\times n}$, $t \in \mathbb{C}$, we write

$$C_m(I_n + tA) = I_{\binom{n}{m}} + tD_m^{(1)}(A) + t^2 D_m^{(2)}(A) + \cdots + t^m D_m^{(m)}(A).$$

The matrix $D_m^{(r)}(A)$ is called the rth **derivative** of $C_m(A)$. In particular $D_m^{(1)}(A)$ is the directional derivative of $C_m(\cdot)$ at the identity I_n in the direction A. It is viewed as the linearization of $C_m(A)$

Theorem 5.1.6. If $A \in \mathbb{C}_{n\times n}$ is upper triangular, then $D_m^{(r)}(A)$ is upper triangular.

Proof. If A is upper triangular, so is $I_n + tA$. Thus $C_m(I_n + tA)$ is upper triangular. □

Theorem 5.1.7. (a) $D_m^{(r)}(SAS^{-1}) = C_m(S)D_m^{(r)}(A)C_m(S)^{-1}$ for all $1 \le r \le m$. So, if A and B are similar, then $C_m(A)$ and $C_m(B)$ are similar.

(b) Let $A \in \mathbb{C}_{n\times n}$ with eigenvalues $\lambda_1, \ldots, \lambda_n$. Then the eigenvalues of $D_m^{(r)}(A)$ are

$$\sum_{\omega\in Q_{r,m}} \prod_{i=1}^{r} \lambda_{\alpha\omega(i)}, \quad \alpha \in Q_{m,n}.$$

In particular $\lambda_{\alpha(1)} + \cdots + \lambda_{\alpha(m)}$, $\alpha \in Q_{m,n}$, are the eigenvalues of $D_m^{(1)}(A)$.

(c) $D_m^{(r)}(A^*) = D_m^{(r)}(A)^*$.

Proof. (a)

$$\begin{aligned}
& C_m(I_n + tSAS^{-1}) \\
= \ & C_m(S(I_n + tA)S^{-1}) \\
= \ & C_m(S)C_m(I_n + tA)C_m(S^{-1}) \\
= \ & C_m(S)C_m(I_n + tA)C_m(S)^{-1} \\
= \ & C_m(S)(I_{\binom{n}{m}} + tD_m^{(1)}(A) + t^2 D_m^{(2)}(A) + \cdots + t^m D_m^{(m)}(A))C_m(S)^{-1} \\
= \ & I_{\binom{n}{m}} + tC_m(S)D_m^{(1)}(A)C_m(S)^{-1} + t^2 C_m(S)D_m^{(2)}(A)C_m(S)^{-1} \\
& \quad + \cdots + t^m C_m(S)D_m^{(m)}(A)C_m(S)^{-1}
\end{aligned}$$

On the other hand

$$C_m(I_n+tSAS^{-1}) = I_{\binom{n}{m}} + tD_m^{(1)}(SAS^{-1}) + t^2 D_m^{(2)}(SAS^{-1}) + \cdots + t^m D_m^{(m)}(SAS^{-1}).$$

Comparing coefficients of t^r of both equations yields the desired result.

(b) By Schur triangularization theorem, there is $U \in \mathrm{U}(n)$ such that $UAU^* = T$ where T is upper triangular and $\operatorname{diag} T = (\lambda_1, \ldots, \lambda_n)$. By (a)

$$D_m^{(r)}(A) = C_m(U^*)D_m^{(r)}(T)C_m(U)$$

and hence $D_m^{(r)}(A))$ and $D_m^{(r)}(T))$ have the same eigenvalues. From Theorem 5.1.6 $D_m^{(r)}(T)$ is upper triangular so the diagonal entries are the eigenvalues of

A. Now

$$\begin{aligned}
& C_m(I+tT)_{\alpha,\alpha} \\
= \; & \det(I_n+tT)[\alpha|\alpha] \\
= \; & \prod_{i=1}^{m}(1+t\lambda_{\alpha(i)}) \\
= \; & 1+t(\lambda_{\alpha(1)}+\cdots+\lambda_{\alpha(m)})+t^2\sum_{\omega\in Q_{2,m}}\prod_{i=1}^{2}\lambda_{\alpha\omega(i)} \\
& \quad +\cdots+t^r\sum_{\omega\in Q_{r,m}}\prod_{i=1}^{r}\lambda_{\alpha\omega(i)}+\cdots
\end{aligned}$$

for all $\alpha \in Q_{m,n}$. So the eigenvalues of $D_m^{(r)}(A)$ are $\sum_{\omega\in Q_{r,m}}\prod_{i=1}^{r}\lambda_{\alpha\omega(i)}$, $\alpha \in Q_{m,n}$.

(c) For any $t \in \mathbb{R}$,

$$\begin{aligned}
& I_{\binom{n}{m}}+tD_m^{(1)}(A^*)+t^2D_m^{(2)}(A^*)+\cdots+t^mD_m^{(m)}(A^*) \\
= \; & C_m(I+tA^*) \\
= \; & C_m(I+tA)^* \\
= \; & I_{\binom{n}{m}}+tD_m^{(1)}(A)^*+t^2D_m^{(2)}(A)^*+\cdots+t^mD_m^{(m)}(A)^*
\end{aligned}$$

Hence $D_m^{(r)}(A^*) = D_m^{(r)}(A)^*$ for all r as desired. □

Let $A \in \mathbb{C}_{n\times n}$ and set $\Delta_m(A) := D_m^{(1)}(A)$. It is called the mth **additive compound** of A.

Theorem 5.1.8. $\Delta_m : \mathbb{C}_{n\times n} \to \mathbb{C}_{\binom{n}{m}\times\binom{n}{m}}$ is linear, i.e., $\Delta_m(A+B) = \Delta_m(A)+\Delta_m(B)$ for all $A, B \in \mathbb{C}_{n\times n}$ and $\Delta_m(cA) = c\Delta_m(A)$.

Proof. By Theorem 5.1.3(a)

$$C_m((I_n+tA)(I+tB)) = C_m(I_n+tA)C_m(I_n+tB).$$

Now

$$\begin{aligned}
C_m(I_n+tA)C_m(I_n+tB) &= (I_{\binom{n}{m}}+t\Delta_m(A)+o(t^2))(I_{\binom{n}{m}}+t\Delta_m(B)+o(t^2)) \\
&= I_{\binom{n}{m}}+t(\Delta_m(A)+\Delta_m(B))+o(t^2),
\end{aligned}$$

and

$$C_m((I_n+tA)(I_n+tB)) = C_m(I_n+t(A+B)+t^2AB) = I_{\binom{n}{m}}+t(\Delta_m(A+B))+o(t^2).$$

Then compare coefficients of t on both sides. The equality $\Delta_m(cA) = c\Delta_m(A)$ is easily deduced as well. □

The following problem is still open

Marcus–de Oliveira Conjecture
Let $A, B \in \mathbb{C}_{n\times n}$ be normal matrices with eigenvalues $\alpha_1, \ldots, \alpha_n$ and $\beta_1, \ldots, \beta_n$. The Marcus–de Oliveira Conjecture claims that

$$\{\det(VAV^{-1} + UBU^{-1}) : V, U \in \mathrm{U}(n)\} \subset \operatorname{conv}\{\prod_{i=1}^{n}(\alpha_i + \beta_{\sigma(i)}) : \sigma \in S_n\},$$

where $\operatorname{conv} S$ denotes the convex hull of the set $S \subset \mathbb{C}$.

