Multilinear Algebra

Anson Li

Contents

Abstract

1. Introduction

We assume the reader is familiar with the notion of a vector space, linear mappings in between them, nonetheless, this section serves to condense a handful of elementary results from linear algebra [**Roman2007Advanced**; Gre12; Gre67; Lan05; Hal12]. We also restrict ourselves to the fields $\mathbb{R}$ or $\mathbb{C}$.

1.1. Notation

Sets $\mathbb{N} = \{0, 1, 2, \ldots\}$ is the set of *natural numbers* (including 0),

$\mathbb{N}^+ = \{1, 2, \ldots\}$ is the set of *counting numbers*,

$\mathbb{Z} = \{0, -1, +1, \ldots\}$ is the set of *integers*,

$\mathbb{R}$ is the set *real numbers*,

$\mathbb{C}$ is the set of *complex numbers*.

Indexing Sets
Let $\{J_i\}_{i=1}^N \subseteq \mathbb{N}^+$ be an ordered list of counting numbers, for each $i = 1, \ldots, N$, we write

$$\mathbb{N}^+(J_i) = \{1, \ldots, J_i\}.$$

It is useful to also define

$$\otimes_1^N \mathbb{N}^+(J_i) = \left\{(m_1, \ldots, m_n) \in (\mathbb{N}^+)^N, \begin{array}{l} 1 \le m_i \le J_i, \\ i = 1, \ldots, N \end{array}\right\}. \tag{1}$$

Shown below is the disjoint union of the initial segments $\mathbb{N}^+(J_i)$:

$$\oplus_1^N \mathbb{N}^+(J_i) = \left\{(m_1, \ldots, m_N) \in (\mathbb{N})^N, \begin{array}{l} \exists j \in \mathbb{N}^+(N), \\ m_j \in \mathbb{N}^+(J_j), \\ \forall k \neq j,\ m_k = 0 \end{array}\right\}. \tag{2}$$

The symbol $\mathbf{1}_N$ will we used to denote the vector with all ones, that is

$$\mathbf{1}_N = (1, \ldots, 1) \in (\mathbb{N}^+)^N.$$

Vector Spaces

Let $\{\mathbb{R}(J_i)\}_{i=1}^N$ and $\{\mathbb{R}(I_i)\}_{i=1}^N$ be finite dimensional vector spaces over $\mathbb{R}$. For every $i = 1, \ldots, N$, the dimension of $\mathbb{R}(J_i)$ is the number in the brackets, J_i. The basis elements of $\mathbb{R}(J_i)$ are indexed by the symbol $\alpha_i = 1, \ldots, J_i$, and

$$\left\{e^{\alpha_i}_{\mathbb{R}(J_i)},\ 1 \le \alpha_i \le J_i\right\}$$

forms a basis of $\mathbb{R}(J_i)$. Every *vector* $x \in \mathbb{R}(J_i)$ admits a representation with respect to this basis with x_{α_i} ($1 \le \alpha_i \le J_i$) being the α_ith coordinate of x so that

$$x = \sum_{1 \le \alpha_i \le J_i} x_{\alpha_i} e^{\alpha_i}_{\mathbb{R}(J_i)}.$$

The dual space of $\mathbb{R}(J_i)$ is denoted by $\mathbb{R}(J_i)^*$, and the dual basis associated with $\{e^{\alpha_i}_{\mathbb{R}(J_i)}\}_{\alpha_i=1}^{J_i}$ is

$$\left\{e^{1,*}_{\mathbb{R}(J_i)}, \ldots, e^{J_i,*}_{\mathbb{R}(J_i)}\right\} = \left\{e^{\alpha_i,*}_{\mathbb{R}(J_i)},\ 1 \le \alpha_i \le J_i\right\}.$$

The scalar product between $\mathbb{R}(J_i)$ and $\mathbb{R}(J_i)^*$ is always denoted by $\langle \cdot, \cdot \rangle_{\mathbb{R}(J_i)^*,\,\mathbb{R}(J_i)}$, where

$$\langle e^{\gamma_i,*}_{\mathbb{R}(J_i)},\ e^{\alpha_i,*}_{\mathbb{R}(J_i)} \rangle_{\mathbb{R}(J_i)^*,\,\mathbb{R}(J_i)} = \overline{\delta}(\gamma_i,\ \alpha_i) = \begin{cases} 1 & \gamma_i = \alpha_i \\ 0 & \gamma_i \neq \alpha_i, \end{cases}$$

and is extended by linearity — since, as is well known that linear mappings characterized by its values on a basis.

Inner Product Spaces

If $\mathbb{R}(J_i)$ is an inner product space, and x, y are elements of $\mathbb{R}(J_i)$, the inner product between them is $\langle x,\ y \rangle_{\mathbb{R}(J_i)}$. We assume all bases are orthonormal unless otherwise specified.

The Riesz map is of $\mathbb{R}(J_i)$ is the mapping $\wp_{\mathbb{R}(J_i)} : \mathbb{R}(J_i)^* \to \mathbb{R}(J_i)$ that associates every covector $\mu \in \mathbb{R}(J_i)^*$ to a vector $\mu^\wedge \in \mathbb{R}(J_i)$ where

$$\langle \mu,\ x \rangle_{\mathbb{R}(J_i)^*,\,\mathbb{R}(J_i)} = \langle \mu^\wedge,\ x \rangle_{\mathbb{R}(J_i)} \quad \forall \mu \in \mathbb{R}(J_i)^*.$$

Operations on Vector Spaces

If $\mathbb{R}(J_i)$ is the direct sum of a family of vector spaces $\{\mathbb{R}(J_i,\ q_i)\}_{q_i=1}^{Q_i}$, where $Q_i \in \mathbb{N}^+$, and each vector space $\mathbb{R}(J_i,\ q_i)$ has dimension $(J_i,\ q_i)$, we write

$$\mathbb{R}(J_i) = \oplus_{q_i=1}^{Q_i} \mathbb{R}(J_i,\ q_i).$$

We recall that

$$J_i = \sum_{q_i=1}^{Q_i} (J_i,\ q_i).$$

Linear Mappings

The space of linear mappings from $\mathbb{R}(J_i)$ into $\mathbb{R}(I_i)$ will be denoted by $L(\mathbb{R}(J_i);\ \mathbb{R}(I_i))$, and we use the symbol $L(\mathbb{R}(J_i))$ to refer to $L(\mathbb{R}(J_i);\ \mathbb{R}(J_i))$.

1.1.1 Enumeration of Lists

We use the following notation to simplify computations with multilinear maps. Let E and F be sets, elements $v_1, \ldots, v_k \in E$, and a mapping $f : E \to F$.

Individual elements

$v_{\underline{k}}$ means $v_1, \ldots, v_k$ as separate elements.

Creating a k-list

$(v_{\underline{k}}) = (v_1, \ldots, v_k) \in \prod E_{j \le k}$ if $v_i \in E_i$ for $i = \underline{k}$.

Nested indices

$(v_{\underline{n_{\underline{k}}}}) = (v_{n_{\underline{k}}}) = (v_{n_1}, \ldots, v_{n_k})$, and $(v_{\underline{n_k}}) \neq (v_{n_{(1,\ldots,k)}})$.

Closest bracket

$(v_{(n_{\underline{k}})}) = (v_{(n_1,\ldots,n_k)})$ and $(v_{n_{(\underline{k})}}) = (v_{n_{(1,\ldots,k)}})$.

Empty Lists

$(v_{\underline{0}}, a, b, c) = (a, b, c)$

Skipping an index

$(v_{\underline{i-1}}, v_{i+\underline{k-i}}) = (v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_k)$ for $i = \underline{k}$. Note that we can find the size of any list by summing over all the underlined terms and the number of single terms.

Applying f to an element

$(v_{\underline{i-1}}, f(v_i), v_{i+\underline{k-i}}) = (v_1, \ldots, v_{i-1}, f(v_i), v_{i+1}, \ldots, v_k)$. Of course, if $i = 1$, then the above expression reads $(f(v_1), v_2 \ldots, v_k)$ by the $\underline{0}$ interpretation.

Mapping over lists

If $\wedge : E \times E \to F$ is any associative law we write $\bigcirc\!\!\!\!\wedge\,(v_{\underline{k}}) = v_1 \wedge \cdots \wedge v_k$.

1.2. Linear Algebra

The definitions and properties which follow correspond to [**Roman2007Advanced**].

A *basis* (of a vector space V) is a linearly-independent spanning subset $\mathbb{B} \subseteq V$. Two bases of V have the same cardinality, and if $V = \{0\}$, then $\mathbb{B} = \varnothing$. Every vector space has a basis, and although we have yet to introduce them: every Hilbert space has an orthonormal basis.

An *ordered basis* of a finite-dimensional vector space V is $\{e^i_V\}_1^n = (e^1_V, \ldots, e^n_V)$. The *coordinate map* is the isomorphism $\varphi_{e_V} \in L(V, \mathbb{C}^n)$, where

$$[x]_{e_V} = \varphi_{e_V}(\sum_1^n x^i e^i_V) = (x^1, \ldots, x^n) \in \mathbb{C}^n,$$

we call $[x]_{e_V}$ the *coordinate representation* of x in $\{e^i_V\}_1^n$. If $\{\varepsilon^i_V\}_1^n$ is another ordered basis of V, the *transition map* from $\{e_V^i\}_1^n$ to $\{\varepsilon_V^i\}_1^n$ is the linear automorphism

$$\varphi_{e_V,\varepsilon_V} \in GL(n,\mathbb{C}), \quad \varphi_{e_V,\varepsilon_V} = \varphi_{\varepsilon_V} \circ \varphi_{e_V}^{-1},$$

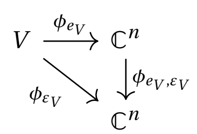
which takes $[x]_{e_V}$ to $[x]_{\varepsilon_V}$. For any fixed $n \geq 1$, the *standard basis* of $\mathbb{C}^n$ is $\{e_{\mathbb{C}^n}^j = (0,\dots,0,1,0,\dots,0)\}_1^n$. Suppose $\mathbb{C}^m$ has standard basis $\{e_{\mathbb{C}^m}^i\}_1^m$, and if $A \in L(\mathbb{C}^n,\mathbb{C}^m)$, its *standard matrix representation* is the matrix

$$[A]_{\varepsilon_V,e_V} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \quad \text{where} \quad \begin{bmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{bmatrix} = \varphi_{\varepsilon_V} A e_V^j \quad \forall 1 \leq j \leq n. \tag{3}$$

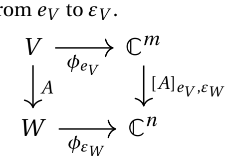
If V and W have ordered bases $\{e_V^j\}_1^n$, and $\{\varepsilon_W^i\}_1^m$, and $A \in L(V,W)$; the *matrix representation* of A with respect to e_V and ε_W is

$$[A]_{e_V,\varepsilon_W} = [\varphi_{\varepsilon_W} A \varphi_{e_V}^{-1}]_{e_V,\varepsilon_W}.$$

More concretely, if $[A]_{e_V,\varepsilon_W}$ is given by the array of numbers in Eq. (3), then $\begin{bmatrix} a_{1j} & \cdots & a_{mj} \end{bmatrix}^T = \varphi_{\varepsilon_W} A e_V^j$. We summarize the matrix representations discussed using fig. 1.



