Tensor spaces – the basics

S. Gill Williamson

Abstract

We present the basic concepts of tensor products of vectors spaces, exploiting the special properties of vector spaces as opposed to more general modules: Introduction (1), Basic multilinear algebra (2), Tensor products of vector spaces (3), Tensor products of matrices (4), Inner products on tensor spaces (5), Direct sums and tensor products (6), Background concepts and notation (7), Discussion and acknowledgements (8). This material draws upon [Mar73] and relates to subsequent work [Mar75].

1. Introduction

We start with an example. Let $x = \binom{a}{b} \in \mathbf{M}_{2,1}$ and $y = (\,c\;\; d\,) \in \mathbf{M}_{1,2}$ where $\mathbf{M}_{m,n}$ denotes the $m \times n$ matrices over the real numbers, $\mathbb{R}$.

The set of matrices $\mathbf{M}_{m,n}$ forms a *vector space* under matrix addition and multiplication by real numbers, $\mathbb{R}$. The *dimension* of this vector space is mn. We write $\dim(\mathbf{M}_{m,n}) = mn$.

Define a function $\nu(x,y) = xy = \begin{pmatrix} ac & ad \\ bc & bd \end{pmatrix} \in \mathbf{M}_{2,2}$ (matrix product of x and y). The function ν has *domain* $V_1 \times V_2$ where $V_1 = \mathbf{M}_{2,1}$ and $V_2 = \mathbf{M}_{1,2}$ are vector spaces of dimension, $\dim(V_i) = 2$, $i = 1, 2$. The range of ν is the vector space $P = \mathbf{M}_{2,2}$ which has $\dim(P) = 4$.

The function ν is *bilinear* in the following sense:

$$\nu(r_1x_1 + r_2x_2, y) = r_1\nu(x_1, y) + r_2\nu(x_2, y) \tag{1.1}$$

$$\nu(x, r_1y_1 + r_2y_2) = r_1\nu(x, y_1) + r_2\nu(x, y_2) \tag{1.2}$$

for any $r_1, r_2 \in \mathbb{R}$, $x, x_1, x_2 \in V_1$, and $y, y_1, y_2 \in V_2$. We denote the set of *all* such bilinear functions by $M(V_1, V_2 : P)$. Recall that the *image* of ν is the set $\mathrm{Im}(\nu) := \{\nu(x,y) \,|\, (x,y) \in V_1 \times V_2\}$ and the *span* of the image of ν, denoted by $\langle \mathrm{Im}(\nu) \rangle$, is the set of all linear combinations of vectors in $\mathrm{Im}(\nu)$.

Let $E_1 = \{e_{11}, e_{12}\}$ be an ordered basis for V_1 and $E_2 = \{e_{21}, e_{22}\}$ be an ordered basis for V_2 specified as follows:

$$e_{11} = \begin{pmatrix} 1 \\ 0 \end{pmatrix},\ e_{12} = \begin{pmatrix} 0 \\ 1 \end{pmatrix},\ e_{21} = \begin{pmatrix} 1 & 0 \end{pmatrix},\ e_{22} = \begin{pmatrix} 0 & 1 \end{pmatrix}. \tag{1.3}$$

Using matrix multiplication, we compute

$$\begin{aligned} \nu(e_{11}, e_{21}) &= e_{11}e_{21} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, & \nu(e_{11}, e_{22}) &= e_{11}e_{22} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \\ \nu(e_{12}, e_{21}) &= e_{12}e_{21} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, & \nu(e_{12}, e_{22}) &= e_{12}e_{22} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}. \end{aligned} \tag{1.4}$$

Note that the set $\{\nu(e_{11}, e_{21}), \nu(e_{11}, e_{22}), \nu(e_{12}, e_{21}), \nu(e_{12}, e_{22})\} \subset \operatorname{Im}(\nu)$ is a basis for P. Since $\operatorname{Im}(\nu)$ contains a basis for P, we have $\langle \operatorname{Im}(\nu) \rangle = P$ (i.e., the span of the image of ν equals P). Here is our basic definition of the tensor product of two vector spaces:

Definition 1.5 (The tensor product of two vector spaces). Let V_1 and V_2 be vector spaces over $\mathbb{R}$ with $\dim(V_1) = n_1$ and $\dim(V_2) = n_2$. Let P be a vector space over $\mathbb{R}$ and $\nu \in M(V_1, V_2 : P)$. Then, (P, ν) is a tensor product of the V_1, V_2 if

$$\langle \operatorname{Im}(\nu) \rangle = P \quad \text{and} \quad \dim(P) = \prod_{i=1}^{2} \dim(V_i). \tag{1.6}$$

We sometimes use $x_1 \otimes x_2 := \nu(x_1, x_2)$ and call $x_1 \otimes x_2$ the *tensor product* of x_1 and x_2. We sometimes use $V_1 \otimes V_2$ for P.

1.7 (**Remarks and intuition about tensor products**). Note that our example, $\nu(x, y) = xy = \begin{pmatrix} ac & ad \\ bc & bd \end{pmatrix} \in \mathbf{M}_{2,2}$ defines a tensor product $(\mathbf{M}_{2,2}, \nu)$ by definition 1.5. Using "tensor" notation, we could write $\begin{pmatrix} a \\ b \end{pmatrix} \otimes \begin{pmatrix} c & d \end{pmatrix} = \begin{pmatrix} ac & ad \\ bc & bd \end{pmatrix}$ and $\mathbf{M}_{2,1} \otimes \mathbf{M}_{1,2} = \mathbf{M}_{2,2}$. A tensor product is simply a special bilinear (multilinear) function ν. To specify any function you must specify its domain and range. Thus the V_1, V_2, $\operatorname{Im}(\nu)$ and, hence, $\langle \operatorname{Im}(\nu) \rangle$ are defined when you specify ν. In this sense the definition of a tensor needs only ν, not the pair (P, ν). It is useful, however, to display P explicitly. The condition $\langle \operatorname{Im}(\nu) \rangle = P$ states that P isn't too big in some sense—every bit of P has some connection to ν. The condition $\dim(P) = \prod_{i=1}^{2} \dim(V_i)$ states that P isn't too small in some sense. The function ν has enough room to reveal all of its tricks. The notation $x_1 \otimes x_2$ is a convenient way to work with ν and not have to keep writing "ν" over and over again. The choice of the *model* P is important in different applications. The subject of multilinear algebra (or tensors) contains linear algebra as a special case, so it is not a trivial subject.

2. Basic multilinear algebra

Section 7 gives background material for review as needed. We focus on multilinear algebra over finite dimensional vector spaces. The default field, $\mathbb{F}$, has characteristic zero although much of the material is valid for finite fields. Having committed to this framework, we make use of the fact that vector spaces have bases. Tensors are special types of multilinear functions so we need to get acquainted with multilinear functions.

Definition 2.1 (Summary of vector space and algebra axioms). Let $\mathbb{F}$ be a field and let $(M,+)$ be an abelian group with identity θ. Assume there is an operation, $\mathbb{F} \times M \to M$, which takes (r,x) to rx (juxtaposition of r and x). To show that $(M,+)$ is a *vector space over* $\mathbb{F}$, we show the following four things hold for every $r,s \in \mathbb{F}$ and $x,y \in M$:

(1) $r(x+y) = rx + ry$ **(2)** $(r+s)x = rx + sx$ **(3)** $(rs)x = r(sx)$ **(4)** $1_{\mathbb{F}}\, x = x$

where $1_{\mathbb{F}}$ is the multiplicative identity in $\mathbb{F}$. If $(M,+,\cdot)$ is a ring for which $(M,+)$ is a vector space over $\mathbb{F}$, then $(M,+,\cdot)$ is an *algebra over* $\mathbb{F}$ if the following scalar rule holds: **(5)** for all $\alpha \in \mathbb{F}$, $x,y \in M$, $\alpha(xy) = (\alpha x)y = x(\alpha y)$.

2.2 (**The free $\mathbb{F}$-vector space over a finite set** K). We use the notation A^B to denote all functions with domain B and range A. Let $\mathbb{F}$ be a field and K a finite set. For $f,g \in \mathbb{F}^K$, $x \in K$, define $f+g$ by $(f+g)(x) := f(x) + g(x)$. For $r \in \mathbb{F}$, $f \in \mathbb{F}^K$, $x \in K$, define $(rf)(x) := rf(x)$. It is easy to check that $(F^K,+)$ is a vector space over $\mathbb{F}$ (see 2.1). The set of *indicator functions* $\{\iota_k \,|\, k \in K\}$ where $\iota_k(x) = 1$ if $x = k$, 0 otherwise, forms a basis for $\mathbb{F}^K$. Thus, $\mathbb{F}^K$ is a vector space of dimension $|K|$ (the cardinality of K). If a set K has some special interest for us (e.g, elements of a finite group, finite sets of graphs or other combinatorial objects) then studying the F-vector space over K sometimes yields new insights about K.

2.3 (**Frequently used notations**). We use $\mathcal{X}(statement) = 1$ if $statement$ is true, 0 if false. For short, $\delta_{\alpha,\gamma} := \mathcal{X}(\alpha = \gamma)$. Let $E_i = \{e_{i1}, e_{i2}, \ldots, e_{i\,n_i}\}$, $i = 1,\ldots,m$, be linearly ordered sets. The order is specified by the second indices, $1,2,\ldots,n_i$, thus, alternatively $E_i = (e_{i1}, e_{i2}, \ldots, e_{i\,n_i})$. For any integer m, we define "underline" notation by $\underline{m} := \{1,\ldots,m\}$. Let

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

The (Cartesian) product of the sets E_i is

$$E_1 \times \cdots \times E_m := \{(e_{1\gamma(1)},\ldots,e_{m\gamma(m)}) \,|\, \gamma \in \Gamma(n_1,\ldots,n_m)\} \tag{2.4}$$

where $\Gamma(n_1,\ldots,n_m) := \underline{n_1} \times \cdots \times \underline{n_m}$. We use the notations $\Gamma(n_1,\ldots,n_m)$ and $\underline{n_1} \times \cdots \times \underline{n_m}$ as convenience dictates. Using this notation, we have

$$\prod_{i=1}^{2}\left(\sum_{k=1}^{2} a_{ik}\right) = \sum_{\gamma \in \underline{2}\times\underline{2}} \prod_{i=1}^{2} a_{i\,\gamma(i)} = \sum_{\gamma \in \Gamma(2,2)} \prod_{i=1}^{2} a_{i\,\gamma(i)}. \tag{2.5}$$

The general form of this identity is

$$\prod_{i=1}^{m}\left(\sum_{k=1}^{n_i} a_{ik}\right) = \sum_{\gamma \in \Gamma(n_1,\ldots,n_m)} \prod_{i=1}^{m} a_{i\,\gamma(i)}. \tag{2.6}$$

This product-sum-interchange, identity 2.6, will be used frequently in what follows.

Definition 2.7. Let V_i, $i = 1,\dots,m$, and U be vector spaces over a field $\mathbb{F}$. Let $\times_1^m V_i$ be the Cartesian product of the V_i. A function $\varphi\colon \times_1^m V_i \to U$ is *multilinear* if for all $t \in \underline{m}$.

$$\varphi(\dots, cx_t + dy_t, \dots) = c\varphi(\dots, x_t, \dots) + d\varphi(\dots, y_t, \dots)$$

where $x_i, y_i \in V_i$ and $c, d \in \mathbb{F}$. The set of all such multilinear functions is denoted by $M(V_1, \dots, V_m : U)$. If $m = 1$, $M(V_1 : U) := \mathbb{L}(V_1, U)$, the set of linear functions from V_1 to U.

2.8 (**Multilinear expansion formula**). Let $E_i = \{e_{i1}, \dots, e_{in_i}\}$ be an ordered basis for V_i, $i = 1, \dots, m$. If $\varphi \in M(V_1, \dots, V_m : U)$ and $x_i = \sum_{k=1}^{n_i} c_{ik} e_{ik}$, $i = 1, \dots m$, we can, using the rules of multilinearity, compute $\varphi(x_1, \dots, x_m) =$

$$\varphi\left(\sum_{k=1}^{n_1} c_{1k}e_{1k}, \dots, \sum_{k=1}^{n_m} c_{mk}e_{mk}\right) = \sum_{\gamma\in\Gamma}\left(\prod_{i=1}^{m} c_{i\gamma(i)}\right)\varphi(e_{1\gamma(1)}, \dots, e_{m\gamma(m)}) \tag{2.9}$$

where $\Gamma := \Gamma(n_1, \dots, n_m)$ (see 2.3). Formula 2.9 can be proved by induction. It is a generalization of the standard algebraic rules for interchanging sums and products (2.6).

A function with domain $\Gamma(n_1, \dots, n_m)$ and range S is sometimes denoted by "maps to" notation: $\gamma \mapsto f(\gamma)$ or $\gamma \mapsto s_\gamma$. In the former case, f is the name of the function so we write $\mathrm{Im}(f)$ for the set $\{f(\gamma) \mid \gamma \in \Gamma\}$. In the latter case (index notation), the s is not consistently interpreted as the name of the function.