Since

$$\begin{aligned}\det(VAV^{-1} + UBU^{-1}) &= \det(V(A + V^{-1}UBU^{-1}V)V^{-1}) \\ &= \det V \det(A + V^{-1}UBU^{-1}V) \det V^* \\ &= \det(A + V^{-1}UBU^{-1}V),\end{aligned}$$

by using spectral theorem for normal matrices, Marcus–de Oliveira Conjecture can be restated as

$$\begin{aligned}&\{\det(\operatorname{diag}(\alpha_1, \ldots, \alpha_n) + U\operatorname{diag}(\beta_1, \ldots, \beta_n)U^{-1}) : U \in \mathrm{U}(n)\} \\ \subset\ &\operatorname{conv}\{\prod_{i=1}^{n}(\alpha_i + \beta_{\sigma(i)}) : \sigma \in S_n\},\end{aligned}$$

where $\operatorname{conv} S$ denotes the convex hull of the set $S \subset \mathbb{C}$.

When A and B are Hermitian, the inclusion becomes an equality, a result of Fielder [2]. See the generalization of Fielder's result in the context of real semisimple Lie algebras [28, 29, 30].

For $A, B \in \mathbb{C}_{n\times n}$, the Cauchy-Binet formula yields [21]

$$\det(A + B) = \det((A\ I_n)\begin{pmatrix} I_n \\ B\end{pmatrix}) = C_n(A\ I_n)C_n\begin{pmatrix} I_n \\ B\end{pmatrix}.$$

Thus

$$\det(A+UBU^{-1}) = C_n(A\ I_n)C_n\begin{pmatrix} I_n \\ UBU^{-1}\end{pmatrix} = C_n(A\ I_n)C_n((U\oplus U)\begin{pmatrix} I_n \\ B\end{pmatrix}U^{-1}).$$

Problems

1. Show that for $\wedge^m V$ $(m \le n = \dim V)$, $\Omega = D_{m,n}$.
2. Let $T \in \operatorname{End} V$ and $\dim V = n$. Prove that for all $v_1, \ldots, v_n \in V$, $C_n(T)v_1 \wedge \cdots \wedge v_n = (\det T)v_1 \wedge \cdots \wedge v_n$.
3. Suppose $V_\chi(G) \ne 0$ and $v_1 * \cdots * v_m = 0$ whenever $v_1, \ldots, v_m$ are linearly dependent. Prove that $V_\chi(G) = \wedge^m V$, i.e., $G = S_m$ and $\chi = \varepsilon$.

4. (Cartan Lemma) Let $e_1, \ldots, e_k$ be linear independent and $\sum_{i=1}^k e_i \wedge v_i = 0$. Prove that there is a symmetric matrix $A \in \mathbb{C}_{k\times k}$ such that $v_i = \sum_{i=1}^k a_{ij} e_j$, $i = 1, \ldots, k$. Hint: Extend $e_1, \ldots, e_k$ to a basis $e_1, \ldots, e_n$ of V. Notice that $\{e_i \wedge e_j : 1 \le i < j \le n\}$ is a basis of $\wedge^2 V$ and $\sum_{i,j=1}^n a_{ij} e_i \wedge e_j = \sum_{i<j} (a_{ij} - a_{ji}) e_i \wedge e_j$.

5. Let $A \in \mathbb{C}_{n\times n}$ and $E_i = \sqrt{\sum_{j=1}^n |a_{ij}|^2}$ and use similar notations as in Theorem 5.1.4. Prove that

$$\prod_{i=1}^m |\lambda_i| \le \binom{n}{m} \prod_{i=1}^m E_{[i]}, \quad m = 1, \ldots, n$$

and

$$|\lambda_n| \ge \frac{|\det A|}{n \prod_{i=1}^{n-1} E_{[i]}}.$$

Hint: Use Problem 4.6 #2.

6. (Weyl) Let $A \in \mathbb{C}_{n\times n}$. Let $\lambda_1, \ldots, \lambda_n$ with the ordering $|\lambda_1| \ge \cdots | \ge \lambda_n|$ and $s_1 \ge \cdots \ge s_n$ be the singular values of A, respectively. Prove that

$$\prod_{t=1}^m |\lambda_t| \le \prod_{t=1}^m s_t, \quad m = 1, \ldots, n-1, \quad \prod_{t=1}^n |\lambda_t| = \prod_{t=1}^n s_t. \tag{5.6}$$

(Hint: Establish $|\lambda_1| \le s_1$ and use Theorem 5.1.3(d). Remark: The converse is true, i.e., given $\lambda_1, \ldots, \lambda_n$ in nonincreasing moduli and $s_1 \ge \cdots \ge s_n \ge 0$, if (5.6) is true, then there exists $A \in \mathbb{C}_{n\times n}$ with eigenvalues λ's and singular values s's. It is due to A. Horn [7] and see a different proof of T.Y. Tam [27]).

7. Let $A \in \mathbb{C}_{n\times n}$ with distinct eigenvalues $\lambda_1, \ldots, \lambda_n$. Define

$$\mathrm{Sep}(A) = \min_{i \ne j} |\lambda_i - \lambda_j|.$$

Set $L := A \otimes I_n - I_n \otimes A$. Prove

$$\mathrm{Sep}(A) \ge \left(\frac{|\operatorname{tr} C_{n^2-n}(L)|}{\prod_{i=1}^{n^2-n-2} R_{[i]}(L)} \right)^{1/2}.$$

Hint: The eigenvalues of L are $\lambda_i - \lambda_j$ $i, j = 1, \ldots, n$, in which $n^2 - n$ are nonzero so $C_{n^2-n}(L)$ has only one nonzero eigenvalue

$$\prod_{i \ne j} (\lambda_i - \lambda_j) = (-1)^{\frac{n^2-n}{2}} \prod_{i<j} (\lambda_i - \lambda_j)^2 = \operatorname{tr} C_{n^2-n}(L).$$

Then apply (5.3).

8. Let $A \in \mathbb{C}_{n\times n}$ be Hermitian with distinct eigenvalues and set $L := A \otimes I_n - I_n \otimes A$. Prove

$$\mathrm{Sep}(A) \geq \left(\frac{|\det L[\omega|\omega]|}{\prod_{i=1}^{n^2-n-2} R_{[i]}(L)} \right)^{1/2},$$

where $\omega \in Q_{n^2-n,n^2}$. Hint: $|\lambda_1| \geq \max_{i=1,\ldots,n} |a_{ii}|$.

9. Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalue of $A \in \mathbb{C}_{n\times n}$. Prove that $(\lambda_i - \lambda_j)^2$, $1 \leq i < j \leq n$, are the eigenvalues of $M := D_2^{(1)}(A^2) - 2C_2(A)$. When the eigenvalues of A are distinct, use M to give a nonzero lower bound of $\mathrm{Sep}(A)$.

Solutions

1.

2.

3.

4.

5.

6.

7.

5.2 Decomposable elements in the exterior spaces

We will give some characterization of decomposable elements in the exterior space $\wedge^m V$. We know from Theorem 5.1.1(d) that $v^\wedge = v_1 \wedge \cdots \wedge v_m$ is nonzero if and only if $v_1, \ldots, v_m$ are linearly independent. Moreover from Theorem 5.1.1(f) if $v^\wedge \neq 0$, then $u^\wedge = v^\wedge$ if and only if $u_i = \sum_{j=1}^m c_{ij} v_j$, $i = 1, \ldots, m$ and $\det(c_{ij}) = 1$. So the two m-dimensional subspaces $\langle u_1, \ldots, u_m\rangle$ and $\langle v_1, \ldots, v_m\rangle$ of V are the same if and only if $u^\wedge = cv^\wedge \neq 0$. Thus nonzero decomposable $v^\wedge$ are in one-one correspondence (up to a nonzero scalar multiple) with m-dimensional subspaces of V. The space of m-dimensional subspaces of V is called the mth **Grassmannian** of V.

Suppose that $E = \{e_1, \ldots, e_n\}$ is a basis of V and $\{v_1, \ldots, v_m\}$ is a basis for the m-dimensional subspace W of V. Since v's are linear combinations of e's, from Theorem 5.1.1(g)

$$v^\wedge = \sum_{\omega \in Q_{m,n}} a_\omega e_\omega^\wedge.$$

So up to a nonzero scalar multiple, those $\binom{n}{m}$ scalars a_ω are uniquely determined by W. In other words, if we choose another basis $\{u_1, \ldots, u_m\}$ of W so that

$$u^\wedge = \sum_{\omega \in Q_{m,n}} b_\omega e^\wedge_\omega,$$

then

$$a_\omega = c b_\omega, \quad \text{for all } \omega \in Q_{m,n}$$

for some constant $c \neq 0$. The $\binom{n}{m}$ scalars a_ω are called the **Plücker coordinates** of $v^\wedge$. The concept was introduced by Julius Plücker in the 19th century for studying geometry. Clearly not all $\binom{n}{m}$ scalars are Plücker coordinates for some $v^\wedge$ since not all vectors in $\wedge^m V$ are decomposable. We are going to give a necessary and sufficient condition for Plücker coordinates.

By see Theorem 5.1.1(c),

$$\sum_{\omega \in Q_{m,n}} a_{\omega\sigma} e^\wedge_\omega = v^\wedge_\sigma = \varepsilon(\sigma) v^\wedge = \sum_{\omega \in Q_{m,n}} \varepsilon(\sigma) a_\omega e^\wedge_\omega$$

so that

$$a_{\omega\sigma} = \varepsilon(\sigma) a_\omega.$$

So it is a necessary condition and motivates the following definition.

Let $p : \Gamma_{m,n} \to \mathbb{C}$ be a function satisfying

$$p(\omega\sigma) = \varepsilon(\sigma) p(\omega), \quad \sigma \in S_m, \omega \in \Gamma_{m,n}. \tag{5.7}$$

For any element $z = \sum_{\omega \in Q_{m,n}} c_\omega e^\wedge_\omega \in \wedge^m V$, the coordinates c_ω can be viewed as a function $c(\omega)$. This function can be extended to $\Gamma_{m,n}$ such that (5.7) remains true. Conversely, any p satisfying (5.7) gives a one-one correspondence between elements of $\wedge^m V$ and $\sum_{\omega \in Q_{m,n}} p(\omega) e^\wedge_\omega$.

For any $\alpha, \beta \in \Gamma_{m,n}$ and p satisfying (5.7), define $P(\alpha, \beta) \in \mathbb{C}_{m\times m}$ such that

$$P(\alpha, \beta)_{i,j} = p(\alpha[i, j : \beta]), \quad 1 \leq i, j \leq m,$$

where $\alpha[i, j : \beta] \in \Gamma_{m,n}$ is obtained from α by replacing $\alpha(i)$ by $\beta(j)$, i.e.,

$$\alpha[i, j : \beta] = (\alpha(1), \ldots, \alpha(i-1), \beta(j), \alpha(i+1), \ldots, \alpha(m)) \in \Gamma_{m,n}.$$

The following result gives the relationship between decomposability and subdeterminants.

Theorem 5.2.1. Let $1 \leq m \leq n$. If p satisfies (5.7) and is not identically zero, then the following are equivalent.

1. $z = \sum_{\omega \in Q_{m,n}} p(\omega) e^\wedge_\omega$ is decomposable.
2. There is $A \in \mathbb{C}_{m\times n}$ such that $p(\omega) = \det A[1, \ldots, m|\omega]$, $\omega \in Q_{m,n}$.
3. $\det P(\alpha, \beta) = p(\alpha)^{m-1} p(\beta)$, $\alpha, \beta \in Q_{m,n}$.

Proof. □

(a) ⇒ (b). Suppose $z = \sum_{\omega\in Q_{m,n}} p(\omega)e^\wedge_\omega$ is decomposable, i.e., $z = v_1 \wedge \cdots \wedge v_m$. Write $v_i = \sum_{j=1}^n a_{ij}e_j$, $i = 1, \ldots, m$, where $E = \{e_1, \ldots, e_n\}$ is a basis of V. By Theorem 5.1.1(g),

$$z = v^\wedge = \sum_{\omega\in Q_{m,n}} \det A[1, \ldots, m|\omega]e^\wedge_\omega$$

Comparing coefficients yields $p(\omega) = \det A[1, \ldots, m|\omega]$, $\omega \in Q_{m,n}$.

(b) ⇒ (a). Suppose that (b) is true. Use A to define $v_i = \sum_{j=1}^n a_{ij}e_j$, $i = 1, \ldots, m$. So

$$z = \sum_{\omega\in Q_{m,n}} p(\omega)e^\wedge_\omega = \sum_{\omega\in Q_{m,n}} \det A[1, \ldots, m|\omega]e^\wedge_\omega = v_1 \wedge \cdots \wedge v_m$$

is decomposable.