(a) Transition map from e_V to ε_V.



(b) Matrix of $A \in L(V,W)$ with respect to bases e_V and ε_W.

Figure 1: Diagrams representing matrix representations.

1.3. Kernel and Range

If $u_i \in L(\mathbb{R}(J_i);\ \mathbb{R}(I_i))$, its *kernel* is $\operatorname{Ker}(u) = \{x \in \mathbb{R}(J_i),\ ux = 0\}$, and its *range* is $\operatorname{Im}(u) = \{ux,\ x \in \mathbb{R}(J_i)\}$, we remark that both are subspaces (of $\mathbb{R}(J_i)$ and $\mathbb{R}(I_i)$ respectively).

On one hand, u_i is an injection iff it has a trivial kernel iff it admits a left inverse $\sigma \in L(\mathbb{R}(I_i);\ \mathbb{R}(J_i))$ where $\sigma u = \mathrm{id}_{\mathbb{R}(J_i)}$. On the other hand, u_i is a surjection iff its range is V iff it admits a right inverse $\sigma \in L(\mathbb{R}(I_i);\ \mathbb{R}(J_i))$ where $u_i \sigma = \mathrm{id}_{\mathbb{R}(I_i)}$.

Injective linear maps are sometimes referred to as *linear embeddings*. If u_i is a bijection, then we say u_i is a *vector space isomorphism*, $\mathbb{R}(J_i)$ and $\mathbb{R}(I_i)$ are *isomorphic* and we write $\mathbb{R}(J_i) \approx \mathbb{R}(I_i)$. A linear *endomorphism* on $\mathbb{R}(I_i)$ is a linear map $u_i \in L(\mathbb{R}(I_i))$, and u_i is an *automorphism* on V whenever it is invertible.

For any subspace $\mathbb{R}(J_i,\ q_i) \subseteq \mathbb{R}(J_i)$, the *quotient space* of $\mathbb{R}(J_i,\ q_i)$ is defined as the set of cosets of $\mathbb{R}(J_i,\ q_i)$ and is denoted by $\mathbb{R}(J_i)/\mathbb{R}(J_i,\ q_i)$. It is a vector space under elementwise addition and scalar multiplication. We recall

$$x + \mathbb{R}(J_i,\ q_i) = y + \mathbb{R}(J_i,\ q_i) \quad \text{iff} \quad x - y \in \mathbb{R}(J_i,\ q_i).$$

If $u_i \in L(\mathbb{R}(J_i);\ \mathbb{R}(I_i))$, we can decompose $\mathbb{R}(J_i) = \operatorname{Ker}(u) \oplus \mathbb{R}(J_i,\ q_i)_{\mathbb{R}(J_i)}$, and for some subspace $\mathbb{R}(I_i,\ p_i) \subseteq \mathbb{R}(I_i)$, $\mathbb{R}(I_i) = \operatorname{Im}(u) \oplus \mathbb{R}(I_i,\ p_i)$. The *universal property of the quotients* states that $\mathbb{R}(J_i)/\operatorname{Ker}(u) \approx \operatorname{Im}(u)$.

1.4. Multilinear mappings

A mapping $X : \mathbb{R}(J_1) \times \mathbb{R}(J_2) \to W$ into a vector space W is *bilinear* if for every $x_1 \in \mathbb{R}(J_1)$, and $x_2 \in \mathbb{R}(J_2)$,

$$X(x_1,\cdot) : \mathbb{R}(J_2) \to W \quad \text{and} \quad X(\cdot,x_2) : \mathbb{R}(J_1) \to W$$

are linear mappings on the indicated domains. Similarly, a mapping $X : \prod_1^N \mathbb{R}(J_i) \to W$ is N-linear, or muiltilinear in its arguments whenever

$$X(x_{\underline{n-1}},\ \cdot\ , x_{n+\underline{N-n}}) \text{ is } N-1 \text{ linear for all } x_i \in \mathbb{R}(J_i),$$
$$1 \leq i \leq N,\ i \neq n, \text{ and } 1 \leq n \leq N.$$

Analogous to the case where a linear mapping between two vector spaces $\mathbb{R}(J_i)$ and $\mathbb{R}(I_i)$ is determined by its action on the basis vectors $\{e_{\mathbb{R}(J_i)}^{\alpha_i}\}_{\alpha_i=1}^{J_i}$, a multilinear mapping X is characterized by its action on the following set of **elements**

$$\left\{(e_{\mathbb{R}(J_1)}^{\alpha_1},\ \dots\ , e_{\mathbb{R}(J_N)}^{\alpha_N}), \quad 1 \leq \alpha_i \leq J_i,\ 1 \leq i \leq N\right\}. \tag{4}$$

A common misconception is that multilinear maps can be defined by specifying values on

$$\left\{(0,\ \dots\ , e_{\mathbb{R}(J_i)}^{\alpha_i}, 0,\ \dots\ ,0),\ 1 \leq \alpha_i \leq J_i,\ 1 \leq i \leq N\right\},$$

this is **not true**. Take $X(x,y) = e_{\mathbb{R}(J_1)}^{1,*} e_{\mathbb{R}(J_2)}^{2,*}$, then $X(e_{\mathbb{R}(J_1)}^1, 0) = 0$, and $X(0, e_{\mathbb{R}(J_2)}^2) = 0$, however $X(e_{\mathbb{R}(J_1)}^1, e_{\mathbb{R}(J_2)}^2) = 1$.

This is because the notion of multilinearity cannot be described by any external direct sums of the vector spaces involved, and to do so we must introduce the *tensor product.*

1.5. Notes and References

1. Undergraduate texts: [Joh21a; Joh21b; Lan12; Str93; Axl97].

2. Advanced undergraduate: [Hal12].

3. Graduate level texts: [**Roman2007Advanced**; Lan05; Hun03]

4. Matrix Analysis: [HJ90; HHJ94]

2. Metric Vector Spaces

2.1. Projection Mappings

We turn our attention to ways of 'factorizing' linear mappings. Suppose $V = S \oplus T$, a *projection* onto S is an operator ρ where $\rho(s+t) = s$ for all $s \in S$ and $t \in T$. It follows that $V = \rho(V) \oplus \mathrm{Ker}(\rho)$, and $x \in S$ iff $\rho x = x$.

Two projections ρ_1, ρ_2, with images S_1, S_2 are *orthogonal* whenever $\rho_1\rho_2 = 0 = \rho_2\rho_1$.[1] A *resolution of the identity* on V is a sum of projection operators $\{\rho_i\}_1^k$

$$\sum_1^k \rho_i = \mathrm{id}_V \quad \text{where} \quad \rho_i\rho_j = 0 = \rho_j\rho_i.$$

It can be shown that the images of the projections form a direct sum decomposition of V, in symbols $V = \oplus_1^k \rho_i(V)$. Conversely, if $V = \oplus_1^k S_i$, the projections ρ_i onto S_i form a resolution of the identity.

2.2. Eigenvalues and eigenvectors

An *eigenvector* for a linear mapping $u \in L(V, W)$ is a non-zero element $x \in V$ where $ux = \lambda x$ for some $\lambda \in \mathbb{C}$; and λ is the *eigenvalue* associated to x. The *eigenspace* is the subspace of V containing all vectors x that satisfy the equation $ux = \lambda x$.

An operator $u \in L(V)$ is said to be *diagonalizable* if there exists a *spectral resolution* of its eigenvalue:

$$u = \sum_1^k \lambda_i \rho_{\mathcal{E}_i} \quad \text{where} \quad \{\lambda_i\}_1^k \text{ are distinct eigenvalues of } u,$$

and $\{\rho_{\mathcal{E}_i}\}_1^k$ are mutually orthogonal projections onto subspaces of V.

2.3. Inner Product Spaces

Geometry is the most potent on *inner product spaces*, which are vector spaces equipped with an *inner product*. A scalar-valued function $\omega : V \times V \to \mathbb{F}$ where $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$ is an inner product if it is

positive definite
$\omega(x,x) \geq 0$ for all x, and $\{0\} = \{x,\ \omega(x,x) = 0\}$.

sesquilinear
if $\mathbb{F} = \mathbb{C}$, then ω is linear (resp. conjugate-linear) in its first (resp. second) coordinate.

symmetric
if $\mathbb{F} = \mathbb{R}$, then ω is a symmetric bilinear form.

We assume for the rest of this section that all vector spaces are finite-dimensional inner product spaces and we use $\langle \cdot,\ \cdot \rangle_V$ in favour of ω. If U is another inner product space, and $u \in L(V, W)$. We say u is an *isometry* if it relates the inner product on V to that on U, this means for all $x, y \in U$, $\langle ux,\ uy \rangle_V = \langle x,\ y \rangle_U$; and if this is the case, u is an *isometric isomorphism* if it is also a vector space isomorphism.

Every inner product space is normed, with $x \mapsto \|x\| = \sqrt{\langle x,\ x \rangle_V}$ for all $x \in V$. Finally, a *Hilbert space* is an inner product space that is Cauchy-complete with respect to its norm topology.

Two vectors $x, y \in V$ are *orthogonal* to each other, written $x \perp y$ whenever $\langle x,\ y \rangle_V = 0$. If $E \subseteq V$ is any subset, its *orthogonal complement* is $E^\perp = \{x \in V,\ x \perp y\ \forall y \in V\}$.