Definition 2.10 (Multilinear function defined by extension of $\gamma \mapsto s_\gamma$). Let V_i, $i = 1, \dots, m$, be finite dimensional vector spaces over a field F. Let $E_i = \{e_{i1}, \dots, e_{in_i}\}$ be an ordered basis for V_i, $i = 1, \dots, m$. Let U be a vector space over $\mathbb{F}$ and let $\gamma \mapsto s_\gamma$ be any function with domain $\Gamma(n_1, \dots, n_m)$ and range U. If $x_i = \sum_{k=1}^{n_i} c_{ik}e_{ik}$, $i = 1, \dots m$, then we can use 2.9 to define $\nu \in M(V_1, \dots, V_m : U)$ by

$$\nu(x_1, \dots, x_m) := \sum_{\gamma\in\Gamma}\left(\prod_{i=1}^{m} c_{i\gamma(i)}\right) s_\gamma. \tag{2.11}$$

We say $\nu \in M(V_1, \dots, V_m : U)$ is *defined by multilinear extension from $\gamma \mapsto s_\gamma$ and the bases E_i, $i = 1, \dots, m$.*

2.12 (**Examples of multilinear functions defined by bases**). We follow definition 2.10. Let $V_1 = \mathbb{R}^2$ and $V_2 = \mathbb{R}^3$, $\mathbb{R}$ the real numbers. Let $E_1 = \{e_{11}, e_{12}\}$ be the ordered basis for $\mathbb{R}^2$, $e_{11} = (1, 0)$, $e_{12} = (0, 1)$. Let $E_2 = \{e_{21}, e_{22}, e_{23}\}$ be the ordered basis for $\mathbb{R}^3$, $e_{21} = (1, 0, 0)$, $e_{22} = (0, 1, 0)$, $e_{23} =$

$(0,0,1)$. From equation 2.4, we have $E_1 \times E_2 = \{(e_{1\gamma(1)}, e_{2\gamma(2)}) \,|\, \gamma \in \Gamma(2,3)\}$. Let $U = \mathbf{M}_{2,3}(\mathbb{R})$. Define an injection $\gamma \mapsto b_\gamma \in U$ as follows:

$$\left\{ \begin{pmatrix} 1&0&0\\0&0&0 \end{pmatrix}_{11}, \begin{pmatrix} 0&1&0\\0&0&0 \end{pmatrix}_{12}, \begin{pmatrix} 0&0&1\\0&0&0 \end{pmatrix}_{13}, \begin{pmatrix} 0&0&0\\1&0&0 \end{pmatrix}_{21}, \begin{pmatrix} 0&0&0\\0&1&0 \end{pmatrix}_{22}, \begin{pmatrix} 0&0&0\\0&0&1 \end{pmatrix}_{23} \right\}. \tag{2.13}$$

The elements of the set 2.13 correspond to ordered pairs (γ, b_γ) in the form $(b_\gamma)_\gamma$ and thus define $\gamma \mapsto b_\gamma$. Using lexicographic order (7.2) on $\Gamma(2,3)$, the set 2.13 defines an ordered basis for U:

$$B = \left(\begin{pmatrix} 1&0&0\\0&0&0 \end{pmatrix}, \begin{pmatrix} 0&1&0\\0&0&0 \end{pmatrix}, \begin{pmatrix} 0&0&1\\0&0&0 \end{pmatrix}, \begin{pmatrix} 0&0&0\\1&0&0 \end{pmatrix}, \begin{pmatrix} 0&0&0\\0&1&0 \end{pmatrix}, \begin{pmatrix} 0&0&0\\0&0&1 \end{pmatrix} \right). \tag{2.14}$$

Write $B = (b_{11}, b_{12}, b_{13}, b_{21}, b_{22}, b_{23})$ where the b_{ij} are the corresponding entries in the list B of 2.14. Take $x_1 = c_{11}e_{11} + c_{12}e_{12} = (c_{11}, c_{12}) \in \mathbb{R}^2$. Take $x_2 = c_{21}e_{21} + c_{22}e_{22} + c_{23}e_{23} = (c_{21}, c_{22}, c_{23}) \in \mathbb{R}^3$. By equation 2.11 we get $\nu_B(x_1, x_2) =$

$$c_{11}c_{21}b_{11} + c_{11}c_{22}b_{12} + c_{11}c_{23}b_{13} + c_{12}c_{21}b_{21} + c_{12}c_{22}b_{22} + c_{12}c_{23}b_{23}. \tag{2.15}$$

Using the elements of B from equation 2.14, the sum shown in 2.15 becomes

$$\nu_B(x_1, x_2) = \begin{pmatrix} c_{11}c_{21} & c_{11}c_{22} & c_{11}c_{23} \\ c_{12}c_{21} & c_{12}c_{22} & c_{12}c_{23} \end{pmatrix}.$$

The nice pattern here where the first row is $c_{11}(c_{21}, c_{22}, c_{23})$ and the second row is $c_{12}(c_{21}, c_{22}, c_{23})$ is a consequence of the choice of correspondence between the elements of the basis B and Γ. For example, instead of equation 2.13 use

$$B = \left\{ \begin{pmatrix} 1&0&0\\0&0&0 \end{pmatrix}_{11}, \begin{pmatrix} 0&1&0\\0&0&0 \end{pmatrix}_{12}, \begin{pmatrix} 0&0&0\\1&0&0 \end{pmatrix}_{13}, \begin{pmatrix} 0&0&1\\0&0&0 \end{pmatrix}_{21}, \begin{pmatrix} 0&0&0\\0&1&0 \end{pmatrix}_{22}, \begin{pmatrix} 0&0&0\\0&0&1 \end{pmatrix}_{23} \right\}. \tag{2.16}$$

We then have the less attractive matrix representation

$$\nu_B(x_1, x_2) = \begin{pmatrix} c_{11}c_{21} & c_{11}c_{22} & c_{12}c_{21} \\ c_{11}c_{23} & c_{12}c_{22} & c_{12}c_{23} \end{pmatrix}.$$

The choice of indexing (i.e., bijective correspondence to $\Gamma(2,3)$) matters.

2.17 (**Component spaces of** $M(V_1, \ldots, V_m : U)$ **are** $M(V_1, \ldots, V_m : \mathbb{F})$**).** Suppose that $U = \langle u_1, \ldots, u_n \rangle$, where angle brackets, "$\langle\rangle$" denote the span of the basis $\{u_1, \ldots, u_n\}$ for U. Note that U is the direct sum $\oplus^n \langle u_i \rangle$ of the one-dimensional subspaces $\langle u_i \rangle$. The vector space U is isomorphic to the vector space $\mathbb{F}^{\underline{n}}$ and the $\langle u_i \rangle$ are isomorphic (as vector spaces over $\mathbb{F}$) to $\mathbb{F}$. The function $\varphi \in M(V_1, \ldots, V_m : U)$ has component functions $\varphi^{(i)}$:

$$\varphi(x_1, \ldots, x_m) = \sum_{i=1}^{n} \varphi^{(i)}(x_1, \ldots, x_m) u_i. \tag{2.18}$$

Each $\varphi^{(i)} \in M(V_1, \ldots, V_m : \langle u_i \rangle) \equiv M(V_1, \ldots, V_m : \mathbb{F})$.

2.19 (**Vector space of multilinear functions**). The set of all multilinear functions, $M(V_1,\dots,V_m : U)$, is a vector space over $\mathbb{F}$ in the standard way:

$$(\varphi_1+\varphi_2)(x_1,\dots,x_m) = \varphi_1(x_1,\dots,x_m)+\varphi_2(x_1,\dots,x_m)$$

and $(c\varphi)(x_1,\dots,x_m) = c\varphi(x_1,\dots,x_m)$. Suppose, for $i \in \underline{m}$, $V_i = \langle E_i\rangle$, $E_i = \{e_{i1},\dots,e_{in_i}\}$ a basis for V_i. Let $U = \langle u_1,\dots,u_n\rangle$ where $\{u_i \,|\, i \in \underline{n}\}$ is a basis. Define the component functions (2.18) $\varphi^{(i)} \in M(V_1,\dots,V_m : \langle u_i\rangle)$ by

$$\varphi^{(j)}_\alpha(e_{1\gamma(1)},\dots,e_{m\gamma(m)}) := \delta_{\alpha,\gamma}u_j. \tag{2.20}$$

The set of $n\prod_1^m n_i$ multilinear functions

$$\{\varphi^{(j)}_\alpha \,|\, \alpha \in \Gamma(n_1,\dots,n_m),\ j \in \underline{n}\} \tag{2.21}$$

is a basis for the vector space $M(V_1,\dots,V_m : U)$. See 2.23 for discussion. Note that as vector spaces

$$M(V_1,\dots,V_m : U) = \bigoplus_{j=1}^{n} M(V_1,\dots,V_m : \left\langle u_j\right\rangle). \tag{2.22}$$

2.23 (**The multilinear function basis**). It suffices to take $U = \mathbb{F}$ (we work with the component functions, see 2.17). We show $\{\varphi_\alpha \,|\, \alpha \in \Gamma(n_1,\dots,n_m)\}$ is a basis for $M(V_1,\dots,V_m : \mathbb{F})$ where $\varphi_\alpha(e_{1\gamma(1)},\dots,e_{m\gamma(m)}) = \delta_{\alpha,\gamma}1_{\mathbb{F}}$.

Here, $E_i = \{e_{i1},\dots,e_{in_i}\}$ as in 2.19. If $\varphi \in M(V_1,\dots,V_m : \mathbb{F})$ and $x_i = \sum_{k=1}^{n_i} c_{ik}e_{ik}$ we have

$$\varphi(x_1,\dots,x_m) = \sum_{\gamma\in\Gamma}\left(\prod_{i=1}^{m} c_{i\gamma(i)}\right)\varphi(e_{1\gamma(1)},\dots,e_{m\gamma(m)}). \tag{2.24}$$

Replacing φ by φ_α in equation 2.24 we obtain

$$\varphi_\alpha(x_1,\dots,x_m) = \sum_{\gamma\in\Gamma}\left(\prod_{i=1}^{n_i} c_{i\gamma(i)}\right)\varphi_\alpha(e_{1\gamma(1)},\dots,e_{m\gamma(m)}) = \prod_{i=1}^{m} c_{i\alpha(i)}. \tag{2.25}$$

From equations 2.24 and 2.25 we obtain

$$\varphi(x_1,\dots,x_m) = \sum_{\alpha\in\Gamma}\varphi_\alpha(x_1,\dots,x_m)\varphi(e_{1\alpha(1)},\dots,e_{m\alpha(m)}). \tag{2.26}$$

Thus, from 2.26, we obtain

$$\varphi = \sum_{\alpha\in\Gamma}\varphi(e_{1\alpha(1)},\dots,e_{m\alpha(m)})\varphi_\alpha \tag{2.27}$$

which can be proved by evaluating both sides at $(x_1,\dots,x_m)$. It is obvious that $\{\varphi_\alpha \,|\, \alpha \in \Gamma(n_1,\dots,n_m)\}$ is linearly independent for if $\varphi = \sum_\alpha d_\alpha\varphi_\alpha = 0$ then $\varphi(e_{1\gamma(1)},\dots,e_{m\gamma(m)}) = d_\gamma = 0$ for all γ.

Definition 3.1 (Tensor product). Let $V_1, \ldots, V_m$, be vector spaces over $\mathbb{F}$ of dimension $\dim(V_i)$, $i \in \underline{m}$. Let P be a vector space over $\mathbb{F}$. Then the pair (P, ν) is a *tensor product* of the V_i if

$$\nu \in M(V_1, \ldots, V_m : P) \text{ and } \langle \mathrm{Im}(\nu) \rangle = P \text{ and } \dim(P) = \prod_{i=1}^{m} \dim(V_i) \tag{3.2}$$

where $\langle \mathrm{Im}(\nu) \rangle := \left\langle \nu(x_1, \ldots, x_m) \,|\, (x_1, \ldots, x_m) \in \times_{i=1}^{m} V_i \right\rangle$ is the span of $\mathrm{Im}(\nu)$, the *image* of ν. The vector space P is denoted by $V_1 \otimes \cdots \otimes V_m$. The vectors $\nu(x_1, \ldots, x_m) \in P$ are designated by $x_1 \otimes \cdots \otimes x_m$ and are called the *homogeneous* tensors. The notation $\otimes(x_1, \ldots, x_m) := \nu(x_1, \ldots, x_m)$ is sometimes used.

Lemma 3.3 (Canonical bases). *Let $V_1, \ldots, V_m$, $\dim(V_i) = n_i$, $i \in \underline{m}$, be vector spaces over $\mathbb{F}$. Let P be a vector space over $\mathbb{F}$ and $\nu \in M(V_1, \ldots, V_m : P)$. Suppose, for all $i \in \underline{m}$, $V_i = \langle E_i \rangle$, $E_i = \{e_{i1}, \ldots, e_{in_i}\}$ an ordered basis for V_i. Then, (P, ν) is a tensor product of the V_i if and only if*

$$\mathbb{B} := \{\nu(e_{1\gamma(1)}, \ldots, e_{m\gamma(m)}) \,|\, \gamma \in \underline{n_1} \times \cdots \times \underline{n_m}\} \text{ is a basis for } P. \tag{3.4}$$

The standard convention is that $\mathbb{B}$ is ordered using lexicographic order on Γ.