(b) ⇒ (c). For any $A \in \mathbb{C}_{m\times n}$ and $\alpha, \beta \in Q_{m,n}$, let

$$S := A[1, \ldots, m|\alpha] = (A_{\alpha(1)} \ \ldots \ A_{\alpha(m)}) \in \mathbb{C}_{m\times m}$$

where $A_{\alpha(i)}$ denotes the ith column of S, $i = 1, \ldots, m$. Similarly let

$$T := A[1, \ldots, m|\beta] = (A_{\beta(1)} \ \ldots \ A_{\beta(m)}) \in \mathbb{C}_{m\times m}.$$

If we define $X \in \mathbb{C}_{m\times m}$ by

$$x_{ij} = \det(A_{\alpha(1)} \ \ldots \ A_{\alpha(i-1)} \ A_{\beta(j)} \ A_{\alpha(i+1)} \ \ldots \ A_{\alpha(m)}), \quad 1 \le i, j \le m,$$

then by Cramer rule

$$SX = (\det S)T.$$

Taking determinants on both sides

$$\det S \det X = \det T(\det S)^m.$$

If S is invertible, then

$$\det X = \det T(\det S)^{m-1}. \tag{5.8}$$

If S is not invertible, then replace S by $S + tI_m$. Since $\det B$ is a polynomial function of its entries b_{ij}, we have (5.8) by continuity argument.

Suppose that (b) is true. Then

$$\begin{aligned} \det S &= \det A[1, \ldots, m|\alpha] = p(\alpha), \\ \det T &= \det A[1, \ldots, m|\beta] = p(\beta) \end{aligned}$$

and

$$\begin{aligned} x_{ij} &= \det(A_{\alpha(1)} \ \ldots \ A_{\alpha(i-1)} \ A_{\beta(j)} \ A_{\alpha(i+1)} \ \ldots \ A_{\alpha(m)}) \\ &= \det A[1, \ldots, m|\alpha(1), \ldots, \alpha(i-1), \beta(j), \alpha(i+1), \ldots, \alpha(m)] \\ &= p(\alpha(1), \ldots, \alpha(i-1), \beta(j), \alpha(i+1), \ldots, \alpha(m)) \\ &= p(\alpha[i, j : \beta]). \end{aligned}$$

Then $X = (x_{ij}) = P(\alpha, \beta)$ and (5.8) becomes

$$\det P(\alpha, \beta) = p(\alpha)^{m-1} p(\beta).$$

Suppose (c) is true. Since p is not a zero polynomial, there is some $\alpha \in Q_{m,n}$ such that $p(\alpha) \neq 0$. Define $A \in \mathbb{C}_{m\times n}$ where

$$a_{ij} = p(\alpha)^{\frac{1}{m}-1} p(\alpha(1), \ldots, \alpha(i-1), \beta(j), \alpha(i+1), \ldots, \alpha(m)), \quad 1 \le i \le m, 1 \le j \le n.$$

Then for any $\omega \in Q_{m,n}$,

$$\begin{aligned}
& \det A[1, \ldots, m|\omega] \\
= \; & \det(a_{i\omega(j)}) \\
= \; & \det(p(\alpha)^{\frac{1}{m}-1} p(\alpha(1), \ldots, \alpha(i-1), \omega(j), \alpha(i+1), \ldots, \alpha(m)) \\
= \; & ((p(\alpha)^{\frac{1}{m}-1})^m \det(p(\alpha[i, j : \omega])_{1\le i,j\le m} \\
= \; & p(\alpha)^{1-m} \det P(\alpha, \omega) \\
= \; & (p(\alpha)^{1-m} p(\alpha)^{m-1} p(\omega) \quad \text{by } (c) \\
= \; & p(\omega).
\end{aligned}$$

There are many equivalent conditions for decomposability, for example

$$P(\alpha, \beta)P(\beta, \gamma) = p(\beta)P(\alpha, \gamma), \quad \alpha, \beta, \gamma \in \Gamma_{m,n}$$

is also necessary and sufficient condition (but we will not discuss it in detail).

We now use Theorem 5.2.1 to deduce some results.

Let $\alpha \in D_{m,n}$. Denote by $\operatorname{Im}\alpha$ denotes the set of the components of α. We say that $\alpha, \beta \in D_{m,n}$ differ by k components if $|\operatorname{Im}\alpha \cap \operatorname{Im}\beta| = m - k$. For example $\alpha = (1, 3, 5)$ and $\beta = (3, 5, 6)$ differ by one component. We say that $\alpha_1, \ldots, \alpha_k \in Q_{m,n}$ form a **chain** if all consecutive sequences differ by one component. The following is an important necessary condition for decomposability.

Theorem 5.2.2. Let $z = \sum_{\omega \in Q_{m,n}} p(\omega) e^\wedge_\omega \in \wedge^k V$ (with $\dim V = n$) be decomposable. If $p(\alpha) \neq 0$, $p(\beta) \neq 0$ and α and β differ by k components, then there are $k - 1$ different sequences $\alpha_1, \ldots, \alpha_{k-1} \in Q_{m,n}$ such that $p(\alpha_i) \neq 0$, $i = 1, \ldots, k - 1$, and $\alpha, \alpha_1, \ldots, \alpha_{k-1}, \beta$ form a chain. In other words, any two nonzero coordinates of a decomposable $v^\wedge$ are connected by a nonzero chain.

Proof. When $k = 1$, the result is trivial. Suppose $k > 1$. Use (5.7) to extend the domain of p to $\Gamma_{m,n}$. By Theorem 5.2.1

$$\det P(\alpha, \beta) = p(\alpha)^{m-1} p(\beta) \neq 0$$

so one term in the expansion of the left side must be nonzero, i.e., there is $\sigma \in S_m$ such that

$$\prod_{i=1}^m p(\alpha[i, \sigma(i) : \beta]) \neq 0.$$

Since α and β differ by $k > 1$ components, there is $i_0 (1 \le i_0 \le m)$ such that $\alpha(i_0) \notin \operatorname{Im}\beta$. From $p(\alpha[i_0, \sigma(i_0) : \beta]) \neq 0$ we know that $\beta(\sigma(i_0)) \notin \operatorname{Im}\alpha$ and $\alpha[i_0, \sigma(i_0) : \beta] \in D_{m,n}$. So there is $\theta \in S_m$ such that $\alpha[i_0, \sigma(i_0) : \beta]\theta = \alpha_1 \in Q_{m,n}$. By (5.7),

$$p(\alpha_1) = \varepsilon(\theta)p(\alpha[i_0, \sigma(i_0) : \beta]) \neq 0.$$

Clearly α and α_1 differ by one component and α_1 and β differ by $k-1$ components. Repeat this process for α_1 and β to have $\alpha_2, \ldots, \alpha_{k-1}$. □

Example: Theorem 5.2.2 can be used to determine indecomposable elements, i.e., if $z = \sum_{\omega \in Q_{m,n}} p(\omega)e^{\wedge}_{\omega}$ has two nonzero unchained coordinates, then z is indecomposable. For example $z = e_1 \wedge e_2 + e_3 \wedge e_4$ is indecomposable since $p(1, 2)$ and $p(3, 4)$ are nonzero but not chained.

Theorem 5.2.3. Let $\dim V = n$ and $1 \le m \le n$. Then all elements in $\wedge^m V$ is decomposable if and only if $m = 1$, $m = n - 1$ or $m = n$.