If V is an inner product space, and $S \subseteq V$ is a subset, it is *orthogonal* whenever $x \perp y$ for all $x \neq y$ in S. If $\|x\| = 1$ for all x, it is *orthonormal*.

For a given subspace $S \subseteq V$, its linear complement T can be chosen such that $V = S \oplus T$, and $S \perp T$. [2]

A projection ρ is *orthogonal* whenever $\mathrm{Ker}(\rho) \perp \rho(V)$, not to be confused with two (mutually) orthogonal projections. A resolution of the identity is *orthogonal* whenever each projection ρ_i involved is orthogonal.

The Hilbert space adjoint of u is instead denoted by $u^* \in L(W, V)$, which is the unique linear mapping that satisfies

$$\langle ux,\ y \rangle_V = \langle x,\ u^* y \rangle_U.$$

We note that $u^{**} = u$, and u is a surjection (resp. an injection) iff u^* is an injection (resp. a surjection). The matrix representation of the Hilbert space adjoint is the *Hermitian transpose*[3] of the matrix representation.

We mention several important classes of operators.

[1] equivalently: $S_2 \subseteq \mathrm{Ker}(\rho_1)$ and $S_1 \subseteq \mathrm{Ker}(\rho_2$.

[2] If V is infinite-dimensional, assume that S is closed.

[3] also called the conjugate transpose

Unitary
A *unitary* operator $u \in L(V)$ is an isometric isomorphism on V. Equivalent characterizations: $u^* u = \mathrm{id}_V$, columns (resp. rows) of any matrix representation of u form orthonormal basis in $\mathbb{C}^n$.

Normal
A *normal* operator $u_i \in L(\mathbb{R}(J_i))$ is an operator that commutes with its adjoint: $u^* u = uu^*$. Equivalent characterization: u can be *unitarily diagonalized*, that is: there exists orthogonal eigenspaces $\{\mathcal{E}_i\}_1^k$ corresponding to the eigenvalues $\{\lambda_i\}_1^k$, such that $u = \sum_1^k \lambda_i \rho_{\mathcal{E}_i}$,

$$V = \oplus \mathcal{E}_i, \quad \mathcal{E}_i \perp \mathcal{E}_j \text{ if } i \neq j.$$

Note that unitarily diagonalizability is the same as saying there exists an orthonormal basis $\{e_V^i\}_1^n$ where $[u]_{b,b} = \mathrm{diag}(\lambda_1, \ldots, \lambda_k)$.

Self-adjoint
An operator u is *self-adjoint* whenever $u^* = u$. Some refer to self-adjoint operators as *Hermitian* when the base field is $\mathbb{C}$ or *symmetric* if the base field is $\mathbb{R}$.

If $u^* = -u$, it is said to be *skew self-adjoint* (resp. skew Hermitian, and skew-symmetric if the base field is $\mathbb{R}$.)

Positive (semi-)definite
A self-adjoint operator u is *positive semi-definite* if $Q_u(x) = \langle ux, x\rangle_V \geq 0$ for all $x \in V$, it is *positive definite* whenever $\mathrm{Ker}(Q_u) = \{0\}$.

Equivalently, u is positive semi-definite iff all of its eigenvalues are non-negative, and positive definite iff all of its eigenvalues are strictly positive.

If $u \in L(V, W)$, then $\mathrm{Ker}(u^*) = \mathrm{Im}(U)^\perp$, and $u^*(V) = \mathrm{Ker}(u)^\perp$. It is clear that u is unitary (resp. normal, positive semi-definite, positive definite) iff u^* is.

2.4. Grammian and its applications

The *Grammian* [AM06] of a linear mapping $u \in L(V, W)$ is the operator $u^* u \in L(U)$. Properties of the Grammian include: $\mathrm{Ker}(u^* u) = \mathrm{Ker}(u)$, $\mathrm{Ker}(uu^*) = \mathrm{Ker}(u^*)$ and also $u^* \mathrm{Im}(U) = \mathrm{Im}(U)$, $uu^*(V) = u^*(V)$, see [**Roman2007Advanced**]. A special case worth mentioning is when u is an embedding, then $\mathrm{Ker}(u^* u) = \{0\}$, whence the Grammian is unitary on U.

If $u \in L(V, W)$ is arbitrary, we can decompose $U = \mathrm{Im}(u^*) \oplus \mathrm{Ker}(u)$, and $V = \mathrm{Im}(u) \oplus \mathrm{Ker}(u^*)$. The Grammian $u^* u$ is positive semi-definite, and can be unitarily diagonalized by basis $\{b_j\}_1^n$. By reordering b, we can assume that

$$[u^* u]_{b,b} = \mathrm{diag}(\sigma_1^2, \ldots, \sigma_r^2, 0, \ldots, 0), \quad \sigma_1^2 \geq \cdots \geq \sigma_r^2 > 0.$$

The numbers $\{\sigma_i\}_1^r$ are the *singular values* of u.
Matrix Norms
We turn our attention to $m \times n$ complex matrices, which we will denote by $\mathcal{M}_{m,n}$. Let $A \in \mathcal{M}_{m,n}$ with entries as in Eq. (3). If $p \in (0, +\infty]$, the entrywise l^p *norm* of A is

$$\|A\|_p = \begin{cases} \left(\sum_{i,j} |a_{ij}|^p\right)^{1/p} & 0 < p < +\infty \\ \max_{i,j} |a_{ij}| & p = +\infty \end{cases},$$

[4] The *Frobenius norm* of A is the entrywise 2-norm of A, is usually denoted by $\|A\|_F$ instead of $\|A\|_2$,

$$\|A\|_F = \sum_i |\sigma_i(A)|^2 \quad \text{is the sum of the squares of the singular values of}$$

The *operator norm* on A is $\|A\|_{op} = \sup_{|x| \leq 1} |Ax|$, with $|\cdot|$ norms induced by the inner products on $\mathbb{C}^n$, and $\mathbb{C}^m$. The *spectral norm* coincides with the 2-norm, the former defined by

$$\|A\|_{spec} = \sigma_1(A) = \|A\|_{op},$$

where $\sigma_1(A)$ is the largest singular value of A. One also has the estimate

$$\|A\|_{op} \leq \|A\|_F.$$

3. Tensors

This section follows [Gre67, Chp 1-3] closely, and summarizes the definitions and theorems that we will need for the discussions that follow. We will shy away from defining, proving, or discussing the notions of universality that is prevalent in category theory [Rie17; Lan05] and the generalization to tensor products of modules and algebras [Lan05], nor will we discuss its implications in differential geometry [**Lee2013Introduction**; **Lee2019Introduction**].

3.1. Tensor Product

The *tensor product* of the vector spaces $\{\mathbb{R}(J_i),\ i = 1, \ldots, N\}$ is **any pair** of the form $(\mathcal{T}, \bigotimes)$ such that $\mathcal{T}$ is a real vector space and $\bigotimes$ is a N-linear map

$$\bigotimes : \prod_1^N \mathbb{R}(J_i) \to \mathcal{T}, \quad \bigotimes(x_1, \ldots, x_N) = x_1 \otimes \cdots x_N.$$

The mapping $\bigotimes$ is required to satisfy the *universal property of multilinearity* (see fig. 2 also):

[4] we note that the l^0 norm is sometimes used to refer to the number of non-zero entries of A.

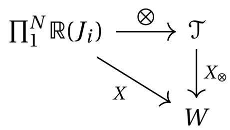


Figure 2: Universal Property of Multilinearity

If $X : \prod_1^N \mathbb{R}(J_i) \to W$ is **any multilinear map** from $\prod_1^N$, there exists a unique **linear map** $X_\otimes \in L(\mathcal{T}; W)$ such that

$$X_\otimes \circ \bigotimes = X.$$

Several points that are worth mentioning.

Not Unique
The tensor product is not unique, it is defined up to a linear isomorphism. Any pair that satisfies the two properties is called a tensor product of $\{\mathbb{R}(J_i)\}_1^N$.

What does $\mathcal{T}$ do?
Roughly speaking, $\mathcal{T}$ is the vector space designed to capture the valuations of multilinear mappings (it is also the leanest possible vector space — thus giving rise to the 'unique' descent of a multilinear map into a linear one [Gre67]), similar to how a multilinear map is characterized by its values on the set given in Eq. (4).

What does $\bigotimes$ do?
The mapping $\bigotimes$ allows one to factor out the multilinearity of N-linear X. The universal property also gives a bijective correspondence between multilinear mappings out of $\prod_1^N$ and linear mappings out of $\mathcal{T}$, **while preserving the multilinear structure of the mappings involved**.

Existence
The tensor product of any finite family of vector spaces always exists, we will not pursue the proof of this here. The reader is encouraged to consult the proof using free vector spaces [**Roman2007Advanced**; **Lee2013Introduction**; **Lee2019Introduction**; Lan05; Gre67].

We will use $(\otimes_1^N \mathbb{R}(J_i), \bigotimes)$ to denote the tensor product of $\{\mathbb{R}(J_i)\}_1^N$, and we may omit the factorization mapping $\bigotimes$ when it is understood. Since every multilinear map gives rise to a unique linear map on $\otimes_1^N \mathbb{R}(J_i)$, we identify $X_\otimes \approx X$.

Multilinear mappings from $\prod_1^N \mathbb{R}(J_i)$ are hereinafter identified as linear maps from $\otimes_1^N \mathbb{R}(J_i)$.

If $(x_1, \dots, x_N) \in \prod_1^N \mathbb{R}(J_i)$ it is customary to write $\otimes_1^N x_i = \bigotimes(x_1, \dots, x_N)$ as an element in $\otimes_1^N \mathbb{R}(J_i)$.

We conclude this section with a list of elementary properties of the tensor product. These properties are proven in [**Roman2007Advanced**; Gre67].

Commutative
If σ is a permutation on $\{1, \dots, N\}$, then $\otimes_1^N \mathbb{R}(J_{\sigma(i)})$ and $\otimes_1^N \mathbb{R}(J_i)$ are linearly isomorphic as vector spaces. Because of this, we sometimes write

$$\bigotimes(x_1, \dots, x_N) = x_1 \otimes \cdots \otimes x_N.$$

Associative
For the case of $N = 3$, $(\mathbb{R}(J_1) \otimes \mathbb{R}(J_2)) \otimes \mathbb{R}(J_3)$ is linearly isomorphic to $\mathbb{R}(J_1) \otimes (\mathbb{R}(J_2) \otimes \mathbb{R}(J_3))$, which is linearly isomorphic to $\otimes_1^3 \mathbb{R}(J_i)$.