Proof. Suppose (P, ν) is a tensor product and assume $x_i = \sum_{k=1}^{n_i} c_{ik} e_{ik}$ so that

$$\nu(x_1, \ldots, x_m) = \sum_{\gamma \in \Gamma} \left(\prod_{i=1}^{m} c_{i\gamma(i)} \right) \nu(e_{1\gamma(1)}, \ldots, e_{m\gamma(m)}). \tag{3.5}$$

By definition 3.1, $\langle \mathrm{Im}(\nu) \rangle = \left\langle \nu(x_1, \ldots, x_m) \,|\, (x_1, \ldots, x_m) \in \times_{i=1}^{m} V_i \right\rangle = P$. Thus, $\mathbb{B} = \{\nu(e_{1\gamma(1)}, \ldots, e_{m\gamma(m)}) \,|\, \gamma \in \underline{n_1} \times \cdots \times \underline{n_m}\}$ spans P. Thus, $|\mathbb{B}| = \dim(P)$ and $\mathbb{B}$ is a basis for P. Conversely, if $\mathbb{B}$ is a basis for P then clearly $\langle \mathrm{Im}(\nu) \rangle = P$ and $\dim(P) = |\mathbb{B}| = \prod_{i=1}^{m} n_i$. □

3.6 (**The space P of (P, ν) can be any P with $\dim(P) = \prod_i \dim(V_i)$**). For short, we write $\Gamma = \Gamma(n_1, \ldots, n_m)$. As before, $\dim(V_i) = n_i$. Let P be any vector space of dimension $\prod_{i=1}^{m} n_i$. Let $\{p_\gamma \,|\, \gamma \in \Gamma\}$ be any basis for P. Let $E_i = \{e_{i1}, \ldots, e_{in_i}\}$ be the bases for V_i, $i = 1, \ldots, m$. Define $\nu(e_{1\gamma(1)}, \ldots, e_{m\gamma(m)}) = p_\gamma$ for $\gamma \in \Gamma$ and define $\nu \in M(V_1, \ldots, V_m : P)$ by multilinear extension:

$$\nu(x_1, \ldots, x_m) = \sum_{\gamma \in \Gamma} \left(\prod_{i=1}^{m} c_{i\gamma(i)} \right) \nu(e_{1\gamma(1)}, \ldots, e_{m\gamma(m)}). \tag{3.7}$$

By lemma 3.3 the pair (P, ν) is a tensor product, $V_1, \ldots, V_m$. The homogenous tensors are, by definition, $\nu(x_1, \ldots, x_m) = x_1 \otimes \cdots \otimes x_m$ and

$$\nu(e_{1\gamma(1)}, \ldots, e_{m\gamma(m)}) = e_{1\gamma(1)} \otimes \cdots \otimes e_{m\gamma(m)} = p_\gamma.$$

In terms of homogeneous tensors, equation 3.7 becomes

$$x_1 \otimes \cdots \otimes x_m = \sum_{\gamma \in \Gamma} \left(\prod_{i=1}^{m} c_{i\gamma(i)} \right) e_{1\gamma(1)} \otimes \cdots \otimes e_{m\gamma(m)}. \tag{3.8}$$

Lemma 3.9 (Universal factorization (UF) property). *Let P and $V_1, \ldots, V_m$, $\dim(V_i) = n_i$, $i = 1, \ldots, m$, be vector spaces over $\mathbb{F}$. The pair (P, ν) is a* tensor product *of the V_i if and only if*

$$\nu \in M(V_1, \ldots, V_m : P) \ \text{ and } \ \langle \mathrm{Im}(\nu) \rangle = P \ \text{ and} \tag{3.10}$$

UF: *For any $\varphi \in M(V_1, \ldots, V_m : \mathbb{F})$ there exists $h_\varphi \in \mathbb{L}(P, \mathbb{F})$ with $\varphi = h_\varphi \nu$. The statement* **UF** *is called the* Universal Factorization *property.*

Proof. First assume (P, ν) is a *tensor product* of the V_i. The set

$$\mathbb{B} = \{\nu(e_{1\gamma(1)}, \ldots, e_{m\gamma(m)}) \,|\, \gamma \in \Gamma\}$$

is a basis for P by lemma 3.3. For $\varphi \in M(V_1, \ldots, V_m : F)$, define

$$h_\varphi \left(\nu(e_{1\gamma(1)}, \ldots, e_{m\gamma(m)}) \right) = \varphi(e_{1\gamma(1)}, \ldots, e_{m\gamma(m)}). \tag{3.11}$$

Thus, φ and $h_\varphi \nu \in M(V_1, \ldots, V_m : \mathbb{F})$ agree on the $(e_{1\gamma(1)}, \ldots, e_{m\gamma(m)}) \in \times_{i=1}^{m} V_i$ and, by multilinear extension, $h_\varphi \nu = \varphi$. The function h_φ is defined on the basis $\mathbb{B}$ and thus is in $\mathbb{L}(P, \mathbb{F})$ by linear extension.

Next, assume that conditions 3.10 and **UF** hold. Let $\{\varphi_\alpha \,|\, \alpha \in \Gamma(n_1, \ldots, n_m)\}$ be the standard basis for $M(V_1, \ldots, V_m : \mathbb{F})$ (2.23). By assumption, for each φ_α there exists $h_{\varphi_\alpha} \in \mathbb{L}(P, \mathbb{F}) := P^*$ (dual space to P) such that

$$\delta_{\alpha,\gamma} = \varphi_\alpha(e_{1\gamma(1)}, \ldots, e_{m\gamma(m)}) = h_{\varphi_\alpha} \nu(e_{1\gamma(1)}, \ldots, e_{m\gamma(m)}). \tag{3.12}$$

To show that $\mathbb{B} = \{\nu(e_{1\gamma(1)}, \ldots, e_{m\gamma(m)}) \,|\, \gamma \in \Gamma(n_1, \ldots, n_m)\}$ is linearly independent, assume

$$\sum_{\gamma \in \Gamma} k_\gamma \nu(e_{1\gamma(1)}, \ldots, e_{m\gamma(m)}) = 0.$$

For each $\alpha \in \Gamma$ we use 3.12 to obtain $h_{\varphi_\alpha} \left(\sum_{\gamma \in \Gamma} k_\gamma \nu(e_{1\gamma(1)}, \ldots, e_{m\gamma(m)} \right) =$

$$\sum_{\gamma \in \Gamma} k_\gamma h_{\varphi_\alpha} \nu(e_{1\gamma(1)}, \ldots, e_{m\gamma(m)}) = \sum_{\gamma \in \Gamma} k_\gamma \delta_{\alpha,\gamma} = k_\alpha = 0.$$

Thus,

$$\mathbb{B} = \{\nu(e_{1\gamma(1)}, \ldots, e_{m\gamma(m)}) \,|\, \gamma \in \Gamma(n_1, \ldots, n_m)\} \tag{3.13}$$

is linearly independent and, by the condition $\langle \mathrm{Im}(\nu) \rangle = P$, spans P. Thus, $\mathbb{B}$ is a basis and (P, ν) is a tensor product by lemma 3.3. The set $\{h_{\varphi_\alpha} \,|\, \alpha \in \Gamma\}$ is the dual basis to $\mathbb{B}$. □

Remark 3.14 (**Remarks about the universal factorization property**). Conditions similar to those of lemma 3.9 are commonly used to define tensor products on more general algebraic structures. Our proof, together with discussion 2.19, shows that for finite dimensional vector spaces, $\mathbb{F}$ can be replaced by U in lemma 3.9 to give an equivalent (for vector spaces) form of **UF** which is summarized by commutative diagram 3.15.

(3.15) **Universal factorization diagram**

$$\begin{tikzcd} \times_1^m V_i \arrow[r, "\nu"] \arrow[dr, "\varphi"'] & \mathrm{P} \arrow[d, "h_\varphi"] \\ & \mathrm{U} \end{tikzcd}$$

Remark 3.16 (**Equivalent definitions of tensor products**). Let $V_1, \ldots, V_m$, $\dim(V_i) = n_i$, be vector spaces over $\mathbb{F}$. Let P be a vector space over $\mathbb{F}$ and $\nu \in M(V_1, \ldots, V_m : P)$. Suppose, for all $i = 1, \ldots m$, $E_i = \{e_{i1}, \ldots, e_{in_i}\}$ is a basis for V_i. Let $\Gamma = \Gamma(n_1, \ldots, n_m)$ (notation 2.3). Then, the following statements are equivalent:

$$\langle \mathrm{Im}(\nu) \rangle = P \quad \text{and} \quad \dim(P) = \prod_{i=1}^{m} \dim(V_i) \tag{3.17}$$

$$\mathbb{B} = \{\nu(e_{1\gamma(1)}, \ldots, e_{m\gamma(m)}) \mid \gamma \in \Gamma\} \textit{ is a basis for } P \tag{3.18}$$

$$\langle \mathrm{Im}(\nu) \rangle = P \quad \text{and} \quad \textbf{Universal Factorization}: \tag{3.19}$$

(**UF**) For any $\varphi \in M(V_1, \ldots, V_m : \mathbb{F})$ there exists a linear function $h_\varphi \in \mathbb{L}(P, \mathbb{F})$ with $\varphi = h_\varphi \nu$.

Statement 3.17 is our definition of tensor product (definition 3.1). The equivalence of statement 3.18 and 3.17 follows from lemma 3.3. The equivalence of statements 3.19 and 3.17 follows from lemma 3.9. (Note: The proof of lemma 3.9 shows that h_φ can be assumed to be unique in this lemma.)

Definition 3.20 (Subspace tensor products). Let $V_1, \ldots, V_m$, $\dim(V_i) = n_i$, be vector spaces. Let $W_1 \subseteq V_1, \ldots, W_m \subseteq V_m$, be subspaces. Let (P, ν) be a tensor product of $V_1, \ldots, V_m$. If (P_w, ν_w) is a tensor product of $W_1, \ldots, W_m$ such that $\nu_w(x_1, \ldots, x_n) = \nu(x_1, \ldots, x_m)$ for $(x_1, \ldots, x_m) \in \times_{i=1}^m W_i$ then (P_w, ν_w) is a *subspace tensor product* of (P, ν).

3.21 (**Subspace tensor products–constructions**). Let $V_1, \ldots, V_m$ with dimensions $\dim(V_i) = n_i$ be vector spaces. Let $W_1 \subseteq V_1, \ldots, W_m \subseteq V_m$ be subspaces and (P, ν) be a tensor product of $V_1, \ldots, V_m$. Construct ordered bases $E_i' = \{e_{i1}, \ldots, e_{ir_i}\}$ for the W_i, $i = 1, \ldots, m$. Extend these bases to $E_i = \{e_{i1}, \ldots, e_{in_i}\}$, ordered bases for V_i, $i = 1, \ldots, m$. By 3.18

(3.22) $\quad \mathbb{B} = \{\nu(e_{1\gamma(1)}, \ldots, e_{m\gamma(m)}) \,|\, \gamma \in \Gamma(n_1, \ldots, n_m)\}$ is a basis for P.

Thus,

$$P_w = \langle\{\nu(e_{1\gamma(1)}, \ldots, e_{m\gamma(m)}) \,|\, \gamma \in \Gamma(r_1, \ldots, r_m)\}\rangle \tag{3.23}$$

defines a subspace $P_w \subseteq P$ of dimension $\prod_{i=1}^m \dim(W_i)$. Note that for $\alpha \in \Gamma(r_1, \ldots, r_m)$, $(e_{1\alpha(1)}, \ldots, e_{m\alpha(m)}) \in \times_{i=1}^m W_i$. Define $\nu_w(e_{1\alpha(1)}, \ldots, e_{m\alpha(m)}) := \nu(e_{1\alpha(1)}, \ldots, e_{m\alpha(m)})$. By multilinear extension, $\nu_w \in M(W_1, \ldots, W_m : P_w)$. By identity 3.17 or 3.18, (P_w, ν_w) is a tensor product of $W_1, \ldots, W_m$ where P_w is a subspace of P and ν_w is the restriction of ν to $\times_{i=1}^m W_i$.

3.24 (**It's not hard to be a tensor product**). Equation 3.12 and the related discussion shows the connection between the **UF** property and the dual space, P^*, of P. Referring to figure 3.15, take $m = 1$ and $U = \mathbb{F}$. Thus, $\times_1^m V_i = V_1$. Take $P = V_1$, $\nu = id$ (the identity function in $\mathbb{L}(V_1, V_1)$). Let $\varphi \in M(V_1 : \mathbb{F}) = \mathbb{L}(V_1, \mathbb{F}) = V_1^*$ (the dual space of V_1). Take $h_\varphi = \varphi$. Thus, (V_1, id) is a tensor product for V_1 (see figure 3.25).

(3.25)

$$\begin{tikzcd}
V_1 \arrow[r, "\nu=id"] \arrow[dr, "\varphi"'] & V_1 \arrow[d, "h_\varphi=\varphi"] \\
& \mathbb{F}
\end{tikzcd}$$

3.26 (**The dual space model for tensors**). Referring to definition 3.1, suppose that (P, ν) is a tensor product for $V_1, \ldots, V_m$. By 3.6, the vector space P can be *any* vector space over $\mathbb{F}$ with dimension $N = \prod_{i=1}^m \dim(V_i)$.