Proof. $\Leftarrow$ Clearly if $m = 1$ or $m = n$, all elements of $\wedge^m V$ are decomposable. So we only need to consider $m = n - 1$. Suppose $z = \sum_{\omega \in Q_{m,n}} e^{\wedge}_{\omega} \in \wedge^m V$. By (5.7) extend the domain of p to $\Gamma_{m,n}$ and we are going to show that p satisfies Theorem 5.2.1(c).

For any $\alpha, \beta \in Q_{m,n}$, if $\alpha = \beta$, then $P(\alpha, \alpha) = p(\alpha)I_m$. Hence (c) is valid. If $\alpha \neq \beta$, then from $m = n - 1$, α and β differ by one component. So α has a unique component $\alpha(t) \notin \operatorname{Im}\beta$ and there is $\sigma \in S_m$ such that $\alpha(j) = \beta\sigma(j)$ unless $j = t$. So

$$\begin{aligned} p([t, t : \beta\sigma]) &= p(\alpha(1), \ldots, \alpha(t-1), \beta\sigma(t), \alpha(t+1), \ldots, \alpha(m)) \\ &= p(\beta\alpha(1), \ldots, \beta\alpha(t-1), \beta\sigma(t), \beta\alpha(t+1), \ldots, \beta\alpha(m)) \\ &= p(\beta\sigma). \end{aligned}$$

When $j \neq t$,

$$p(\alpha[i, j : \beta\sigma]) = p(\alpha[i, j : \alpha]) = p(\alpha)\delta_{ij}.$$

So the $(m-1)\times(m-1)$ principal submatrix $P(\alpha, \beta\sigma)(t|t)$ of $P(\alpha, \beta\sigma)$ is $p(\alpha)I_{m-1}$ and the tth diagonal entry of $P(\alpha, \beta\sigma)$ is $p(\beta\sigma)$. So

$$\det P(\alpha, \beta\sigma) = p(\alpha)^{m-1}p(\beta\sigma) = \varepsilon(\sigma)p(\alpha)^{m-1}p(\beta).$$

By the property of determinant,

$$\det P(\alpha, \beta\sigma) = \det p(\alpha[i, \sigma(j) : \beta]) = \varepsilon(\sigma) \det p(\alpha[i, j : \beta]) = \varepsilon(\sigma) \det P(\alpha, \beta).$$

Hence $\det P(\alpha, \beta) = p(\alpha)^{m-1}p(\beta)$, i.e., Theorem 5.2.1(c) is satisfied. Hence all elements of $\wedge^{n-1}V$ are decomposable.

Conversely if $2 \le m \le n - 2$, let

$$\alpha = (1, \ldots, m), \quad \beta = (1, \ldots, m-2, m+1, m+2).$$

Then consider $z = e^{\wedge}_{\alpha} + e^{\wedge}_{\beta}$. Notice that α and β differ by two components and are linked by a nonzero chain. By Theorem 5.2.2, z is indecomposable. So not all elements of $\wedge^m V$ are decomposable. □

Theorem 5.2.4. Let $v_1 \wedge \cdots \wedge v_m = \sum_{\omega \in Q_{m,n}} p(\omega) e^\wedge_\omega \neq 0$. If $p(\alpha) \neq 0$, let $u_i = \sum_{j=1}^n p(\alpha[i,j]) e_j$, $i = 1, \dots, m$. Then $u_1 \wedge \cdots \wedge u_m = p(\alpha)^{m-1} v_1 \wedge \cdots \wedge v_m$. Thus $\langle u_1, \dots, v_m \rangle = \langle v_1, \dots, v_m \rangle$. Here

$$\alpha[i,j] := (\alpha(1), \dots, \alpha(i-1), j, \alpha(i+1), \dots, \alpha(m)) \in \Gamma_{m,n}.$$

Proof. Let $a_{ij} := p(\alpha[i,j]) = p(\alpha(1), \dots, \alpha(i-1), j, \alpha(i+1), \dots, \alpha(m))$. By Theorem 5.1.1,

$$u^\wedge = \sum_{\omega \in Q_{m,n}} \det A[1, \dots, m|\omega] e^\wedge_\omega.$$

By Theorem 5.2.1

$$\det A[1, \dots, m|\omega] = \det(a_{i\omega(j)} = \det(p(\alpha[i, j : \omega]) = p(\alpha)^{m-1} p(\omega).$$

So $u^\wedge = p(\alpha)^{m-1} v^\wedge$. Then by $p(\alpha) \neq 0$, $v^\wedge \neq 0$ to have $\langle u_1, \dots, , u_m \rangle = \langle v_1, \dots, v_m \rangle$. □

By Theorem 5.2.4, it is easy to find a basis of the subspace that corresponds to a given decomposable element. For example, from Theorem 5.2.3,

$$z = e_1 \wedge e_2 \wedge e_3 + 2e_1 \wedge e_2 \wedge e_4 - e_1 \wedge e_3 \wedge e_4 - 3e_2 \wedge e_3 \wedge e_4 \in \wedge^3 V$$

is decomposable with $\dim V = 4$. It is not hard to find $u_1, u_2, u_3 \in V$ such that $z = u_1 \wedge u_2 \wedge u_3$: by Theorem 5.2.4, let $\alpha = (1, 2, 3)$ so that $p(\alpha) = 1$. Then

$$u_1 = p(1, 2, 3)e_1 + p(2, 2, 3)e_2 + p(3, 2, 3)e_3 + p(4, 2, 3)e_4 = e_1 - 3e_4$$

because $p(2, 2, 3) = p(3, 2, 3) = 0$ and $p(4, 2, 3) = p(2, 3, 4)$ by (5.7). Similarly $u_2 = e_2 + e_4$, $u_3 = e_3 + 2e_4$. So

$$z = (e_1 - 3e_4) \wedge (e_2 + e_4) \wedge (e_3 + 2e_4) = -3.$$

Problems

1. Let $v_1 \wedge \cdots \wedge v_m = \sum_{\omega \in W_{m,n}} p(\omega) e^\wedge_\omega \neq 0$. Prove that for any $\alpha \in \Gamma_{m,n}$,

$$u_i = \sum_{j=1}^n p(\alpha[i,j]) e_j \in \langle v_1, \dots, v_m \rangle, \quad i = 1, \dots, m.$$

2. It is known that $z = 3e_1 \wedge e_2 + 5e_1 \wedge e_3 + 5e_1 \wedge e_4 + e_2 \wedge e_3 - 2e_2 \wedge e_4 + 5e_3 \wedge e_4$ is decomposable, find $u_1, u_2 \in V$ such that $z = u_1 \wedge u_2$.

3. Let $p : \Gamma_{m,n} \to \mathbb{C}$ not identically zero. Suppose that p satisfies (5.7). Prove that the following are equivalent.

   (a) $P(\alpha, \beta) P(\beta, \gamma) = p(\beta) P(\alpha, \gamma)$, $\alpha, \beta, \gamma \in \Gamma_{m,n}$.

(b) $P(\alpha,\beta)P(\beta,\alpha) = p(\alpha)p(\beta)I_m$, $\alpha, \beta \in Gmn$.