Tensor Product of Base Field
For every N, the N-fold tensor product of the base field N isomorphic to the base field $\otimes_1^N \mathbb{R} \cong \mathbb{R}$.

Tensor Product of Dual Spaces
The tensor product of the dual spaces is the dual space of the tensor product:

$$\left[\otimes_1^N \mathbb{R}(J_i)\right]^* \cong \otimes_1^N \mathbb{R}(J_i)^*.$$

Basis of the Tensor Product
The mapping $\bigotimes$ of the set in Eq. (4) forms a basis of $\otimes_1^N \mathbb{R}(J_i)$.

$$\left\{ e^{\alpha}_{\otimes_1^N \mathbb{R}(J_i)} = e^{\alpha_1}_{\mathbb{R}(J_1)} \otimes \cdots \otimes e^{\alpha_N}_{\mathbb{R}(J_N)},\ \alpha \in \otimes_{i=1}^N \mathbb{N}^+(J_i) \right\}; \quad (5)$$

An element $x \in \otimes_1^N \mathbb{R}(J_i)$ can be expressed as a sum

$$\begin{aligned} x &= \sum_{1 \le \alpha_1 \le J_1} \cdots \sum_{1 \le \alpha_N \le J_N} x_{\alpha_1,\dots,\alpha_N} e^{\alpha_1}_{\mathbb{R}(J_1)} \otimes \cdots \otimes e^{\alpha_N}_{\mathbb{R}(J_N)} \\ &= \sum_{\alpha \in \otimes_{i=1}^N \mathbb{N}^+(J_i)} x_\alpha e^{\alpha}_{\otimes_1^N \mathbb{R}(J_i)}, \end{aligned} \quad (6)$$

where $x_\alpha \in \mathbb{R}$ corresponds to the coefficient of $e^{\alpha_i}_{\mathbb{R}(J_i)}$. When we discuss linear mappings between vector spaces (and tensor products), we will often use the symbols I and β.

The tensor product $\otimes_1^N \mathbb{R}(I_i)$ has basis

$$\left\{ e^{\beta}_{\otimes_1^N \mathbb{R}(I_i)} = e^{\beta_1}_{\mathbb{R}(I_1)} \otimes \cdots \otimes e^{\beta_N}_{\mathbb{R}(I_N)},\ \beta \in \otimes_{i=1}^N \mathbb{N}^+(I_i) \right\}.$$

Tensor Product is not a Direct Sum
For simplicity we restrict ou r study to the tensor product $\mathbb{R}(J_1) \otimes \mathbb{R}(J_2)$. The observation that $\bigotimes$ is bilinear leads us to

$$x_1 \otimes 0 + 0 \otimes x_2 \quad \text{does not equal} \quad x_1 \otimes x_2,$$

unless $x_1 = 0$ or $x_2 = 0$, and

$$a_1 x_1 \otimes y + a_2 x_1' \otimes y = (a_1 x_1 + a_1 x_1') \otimes y,$$

for all $a_1, a_2 \in \mathbb{R}$, $x_1,\ x_1' \in \mathbb{R}(J_1)$; similarly for the second coordinate.

Dimension of the Tensor Product
The dimension of the tensor product $\otimes_1^N \mathbb{R}(J_i)$ is the product of the dimensions $\prod_1^N J_i$. The same is not true for external direct sums.

In fact, $\prod_1^N \mathbb{R}(J_i)$ has dimension $\sum_1^N J_i$, this is the principal reason why the external direct sum **always** fail to capture multilinearity when $N \geq 2$ and $J_i \geq 2$.

3.2. Direct Sums (Optional)

Direct Sums (highly related to Tracy Singh decomposition), indexing conventions, and matrix unfoldings. Idea is that block matrices factorize a linear map into smaller linear maps that operate between subspaces which sum (directly) to the original space.

Suppose for each $i = 1, \ldots, N$, the direct sums

$$\mathbb{R}(J_i) = \oplus_{q_i=1}^{Q_i} \mathbb{R}(J_i, q_i) \quad \text{and} \quad \mathbb{R}(I_i) = \oplus_{p_i=1}^{P_i} \mathbb{R}(I_i, p_i), \tag{7}$$

where each of the subspaces $\mathbb{R}(J_i, q_i)$ (resp. $\mathbb{R}(I_i, p_i)$ have dimension $\overline{q}_i$ (resp. $\overline{p}_i$), and $J_i = \sum_{q_i=1}^{Q_i} \overline{q}_i$ (resp. $I_i = \sum_{p_i=1}^{P_i} \overline{p}_i$. Then,

$$\otimes_1^N \mathbb{R}(J_i) = \oplus_{q_1=1}^{Q_1} \cdots \oplus_{q_N=1}^{Q_N} \otimes_1^N \mathbb{R}(J_i, q_i)$$
$$\otimes_1^N \mathbb{R}(I_i) = \oplus_{p_1=1}^{P_1} \cdots \oplus_{p_N=1}^{p_N} \otimes_1^N \mathbb{R}(I_i, p_i)$$

Proposition 3.1 Direct Sums of Tensor Products

If $\mathbb{R}(J_i)$ is given by Eq. (7), there exists a linear isomorphism such that

$$\otimes_1^N \mathbb{R}(J_i) = \otimes_{i=1}^N \oplus_{q_i=1}^{Q_i} [\mathbb{R}(J_i,\ q_i)] = \oplus_{q \in \otimes_1^N Q_i} \otimes_{i=1}^N \mathbb{R}(J_i,\ q_i)$$

See [Gre67, chp 1.9-1.12] for the proof.

For each $\mathbb{R}(J_i)$, we can enumerate its basis vectors in a manner that respects the direct sum structure.

Dimension

$$\dim \mathbb{R}(J_i,\ p_i) = (J_i;\ q_i) \quad \text{and} \quad \sum_{p_i=1}^{P_i} (J_i;\ p_i) = J_i.$$

Restricted / Projected Maps
We want a good way of unfolding the subspaces, fix $(I_i,\ q_i)$ and $(J_i,\ p_i)$, referring to the $(\beta_i,\ q_i)$th (resp. $(\alpha_i,\ p_i)$th) basis vector in the subspace $(I_i,\ q_i)$ (resp. $(J_i,\ p_i)$)) of $\mathbb{R}(I_i)$ (resp. $\mathbb{R}(J_i)$).

$$\sum_{w_i=1}^{q_i-1} (I_i;\ w_i) + (\beta_i,\ q_i)$$

Lexicographical ordering

$$\left[\oplus_{q_1=1}^{Q_1} \mathbb{R}(J_1, q_1)\right] \otimes \cdots \otimes \left[\oplus_{q_N=1}^{Q_N} \mathbb{R}(J_N, q_N)\right]$$

Tracy-Singh Ordering

$$\oplus_{q_1=1}^{Q_1} \cdots \oplus_{q_N=1}^{Q_N} \left[\mathbb{R}(J_1, q_1) \otimes \cdots \otimes \mathbb{R}(J_N, q_N)\right]$$

Definition 3.2 Tracy Singh Product

Also called partitioned Kronecker product, block Kronecker product, generalized Rao product.
[TJ89] contains proofs of various properties, of the partitioned Kronecker product. It goes over a special case that is of interest in statistics: the balanced partitioned product. This is when $\overline{q}_i = \overline{p}_i = \overline{r}_i = \overline{s}_i$, and provides transformation formulas.

[Liu99] contain properties for when the A and B are matrices, where

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}. \tag{8}$$

The matrices in the form of Eq. (8) is of particular interest, because of SVD.

The *lexicographical ordering* (or dictionary ordering) of basis vectors [Tuc66, pp.281] in $\otimes_1^N \mathbb{R}(J_i)$ is the correspondence between

$$\{1, \ldots, \prod_1^N J_i\} \quad \text{and} \quad \left\{\otimes_1^N e_{\mathbb{R}(J_i)}^{\alpha_i},\ 1 \leq \alpha_i \leq J_i,\ 1 \leq i \leq N\right\},$$

such that indices α_i corresponding to a larger i value (closer to N) take precedence during enumeration from 1 to $\prod_1^N J_i$. If $\alpha = (\alpha_1, \ldots, \alpha_N) \in \otimes_{i=1}^N \mathbb{N}^+(J_i)$,

$$\begin{aligned}\alpha &\cong (\alpha_1 - 1)\prod_2^N J_i + (\alpha_2 - 1)\prod_3^N J_i + \cdots + (\alpha_{N-1} - 1)J_N + \alpha_N \\ &= \sum_{i=1}^N (\alpha_i - 1) \prod_{j=i+1}^N J_j.\end{aligned} \tag{9}$$

Given an integer $a \in [1, \prod_1^N J_i]$, we can find its corresponding $\alpha \in \otimes_{i=1}^N \mathbb{N}^+(J_i)$ by taking the inverse of Eq. (9). An explicit construction of this inverse is found in **??**.