Choose $P = M^*$, the dual space to the vector space $M := M(V_1, \ldots, V_m : \mathbb{F})$. Define a tensor product of $V_1, \ldots, V_m$ to be (M^*, ν). For $\varphi \in M(V_1, \ldots, V_m : \mathbb{F})$, ν is defined by $\nu(x_1, \ldots, x_m)(\varphi) = \varphi(x_1, \ldots, x_m)$. It is easy to see that ν is multilinear and well defined. From 2.23, $\{\varphi_\alpha \,|\, \alpha \in \Gamma(n_1, \ldots, n_m)\}$, where $\varphi_\alpha(e_{1\gamma(1)}, \ldots, e_{m\gamma(m)}) = \delta_{\alpha,\gamma}$, is a basis for $M(V_1, \ldots, V_m : \mathbb{F})$. Thus,

$$\{\nu(e_{1\gamma(1)}, \ldots, e_{m\gamma(m)}) \,|\, \gamma \in \Gamma(n_1, \ldots, n_m)\}$$

is the basis for M^* dual to the basis φ_α for M. The tensor product (M^*, ν) of $V_1, \ldots, V_n$ is called the *dual tensor product* of $V_1, \ldots, V_m$. As usual, the homogeneous tensors are $\nu(x_1, \ldots, x_m) := x_1 \otimes \cdots \otimes x_m$. These homogeneous tensors satisfy $x_1 \otimes \cdots \otimes x_m(\varphi) = \varphi(x_1, \ldots, x_m)$.

Definition 3.27 (Matrix of a pair of bases). Let $\mathbf{v} = (v_1, \ldots v_q)$ and $\mathbf{w} = (w_1, \ldots, w_r)$ be ordered bases for V and W respectively. Suppose for each j,

$1 \le j \le q$, $T(v_j) = \sum_{i=1}^{r} a_{ij} w_i$. The matrix $A = (a_{ij})$ is called the matrix of T with respect to the base pair $(\mathbf{v}, \mathbf{w})$. We write $[T]_{\mathbf{v}}^{\mathbf{w}}$ for A.

For example, let $V = \mathbb{R}^2$ and $W = \mathbb{R}^3$. Let $\mathbf{w} = \{w_1, w_2, w_3\}$ and $\mathbf{v} = \{v_1, v_2\}$. Define T by $T(v_1) = 2w_1 + 3w_2 - w_3$ and $T(v_2) = w_1 + 5w_2 + w_3$. Then

$$[T]_{\mathbf{v}}^{\mathbf{w}} = \begin{bmatrix} 2 & 1 \\ 3 & 5 \\ -1 & 1 \end{bmatrix}$$

is the matrix of T with respect to the base pair $(\mathbf{v}, \mathbf{w})$.

3.28 (**Obligatory observations about tensor products of** $V_1, \ldots, V_m$**.**). Let (P, ν) and (Q, μ) be two tensor products of $V_1, \ldots, V_m$. Both P and Q are vector spaces over $\mathbb{F}$ of the same dimension, $N = \prod_{i=1}^{m} \dim(V_i)$. Thus, they are isomorphic as vector spaces. By lemma 3.3, $\mathbb{B}_\nu := \{\nu(e_{1\gamma(1)}, \ldots, e_{m\gamma(m)}) \mid \gamma \in \Gamma\}$ is a basis for P, and $\mathbb{B}_\mu := \{\mu(e_{1\gamma(1)}, \ldots, e_{m\gamma(m)}) \mid \gamma \in \Gamma\}$ is a basis for Q. The isomorphism $T \in \mathbb{L}(P, Q)$ defined by these bases:

$$T(\nu(e_{1\gamma(1)}, \ldots, e_{m\gamma(m)}) := \mu(e_{1\gamma(1)}, \ldots, e_{m\gamma(m)}) \tag{3.29}$$

is a natural choice of correspondence between these tensor products. By multilinear extension, equation 3.29 implies $T\nu = \mu \in M(V_1, \ldots, V_m : Q)$. In matrix terms, ordering $\mathbb{B}_\nu$and $\mathbb{B}_\mu$ lexicographically based on Γ gives

$$\Big[T\Big]_{\mathbb{B}_\nu}^{\mathbb{B}_\mu} = I_N.$$

Apply lemma 3.9 to (P, ν) with $U = Q$ and $\varphi = \mu \in M(V_1, \ldots, V_m : Q)$. Let $h_{\nu,\mu} \in \mathbb{L}(P, Q)$ be such that $h_{\nu,\mu}\nu = \mu$. Thus, $h_{\nu,\mu} = T$ of equation 3.29. Reversing the roles of P, Q and ν, μ, let $h_{\mu,\nu}$ be such that $h_{\mu,\nu}\mu = \nu$. Thus, $h_{\mu,\nu} = T^{-1} \in \mathbb{L}(Q, P)$.

3.30 (**Associative laws for tensor products**). Suppose, for all $i \in \underline{m}$, $V_i = \langle E_i \rangle$ is the space spanned by the ordered basis $E_i = \{e_{i1}, \ldots, e_{in_i}\}$. Form a tensor product

$$Z = (V_1 \otimes \cdots \otimes V_p) \otimes (V_{p+1} \otimes \cdots \otimes V_{p+q})$$

where $p + q = m$. Thus, we have a tensor product of two finite dimensional vector spaces $P_1 = V_1 \otimes \cdots \otimes V_p$ and $P_2 = V_{p+1} \otimes \cdots \otimes V_{p+q}$. We have $\dim(P_1) = \prod_{i=1}^{p} \dim(V_i)$ and $\dim(P_2) = \prod_{i=p+1}^{m} \dim(V_i)$. Thus,

$$\dim(Z) = \dim(P_1)\dim(P_2) = \prod_{1}^{m} \dim(V_i) = \dim(V_1 \otimes \cdots \otimes V_m)$$

so Z and $(V_1 \otimes \cdots \otimes V_m)$ are isomorphic vector spaces. There are many isomorphisms, but some are more "natural" than others. Take $\Gamma_1(n_1, \ldots, n_p) :=$

$\underline{n_1} \times \cdots \times \underline{n_p}$ and $\Gamma_2(n_{n_{p+1}}, \ldots, n_m) \coloneqq \underline{n_{p+1}} \times \cdots \times \underline{n_m}$ and order each lexicographically. The natural basis for Z is $e_\alpha \otimes e_\beta$ where

$$e_\alpha \otimes e_\beta = (e_{1\alpha(1)} \otimes \cdots \otimes e_{p\alpha(p)}) \otimes (e_{(p+1)\beta(p+1)} \otimes \cdots \otimes e_{m\beta(m)})$$

where $(\alpha, \beta) \in \Gamma_1(n_1, \ldots, n_p) \times \Gamma_2(n_{n_{p+1}}, \ldots, n_m)$ is ordered lexicographically based on the lexicographic orders on Γ_1 and Γ_2. A pair, (α, β) defines a $\gamma \in \Gamma(n_1, \ldots, n_m)$ by concatenation, $(\alpha, \beta) \mapsto \alpha\beta \coloneqq \gamma$. This correspondence is order preserving if we assume lexicographic order on $\Gamma(n_1, \ldots, n_m)$.

4. Tensor products of matrices

4.1 (**Tensor product of matrices example**). Let $V_1 = \mathbf{M}_{p,q}$ ($p \times q$ matrices over $\mathbb{F}$), and let $V_2 = \mathbf{M}_{r,s}$. Let $A = (a_{ij}) \in \mathbf{M}_{p,q}$ and $B = (b_{ij}) \in \mathbf{M}_{r,s}$. A tensor product $(P, \nu) = V_1 \otimes V_2$ can be constructed following the general approach of 3.6. Let P be any vector space over $\mathbb{F}$ of dimension $pqrs$. Let $\{p_{ijkl} \,|\, (i,j,k,l) \in \underline{p} \times \underline{q} \times \underline{r} \times \underline{s}\}$ be a basis for P. Let $\{E^{(1)}_{ij} \,|\, (i,j) \in \underline{p} \times \underline{q}\}$ be the standard basis for $\mathbf{M}_{p,q}$ (i.e., $E^{(1)}_{ij}(i',j') = \mathcal{X}((i,j) = (i',j'))$). Let $\{E^{(2)}_{ij} \,|\, (i,j) \in \underline{r} \times \underline{s}\}$ be the standard basis for $\mathbf{M}_{r,s}$. Define $\nu(E^{(1)}_{ij}, E^{(2)}_{kl}) \coloneqq E^{(1)}_{ij} \otimes E^{(2)}_{kl} = p_{ijkl}$, and by multilinear (bilinear here) extension

$$\nu(A,B) \coloneqq A \otimes B = \sum_{(i,j,k,l)} a_{ij} b_{kl} E^{(1)}_{ij} \otimes E^{(2)}_{kl} \tag{4.2}$$

where

$$A = \left(\sum_{(i,j) \in \underline{p} \times \underline{q}} a_{ij} E^{(1)}_{ij} \right) \text{ and } B = \left(\sum_{(i,j) \in \underline{r} \times \underline{s}} b_{kl} E^{(2)}_{kl} \right). \tag{4.3}$$

Note that although A and B are matrices, the $E^{(1)}_{ij} \otimes E^{(2)}_{kl} = p_{ijkl}$ are basis elements of P. The matrix structure is not utilized here except in the indexing of coefficients.

4.4 (**Comments about notation for standard bases $E_i = \{e_{i1}, \ldots e_{in_i}\}$**). We have stated our standard assumptions for bases as follows: "Let $V_1, \ldots, V_m$, $\dim(V_i) = n_i$, $i \in \underline{m}$, be vector spaces over $\mathbb{F}$. Suppose, for all $i \in \underline{m}$, $V_i = \langle E_i \rangle$, $E_i = \{e_{i1}, \ldots, e_{in_i}\}$ is an ordered basis for V_i." Suppose, analogous to 4.1, the $V_i = M_{p_i,q_i}$, $i = 1, \ldots, m$, are vector spaces of matrices. The standard ordered basis for V_i is now

$$E_i = \{E^{(i)}_{11}, E^{(i)}_{12}, \ldots, E^{(i)}_{p_i q_i}\}.$$

The basis element e_{ij} has been replaced by $E^{(i)}_{st}$ where (s,t) is the element in position j in the list $\underline{p_i} \times \underline{q_i}$ in lexicographic order. We don't attempt to formalize this type of variation from the standard notation. The general case will involve ordered bases where the order is specified by a linear order on indices in some manner (usually some type of lex order).

4.5 (**Universal factorization property: example**). This example continues the discussion of 4.1 in order to illustrate lemma 3.9, the universal factorization property. Let $V_1 = \mathbf{M}_{p,q}$ ($p \times q$ matrices over $\mathbb{F}$), and let $V_2 = \mathbf{M}_{q,s}$. Let $A = (a_{ij}) \in \mathbf{M}_{p,q}$ and $B = (b_{ij}) \in \mathbf{M}_{q,s}$. Here we have set $r = q$ in 4.1. Take the φ of lemma 3.9 to be the bilinear function $\varphi(A,B) = AB$, the matrix product of A and B. Thus, $\varphi \in M(V_1, V_2 : U)$ where $U = \mathbf{M}_{p,s}$. We want to construct the linear function

$$h_\varphi \in \mathbb{L}(\mathbf{M}_{p,q} \otimes \mathbf{M}_{q,s}, \mathbf{M}_{p,s})$$

such that $h_\varphi(A \otimes B) = \varphi(A,B)$. The set of matrices $\{E^{(3)}_{cf} \,|\, (c,f) \in \underline{p} \times \underline{s}\}$ is the standard basis for U (notation as in 4.1). Following discussion 2.17, we work with the component functions $\varphi^{(i,j)} \in M(V_1, V_2 : \langle E^{(3)}_{ij} \rangle)$ for each i,j fixed:

$$h_{\varphi^{(i,j)}} \in \mathbb{L}(\mathbf{M}_{p,q} \otimes \mathbf{M}_{q,s}, \langle E^{(3)}_{ij} \rangle) \equiv \mathbb{L}(\mathbf{M}_{p,q} \otimes \mathbf{M}_{q,s}, \mathbb{F}).$$

Define $h_{\varphi^{(i,j)}}$ on the basis elements

$$\{E^{(1)}_{cd} \otimes E^{(2)}_{ef} \,|\, (c,d) \in \underline{p} \times \underline{q}, (e,f) \in \underline{q} \times \underline{s}\}$$

analogous to equation 3.11:

$$h_{\varphi^{(i,j)}}(E^{(1)}_{cd} \otimes E^{(2)}_{ef}) = \varphi^{(i,j)}(E^{(1)}_{cd}, E^{(2)}_{ef}) = E^{(1)}_{cd} E^{(2)}_{ef}(i,j) \tag{4.6}$$

where $E^{(1)}_{cd} E^{(2)}_{ef}$ denotes the matrix product of these two basis elements. An elementary result from matrix theory states $E^{(1)}_{cd} E^{(2)}_{ef} = \mathcal{X}(d = e) E^{(3)}_{cf}$. Thus,

$$E^{(1)}_{cd} E^{(2)}_{ef}(i,j) = \mathcal{X}\,(d = e)\, E^{(3)}_{cf}(i,j) = \mathcal{X}\,(d = e)\, \mathcal{X}\,((c,f) = (i,j))\,.$$

Defining $h_{\varphi^{(i,j)}}$ on $A \otimes B$ by linear extension gives

$$h_{\varphi^{(i,j)}}\,(A \otimes B) := \sum_{(c,d,e,f)} a_{cd} b_{ef} h_{\varphi^{(i,j)}}(E^{(1)}_{cd} \otimes E^{(2)}_{ef}) \tag{4.7}$$

Thus,

$$h_{\varphi^{(i,j)}}\,(A \otimes B) = \sum_{(c,d,e,f)} a_{cd} b_{ef} \mathcal{X}\,(d = e)\, \mathcal{X}\,((c,f) = (i,j))\,. \tag{4.8}$$

Thus,

$$h_{\varphi^{(i,j)}}(A \otimes B) = \left(\sum_{d \in \underline{q}} a_{id} b_{dj} \right) = AB(i,j). \tag{4.9}$$

Thus,

$$h_\varphi = \bigoplus_{(i,j) \in \Gamma(p,s)} h_{\varphi^{(i,j)}}$$

is in

$$\mathbb{L}(\mathbf{M}_{p,q} \otimes \mathbf{M}_{q,s}, \mathbf{M}_{p,s})$$

and satisfies $h_\varphi(A \otimes B) = AB$.