(c) $\sum_{t=1}^{m} p(\alpha[s,t:\beta])p(\beta[t,s:\alpha]) = p(\alpha)p(\beta)$, $\alpha,\beta \in \Gamma_{m,n}$, $s = 1,\dots,m$.

(d) $\sum_{k=1}^{m+1}(-1)^{k-1}p(\alpha(\hat{i}) : \delta(k))p(\delta(k)) = 0$, $\alpha \in \Gamma_{m,n}$, $\delta \in \Gamma_{m+1,n}$, where

$$\begin{aligned} \alpha(\hat{i}) : \delta(k) &= (\alpha(1),\dots,\alpha(i-1),\alpha(i+1),\dots,\alpha(m),\delta(k)) \in \Gamma_{m,n} \\ \delta(\hat{k}) &= (\delta(1),\dots,\delta(k-1),\delta(k+1),\dots,\delta(m+1)) \in \Gamma_{m,n}. \end{aligned}$$

5.3 Completely symmetric spaces

We denote $V_\chi(G)$ by $V_1(S_m) = \bullet^m V$ when $G = S_m$ and $\chi \equiv 1$. The space $\bullet^m V$ is called a **completely symmetric spaces**. The decomposable element v^* is written as $v^\bullet = v_1 \bullet \cdots \bullet v_m$.

Let $A \in \mathbb{C}_{m\times m}$, the permanent

$$\operatorname{per} A = \sum_{\sigma\in S_m}\sum_{t=1}^{m} a_t\sigma(t)$$

plays an important role for the completely symmetric space as the determinant to the exterior space.

Theorem 5.3.1. Let $\dim V = n \geq m$. Denote by $\nu(\alpha) = |G_\alpha|$ $(G = S_m)$. For $\bullet^m V$,

(a) $\Delta = \bar{\Delta} = \hat{\Delta} = G_{m,n}$ and $\dim \bullet^m V = \binom{n+m-1}{m}$.

(b) If $E = \{e_1,\dots,e_n\}$ is a basis of V, then $E_\bullet := \{e^\bullet_\alpha : \alpha \in G_{m,n}\}$ is a basis of $\bullet^m V$. If E is an orthonormal basis of V, then $E'_\bullet := \{\sqrt{\frac{m!}{\nu(\alpha)}}e^\bullet_\alpha : \alpha \in G_{m,n}\}$ is an orthonormal basis of $\bullet^m V$.

(c) For each $\sigma \in S_m$, $v^\bullet_\sigma = v^\bullet$.

(d) $v_1 \bullet \cdots \bullet v_m = 0$ if and only if some $v_i = 0$.

(e) If $u^\bullet = v^\bullet \neq 0$, then there is $\sigma \in S_m$ and $d_i \neq 0$ such that $u_i = d_i v_{\sigma(i)}$, $i = 1,\dots,m$ and $\prod_{i=1}^{m} d_i = 1$.

(f) If $v_i = \sum_{j=1}^{n} a_{ij}e_j$, $i = 1,\dots,m$, then

$$v^\bullet = \sum_{\alpha\in G_{m,n}} \frac{1}{\nu(\alpha)}\operatorname{per} A[1,\dots,m|\alpha]e^\bullet_\alpha.$$

Proof. (a), (b), (c) are clear from Chapter 4.

(d) It suffices to prove the necessary part and we use contrapositive. Equip V with an inner product. Suppose $v_t \neq 0$ for all $t = 1,\dots,m$. The union

$\cup_{i=1}^m v_i^\perp$ of the finite hyperplanes $v_i^\perp$ is not dense in V. So there is $w \in W$ such that

$$\prod_{t=1}^m (w, v_t) \neq 0.$$

On the other hand, for any $w \in V$

$$(w\otimes\cdots\otimes w, v^\bullet) = (T(G,1)w\otimes\cdots\otimes w, v^\otimes) = (w\otimes\cdots\otimes w, v_1\otimes\cdots\otimes v_m) = \prod_{t=1}^m (w, v_t).$$

So $v^\bullet = 0$ would imply $\prod_{t=1}^m (w, v_t) = 0$ for all $w \in V$. Thus $v^\bullet \neq 0$.

(e) If $u^\bullet = v^\bullet \neq 0$, then for any $w \in V$,

$$\prod_{t=1}^m (w, u_t) = (w \otimes \cdots \otimes w, u^\bullet) = (w \otimes \cdots \otimes w, v^\bullet) = \prod_{t=1}^m (w, v_t). \tag{5.9}$$

Decompose $u_t = x_t + y_t$ such that $x_t \in \langle v_1 \rangle$, $y_t \in v_1^\perp$, $t = 1, \ldots, m$. For any $z \in v_1^\perp$, we have $\prod_{t=1}^m (z, v_t) = 0$. By (5.9)

$$0 = \prod_{t=1}^m (z, v_t) = \prod_{t=1}^m (z, u_t) = \prod_{t=1}^m (z, y_t), \quad \text{for all } z \in v_1^\perp.$$

Applying the claim in the proof of (d) on the space $v_1^\perp$ to have $y_i = 0$ for some i, i.e., $u_i = x_i \in \langle v_1 \rangle$. So there is $d_i \neq 0$ such that $u_i = d_i v_1$. Substitute it into (5.9)

$$0 = (w, v_1)(\prod_{i=2}^m (w, v_t) - d_i \prod_{t\neq i} (w, u_t)). \tag{5.10}$$

Since $v_1 \neq 0$ and $w \in V$ is arbitrary, we have

$$\prod_{i=2}^m (w, v_t) = d_i \prod_{t\neq i} (w, u_t), \quad \text{for all } w \in v_1^\perp$$

Then apply continuity argument, so that the above equality is valid for $w \in V$. Then use induction.

(f)

$$\begin{aligned} v^\bullet &= (\sum_{j=1}^n a_{1j} e_j) \cdots (\sum_{j=1}^n a_{mj} e_j) \\ &= \sum_{\alpha \in \Gamma_{m,n}} \prod_{i=1}^m a_{i\alpha(i)} e_\alpha^\bullet \\ &= \sum_{\alpha \in G_{m,n}} \frac{1}{\nu(\alpha)} \sum_{\sigma \in S_m} \prod_{i=1}^m a_{i\alpha\sigma(i)} e_{\alpha\sigma}^\bullet \quad \text{(Theorem 4.4.3)} \\ &= \sum_{\alpha \in G_{m,n}} \frac{1}{\nu(\alpha)} \operatorname{per} A[1, \ldots, m|\alpha] e_\alpha^\bullet. \end{aligned}$$

□

We denoted by $P_m(T)$ for $K(T)$ if $T \in \operatorname{End} V$ and $P_m(A)$ for $K(A)$ if $A \in \mathbb{C}_{n\times n}$.