3.3. Linear Maps as Tensors

Given $\mathbb{R}(J_i)$ and $\mathbb{R}(I_i)$ and their bases

$$\left\{e^{\beta}_{\otimes_1^N \mathbb{R}(I_i)} = e^{\beta_1}_{\mathbb{R}(I_1)} \otimes \cdots \otimes e^{\beta_N}_{\mathbb{R}(I_N)},\ \beta \in \otimes_{i=1}^N \mathbb{N}^+(I_i)\right\}$$
$$\left\{e^{\alpha}_{\otimes_1^N \mathbb{R}(J_i)} = e^{\alpha_1}_{\mathbb{R}(J_1)} \otimes \cdots \otimes e^{\alpha_N}_{\mathbb{R}(J_N)},\ \alpha \in \otimes_{i=1}^N \mathbb{N}^+(J_i)\right\} \tag{10}$$

the tensor product (cf. Eq. (5)) induces a basis of $\mathbb{R}(I_i) \otimes \mathbb{R}(J_i)^*$

$$\left\{e^{\beta_i}_{\mathbb{R}(I_i)} \otimes e^{\alpha_i,*}_{\mathbb{R}(J_i)},\ 1 \le \beta_i \le I_i,\ 1 \le \alpha_i \le J_i\right\}. \tag{11}$$

The space $L(\mathbb{R}(J_i);\ \mathbb{R}(I_i))$ is *naturally isomorphic* to $\mathbb{R}(I_i) \otimes \mathbb{R}(J_i)^*$. To see this, fix $i = 1, \ldots, N$, and we define the map

$$\mathcal{T} : \mathbb{R}(I_i) \otimes \mathbb{R}(J_i)^* \to L(\mathbb{R}(J_i);\ \mathbb{R}(I_i)),$$

by specifying its action on basis vectors of $\mathbb{R}(I_i) \otimes \mathbb{R}(J_i)^*$:

$$\mathcal{T}(y_{\mathbb{R}(I_i)} \otimes y^*_{\mathbb{R}(J_i)})(x) = y_{\mathbb{R}(I_i)} \langle y^*_{\mathbb{R}(J_i)},\ x\rangle_{\mathbb{R}(J_i)^*,\,\mathbb{R}(J_i)}. \tag{12}$$

If $y_{\mathbb{R}(I_i)} \otimes y^*_{\mathbb{R}(J_i)} \in \mathrm{Ker}(\mathcal{T})$, then either $y_{\mathbb{R}(I_i)} = 0$ or $y^*_{\mathbb{R}(J_i)}$, and in both cases $y_{\mathbb{R}(I_i)} \otimes y^*_{\mathbb{R}(J_i)} = 0 \otimes 0 = 0$.

Conversely, given $u_i \in L(\mathbb{R}(J_i);\ \mathbb{R}(I_i))$, we claim that the numbers

$$\left\{u_i(\beta_i,\ \alpha_i) \in \mathbb{R},\ 1 \le \alpha_i \le J_i,\ 1 \le \beta_i \le I_i\right\}$$

obtained through the scalar product of $\mathbb{R}(I_i)$ with its dual $\mathbb{R}(I_i)^*$ completely characterizes u_i:

$$u_i(\beta_i,\ \alpha_i) = \langle e^{\beta_i,*}_{\mathbb{R}(I_i)},\ u_i(e^{\alpha_i}_{\mathbb{R}(J_i)})\rangle_{\mathbb{R}(I_i)^*} \quad \begin{matrix} 1 \le \alpha_i \le J_i \\ 1 \le \beta_i \le I_i \end{matrix}\ . \tag{13}$$

Fix $u'_i \in \mathbb{R}(I_i)^* \otimes \mathbb{R}(J_i)$ with

$$u'_i = \sum_{\substack{1 \le \beta_i \le I_i, \\ 1 \le \alpha_i \le J_i}} u_i(\beta_i,\ \alpha_i)(e^{\beta_i}_{I_i} \otimes e^{\alpha_i,*}_{J_i}). \tag{14}$$

For any $x_i = \sum_{1 \le \alpha_i \le J_i} x_i^{\alpha_i} e^{\alpha_i}_{\mathbb{R}(J_i)} \in \mathbb{R}(J_i)$,

$$\begin{aligned}
\mathcal{T}(u'_i)(x_i) &= \sum_{\substack{1 \le \alpha_i \le J_i, \\ 1 \le \beta_i \le I_i}} u_i(\beta_i,\ \alpha_i)\mathcal{T}(e^{\beta_i}_{\mathbb{R}(I_i)} \otimes e^{\alpha_i,*}_{\mathbb{R}(J_i)})(x_i) \\
&= \sum_{\substack{1 \le \alpha_i \le J_i, \\ 1 \le \beta_i \le I_i}} u_i(\beta_i,\ \alpha_i)\langle e^{\alpha_i,*}_{\mathbb{R}(J_i)},\ x_i\rangle_{\mathbb{R}(J_i)^*,\,\mathbb{R}(J_i)} e^{\beta_i}_{\mathbb{R}(I_i)} \\
&= \sum_{\substack{1 \le \alpha_i \le J_i, \\ 1 \le \beta_i \le I_i}} u_i(\beta_i,\ \alpha_i)\langle e^{\alpha_i,*}_{\mathbb{R}(J_i)},\ \sum_{1 \le \gamma_i \le J_i} x_i^{\gamma_i} e^{\gamma_i}_{\mathbb{R}(J_i)}\rangle_{\mathbb{R}(J_i)^*,\,\mathbb{R}(J_i)} e^{\beta_i}_{\mathbb{R}(I_i)} \\
&= \sum_{\substack{1 \le \alpha_i \le J_i, \\ 1 \le \beta_i \le I_i}} u_i(\beta_i,\ \alpha_i) x_i^{\alpha_i} e^{\beta_i}_{\mathbb{R}(I_i)}.
\end{aligned} \tag{15}$$

Which proves that $\mathcal{T}(u'_i) = u_i$, and we identify u'_i with its image under $\mathcal{T}$ (and conversely)

$$u_i = \sum_{\substack{1 \le \beta_i \le I_i, \\ 1 \le \alpha_i \le J_i}} u_i(\beta_i,\ \alpha_i)(e^{\beta_i}_{I_i} \otimes e^{\alpha_i,*}_{J_i}). \tag{16}$$

Adjoint Mapping in Coordinates

If $i = 1, \ldots, N$, the adjoint map $u_i^* \in L(\mathbb{R}(I_i)^*;\ \mathbb{R}(J_i)^*)$ has coordinate representation

$$u_i^*(\alpha_i, \beta_i) = \langle e^{\beta_i,*}_{\mathbb{R}(I_i)},\ u_i(e^{\alpha_i}_{\mathbb{R}(J_i)})\rangle_{\mathbb{R}(I_i)^*,\,\mathbb{R}(I_i)} = u_i(\beta_i,\ \alpha_i) \qquad \forall 1 \le \alpha_i \le J_i,\ 1 \le \beta_i \le I_i.$$

Scalar Product on Linear Endomorphisms

Composition Algebra. Because $\mathbb{R}(J_i)$ and $\mathbb{R}(J_i)^*$ are in duality, and since

$$\mathbb{R}(J_i) \otimes \mathbb{R}(J_i)^* \cong \mathbb{R}(J_i)^* \otimes \mathbb{R}(J_i),$$

we consider $\mathbb{R}(J_i) \otimes \mathbb{R}(J_i)^*$ to be in self-duality, and its scalar product exhibits a certain cross-multiplication structure:

$$\begin{aligned}
&\langle y_{\mathbb{R}(J_i)} \otimes y^*_{\mathbb{R}(J_i)},\ x_{\mathbb{R}(J_i)} \otimes x^*_{\mathbb{R}(J_i)}\rangle_{\mathbb{R}(J_i) \otimes \mathbb{R}(J_i)^*} \\
&\quad = \langle y^*_{\mathbb{R}(J_i)},\ x_{\mathbb{R}(J_i)}\rangle_{\mathbb{R}(J_i)^*,\,\mathbb{R}(J_i)} \langle x^*_{\mathbb{R}(J_i)},\ y_{\mathbb{R}(J_i)}\rangle_{\mathbb{R}(J_i)^*,\,\mathbb{R}(J_i)}.
\end{aligned}$$

3.4. Tensor Product of Linear Mappings

Let u_i be linear mappings given as in Eq. (16) for $i = 1, \ldots, N$, we define its *tensor product* [**Roman2007Advanced**; **Lee2013Introduction**; **Lee2019Introduction**; Gre67, Chp 14, Chp 1-3, Chp 12, App. D] as the map

$$U : \otimes_1^N \mathbb{R}(J_i) \to \otimes_1^N \mathbb{R}(I_i), \quad \text{we write} \quad U = \otimes_1^N u_i,$$

by specifying its values on the basis of $\otimes_1^N \mathbb{R}(J_i)$:

$$U(e^{\alpha}_{\otimes_1^N \mathbb{R}(J_i)}) = \otimes_1^N u_i(e^{\alpha_i}_{\mathbb{R}(J_i)}). \tag{17}$$

We remark that Equation (17) implies

$$U(\otimes_1^N x_i) = \otimes_1^N u_i(x_i), \quad \forall x_i \in \mathbb{R}(J_i),\ 1 \le i \le N.$$

Since U is an element in $L(\otimes_1^N \mathbb{R}(J_i);\ \otimes_1^N \mathbb{R}(I_i)) \cong \otimes_1^N \mathbb{R}(J_i) \otimes \otimes_1^N \mathbb{R}(I_i)^*$ and thus admits a basis expansion in the form of Eq. (16) — which we will now compute. If $\gamma = (\gamma_1, \ldots, \gamma_N) \in \otimes_{i=1}^N \mathbb{N}^+(J_i)$,

$$\begin{aligned}
U(e^{\gamma}_{\otimes_1^N \mathbb{R}(J_i)}) &= \bigotimes_{i=1,\ldots,N} \sum_{\substack{1 \le \alpha_i \le J_i, \\ 1 \le \beta_i \le I_i}} u_i(\beta_i,\ \alpha_i) e^{\beta_i}_{\mathbb{R}(I_i)} \langle e^{\alpha_i,*}_{\mathbb{R}(J_i)},\ e^{\gamma_i}_{\mathbb{R}(J_i)}\rangle_{\mathbb{R}(J_i)^*} \\
&= \bigotimes_{i=1,\ldots,N} \sum_{\substack{1 \le \alpha_i \le J_i, \\ 1 \le \beta_i \le I_i}} u_i(\beta_i,\ \alpha_i) e^{\beta_i}_{\mathbb{R}(I_i)} \overline{\delta}(\alpha_i,\ \gamma_i) \\
&= \bigotimes_{i=1,\ldots,N} \sum_{1 \le \beta_i \le I_i} u_i(\beta_i,\ \gamma_i) e^{\beta_i}_{\mathbb{R}(I_i)} \\
&= \sum_{1 \le \beta_1 \le I_1} \cdots \sum_{1 \le \beta_N \le I_N} \prod_{i=1,\ldots,N} u_i(\beta_i,\ \gamma_i) \bigotimes_{i=1,\ldots,N} e^{\beta_i}_{\mathbb{R}(I_i)} \\
&= \sum_{1 \le \beta_1 \le I_1} \cdots \sum_{1 \le \beta_N \le I_N} \prod_{i=1,\ldots,N} u_i(\beta_i,\ \gamma_i) e^{\beta}_{\otimes_1^N \mathbb{R}(I_i)}\ .
\end{aligned} \tag{18}$$