4.10 (**Tensor products of matrices as matrices: a bad choice**). Let $V_1 = \mathbf{M}_{p_1,q_1}$, and let $V_2 = \mathbf{M}_{p_2,q_2}$. We construct a tensor product (P,ν) for V_1,V_2. Let $n_1 = p_1q_1$ and $n_2 = p_2q_2$ and choose $P = \mathbf{M}_{n_1,n_2}$. Define $\nu(E^{(1)}_{ij}, E^{(2)}_{kl}) := E^{(1)}_{ij} \otimes E^{(2)}_{kl} = p_{ijkl}$ where $p_{ijkl} = E^{(3)}_{pq}$ is the standard basis element of $\mathbf{M}_{n_1,n_2}$ with p the position of (i,j) in the lexicographic list of $\underline{p_1} \times \underline{q_1}$ and q defined similarly for $\underline{p_2} \times \underline{q_2}$. By the construction of 3.6 such a choice is possible. We have represented $\mathbf{M}_{p_1,q_1} \otimes \mathbf{M}_{p_2,q_2}$ as matrices so that $A_1 \otimes A_2$ is an $n_1 \times n_2$ matrix where n_1 is the number of entries in A_1 and n_2 the number of entries in A_2. The problem with this is that in the obvious extension to $V_1 \otimes V_2 \otimes V_3$ we would have $(V_1 \otimes V_2) \otimes V_3$ and $V_1 \otimes (V_2 \otimes V_3)$ isomorphic as tensor spaces (see 3.30) but, in general, $(A_1 \otimes A_2) \otimes A_3$ not equal to $A_1 \otimes (A_2 \otimes A_3)$ as matrices. (the former has n_1n_2 rows and the latter n_1 rows).

4.11 (**Tensor products of matrices, a good choice: Kronecker product**). Let $V_1 = \mathbf{M}_{p_1,q_1}$, and let $V_2 = \mathbf{M}_{p_2,q_2}$. As in 4.10, we construct a tensor product (P,ν) for V_1,V_2. Choose $P := \mathbf{M}_{p,q}$ where $p = p_1p_2$ and $q = q_1q_2$. Let $\mu \in \underline{p_1} \times \underline{p_2}$ and $\kappa \in \underline{q_1} \times \underline{q_2}$ (both with lexicographic order). We use these two ordered sets to index the rows and columns of the matrices in P. The matrices

$$\{E_{\mu,\kappa} \,|\, \mu \in \underline{p_1} \times \underline{p_2}\,,\ \kappa \in \underline{q_1} \times \underline{q_2}\,\} \tag{4.12}$$

where $E_{\mu,\kappa}(\alpha,\beta) = \mathcal{X}((\alpha,\beta) = (\mu,\kappa))$ are the standard basis elements of $\mathbf{M}_{p,q}$. Let

$$\{E^{(1)}_{\mu(1),\kappa(1)} \,|\, (\mu(1),\kappa(1)) \in \underline{p_1} \times \underline{q_1}\}$$

and

$$\{E^{(2)}_{\mu(2),\kappa(2)} \,|\, (\mu(2),\kappa(2)) \in \underline{p_2} \times \underline{q_2}\},$$

each ordered lexicographically, denote the standard ordered bases for $\mathbf{M}_{p_1,q_1}$ and $V_2 = \mathbf{M}_{p_2,q_2}$.

Define (see 3.3) the function ν of (P,ν) by

$$\nu(E^{(1)}_{\mu(1),\kappa(1)}, E^{(2)}_{\mu(2),\kappa(2)}) := E^{(1)}_{\mu(1),\kappa(1)} \otimes E^{(2)}_{\mu(2),\kappa(2)} = E_{\mu,\kappa}. \tag{4.13}$$

Let $A_i = (a_i(s,t)) \in \mathbf{M}_{p_i,q_i}$, $i = 1,2$. Applying multilinear extension 3.7, we get $\nu(A_1,A_2) = A_1 \otimes A_2 =$

$$\nu\left(\sum_{\mu(1),\kappa(1)} a_1(\mu(1),\kappa(1))E^{(1)}_{\mu(1),\kappa(1)}, \sum_{\mu(2),\kappa(2)} a_2(\mu(2),\kappa(2))E^{(2)}_{\mu(2),\kappa(2)}\right) = \tag{4.14}$$

$$\sum_{\mu,\kappa}\prod_{i=1}^{2} a_i(\mu(i),\kappa(i))E_{\mu,\kappa}.$$

Equation 4.13 guarantees that (P,ν) is a tensor product of V_1 and V_2 (3.16). The notation of 4.14 is chosen to extend from $m = 2$ to the general case. The tensor product, (P,ν), $P = \otimes_{i=1}^{2}\mathbf{M}_{p_i,q_i}$ and ν defined by equation 4.13 and equation 4.14 is called the *Kronecker tensor product* of the vector spaces

$\mathbf{M}_{p_1,q_1}$ and $\mathbf{M}_{p_2,q_2}$. From equation 4.14, we see that the homogeneous tensors, $\nu(A_1, A_2) := A_1 \otimes A_2$ can be regarded as matrices in $\mathbf{M}_{p,q}$ and that the μ, κ entry of $A_1 \otimes A_2$ is

$$(A_1 \otimes A_2)\,(\mu, \kappa) = \prod_{i=1}^{2} A_i(\mu(i), \kappa(i)). \tag{4.15}$$

These homogeneous elements are are sometimes defined without reference to the general theory of tensor spaces and are called "Kronecker products of matrices."

4.16 (**Kronecker product: example of homogeneous tensors**). For

$$A_1 = (a_1(i,j)),\ A_2 = (a_2(i,j)) \in \mathbf{M}_{2,2}$$

the Kronecker product $A_1 \otimes A_2 =$

$$\begin{array}{cc} & \begin{array}{cccc} \quad 1\,1 \qquad\qquad & \qquad 1\,2 \qquad\qquad & \qquad 2\,1 \qquad\qquad & \qquad 2\,2 \quad \end{array} \\ \begin{array}{c} \underline{1}\,\underline{1} \\ \underline{1}\,\underline{2} \\ \underline{2}\,\underline{1} \\ \underline{2}\,\underline{2} \end{array} & \begin{pmatrix} a_1(\underline{1},1)a_2(\underline{1},1) & a_1(\underline{1},1)a_2(\underline{1},2) & a_1(\underline{1},2)a_2(\underline{1},1) & a_1(\underline{1},2)a_2(\underline{1},2) \\ a_1(\underline{1},1)a_2(\underline{2},1) & a_1(\underline{1},1)a_2(\underline{2},2) & a_1(\underline{1},2)a_2(\underline{2},1) & a_1(\underline{1},2)a_2(\underline{2},2) \\ a_1(\underline{2},1)a_2(\underline{1},1) & a_1(\underline{2},1)a_2(\underline{1},2) & a_1(\underline{2},2)a_2(\underline{1},1) & a_1(\underline{2},2)a_2(\underline{1},2) \\ a_1(\underline{2},1)a_2(\underline{2},1) & a_1(\underline{2},1)a_2(\underline{2},2) & a_1(\underline{2},2)a_2(\underline{2},1) & a_1(\underline{2},2)a_2(\underline{2},2) \end{pmatrix} \end{array}$$

Note that $A_1 \otimes A_2$ can be constructed by starting with a 2×2 matrix with entries in $\mathbf{M}_{2,2}$,

$$\begin{pmatrix} a_1(1,1)A_2 & a_1(1,2)A_2 \\ a_1(2,1)A_2 & a_1(2,2)A_2 \end{pmatrix},$$

and doing the indicated multiplications of entries from A_1 with A_2 to construct the blocks of $A_1 \otimes A_2$. This pleasing structure is a consequence of using lexicographic order as done in 4.11. Other orders would work just as well, but the result might be a mess to the human eye.

4.17 (**General Kronecker products of vector spaces of matrices**). We follow the discussion 4.11, developing the notation for the general case. Define a tensor product (P, ν) for $V_1, \ldots, V_m$ where $V_1 = \mathbf{M}_{p_1,q_1}, \ldots, V_m = \mathbf{M}_{p_m,q_m}$. Choose $P := \mathbf{M}_{p,q}$ where $p = \prod_{i=1}^m p_i$ and $q = \prod_{i=1}^m q_i$. Let $\mu \in \underline{p_1} \times \cdots \times \underline{p_m}$ and $\kappa \in \underline{q_1} \times \cdots \times \underline{q_m}$ (both with lexicographic order). We use these two ordered sets to index, respectively, the rows and columns of the matrices in P. We choose the matrices

$$\{E_{\mu,\kappa} \,|\, \mu \in \underline{p_1} \times \cdots \times \underline{p_m}\,,\ \kappa \in \underline{q_1} \times \cdots \times \underline{q_m}\,\} \tag{4.18}$$

where $E_{\mu,\kappa}(\alpha, \beta) = \mathcal{X}((\alpha, \beta) = (\mu, \kappa))$ to be the basis elements of $\mathbf{M}_{p,q}$. Let

$$\{E^{(i)}_{\mu(i),\kappa(i)} \,|\, (\mu(i), \kappa(i)) \in \underline{p_i} \times \underline{q_i}\},\ \ i = 1, \ldots, m,$$

be the standard bases for the $V_i = \mathbf{M}_{p_i,q_i}$. We define $\nu \in M(V_1, \ldots, V_m : P)$ by

$$\nu(E^{(1)}_{\mu(1),\kappa(1)}, \ldots, E^{(m)}_{\mu(m),\kappa(m)}) := E^{(1)}_{\mu(1),\kappa(1)} \otimes \cdots \otimes E^{(m)}_{\mu(2),\kappa(m)} = E_{\mu,\kappa}. \tag{4.19}$$

Let $A_i = (a_i(s,t)) \in \mathbf{M}_{p_i,q_i}$, $i = 1,\ldots,m$. Applying multilinear extension 3.7, we get $\nu(A_1,\ldots,A_m) = A_1 \otimes \cdots \otimes A_m =$

(4.20)

$$\nu\left(\sum_{\mu(1),\kappa(1)} a_1(\mu(1),\kappa(1))E^{(1)}_{\mu(1),\kappa(1)},\ldots,\sum_{\mu(m),\kappa(m)} a_m(\mu(m),\kappa(m))E^{(m)}_{\mu(m),\kappa(m}\right) =$$

$$\sum_{\mu,\kappa}\prod_{i=1}^{m} a_i(\mu(i),\kappa(i))E_{\mu,\kappa}.$$

Definition 4.21 (Kronecker product general definition). Consider the vector spaces of matrices $\mathbf{M}_{p_1,q_1},\ldots,\mathbf{M}_{p_m,q_m}$ and $\mathbf{M}_{p,q}$ where $p = \prod_{i=1}^m p_i$, $q = \prod_{i=1}^m q_i$, $i = 1,\ldots,m$. Order the rows of the matrices in $\mathbf{M}_{p,q}$ with the $\{\mu\,|\,\mu \in \underline{p_1} \times \cdots \times \underline{p_m}\}$ and the columns with $\{\kappa\,|\,\kappa \in \underline{q_1} \times \cdots \times \underline{q_m}\}$, each set ordered lexicograhically. The pair $(\mathbf{M}_{p,q},\nu)$ is a tensor product of these vector spaces where $p = \prod_{i=1}^m p_i$, $q = \prod_{i=1}^m q_i$, $i = 1,\ldots,m$, and $\nu \in M(\mathbf{M}_{p_1,q_1},\ldots,\mathbf{M}_{p_m,q_m} : \mathbf{M}_{p,q})$ is defined by $\nu(A_1,\ldots,A_m) := A_1 \otimes \cdots \otimes A_m$ where

$$(A_1 \otimes \cdots \otimes A_m)(\mu,\kappa) = \prod_{i=1}^{m} A_i(\mu(i),\kappa(i)).$$

4.22 (**Is the Kronecker tensor product a special case?**). In 3.16 we give three equivalent definitions of a tensor product of vector spaces. The setup for these definitions is as follows:

"Let $V_1,\ldots,V_m$, $\dim(V_i) = n_i$, be vector spaces over $\mathbb{F}$. Let P be a vector space over $\mathbb{F}$ and $\nu \in M(V_1,\ldots,V_m : P)$. Suppose, for all $i = 1,\ldots m$, $E_i = \{e_{i1},\ldots,e_{in_i}\}$ is a basis for V_i."