Theorem 5.3.2. Let $T \in \operatorname{End} V$ with $\dim V = n$. Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues and $s_1, \ldots, s_n$ be the singular values of A. Let $A = [T]_E^E \in \mathbb{C}_{n\times n}$ where E is a basis of V. Then

(a) Let $E'_\bullet$ be the basis as in Theorem 5.3.1. Then

$$\left([P_m(T)]_{E'_\bullet}^{E'_\bullet}\right)_{\alpha,\beta} = P_m(A)_{\alpha,\beta} = \frac{\operatorname{per} A[\alpha|\beta]}{\sqrt{\nu(\alpha)\nu(\beta)}}.$$

(b) $P_m(A^*) = P_m(A)^*$, $P_m(A^{-1}) = P_m(A)^{-1}$ if A is invertible, $P_m(AB) = P_m(A)P_m(B)$.

(c) If $\operatorname{rank} A = r$, then $\operatorname{rank} P_m(A) = \binom{r+m-1}{m}$.

(d) If A is upper triangular, so is $P_m(A)$.

(e) $\prod_{t=1}^m \lambda_{\alpha(t)}$ are the eigenvalues and $\prod_{t=1}^m s_{\alpha(t)}$ are the singular values, $\alpha \in G_{m,n}$, of $P_m(A)$

(f) $\operatorname{tr} P_m(A) = \sum_{\alpha\in G_{m,n}} \prod_{t=1}^m \lambda_{\sigma(t)} = \sum_{\alpha\in G_{m,n}} \frac{1}{\nu(\alpha)} \operatorname{per} A[\alpha|\alpha]$.

(g) $\det P_m(A) = (\det A)^{\binom{n+m-1}{n}}$.

(h) If A is normal, Hermitian, pd, psd or unitary, so is $P_m(A)$.

Problems

1. Prove the Cauchy-Binet formula for the permanent: Let $A, B \in \mathbb{C}_{n\times n}$. For any $\alpha, \beta \in G_{m,n}$,

$$\operatorname{per}((AB)[\alpha|\beta]) = \sum_{\omega\in G_{m,n}} \frac{1}{\nu(\omega)} \operatorname{per}(A[\alpha|\omega]) \operatorname{per}(B[\omega|\beta]).$$

2. Prove Laplace expansion for permanent: for any $A \in \mathbb{C}_{n\times n}$, $\alpha \in Q_{m,n}$,

$$\operatorname{per} A = \sum_{\beta\in Q_{m,n}} \operatorname{per} A[\alpha|\beta] \operatorname{per} A(\alpha|\beta).$$

3. Let $G < S_m$. Check which properties of Theorem 5.3.1 remain true for $V_1(G)$.

4. Prove that the set $\{v \otimes \cdots \otimes v : v \in V\}$ generates $V^{\bullet m} = V_1(S_m)$.

5. Let $A(t) \in \mathbb{C}_{n\times n}$ be a differentiable matrix function. Prove that

$$\frac{d}{dt}\operatorname{per} A(t) = \sum_{i,j=1}^{n} \operatorname{per} A(i|j)\frac{da_{ij}(t)}{dt}.$$

Chapter 6

Research topics

6.1 Orthonormal decomposable basis

The material in this section is from Holmes [3].

Let V be a finite-dimensional complex inner product space and assume $n := \dim V \geq 2$ (to avoid trivialities). Let $G < S_m$. Fix an orthonormal basis $\{e_1, \ldots, e_n\}$ of V. Given $\gamma \in \Gamma_{m,n}$ and $\chi \in I(G)$, the inner product on V induces an inner product on $\otimes^m V$. If the subspace W of $\otimes^m V$ has a basis consisting of mutually orthogonal standard symmetrized tensors, we will say that W has an *o-basis*.

Given $\gamma \in \Gamma$, set $G_\gamma := \{\sigma \in G \,|\, \gamma\sigma = \gamma\} \leq G$. We have

$$(e^\chi_{\gamma\mu}, e^\chi_{\gamma\tau}) = \frac{\chi(1)}{|G|} \sum_{\sigma \in G_{\gamma\tau}} \chi(\sigma\tau^{-1}\mu) = \frac{\chi(1)}{|G|} \sum_{\sigma \in G_\gamma} \chi(\sigma\mu\tau^{-1}),$$

the first equality from [1, p. 339] and the second from the observations that $\tau G_{\gamma\tau}\tau^{-1} = G_\gamma$ and $\chi(\sigma\tau^{-1}\mu) = \chi(\tau\sigma\tau^{-1}\mu\tau^{-1})$.

We have $\dim V^\chi_\gamma = \chi(1)(\chi, 1)_{G_\gamma}$.

The group G is 2-**transitive** if, with respect to the componentwise action, it is transitive on the set of pairs (i, j), with $i, j = 1, \ldots, m$, $i \neq j$. Note that if G is 2-transitive, then for any $i = 1, \ldots, m$, the subgroup $\{\sigma \in G \,|\, \sigma(i) = i\}$ of G is transitive on the set $\{1, \ldots, \hat{i}, \ldots, m\}$.

Theorem 6.1.1. Assume $m \geq 3$. If G is 2-transitive, then $\otimes^m V$ does not have an o-basis.

Proof. By the remarks above, it is enough to show that V^χ_γ does not have an o-basis of some $\chi \in I(G)$, $\gamma \in \Gamma$.

Let $H = \{\sigma \in G \,|\, \sigma(m) = m\} < G$ and denote by ψ the induced character $(1_H)^G$, so that $\psi(\sigma) = |\{i \,|\, \sigma(i) = i\}|$ for $\sigma \in G$ (see [4, p. 68]).

Let $\rho \in G - H$ and for $i = 1, \ldots, m$, set $R_i = \{\sigma \in H \,|\, \sigma\rho(i) = i\}$. Clearly, $R_i = \emptyset$ if $i \in \{m, \rho^{-1}(m)\}$. Assume $i \notin \{m, \rho^{-1}(m)\}$. Since H acts transitively on $1, \ldots, m-1$, there exists some $\tau \in R_i$. Then $R_i = H_i\tau$, where $H_i := \{\sigma \in H \mid \sigma(i) = i\}$. Now $[H : H_i]$ equals the number of elements in the orbit of i under the action of H, so $[H : H_i] = m-1$. Therefore, $|R_i| = |H_i\tau| = |H_i| = |H|/[H : H_i] = |H|/(m-1)$. We obtain the formula

$$\sum_{\sigma \in H} \psi(\sigma\rho) = \sum_{i=1}^{n} |R_i| = \sum_{i \neq m, \rho^{-1}(m)} |R_i| = \frac{m-2}{m-1}|H|.$$

Since $(\psi, 1)_G = (1, 1)_H = 1$ by Frobenius reciprocity, 1 is a constituent of ψ, whence $\chi := \psi - 1$ is a character of G. Moreover, the 2-transitivity of G implies that $(\psi, \psi)_G = 2$ (see [4, p. 68]). Hence, $(\chi, \chi)_G = 1$, so that χ is irreducible.