Applying the dual vector $e^{\nu,*}_{\otimes_1^N\mathbb{R}(I_i)}$, where $\nu=(\nu_1,\ldots,\nu_N)\in\otimes_{i=1}^N\mathbb{N}^+(I_i)$ to Eq. (18) yields

$$\begin{aligned}
U(\nu,\gamma) &= U((\nu_1,\ldots,\nu_N),(\gamma_1,\ldots\gamma_N))\\
&=\left\langle e^{\nu,*}_{\otimes_1^N\mathbb{R}(I_i)}, U(e^{\gamma}_{\otimes_1^N\mathbb{R}(J_i)})\right\rangle_{\otimes_1^N\mathbb{R}(I_i)^*,\otimes_1^N\mathbb{R}(I_i)}\\
&=\left\langle e^{\nu,*}_{\otimes_1^N\mathbb{R}(I_i)}, \sum_{\beta\in\otimes_{i=1}^N\mathbb{N}^+(I_i)}\prod_{i=1,\ldots,N}u_i(\beta_i,\gamma_i)e^{\beta}_{\otimes_1^N\mathbb{R}(I_i)}\right\rangle_{\otimes_1^N\mathbb{R}(I_i)^*,\otimes_1^N\mathbb{R}(I_i)}\\
&=\sum_{\beta\in\otimes_{i=1}^N\mathbb{N}^+(I_i)}\prod_{i=1,\ldots,N}u_i(\beta_i,\gamma_i)\left\langle e^{\nu,*}_{\otimes_1^N\mathbb{R}(I_i)}, e^{\beta}_{\otimes_1^N\mathbb{R}(I_i)}\right\rangle_{\otimes_1^N\mathbb{R}(I_i)^*,\otimes_1^N\mathbb{R}(I_i)}\\
&=\sum_{\beta\in\otimes_{i=1}^N\mathbb{N}^+(I_i)}\prod_{i=1,\ldots,N}u_i(\beta_i,\gamma_i)\prod_{i=1,\ldots,N}\langle e^{\nu_i,*}_{\mathbb{R}(I_i)^*}, e^{\beta}_{\mathbb{R}(I_i)}\rangle_{\mathbb{R}(I_i)^*,\mathbb{R}(I_i)}\\
&=\sum_{\beta\in\otimes_{i=1}^N\mathbb{N}^+(I_i)}\prod_{i=1,\ldots,N}u_i(\beta_i,\gamma_i)\prod_{i=1,\ldots,N}\overline{\delta}(\nu_i,\beta_i)\\
&=\sum_{\beta\in\otimes_{i=1}^N\mathbb{N}^+(I_i)}\prod_{i=1,\ldots,N}u_i(\beta_i,\gamma_i)\overline{\delta}(\nu,\beta)\\
&=\prod_{i=1,\ldots,N}u_i(\nu_i,\gamma_i)
\end{aligned}\tag{19}$$

This leads us to write

$$\begin{aligned}
U &= \sum_{\substack{1\le\alpha_1\le J_1,\\1\le\beta_1\le I_1}}\cdots\sum_{\substack{1\le\alpha_N\le J_N,\\1\le\beta_N\le I_N}}\prod_{i=1,\ldots,N}u_i(\beta_i,\alpha_i)e^{\beta,*}_{\otimes_1^N\mathbb{R}(I_i)}\otimes e^{\alpha}_{\otimes_1^N\mathbb{R}(J_i)}\\
&=\sum_{\substack{\alpha\in\otimes_{i=1}^N\mathbb{N}^+(J_i),\\\beta\in\otimes_{i=1}^N\mathbb{N}^+(I_i)}}U(\beta,\alpha)e^{\beta,*}_{\otimes_1^N\mathbb{R}(I_i)}\otimes e^{\alpha}_{\otimes_1^N\mathbb{R}(J_i)}.
\end{aligned}\tag{20}$$

Next, suppose we are given $x\in\otimes_1^N\mathbb{R}(J_i)$ as in Eq. (6), we compute $U(x)$ in explicitly using Eq. (19)

$$\begin{aligned}
U(x) &= U\left(\sum_{1\le\alpha_1\le J_1}\cdots\sum_{1\le\alpha_N\le J_N}x_\alpha e^{\alpha}_{\otimes_1^N\mathbb{R}(J_i)}\right)\\
&=\sum_{\alpha\in\otimes_{i=1}^N\mathbb{N}^+(J_i)}\sum_{\beta\in\otimes_{i=1}^N\mathbb{N}^+(I_i)}U(\beta,\alpha)x_\alpha e^{\beta,*}_{\otimes_1^N\mathbb{R}(I_i)}\otimes e^{\alpha}_{\otimes_1^N\mathbb{R}(J_i)}\\
&=\sum_{\substack{\alpha\in\otimes_{i=1}^N\mathbb{N}^+(J_i),\\\beta\in\otimes_{i=1}^N\mathbb{N}^+(I_i)}}U(\beta,\alpha)x_\alpha e^{\beta,*}_{\otimes_1^N\mathbb{R}(I_i)}\otimes e^{\alpha}_{\otimes_1^N\mathbb{R}(J_i)}.
\end{aligned}\tag{21}$$

Basis of $L(\otimes_1^N\mathbb{R}(J_i);\otimes_1^N\mathbb{R}(I_i))$

The space $L(\otimes_1^N\mathbb{R}(J_i);\otimes_1^N\mathbb{R}(I_i))$ is spanned by elements

$$\left\{\otimes_1^N u_i,\ u_i\in L(\mathbb{R}(J_i);\mathbb{R}(I_i)),\ 1\le i\le N\right\}.$$

The same does not hold in infinite dimensions [Gre67].

Products of Functionals

If we take $\mathbb{R}(I_i)=\mathbb{R}$ for all $i=1,\ldots,N$, and $\mu_i\in L(\mathbb{R}(J_i);\mathbb{R})=\mathbb{R}(J_i)^*$,

$$\mu_i=\sum_{1\le\alpha_i\le J_i}\mu_i(\alpha_i)e^{\alpha_i,*}_{\mathbb{R}(J_i)}\tag{22}$$

Previous calculations (cf. Eqs. (16) and (19)) show that the tensor product of $\overline{\mu}=\otimes_1^N\mu_i$ is given by

$$\begin{aligned}
\overline{\mu} &= \sum_{\alpha\in\otimes_{i=1}^N\mathbb{N}^+(J_i)}\prod_{i=1,\ldots,N}\mu_i(\alpha_i)e^{\alpha,*}_{\otimes_1^N\mathbb{R}(J_i)}\\
&=\sum_{\alpha\in\otimes_{i=1}^N\mathbb{N}^+(J_i)}\overline{\mu}(\alpha),
\end{aligned}\tag{23}$$

where $\overline{\mu}(\alpha)=\prod_{i=1,\ldots,N}\mu_i(\alpha_i)$ for all $\alpha\in\otimes_{i=1}^N\mathbb{N}^+(J_i)$. The tensor product $\overline{\mu}$ is identified with the N-linear function: $\overline{\mu}:\prod_{i=1,\ldots,N}\mathbb{R}(J_i)\to\mathbb{R}$, with

$$\overline{\mu}(x_1,\ldots,x_n)=\prod_{i=1,\ldots,N}\mu_i(x_i)\in\mathbb{R}\quad\forall x_i\in\mathbb{R}(J_i),\ i=1,\ldots,N.$$

3.5. Multilinear Functions

Let $Y\in\otimes_1^N\mathbb{R}(J_i)^*$, the space of N-linear functions defined on $\prod_{i=1,\ldots,N}\mathbb{R}(J_i)$ and identified as an element in $L(\otimes_1^N\mathbb{R}(J_i);\mathbb{R})=\otimes_1^N\mathbb{R}(J_i)^*$. The latter has basis

$$\left\{e^{\alpha,*}_{\otimes_1^N\mathbb{R}(J_i)}=e^{\alpha_1,*}_{\mathbb{R}(J_1)}\otimes\cdots\otimes e^{\alpha_N,*}_{\mathbb{R}(J_N)},\ \alpha\in\otimes_{i=1}^N\mathbb{N}^+(J_i)\right\}.\tag{24}$$

The set of numbers $\{Y(\gamma)\}_{\gamma\in\otimes_{i=1}^N\mathbb{N}^+(J_i)}$ define Y, where

$$Y=\sum_{\gamma\in\otimes_{i=1}^N\mathbb{N}^+(J_i)}Y(\gamma)e^{\gamma,*}_{\otimes_1^N\mathbb{R}(J_i)}\quad.\tag{25}$$

Suppose $x\in\otimes_1^N\mathbb{R}(J_i)$ is given by Eq. (6), then $Y(x)$ has the following satisfying formula:

$$Y(x)=\sum_{\gamma\in\otimes_{i=1}^N\mathbb{N}^+(J_i)}Y(\gamma)x_\gamma\quad.$$

This forms a scalar product between $\otimes_1^N\mathbb{R}(J_i)$ and $\otimes_1^N\mathbb{R}(J_i)^*$, since $\otimes_1^N\mathbb{R}(I_i)^*\cong\otimes_1^N\mathbb{R}(I_i)^*=\mathbb{R}(I_1)^*\otimes\cdots\otimes\mathbb{R}(I_N)^*$ [Gre67, Chp 1], we have the scalar product

$$\begin{aligned}
&\langle y^{1,*}\otimes\cdots\otimes y^{N,*},\ x_1\otimes\cdots\otimes x_N\rangle_{\otimes_1^N\mathbb{R}(I_i)^*,\otimes_1^N\mathbb{R}(I_i)}\\
&=\prod_{i=1,\ldots,N}\langle y^{i,*},x_i\rangle_{\mathbb{R}(I_i)^*,\mathbb{R}(I_i)}\quad\forall y^{i,*}\in\mathbb{R}(I_i)^*,\ x_i\in\mathbb{R}(J_i),\\
&\hfill\forall i=1,\ldots,N.
\end{aligned}\tag{26}$$

Now, let $Y\in\otimes_1^N\mathbb{R}(I_i)^*$, and $U=\otimes_1^N u_i\in L(\otimes_1^N\mathbb{R}(J_i);\otimes_1^N\mathbb{R}(I_i))$ where each u_i is as in Eq. (16). Suppose that

$$Y(x)=\sum_{\gamma\in\otimes_{i=1}^N\mathbb{N}^+(I_i)}Y(\gamma)e^{\gamma,*}_{\otimes_1^N\mathbb{R}(I_i)}\quad.\tag{27}$$