The Kronecker product is usually described as a "special case" of a tensor product. It is actually equivalent to the definition of a tensor product, differing only by the indexing of the bases of the V_i and P. If we factor the dimensions $n_i = p_iq_i$ (possible in many ways; we allow $p_i = 1$ or $q_i = 1$), then we can replace $E_i = \{e_{i1},\ldots,e_{in_i}\}$ by $E^{(i)} = \{E^{(i)}_{11},\ldots,E^{(i)}_{p_iq_i}\}$ ordered, for example, lexicographically on indices. The vector space P has dimension $\prod_{i=1}^m n_i$ which is the same as pq where $p = \prod_{i=1}^m p_i$ and $q = \prod_{i=1}^m q_i$. The most direct correspondence is to replace $E_i = \{e_{i1},\ldots,e_{in_i}\}$ by $E^{(i)} = \{E^{(i)}_{11},\ldots,E^{(i)}_{1,n_i}\}$. The matrix model for Kronecker products is important for applications to matrix theory that are model specific.

5. Inner products on tensor spaces

5.1 (**Inner products and tensor spaces**). Assume $V_1,\ldots,V_m$, $\dim(V_i) = n_i$, are vector spaces over the complex numbers, $\mathbb{C}$. Let (W,ω) be a tensor product of these vector spaces. As usual, $W = \otimes_{i=1}^m V_i$, and $\omega(x_1,\ldots,x_m) = x_1\otimes\cdots\otimes x_m$

are the homogeneous tensors. Recall that $\varphi \in M(W, W : \mathbb{C})$ is *conjugate bilinear* if for all $a, a', b, b' \in W$ and $c, d \in \mathbb{C}$

$$\begin{aligned} \varphi(ca + da', b) &= c\varphi(a, b) + d\varphi(a', b) \\ \varphi(b, ca + da') &= \overline{c}\varphi(b, a) + \overline{d}\varphi(b, a'). \end{aligned} \tag{5.2}$$

A conjugate bilinear $\varphi \in M(W, W : \mathbb{C})$ is an *inner product* on W if for all $a, b \in W$

$$\begin{gathered} \varphi(a, b) = \overline{\varphi(b, a)} \\ \varphi(a, a) \geq 0 \text{ with } \varphi(a, a) = 0 \text{ iff } a = 0. \end{gathered} \tag{5.3}$$

In that case, the pair (W, φ) is a *unitary space*. The first condition of equation 5.3 is called *conjugate symmetric*, the second is called *positive definite*.

If $\{f_\gamma \mid \gamma \in \Gamma(n_1, \ldots, n_m)\}$ is any basis of W then an inner product on W can be defined by specifying $\varphi(f_\alpha, f_\beta) = \delta_{\alpha,\beta}$ and defining $\varphi \in M(W, W : \mathbb{C})$ by extending these values by (conjugate) bilinear extension. In this case, $\{f_\gamma \mid \gamma \in \Gamma(n_1, \ldots, n_m)\}$ is called an *orthonormal basis* for the inner product φ. If φ is any inner product on W, there exists an orthonormal basis that defines φ in the manner just described (e.g., by using the Gram-Schmidt orthonormalization process).

If $\{f_\gamma \mid \gamma \in \Gamma(n_1, \ldots, n_m)\}$ is an *orthonormal basis* for the inner product φ and $a = \sum_{\alpha \in \Gamma} c_\alpha f_\alpha$ and $b = \sum_{\beta \in \Gamma} d_\beta f_\beta$ then

$$\varphi(a, b) = \sum_{\alpha \in \Gamma} \sum_{\beta \in \Gamma} c_\alpha \overline{d_\beta} \varphi(f_\alpha, f_\beta) = \sum_{\alpha \in \Gamma} c_\alpha \overline{d_\alpha}. \tag{5.4}$$

We want to relate these ideas more closely to the tensor product, (W, ω). Suppose, for $i = 1, \ldots m$, $E_i = \{e_{i1}, \ldots, e_{in_i}\}$ is an orthonormal basis for V_i with inner product $\varphi_i \in M(V_i, V_i : \mathbb{C})$. Assume $x_i = \sum_{k=1}^{n_i} c_{ik} e_{ik}$ and assume $y_i = \sum_{k=1}^{n_i} d_{ik} e_{ik}$. We use the notation

$$x^\otimes := x_1 \otimes \cdots \otimes x_m \text{ and } e_\gamma^\otimes := e_{1\gamma(1)} \otimes \cdots \otimes e_{m\gamma(m)}. \tag{5.5}$$

Lemma 5.6 (Inner product on W as a product). *We refer to 5.1 for notation. Suppose, for $i = 1, \ldots m$, $E_i = \{e_{i1}, \ldots, e_{in_i}\}$ is an orthonormal basis for V_i with inner product $\varphi_i \in M(V_i, V_i : \mathbb{C})$. There exists a unique inner product $\varphi \in M(W, W : \mathbb{C})$, $W = \otimes_{i=1}^m V_i$, such that*

$$\varphi(x^\otimes, y^\otimes) = \varphi(x_1 \otimes \cdots \otimes x_m, y_1 \otimes \cdots \otimes y_m) = \prod_{i=1}^m \varphi_i(x_i, y_i). \tag{5.7}$$

This φ is defined by $\varphi(e_\alpha^\otimes, e_\beta^\otimes) := \prod_{i=1}^m \varphi_i(e_{i\alpha(i)}, e_{i\beta(i)})$ for all $\alpha, \beta \in \Gamma(n_1, \ldots, n_m)$. The basis $\{e_\alpha^\otimes \mid \alpha \in \Gamma(n_1, \ldots, n_m)\}$ is an orthonormal basis for φ.

Proof. Define $\varphi(e_\alpha^\otimes, e_\beta^\otimes) := \prod_{i=1}^m \varphi_i(e_{i\alpha(i)}, e_{i\beta(i)})$ for all $\alpha, \beta \in \Gamma(n_1, \ldots, n_m)$ and define $\varphi \in M(W, W : \mathbb{C})$ by conjugate bilinear extension. Note that

$\prod_{i=1}^m \varphi_i(e_{i\alpha(i)}, e_{i\beta(i)}) = \delta_{\alpha\beta}$ so that the φ so defined has $\{e^{\otimes}_\alpha \,|\, \alpha \in \Gamma(n_1, \ldots, n_m)\}$ as an orthonormal basis. We have

$$\prod_{i=1}^m \varphi_i(x_i, y_i) = \prod_{i=1}^m \varphi_i\left(\sum_{j=1}^{n_i} c_{ij}e_{ik}, \sum_{k=1}^{n_i} d_{ik}e_{ik}\right) =$$

$$\prod_{i=1}^m \left(\sum_{j,k} c_{ij}\overline{d_{ik}}\varphi_i(e_{ij}, e_{ik})\right) = \prod_{i=1}^m \left(\sum_{t=1}^{n_i} c_{it}\overline{d_{it}}\right) =$$

$$\sum_{\alpha\in\Gamma}\prod_{i=1}^m c_{i\alpha(i)}\overline{d_{i\alpha(i)}} = \sum_{\alpha\in\Gamma}\prod_{i=1}^m c_{i\alpha(i)}\prod_{i=1}^m \overline{d_{i\alpha(i)}} = \varphi(x^{\otimes}, y^{\otimes}).$$

The last equality follows from equation 5.4 with $a = x^{\otimes}$ and $b = y^{\otimes}$ noting that $c_\alpha = \prod_{i=1}^m c_{i\alpha(i)}$ and $\overline{d_\alpha} = \prod_{i=1}^m \overline{d_{i\alpha(i)}}$ for these choices. □

We now extend lemma 5.6 to the case of conjugate bilinear functions that need not be inner products.

5.8 (**Conjugate bilinear functions on tensor spaces, general remarks**). We specify some notational conventions to be used in lemma 5.12. Assume $V_1, \ldots, V_m$, $\dim(V_i) = n_i$, and $\hat V_1, \ldots, \hat V_m$, $\dim(\hat V_i) = \hat n_i$, are vector spaces over the complex numbers, $\mathbb{C}$. Let $W = \otimes_{i=1}^m V_i$ and $\hat W = \otimes_{i=1}^m \hat V_i$. Suppose, for $i = 1, \ldots m$, $E_i = \{e_{i1}, \ldots, e_{in_i}\}$ is a basis for V_i and $\hat E_i = \{\hat e_{i1}, \ldots, \hat e_{in_i}\}$ is a basis for $\hat V_i$. Assume $x_i = \sum_{k=1}^{n_i} c_{ik}e_{ik}$ and $y_i = \sum_{k=1}^{\hat n_i} d_{ik}\hat e_{ik}$. We use the notation

$$x^{\otimes} := x_1 \otimes \cdots \otimes x_m \text{ and } e^{\otimes}_\gamma := e_{1\gamma(1)} \otimes \cdots \otimes e_{m\gamma(m)}. \tag{5.9}$$

Thus, $\hat e^{\otimes}_\gamma = \hat e_{1\gamma(1)} \otimes \cdots \otimes \hat e_{m\gamma(m)}$. Let $\Gamma = \Gamma(n_1, \ldots, n_m)$ and $\hat\Gamma = \Gamma(\hat n_1, \ldots, \hat n_m)$ (notation 2.3). Let $\{e^{\otimes}_\gamma \,|\, \gamma \in \Gamma\}$ be the basis of W induced by the bases E_i, and define $\{\hat e^{\otimes}_\gamma \,|\, \gamma \in \hat\Gamma\}$ similarly for $\hat E_i$ and $\hat W$. Assume that $\varphi \in M(W, \hat W : \mathbb{C})$ is conjugate bilinear: for all $a, a', b, b' \in W$ or $\hat W$ (as appropriate) and $c, d \in \mathbb{C}$

$$\begin{aligned}\varphi(ca + da', b) &= c\varphi(a, b) + d\varphi(a', b)\\ \varphi(b, ca + da') &= \bar c\varphi(b, a) + \bar d\varphi(b, a').\end{aligned} \tag{5.10}$$

If $\{f_\gamma \,|\, \gamma \in \Gamma\}$, $\{\hat f_\gamma \,|\, \gamma \in \hat\Gamma\}$ are bases for W, $\hat W$, respectivly and $a = \sum_{\alpha\in\Gamma} c_\alpha f_\alpha$ and $b = \sum_{\beta\in\hat\Gamma} d_\beta \hat f_\beta$ then

$$\varphi(a, b) = \sum_{\alpha\in\Gamma}\sum_{\beta\in\hat\Gamma} c_\alpha\overline{d_\beta}\varphi(f_\alpha, \hat f_\beta). \tag{5.11}$$

Lemma 5.12 (Conjugate bilinear functions as products). *We use the terminology of 5.8. Assume $\varphi_i \in M(V_i, \hat V_i : \mathbb{C})$, $i \in \underline{m}$, is conjugate bilinear. There* *exists a unique conjugate bilinear function $\varphi \in M(W, \hat{W} : \mathbb{C})$ such that*

$$\varphi(x_1 \otimes \cdots \otimes x_m, y_1 \otimes \cdots \otimes y_m) = \prod_{i=1}^{m} \varphi_i(x_i, y_i). \tag{5.13}$$

This φ is defined by $\varphi(e^{\otimes}_{\alpha}, \hat{e}^{\otimes}_{\beta}) := \prod_{i=1}^{m} \varphi_i(e_{i\alpha(i)}, \hat{e}_{i\beta(i)})$ for all $\alpha \in \Gamma(n_1, \ldots, n_m)$ and $\beta \in \Gamma(\hat{n}_1, \ldots, \hat{n}_m)$.