Let $\gamma = (1, \ldots, 1, 2) \in \Gamma$ and note that $G_\gamma = H$. Let μ and τ be representatives of distinct right cosets of H in G. Then $\rho := \mu\tau^{-1} \in G - H$, so the discussion above shows that

$$(e^\chi_{\gamma\mu}, e^\chi_{\gamma\tau}) = \frac{\chi(1)}{|G|} \sum_{\sigma \in H} \chi(\sigma\mu\tau^{-1}) = \frac{\chi(1)}{|G|} \left[\frac{n-2}{m-1}|H| - |H| \right] < 0.$$

It follows that distinct standard symmetrized tensors in V^χ_γ are not orthogonal.

On the other hand,

$$\dim V^\chi_\gamma = \chi(1)(\chi, 1)_H = (m-1)[(\psi, 1)_H - 1],$$

and since $(\psi, 1)_H = (\psi, \psi)_G = 2$ by Frobenius reciprocity, $\dim V^\chi_\gamma = n - 1 > 1$.

Therefore, V^χ_γ does not have an o-basis. This completes the proof. □

Corollary 6.1.2. If $G = S_m$ $(m \geq 3)$ or $G = A_m$ $(m \geq 4)$, then $\otimes^m V$ does not have an o-basis.

2-transitive groups have been studied extensively (see [3, Chapter XII], for example).

Problems

1. Show that if G is abelian, then $\otimes^m V$ has an o-basis.
2. Give a proof of Corollary 6.1.2 by showing that (i) S_m is 2-transitive for all m and (ii) the alternating group A_m is 2-transitive if $m \geq 4$.
3. Show that if G is the dihedral group $D_m \leq S_n$, then $\otimes^m V$ has an o-basis if and only if m is a power of 2.

Bibliography

[1] A.C. Aitken, The normal form of compound and induced matrices, Proc. London Math. Soc. **38** (1934), 354–376.

[2] M. Fiedler, Bounds for the determinant of the sum of hermitian matrices, Proc. Amer. Math. Soc., **30** (1971), 27–31.

[3] R.R. Holmes, Orthogonal bases of symmetrized tensor spaces, Linear and Multilinear Algebra **39** (1995), 241–243.

[4] R.R. Holmes and T.Y. Tam, Symmetry classes of tensors associated with certain groups, Linear and Multilinear Algebra **32** (1992), 21–31.

[5] R.R. Holmes, C.K. Li, T.Y. Tam, Numerical range of the derivation of an induced operator. J. Aust. Math. Soc. **82** (2007), 325-3-43.

[6] C. Gamas, Conditions for a symmetrized decomposable tensor to be zero, Linear Algebra Appl., **108** (1988), 83-119.

[7] A. Horn, On the eigenvalues of a matrix with prescribed singular values, Proc. Amer. Math. Soc., **5** (1954) 4–7.

[8] I.M. Isaacs, *Character Theory of Finite Groups*, Academic Press, New York, 1976.

[9] G. James and A. Kerber, *The Representation Theory of the Symmetric Group*, Encyclopedia of Mathematics and Its Applications, Volume 16, Addison-Wesley Publishing Company, 1981.

[10] G. James and M. Liebeck, *Representations and Characters of Groups*, Cambridge Mathematical Textbooks, Cambridge University Press, 1993.

[11] B. Kostant, On convexity, the Weyl group and Iwasawa decomposition, Ann. Sci. Ecole Norm. Sup. (4), **6** (1973) 413–460.

[12] C.K. Li, and T.Y. Tam, Operator properties of T and $K(T)$, Linear Algebra Appl. **401** (2005), 173–191.

[13] D.E. Littlewood, On induced and compound matrices, Proc. London Math. Soc. **40** (1936), 370–380.

[14] M. Marcus, *Finite Dimensional Multilinear Algebra, Part I*, Marcel Dekker, New York, 1973.

[15] M. Marcus, *Finite Dimensional Multilinear Algebra, Part II*, Marcel Dekker, New York, 1975.

[16] R. Merris, Nonzero decomposable symmetrized tensors, Linear and Multilinear Algebra **5** (1977), 287–292.

[17] R. Merris, On vanishing decomposable symmetrized tensors, Linear and Multilinear Algebra, **5** (1977/78), 79-86.

[18] R. Merris, The structure of higher degree symmetry classes of tensors II, Linear and Multilinear Algebra, **6** (1978), 171–178.

[19] R. Merris, Recent advances in symmetry classes of tensors, Linear and Multilinear Algebra , **7** (1979) 317-328.

[20] R. Merris, *Multilinear Algebra*, Gordon and Breach Science Publishers, Amsterdam, 1997.

[21] U. Prells, M.I. Friswell, S.D. Garvey, Use of geometric algebra: Compound matrices and the determinant of the sum of two matrices, Proc. R. Soc. Lond. A, **459** (2030), 273–285.

[22] J.A. Dias da Silva, Conditions for equality of decomposable symmetric tensors, Linear Algebra and its Appl., **24** (1979), 85–92.

[23] J.A. Dias da Silva, Colorings and equality of tensors, Linear Algebra Appl., **342** (2002) 79–91.

[24] J.A. Dias da Silva, J. A. and M.M. Torres, On the orthogonal dimension of orbital sets, Linear Algebra Appl., **401** (2005), 77–107.

[25] , J.A. Dias da Silva and H.F. da Cruz, Equality of immanantal decomposable tensors, Linear Algebra Appl., **401** (2005), 29–46.

[26] J.P. Serre, *Linear Representations of Finite Groups*, Springer, New York, 1977.

[27] T.Y. Tam, A. Horn's result on matrices with prescribed singular values and eigenvalues, Electronic Linear Algebra, **21** (2010) 25–27.

[28] T.Y. Tam and M.C. Thompson, Determinant and Pfaffian of sum of skew symmetric matrices, Linear Algebra Appl., **433** (2010) 412–423.

[29] T.Y. Tam and M.C. Thompson, Determinants of sum of orbits under compact Lie group, to appear in Linear Algebra Appl.

[30] T.Y. Tam and M.C. Thompson, Determinants of sums of two real matrices and their extensions, submitted for publication.

[31] J.H. van Lint, Notes on Egoritsjev's proof of the van der Waerden conjecture. Linear Algebra Appl., **39** (1981), 1–8.

[32] J.H. van Lint, The van der Waerden conjecture: two proofs in one year. Math. Intelligencer **4** (1982), 72–77.

[33] B.Y. Wang, *Foundation of Multilinear Algebra*, Beijing Normal University Press (in Chinese), 1985.

[34] B.Y. Wang, Wang and M.P. Gong, A high symmetry class of tensors with an orthogonal basis of decomposable symmetrized tensors, Linear and Multilinear Algebra, **30** (1991), 61–64.

[35] B.Y. Wang, Wang and M.P. Gong, The subspaces and orthonormal bases of symmetry classess of tensor, Linear and Multilinear Algebra, **30** (1991), 195–204.

[36] H. Weyl, Inequalities between the two kinds of eigenvalues of a linear transformation, Proc. Nat. Acad. Sci. U. S. A., **35** (1949) 408–411.