We compute the adjoint $U^*Y\in\otimes_1^N\mathbb{R}(J_i)^*$ in coordinates. For any $\alpha\in\otimes_{i=1}^N\mathbb{N}^+(J_i)$ we compute $U^*Y(\alpha)=$

$U^*Y(e^{\alpha}_{\otimes_1^N \mathbb{R}(J_i)})$:

$$
\begin{aligned}
U^*Y(\alpha) &= \langle Y,\, U(e^{\alpha}_{\otimes_1^N \mathbb{R}(J_i)})\rangle_{\otimes_1^N \mathbb{R}(I_i)^*,\, \otimes_1^N \mathbb{R}(I_i)}\\
&= \langle Y, \sum_{\beta\in\otimes_{i=1}^N \mathbb{N}^+(I_i)} U(\beta,\alpha) e^{\beta}_{\otimes_1^N \mathbb{R}(I_i)}\rangle_{\otimes_1^N \mathbb{R}(I_i)^*,\, \otimes_1^N \mathbb{R}(I_i)}\\
&= \sum_{\gamma\in\otimes_{i=1}^N \mathbb{N}^+(I_i)} \sum_{\beta\in\otimes_{i=1}^N \mathbb{N}^+(J_i)} \langle Y(\gamma) e^{\gamma,*}_{\otimes_1^N \mathbb{R}(I_i)},\, U(\beta,\alpha) e^{\beta}_{\otimes_1^N \mathbb{R}(I_i)}\rangle_{\otimes_1^N \mathbb{R}(I_i)^*,\, \otimes_1^N \mathbb{R}(I_i)}\\
&= \sum_{\substack{\gamma\in\otimes_{i=1}^N \mathbb{N}^+(I_i),\\ \beta\in\otimes_{i=1}^N \mathbb{N}^+(J_i)}} Y(\gamma)U(\beta,\alpha)\langle e^{\gamma,*}_{\otimes_1^N \mathbb{R}(I_i)},\, e^{\beta}_{\otimes_1^N \mathbb{R}(I_i)}\rangle_{\otimes_1^N \mathbb{R}(I_i)^*,\, \otimes_1^N \mathbb{R}(I_i)}\\
&= \sum_{\substack{\gamma\in\otimes_{i=1}^N \mathbb{N}^+(I_i),\\ \beta\in\otimes_{i=1}^N \mathbb{N}^+(J_i)}} Y(\gamma)U(\beta,\alpha) \prod_{i=1,\dots,N} \overline{\delta}(\gamma_i,\beta_i)\\
&= \sum_{\substack{\gamma\in\otimes_{i=1}^N \mathbb{N}^+(I_i),\\ \beta\in\otimes_{i=1}^N \mathbb{N}^+(J_i)}} Y(\gamma)U(\beta,\alpha)\overline{\delta}(\gamma,\beta) = \sum_{\gamma\in\otimes_{i=1}^N \mathbb{N}^+(I_i)} Y(\gamma)U(\gamma,\alpha) \quad .
\end{aligned} \tag{28}
$$

From Eq. (28), we can read off the coefficient $(U^*Y)(\alpha)$

$$
(U^*Y)(\alpha) = \sum_{\gamma\in\otimes_{i=1}^N \mathbb{N}^+(I_i)} Y(\gamma)U(\gamma,\alpha), \tag{29}
$$

and the N-linear mapping $U^*Y \in \otimes_1^N \mathbb{R}(I_i)^*$ can be written

$$
U^*Y = \sum_{\substack{\gamma\in\otimes_{i=1}^N \mathbb{N}^+(I_i),\\ \alpha\in\otimes_{i=1}^N \mathbb{N}^+(J_i)}} Y(\gamma)U(\gamma,\alpha) e^{\alpha,*}_{\otimes_1^N \mathbb{R}(J_i)} \quad . \tag{30}
$$

It is fruitful to also consider $U^* \in L(\otimes_1^N \mathbb{R}(I_i)^*;\ \otimes_1^N \mathbb{R}(J_i)^*)$ and its coordinate representation with respect to the basis:

$$
\left\{ e^{\alpha,*}_{\otimes_1^N \mathbb{R}(J_i)} \otimes e^{\gamma}_{\otimes_1^N \mathbb{R}(I_i)},\ \alpha\in\otimes_{i=1}^N \mathbb{N}^+(J_i),\ \gamma\in\otimes_{i=1}^N \mathbb{N}^+(I_i)\right\}.
$$
[5]

We claim that $(\otimes_1^N u_i^*) = (\otimes_1^N u_i)^* = U^*$ and

$$
U^*(\alpha,\beta) = \prod_{i=1,\dots,N} u_i^*(\alpha_i,\beta_i) \quad .
$$

According to Eq. (19), for any $\beta \in \otimes_{i=1}^N \mathbb{N}^+(I_i)$,

$$
\begin{aligned}
U^*(e^{\beta,*}_{\otimes_1^N \mathbb{R}(I_i)}) &= \sum_{\alpha\in\otimes_{i=1}^N \mathbb{N}^+(J_i)} U(\beta,\alpha)\\
&= \sum_{\alpha\in\otimes_{i=1}^N \mathbb{N}^+(J_i)} \prod_{i=1,\dots,N} u_i(\beta_i,\alpha_i)\\
&= \sum_{\alpha\in\otimes_{i=1}^N \mathbb{N}^+(J_i)} \prod_{i=1,\dots,N} u_i^*(\alpha_i,\beta_i)\\
&= (\otimes_1^N u_i^*)(e^{\beta,*}_{\otimes_1^N \mathbb{R}(I_i)}) \quad .
\end{aligned}
$$

[5]Note that $e^{\gamma,**}_{\otimes_1^N \mathbb{R}(I_i)} = e^{\gamma}_{\otimes_1^N \mathbb{R}(I_i)}$ under the canonical identification of a vector space with its bidual, since all vector spaces under consideration are of finite dimension [**Brezis2010Functional**].

3.6. Matrization

If $\mu_i = \sum_{1\le\alpha_i\le J_i} \mu_i(\alpha_i) e^{\alpha_i,*}_{\mathbb{R}(J_i)}$ is an element in $\mathbb{R}(J_i)^*$, its *Riesz vector* is the unique vector in $\mathbb{R}(J_i)$ that satisfies

$$
\langle \mu_i,\, x_i\rangle_{\mathbb{R}(J_i)^*,\,\mathbb{R}(J_i)} = \langle \mu_i^\wedge,\, x_i\rangle_{\mathbb{R}(J_i)} \quad \forall x_i \in \mathbb{R}(J_i).
$$

The basis $\{e^{\alpha_i}_{\mathbb{R}(J_i)}\}_{\alpha_i=1}^{J_i}$ is assumed to be orthonormal, if $x_i = e^{\gamma_i}_{\mathbb{R}(J_i)}$, for $\gamma_i \in \mathbb{N}^+(J_i)$, then

$$
\langle \mu_i,\, e^{\gamma_i}_{\mathbb{R}(J_i)}\rangle_{\mathbb{R}(J_i)^*,\,\mathbb{R}(J_i)} = \mu_i(\gamma_i) = \langle \mu_i^\wedge,\, e^{\gamma_i}_{\mathbb{R}(J_i)}\rangle_{\mathbb{R}(J_i)} = {\mu_i}^\wedge{}_{\gamma_i}.
$$

This transforms a dual vector, or *covector* into a vector in $\mathbb{R}(J_i)$. Similarly, given a covector $Y \in \otimes_1^N \mathbb{R}(J_i)^*$ is given as in Eq. (25), we define the *substitution operator* that holds all but one coordinate fixed in Y. For an arbitrary $j = 1,\dots,N$,

$$
\mathcal{S}_j^* : \left[\otimes_{\substack{i=1,\\ i\ne j}}^N \mathbb{R}(J_i)\right] \times \left[\otimes_1^N \mathbb{R}(J_i)^*\right] \to \mathbb{R}(J_j)^*.
$$

If $x_i \in \mathbb{R}(J_i)$ for $i = 1,\dots,N$,

$$
\left\langle \mathcal{S}_j^*(\otimes_{\substack{i=1,\\ i\ne j}}^N x_i,\, Y),\, x_j \right\rangle_{\mathbb{R}(J_i)^*,\,\mathbb{R}(J_j)} = \left\langle Y,\, \otimes_1^N x_i\right\rangle_{\otimes_1^N \mathbb{R}(J_i)^*,\,\otimes_1^N \mathbb{R}(J_i)}. \tag{31}
$$

Using ${\mathcal{S}_j}^*{}_Y = \mathcal{S}_j^*(\cdot,\, Y) \in L(\otimes_{\substack{i=1,\\ i\ne j}}^N \mathbb{R}(J_i);\ \mathbb{R}(J_j)^*)$, we can compute its using Eq. (31). To wit, let us agree to use λ to index the basis vectors of $\otimes_{\substack{i=1,\\ i\ne j}}^N \mathbb{R}(J_i)$, where

$$
\left\{ e^{\lambda}_{\otimes_{\substack{i=1,\\ i\ne j}}^N \mathbb{R}(J_i)} = \otimes_{\substack{i=1,\\ i\ne j}}^N e^{\lambda_i}_{\mathbb{R}(J_i)},\ \lambda \in \otimes_{\substack{i=1,\\ i\ne j}}^N \mathbb{N}^+(J_i)\right\} \tag{32}
$$

forms a basis of $\otimes_{\substack{i=1,\\ i\ne j}}^N \mathbb{R}(J_i)$; and $\alpha_j \in \mathbb{N}^+(J_i)$ to enumerate the basis vectors of $\mathbb{R}(J_j)$. Suppose that

$$
{\mathcal{S}_j}^*{}_Y = \sum_{\substack{\lambda\in\otimes_{i=1, i\ne j}^N \mathbb{N}^+(J_i),\\ 1\le\alpha_j\le J_j}} {\mathcal{S}_j}^*{}_Y(\alpha_j,\lambda) \left[e^{\alpha_j,*}_{\mathbb{R}(J_j)}\right] \otimes \left[\otimes_{\substack{i=1,\\ i\ne j}}^N e^{\lambda_i}_{\mathbb{R}(J_i)}\right]. \tag{33}
$$