Proof. In equation 5.11, take $f_\gamma := e^{\otimes}_{\gamma}$, $\gamma \in \Gamma$ and $\hat{f}_\gamma := \hat{e}^{\otimes}_{\gamma}$, $\gamma \in \hat{\Gamma}$, and let $a = x^{\otimes}$, $b = y^{\otimes}$ (5.8). We have $x^{\otimes} = \sum_{\alpha\in\Gamma}(\prod_{i=1}^{m} c_{i\alpha(i)})e^{\otimes}_{\alpha}$ and $y^{\otimes} = \sum_{\beta\in\hat{\Gamma}}(\prod_{i=1}^{m} d_{i\beta(i)})\hat{e}^{\otimes}_{\beta}$. Equation 5.11 becomes

$$\varphi(x^{\otimes}, y^{\otimes}) = \sum_{\alpha\in\Gamma}\sum_{\beta\in\hat{\Gamma}}\prod_{i=1}^{m} c_{i\alpha(i)} \prod_{i=1}^{m} \overline{d}_{i\beta(i)} \varphi(e^{\otimes}_{\alpha}, \hat{e}^{\otimes}_{\beta}). \tag{5.14}$$

Define $\varphi(e^{\otimes}_{\alpha}, \hat{e}^{\otimes}_{\beta}) := \prod_{i=1}^{m} \varphi_i(e_{i\alpha(i)}, \hat{e}_{i\beta(i)})$. This specifies φ uniquely as a conjugate bilinear function on W (equation 5.11). We need to show that 5.13 holds. Write 5.14 as

$$\varphi(x^{\otimes}, y^{\otimes}) = \sum_{\alpha\in\Gamma}\prod_{i=1}^{m} c_{i\alpha(i)} \left(\sum_{\beta\in\hat{\Gamma}}\prod_{i=1}^{m} \overline{d}_{i\beta(i)} \varphi_i(e_{i\alpha(i)}, \hat{e}_{i\beta(i)}) \right). \tag{5.15}$$

Next, we interchange product and sum (2.6) on the expression inside the parentheses of 5.15. With $a_{ij} = \overline{d}_{ij}\varphi_i(e_{i\alpha(i)}, \hat{e}_{ij})$ we obtain

$$\varphi(x^{\otimes}, y^{\otimes}) = \sum_{\alpha\in\Gamma}\prod_{i=1}^{m} c_{i\alpha(i)} \prod_{i=1}^{m} \left(\sum_{j=1}^{n_i} \overline{d}_{ij} \varphi_i(e_{i\alpha(i)}, \hat{e}_{ij}) \right). \tag{5.16}$$

Using that the φ_i are conjugate bilinear and recalling that $y_i = \sum_{j=1}^{n_i} d_{ij}\hat{e}_{ij}$, we obtain

$$\varphi(x^{\otimes}, y^{\otimes}) = \sum_{\alpha\in\Gamma}\prod_{i=1}^{m} c_{i\alpha(i)} \varphi_i(e_{i\alpha(i)}, y_i). \tag{5.17}$$

Repeating this sum-interchange process with $a_{ij} = c_{ij}\varphi_i(e_{ij}, y_i)$ gives

$$\varphi(x^{\otimes}, y^{\otimes}) = \prod_{i=1}^{m} \varphi_i(x_i, y_i). \tag{5.18}$$

By definition, $\varphi_i(x_i, y_i)$ are independent of the E_i and $\hat{E}_i$, $i \in \underline{m}$. □

6. Direct sums and tensor products

6.1 (**Partitioning** $\Gamma(n_1, \ldots, n_m) = \underline{n_1} \times \cdots \underline{n_m}$). Let $\mathbb{D}_i = \{D_{i1}, D_{i2}, \ldots D_{ir_i}\}$ be a partition of $\underline{n_i}$ for $i = 1, 2, \ldots, m$ with blocks ordered as indicated. For every $\gamma \in \Gamma(n_1, \ldots, n_m)$ there exists a unique $\alpha \in \underline{r_1} \times \cdots \times \underline{r_m}$ such that $\gamma(i) \in D_{i\alpha(i)}$ for $i = 1, 2, \ldots, m$.

Let $\mathbb{D}_\alpha := \times_{i=1}^m D_{i\alpha(i)} = \{\gamma \,|\, \gamma(i) \in D_{i\alpha(i)},\, i \in \underline{m}\}$. The set $\mathbb{D}_\Gamma = \{\mathbb{D}_\alpha \,|\, \alpha \in \underline{r_1} \times \cdots \underline{r_m}\}$, is a partition of $\Gamma(n_1, \ldots, n_m)$ with blocks $\mathbb{D}_\alpha$. At this point, we haven't specified an order on the blocks of $\mathbb{D}_\Gamma$. A natural choice would be to use lexicographic order on the domain, $\Gamma(r_1, \ldots, r_m)$, of the map $\alpha \mapsto \mathbb{D}_\alpha$.

We introduce the following definition.

Definition 6.2 (Partition of Γ induced by partitions of the $\underline{n_i}$). Let $\mathbb{D}_i = \{D_{i1}, D_{i2}, \ldots D_{ir_i}\}$ be a partition of $\underline{n_i}$ for $i = 1, \ldots, m$ with blocks ordered as indicated. The partition

$$\mathbb{D}_\Gamma = \{\mathbb{D}_\alpha \,|\, \mathbb{D}_\alpha = \times_{i=1}^m D_{i\alpha(i)},\ \alpha \in \underline{r_1} \times \cdots \underline{r_m}\}$$

is the *partition of* $\Gamma(n_1, \ldots, n_m)$ *induced by the partitions* $\mathbb{D}_i$, $i = 1, \ldots, m$.

6.3 (**Examples of partitions of type $\mathbb{D}_\Gamma$**). Let V_1 and V_2 be vectors spaces with ordered bases $E_1 = \{e_{11}, e_{12}, e_{13}\}$ and $E_2 = \{e_{21}, e_{22}, e_{23}, e_{24}\}$ respectively. For convenience, we list the elements of Γ in lexicographic order ($m = 2$ in this case):

$\Gamma(3,4) = (1,1), (1,2), (1,3), (1,4), (2,1), (2,2), (2,3), (2,4), (3,1), (3,2), (3,3), (3,4)$.

Take the ordered partitions D_i to be as follows: $D_1 = \{D_{11}, D_{12}\}$ with $D_{11} = \{1,3\}$ and $D_{12} = \{2\}$; $D_2 = \{D_{21}, D_{22}\}$ with $D_{21} = \{2,4\}$ and $D_{22} = \{1,3\}$.

$$\mathbb{D}_\Gamma = \{\mathbb{D}_\alpha \,|\, \mathbb{D}_\alpha = \times_{i=1}^2 D_{i\alpha(i)},\ \alpha \in \underline{2} \times \underline{2}\}.$$

We construct $\mathbb{D}_\alpha$:

With $\alpha = (1,1)$ we get

$$\mathbb{D}_{(11)} = D_{11} \times D_{21} = \{1,3\} \times \{2,4\} = \{(1,2), (1,4), (3,2), (3,4)\}.$$

With $\alpha = (1,2)$ we get

$$\mathbb{D}_{(12)} = D_{11} \times D_{22} = \{1,3\} \times \{1,3\} = \{(1,1), (1,3), (3,1), (3,3)\}.$$

With $\alpha = (2,1)$ we get

$$\mathbb{D}_{(21)} = D_{12} \times D_{21} = \{2\} \times \{2,4\} = \{(2,2), (2,4)\}.$$

With $\alpha = (2,2)$ we get

$$\mathbb{D}_{(22)} = D_{12} \times D_{22} = \{2\} \times \{1,3\} = \{(2,1), (2,3)\}.$$

Thus, $\mathbb{D}_\Gamma =$

$\{\{(1,2), (1,4), (3,2), (3,4)\}, \{(1,1), (1,3), (3,1), (3,3)\}, \{(2,2), (2,4)\}, \{(2,1), (2,3)\}\}$.

The blocks of $\mathbb{D}_\Gamma$ can be ordered by lexicographically ordering the domain, $\Gamma(2,2)$, of $\alpha \mapsto \mathbb{D}_\alpha$ to get $\mathbb{D}_{(11)}, \mathbb{D}_{(12)}, \mathbb{D}_{(21)}, \mathbb{D}_{(22)}$. If we let $W_{ij} = \langle e_{it} \,|\, t \in D_{ij}\rangle$ then $V_i = \oplus_{j=1}^2 W_{ij}$. For example, $W_{21} = \langle e_{22}, e_{24}\rangle$ and $W_{22} = \langle e_{21}, e_{23}\rangle$, and $V_2 = W_{21} \oplus W_{22}$.

Remark 6.4 (**Direct sums and bases**). The following is a slight generalization of a standard theorem from linear algebra:

Let V_i, $\dim(V_i) = n_i$, $i = 1, \ldots, n_i$ be vector spaces. Let $\mathbb{D}_i = \{D_{i1}, D_{i2}, \ldots D_{ir_i}\}$ be an ordered partition of $\underline{n_i}$ for $i = 1, \ldots, m$. Then $V_i = \oplus_{j=1}^{r_i} W_{ij}$, $\dim(W_{ij}) = |D_{ij}|$, if and only if there exists bases $e_{i1}, \ldots, e_{in_i}$ such that $W_{ij} = \langle e_{it} \,|\, t \in D_{ij} \rangle$.

6.5 (**Tensor products of direct sums**). Let (P, ν) be a tensor product (3.1) of V_i, $\dim(V_i) = n_i$ Let $E_i = \{e_{i1}, \ldots, e_{in_i}\}$ be ordered bases for the V_i, $i = 1, \ldots, m$. Let $\{p_\gamma \,|\, \gamma \in \Gamma(n_1, \ldots, n_m), p_\gamma = \nu(e_{1\gamma(1)}, \ldots, e_{m\gamma(m)})\}$ be the associated basis. As in definition 6.2, examples 6.3 and remark 6.4, let $\mathbb{D}_i = \{D_{i1}, D_{i2}, \ldots D_{ir_i}\}$ be an ordered partition of $\underline{n_i}$ for $i = 1, \ldots, m$. If $W_{ij} = \langle e_{it} \,|\, t \in D_{ij} \rangle$ then $V_i = \oplus_{j=1}^{r_i} W_{ij}$, $i = 1 \ldots m$. From definition 6.2 we have

$$\mathbb{D}_\Gamma = \{\mathbb{D}_\alpha \,|\, \mathbb{D}_\alpha = \times_{i=1}^{m} D_{i\alpha(i)},\ \alpha \in \underline{r_1} \times \cdots \times \underline{r_m}\}.$$

Consider $\times_{i=1}^{m} W_{i\alpha(i)}$. Let $P_\alpha = \langle p_\gamma \,|\, \gamma \in \mathbb{D}_\alpha \rangle$. From the fact that $\mathbb{D}_\Gamma$ is a partition of Γ, we have $P = \oplus_\alpha P_\alpha$. We claim that (P_α, ν_α) is a tensor product of $W_{1\alpha(1)}, \ldots, W_{m\alpha(m)}$ where ν_α is defined by

$$\nu_\alpha(e_{1\gamma(1)}, \ldots, e_{m\gamma(m)}) := \nu(e_{1\gamma(1)}, \ldots, e_{m\gamma(m)}).$$

Thus, $P_\alpha = \langle \nu_\alpha(e_{1\gamma(1)}, \ldots, e_{m\gamma(m)}) \,|\, \gamma \in \mathbb{D}_\alpha \rangle$. Note that

$$\dim(P_\alpha) = \prod_{i=1}^{m} |D_{i\alpha(i)}| = \prod_{i=1}^{m} \dim(W_{i\alpha(i)}).$$

Thus, by definition 3.20, (P_α, ν_α) is a subspace tensor product of (P, ν).

Theorem 6.6. *Let $V_1, \ldots, V_m$ be vector spaces of dimensions $\dim(V_1) = n_i$. Suppose $V_i = \oplus_{t=1}^{r_i} W_{it}$, $i = 1, \ldots, m$, is the direct sum of subspaces W_{it}. Then*

$$\bigotimes_{i=1}^{m} V_i = \bigotimes_{i=1}^{m} \bigoplus_{t=1}^{r_i} W_{it} = \bigoplus_{\alpha \in \Gamma(r_1, \ldots, r_m)} \bigotimes_{i=1}^{m} W_{i\alpha(i)} \tag{6.7}$$

where the $\otimes_{i=1}^{m} W_{i\alpha(i)}$ are subspace tensor products of $\otimes_{i=1}^{m} V_i$.

Proof. Let $E_i = \{e_{i1}, \ldots, e_{in_i}\}$, ordered bases of V_i, $i = 1, \ldots, m$, and $\mathbb{D}_i = \{D_{i1}, \ldots, D_{ir_i}\}$, ordered partitions of $\underline{n_i}$, be such that $W_{ij} = \langle e_{it} \,|\, t \in D_{ij} \rangle$. The associated basis for $\otimes_{i=1}^{m} V_i$ is $\{e_{1,\gamma(1)} \otimes \cdots \otimes e_{m,\gamma(m)} \,|\, \gamma \in \Gamma(n_1, \ldots, n_m)\}$. From definition 6.2, the induced partition of $\Gamma(n_1, \ldots, n_m)$ is

$$\mathbb{D}_\Gamma = \{\mathbb{D}_\alpha \,|\, \mathbb{D}_\alpha = \times_{i=1}^{m} D_{i\alpha(i)},\ \alpha \in \underline{r_1} \times \cdots \underline{r_m}\}.$$

Thus,

$$\bigotimes_{i=1}^{m} V_i = \bigoplus_{\alpha \in \Gamma(r_1, \ldots, r_m)} \langle e_{1,\gamma(1)} \otimes \cdots \otimes e_{m,\gamma(m)} \,|\, \gamma \in \mathbb{D}_\alpha \rangle.$$

Each vector space $P_\alpha = \langle e_{1,\gamma(1)} \otimes \cdots \otimes e_{m,\gamma(m)} \,|\, \gamma \in \mathbb{D}_\alpha \rangle$ is a subspace of $\otimes_{i=1}^{m} V_i$ and has dimension $\prod_{i=1}^{m} |D_{i\alpha(i)}| = \prod_{i=1}^{m} \dim(W_{i\alpha(i)})$. If we define ν_α :

$\times_{i=1}^{m} W_{i\alpha(i)} \to P_\alpha$ by $\nu_\alpha(x_1, \ldots, x_m) := \nu(x_1, \ldots, x_m)$. Then, (P_α, ν_α) is the required construction. □