Let $\lambda \in \otimes_{\substack{i=1,\\ i\ne j}}^N \mathbb{N}^+(J_i)$ and $\alpha_j \in \mathbb{N}^+(J_i)$, we see that

$$
\begin{aligned}
&{\mathcal{S}_j}^*{}_Y(\alpha_j,\lambda)\\
&= \left\langle {\mathcal{S}_j}^*{}_Y(\otimes_{\substack{i=1,\\ i\ne j}}^N e^{\lambda_i}_{\mathbb{R}(J_i)}),\, e^{\alpha_j}_{\mathbb{R}(J_j)} \right\rangle_{\mathbb{R}(J_i)^*,\,\mathbb{R}(J_j)}\\
&= \left\langle Y,\, \left[\otimes_1^{j-1} e^{\lambda_i}_{\mathbb{R}(J_i)}\right] \otimes \left[e^{\alpha_j}_{\mathbb{R}(J_j)}\right] \otimes \left[\otimes_{j+1}^N e^{\lambda_i}_{\mathbb{R}(J_i)}\right]\right\rangle_{\otimes_1^N \mathbb{R}(J_i)^*,\,\otimes_1^N \mathbb{R}(J_i)}\\
&= Y\Big((\lambda_{\underline{j-1}},\, \alpha_j\,,\lambda_{j+\underline{N-j}})\Big).
\end{aligned} \tag{34}
$$

It is convenient to rewrite Equation (34) using less clumsy notation: [6]

$$\mathcal{S}_j{}^*{}_Y(\alpha_j, \lambda) = Y(\alpha_j \lrcorner \lambda) \quad \text{where} \quad \alpha_j \lrcorner \lambda = \left(\lambda_{\underline{j-1}}, \alpha_j, \lambda_{j+\underline{N-j}}\right).$$

We define the *matrization* of a tensor $Y \in \otimes_1^N \mathbb{R}(J_i)^*$ at the j index ($j = 1, \dots, N$) as the **linear map**

$$Y_{(j)} : \otimes_{\substack{i=1,\\ i\neq j}}^N \mathbb{R}(J_i) \to \mathbb{R}(J_j)$$

that post-composes $\mathcal{S}_j{}^*{}_Y$ by the Riesz map of $\mathbb{R}(J_j)$. According to Equations (31) and (34), it has coordinate representation

$$Y_{(j)} = \sum_{\substack{\lambda \in \otimes_{i=1,}^N \mathbb{R}(J_i),\\ i\neq j\\ 1\leq \alpha_j \leq J_j}} Y(\alpha_j \lrcorner \lambda)\left[e^{\alpha_j}_{\mathbb{R}(J_j)}\right] \otimes \left[\otimes_{\substack{i=1,\\ i\neq j}}^N e^{\lambda_i,*}_{\mathbb{R}(J_i)}\right]$$

3.7. Notes and References

In general, the Tracy-Singh ordering prescribes a natural ordering to subsets, suppose $\{(A_i, \leq_i)\}_{i=1}^N$ is a collection of linearly ordered subsets, and $A_i = \bigsqcup_{j=1}^{M_i} A(i,j)$. The usual way of ordering the disjoint union $\mathbf{A} = \bigsqcup_{i=1}^N A_i$ is to write

$$x \leq y \in \mathbf{A} \iff \begin{cases} x, y \in A_i \text{ and } x \leq_i y & \text{or} \\ x \in A_i,\ y \in A_j & 1 \leq i < j \leq N. \end{cases}$$

Alternatively, we can prescribe an ordering to $\mathbf{A}$ that preserves the subset structure.

$$x \leq y \in \mathbf{A} \iff \left\{ x_i \leq y_i \right.$$

The substitution operator $\mathcal{S}_j$ for $j = 1, \dots, N$ is bilinear in its two arguments, and descends into a linear map also denoted by $\mathcal{S}_j$

$$\mathcal{S}_j \in L\left(\left[\otimes_{\substack{i=1,\\ i\neq j}}^N \mathbb{R}(J_i)\right] \otimes \left[\otimes_1^N \mathbb{R}(J_i)^*\right];\ \mathbb{R}(J_i)^*\right).$$

For the case of vector spaces over $\mathbb{C}$, it is customary to suppress the complex conjugate of the Riesz map, so that matrization is a linear correspondence from $\otimes_1^N \mathbb{C}(J_i)^*$ into $L(\otimes_{\substack{i=1,\\ i\neq j}}^N \mathbb{C}(J_i);\ \mathbb{C}(J_j))$. We also have the following relationship between the Riesz adjoint and the tensor product.

> Let $\{\mathbb{R}(J_i)\}_1^N$ be inner product spaces, then the tensor product of the Riesz map is the Riesz map of the tensor product.

[6] In other words, the functional induced by Y with λ held fixed has coordinates that are equal to the coordinates of Y with λ held fixed.

We give an example of the border rank of a tensor, but first an equation for writing the kernel of a tensor. Let $\varphi : \prod_1^N \mathbb{R}(J_i) \to W$ be multilinear, then

$$\varphi(b^{\underline{N}}) - \varphi(a^{\underline{N}}) = \sum_{i=\underline{N}} \varphi(b^{\underline{i-1}}, \Delta_i, a^{i+\underline{N-i}}), \tag{35}$$

for all $(a^{\underline{N}}), (b^{\underline{N}}) \in \prod_1^N \mathbb{R}(J_i)$, and $\Delta^i = b^i - a^i$ for $i = 1, \dots, N$.

Proof. We proceed by induction, and Eq. (35) follows by setting $m = k$ in

$$\varphi(a^{\underline{N}}) = \varphi(b^{\underline{m}}, a^{m+\underline{k-m}}) - \sum_{i=\underline{m}} \varphi(b^{\underline{i-1}}, \Delta_i, a^{i+\underline{k-i}}) \tag{36}$$

Base case: set $m = 1$, by definition of k-linearity (**??**) of φ. Since $a^1 = b^1 - \Delta_1$,

$$\begin{aligned} \varphi(a^{\underline{N}}) &= \varphi(b^1 - \Delta_1, a^{1+\underline{N-1}}) \\ &= \varphi(b^1, a^{1+\underline{N-1}}) - \varphi(\Delta_1, a^{1+\underline{N-1}}). \end{aligned}$$

Suppose Eq. (36) holds for m, since $a^{m+1} = b^{m+1} - \Delta_{m+1}$,

$$\begin{aligned} &\varphi(a^{\underline{N}}) \\ &= \varphi(b^{\underline{m}}, a^{m+\underline{N-m}}) - \sum_{i=\underline{m}} \varphi(b^{\underline{i-1}}, \Delta_i, a^{i+\underline{N-i}}) \\ &= \varphi(b^{\underline{m}}, a^{m+1}, a^{(m+1)+\underline{N-(m+1)}}) - \sum_{i=\underline{m}} \varphi(b^{\underline{i-1}}, \Delta_i, a^{i+\underline{N-i}}) \\ &= \varphi(b^{\underline{m+1}}, a^{(m+1)+\underline{k-(m+1)}}) \\ &\quad - \varphi(b^{\underline{m+1}}, \Delta_{m+1}, a^{(m+1)+\underline{N-(m+1)}}) \\ &\quad - \sum_{i=\underline{m}} \varphi(b^{\underline{i-1}}, \Delta_i, a^{i+\underline{N-i}}) \end{aligned}$$

and this proves Eq. (35). ■

References

[Gre12] Werner H. Greub. *Linear algebra*. [Online; accessed 2024-02-22]. Springer Science & Business Media, Dec. 2012.

[Gre67] Werner Hildbert Greub. *Multilinear algebra [by] W. H. Greub*. [Online; accessed 2024-02-22]. Springer, 1967.

[Lan05] Serge Lang. *Algebra*. Springer Science & Business Media, 2005.

[Hal12] P.R. Halmos. *Finite-Dimensional Vector Spaces*. Springer Science & Business Media, Dec. 2012. ISBN: 9781461263876.

[Joh21a] Nathaniel Johnston. *Introduction to Linear and Matrix Algebra*. Springer Nature, May 2021. ISBN: 9783030528119.

[Joh21b] Nathaniel Johnston. *Advanced Linear and Matrix Algebra*. Springer Nature, May 2021. ISBN: 9783030528157.

[Lan12] Serge Lang. *Introduction to Linear Algebra*. Springer Science & Business Media, Dec. 2012. ISBN: 9781461210702.

[Str93] Gilbert Strang. *Introduction to Linear Algebra*. Wellesley College, 1993. ISBN: STANFORD:36105007523389.

[Axl97] Sheldon Axler. *Linear Algebra Done Right*. Springer Science & Business Media, July 1997. ISBN: 0387982590.

[Hun03] Thomas W. Hungerford. *Algebra*. Springer Science & Business Media, Feb. 2003. ISBN: 0387905189.

[HJ90] Roger A. Horn and Charles R. Johnson. *Matrix analysis*. [Online; accessed 2024-02-22]. Cambridge University Press, Feb. 1990.

[HHJ94] Roger A. Horn, Roger A Horn, and Charles R. Johnson. *Topics in matrix analysis*. [Online; accessed 2024-02-22]. Cambridge University Press, June 1994.

[AM06] Panos J. Antsaklis and Anthony N. Michel. *Linear Systems*. Springer Science & Business Media, Nov. 2006. ISBN: 9780817644352.

[Rie17] Emily Riehl. *Category Theory in Context*. Courier Dover Publications, 2017.

[TJ89] Derrick S. Tracy and Kankanam G. Jinadasa. "Partitioned kronecker products of matrices and applications". In: *Canadian Journal of Statistics* 17.1 (Mar. 1989), pp. 107–120. DOI: 10.2307/3314768.

[Liu99] Shuangzhe Liu. "Matrix results on the Khatri-Rao and Tracy-Singh products". In: *Linear Algebra and its Applications* 289.13 (Mar. 1999), pp. 267–277. DOI: 10.1016/s0024-3795(98)10209-4.

[Tuc66] Ledyard R. Tucker. "Some mathematical notes on three-mode factor analysis". In: *Psychometrika* 31.3 (1966), pp. 279–311. DOI: 10.1007/BF02289464.