6.8 (**Example: tensor products of direct sums**). We follow the discussion 6.5. Let (P, ν) be a tensor product of V_i, $\dim(V_i) = n_i$ with $n_1 = n_3 = 2$ and $n_2 = n_4 = 4$. Let $E_i = \{e_{i1}, \ldots, e_{in_i}\}$ be ordered bases for the V_i, $i = 1, \ldots, 4$. Specifically, take $V_1 = \mathbf{M}_{2,1}$, $V_2 = \mathbf{M}_{2,2}$, $V_3 = \mathbf{M}_{1,2}$ and $V_4 = \mathbf{M}_{2,2}$. Define the bases

$$\begin{gathered} e_{11} = \begin{pmatrix} 1 \\ 0 \end{pmatrix},\ e_{12} = \begin{pmatrix} 0 \\ 1 \end{pmatrix},\ e_{31} = \begin{pmatrix} 1 & 0 \end{pmatrix},\ e_{32} = \begin{pmatrix} 0 & 1 \end{pmatrix} \\ e_{21} = e_{41} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},\quad e_{22} = e_{42} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \\ e_{23} = e_{43} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix},\quad e_{24} = e_{44} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}. \end{gathered} \tag{6.9}$$

Assume the partitions D_i of $\underline{n_i}$ are

$$D_1 = D_3 = \{\{1\}, \{2\}\},\ D_2 = D_4 = \{\{1, 2, 3, 4\}\}. \tag{6.10}$$

$$V_1 = W_{11} \oplus W_{12} \text{ where } W_{11} = \langle e_{11} \rangle,\ W_{12} = \langle e_{12} \rangle \text{ and } V_2 = W_{21} = \mathbf{M}_{2,2} \tag{6.11}$$

$$V_3 = W_{31} \oplus W_{32} \text{ where } W_{31} = \langle e_{31} \rangle,\ W_{12} = \langle e_{32} \rangle \text{ and } V_4 = W_{41} = \mathbf{M}_{2,2}. \tag{6.12}$$

We have

$$\mathbb{D}_\Gamma = \{\mathbb{D}_\alpha \,|\, \mathbb{D}_\alpha = \times_{i=1}^{4} D_{i\alpha(i)},\ \alpha \in \underline{2} \times \underline{1} \times \underline{2} \times \underline{1}\}.$$

Thus, $\otimes_{i=1}^{4} V_i =$

$$\begin{gathered} \langle e_{11} \rangle \otimes \mathbf{M}_{2,2} \otimes \langle e_{31} \rangle \otimes \mathbf{M}_{2,2} \bigoplus \langle e_{11} \rangle \otimes \mathbf{M}_{2,2} \otimes \langle e_{32} \rangle \otimes \mathbf{M}_{2,2} \bigoplus \\ \langle e_{12} \rangle \otimes \mathbf{M}_{2,2} \otimes \langle e_{31} \rangle \otimes \mathbf{M}_{2,2} \bigoplus \langle e_{12} \rangle \otimes \mathbf{M}_{2,2} \otimes \langle e_{32} \rangle \otimes \mathbf{M}_{2,2}. \end{gathered} \tag{6.13}$$

Each of the summands in equation 6.13 is a subspace tensor product (3.21) of $\otimes_{i=1}^{4} V_i$, and each is isomorphic to $\mathbf{M}_{2,2} \otimes \mathbf{M}_{2,2}$, each with a different isomorphism. For example, the basis elements of $\langle e_{11} \rangle \otimes \mathbf{M}_{2,2} \otimes \langle e_{31} \rangle \otimes \mathbf{M}_{2,2}$ are

$$\{e_{11} \otimes e_{2\beta(2)} \otimes e_{31} \otimes e_{4\beta(4)} \,|\, \beta \in \{1, 2\}^{\{2,4\}}\} \tag{6.14}$$

which correspond bijectively to basis elements $\{e_{2\beta(2)} \otimes e_{4\beta(4)} \,|\, \beta \in \{1, 2\}^{\{2,4\}}\}$ of $\mathbf{M}_{2,2} \otimes \mathbf{M}_{2,2}$.

6.15 (**Example: tensor products of direct sums – Kronecker products**). In example 6.8 our model for the tensor products defined the underlying vector spaces as matrices but didn't use any properties of them except for the dimensions. Let (P, ν) be a tensor product of V_i, $i = 1, \ldots, 4$, $\dim(V_i) = n_i$ with $n_1 = n_3 = 2$ and $n_2 = n_4 = 4$. Specifically, take $V_i \in \mathbf{M}_{p_i, q_i}$ as follows: $V_1 = \mathbf{M}_{2,1}$, $V_2 = \mathbf{M}_{2,2}$, $V_3 = \mathbf{M}_{1,2}$ and $V_4 = \mathbf{M}_{2,2}$. Take $P = \mathbf{M}_{p,q}$ where $p = p_1 p_2 p_3 p_4 = 8$ and $q = q_1 q_2 q_3 q_4 = 8$. In this example, we associate the bases of these vector spaces with the standard bases of matrices (as used in discussions 4.1, 4.5, 4.10, and 4.11).

$$\begin{aligned} E^{(1)}_{11} = \left(\begin{smallmatrix}1\\0\end{smallmatrix}\right),\ E^{(1)}_{21} = \left(\begin{smallmatrix}0\\1\end{smallmatrix}\right),\ E^{(3)}_{11} = \left(\begin{smallmatrix}1 & 0\end{smallmatrix}\right),\ E^{(3)}_{12} = \left(\begin{smallmatrix}0 & 1\end{smallmatrix}\right)\\ E^{(2)}_{11} = E^{(4)}_{11} = \left(\begin{smallmatrix}1&0\\0&0\end{smallmatrix}\right),\ \ E^{(2)}_{12} = E^{(4)}_{12} = \left(\begin{smallmatrix}0&1\\0&0\end{smallmatrix}\right),\\ E^{(2)}_{21} = E^{(4)}_{21} = \left(\begin{smallmatrix}0&0\\1&0\end{smallmatrix}\right),\ \ E^{(2)}_{22} = E^{(4)}_{22} = \left(\begin{smallmatrix}0&0\\0&1\end{smallmatrix}\right). \end{aligned} \tag{6.16}$$

Each basis $E^{(i)}$ is ordered lexicographically. Thus, $E^{(2)} = (E^{(2)}_{11}, E^{(2)}_{12}, E^{(2)}_{21}, E^{(2)}_{22})$. In this case, equation 6.13 becomes $\otimes_{i=1}^4 V_i =$

$$\begin{aligned} \langle E^{(1)}_{11}\rangle \otimes \mathbf{M}_{2,2} \otimes \langle E^{(3)}_{11}\rangle \otimes \mathbf{M}_{2,2} \bigoplus \langle E^{(1)}_{11}\rangle \otimes \mathbf{M}_{2,2} \otimes \langle E^{(3)}_{12}\rangle \otimes \mathbf{M}_{2,2} \bigoplus\\ \langle E^{(1)}_{21}\rangle \otimes \mathbf{M}_{2,2} \otimes \langle E^{(3)}_{11}\rangle \otimes \mathbf{M}_{2,2} \bigoplus \langle E^{(1)}_{21}\rangle \otimes \mathbf{M}_{2,2} \otimes \langle E^{(3)}_{12}\rangle \otimes \mathbf{M}_{2,2}. \end{aligned} \tag{6.17}$$

We interpret these tensor products as Kronecker products of matrices as in examples 4.11 and 4.16. Let $A_i = (a_i(s,t)) \in \mathbf{M}_{p_i,q_i}$, $i = 1, \ldots, 4$. We have

$$A_i = \sum_{(\mu(i),\kappa(i))\in\Gamma(p_i,q_i)} a_i(\mu(i),\kappa(i))E^{(i)}_{\mu(i),\kappa(i)} \text{ and}$$

$$A_1 \otimes A_2 \otimes A_3 \otimes A_4 = \sum_{\mu,\kappa} \prod_{i=1}^{4} a_i(\mu(i),\kappa(i))E_{\mu,\kappa} \tag{6.18}$$

where the $E_{\mu,\kappa}(\alpha,\beta) = \mathcal{X}((\alpha,\beta) = (\mu,\kappa))$ are the standard basis elements of $\mathbf{M}_{p,q}$. If $\mathbf{A} = A_1 \otimes \cdots \otimes A_4$, the entry $\mathbf{A}(\mu,\kappa) = \prod_{i=1}^4 a_i(\mu(i),\kappa(i))$. Denote the sequences of possible row values by $\Gamma_{\mathbf{r}} := \Gamma(p_1,p_2,p_3,p_4) = \Gamma(2,2,1,2)$ and the sequence of possible column values by $\Gamma_{\mathbf{c}} := \Gamma(q_1,q_2,q_3,q_4) = \Gamma(1,2,2,2)$. Order both $\Gamma_{\mathbf{r}}$ and $\Gamma_{\mathbf{c}}$ lexicographically. The basis matrices

$$\{E_{\mu,\kappa} \mid \mu \in \Gamma_{\mathbf{r}}\,,\ \kappa \in \Gamma_{\mathbf{c}}\} \tag{6.19}$$

are ordered by lexicographic order on $(\mu,\kappa) \in \Gamma_{\mathbf{r}} \times \Gamma_{\mathbf{c}}$. In matrix 6.20, the rows are shown indexed lexicographically by $\Gamma_{\mathbf{r}}$ and the columns by $\Gamma_{\mathbf{c}}$ (corresponding to $\Phi = \Gamma_{\mathbf{r}}$ and $\Lambda = \Gamma_{\mathbf{c}}$ in notation 7.16 below). Note that if

$$A_1 \otimes \cdots \otimes A_4 = \mathbf{A} \in \langle E^{(1)}_{11}\rangle \otimes \mathbf{M}_{2,2} \otimes \langle E^{(3)}_{11}\rangle \otimes \mathbf{M}_{2,2}$$

then the support of A (7.15) is in the submatrix of 6.20 labeled with "a". Similarly for

$$B_1 \otimes \cdots \otimes B_4 = \mathbf{B} \in \langle E^{(1)}_{11}\rangle \otimes \mathbf{M}_{2,2} \otimes \langle E^{(3)}_{12}\rangle \otimes \mathbf{M}_{2,2} \text{ (labeled } b),$$

$$C_1 \otimes \cdots \otimes C_4 = \mathbf{C} \in \langle E^{(1)}_{21}\rangle \otimes \mathbf{M}_{2,2} \otimes \langle E^{(3)}_{11}\rangle \otimes \mathbf{M}_{2,2} \text{ (labeled } c),$$

$$D_1 \otimes \cdots \otimes D_4 = \mathbf{D} \in \langle E^{(1)}_{21}\rangle \otimes \mathbf{M}_{2,2} \otimes \langle E^{(3)}_{12}\rangle \otimes \mathbf{M}_{2,2} \text{ (labeled } d).$$

$$
(6.20) \qquad \begin{array}{c} \\ 1111 \\ 1112 \\ 1211 \\ 1212 \\ 2111 \\ 2112 \\ 2211 \\ 2212 \end{array}
\begin{array}{c}
\begin{array}{cccccccc} 1111 & 1112 & 1121 & 1122 & 1211 & 1212 & 1221 & 1222 \end{array} \\
\begin{pmatrix}
a & a & b & b & a & a & b & b \\
a & a & b & b & a & a & b & b \\
a & a & b & b & a & a & b & b \\
a & a & b & b & a & a & b & b \\
c & c & d & d & c & c & d & d \\
c & c & d & d & c & c & d & d \\
c & c & d & d & c & c & d & d \\
c & c & d & d & c & c & d & d
\end{pmatrix}
\end{array}
$$

7. Background concepts and notation

In this section we review concepts and notation from basic discrete mathematics courses. This section can be skipped and reviewed as needed. Wikipedia is a good source.

7.1 (**Sets and lists**). The empty set is denoted by $\emptyset$. Sets are specified by braces: $A = \{1\}$, $B = \{1,2\}$. They are unordered, so $B = \{1,2\} = \{2,1\}$. Sets $C = D$ if $x \in C$ implies $x \in D$ (equivalently, $C \subseteq D$) and $x \in D$ implies $x \in C$. If you write $C = \{1,1,2\}$ and $D = \{1,2\}$ then, by the definition of set equality, $C = D$. If A is a set then $\mathbb{P}(A)$ is all subsets of A and $\mathbb{P}_k(A)$ is all subsets of cardinality, $|A| = k$.

A *list*, *vector* or *sequence* (specified by parentheses) is ordered: $C' = (1,1,2)$ is not the same as $(1,2,1)$ or $(1,2)$. Two lists (vectors, sequences), $(x_1, x_2, \ldots, x_n) = (y_1, y_2, \ldots, y_m)$, are equal if and only if $n = m$ and $x_i = y_i$ for $i = 1, \ldots, n$. A list such as $(x_1, x_2, \ldots, x_n)$ is also written $x_1, x_2, \ldots, x_n$, without the parentheses. Sometimes a list L will be specified as a set, S with additional information defining the linear order $<_S$. Lex order 7.2 is an example.

Definition 7.2 (Lexicographic, "lex," order). Let C_i, $i = 1, \ldots, k$, be lists of distinct elements, and let $L = C_1 \times \cdots \times C_k$ be the their product as sets. Define lexicographic order on L, indicated by $<_L$, by $(a_1, \ldots, a_k) <_L (b_1, \ldots, b_k)$ if $a_1 < b_1$ or if there is some $t \leq k$ such that $a_i = b_i$, $i < t$, but $a_t < b_t$. We have used "$<$" for the various linear orders on the C_i and will usually do the same for lexicographic ("lex") order, replacing $<_L$ by simply $<$.

Definition 7.3 (Position and rank functions for linear orders). Let $(\Lambda, \leq)$ be a finite linearly ordered set. For $x \in \Lambda$ and $S \subseteq \Lambda$ let

$$
(7.4) \qquad \pi_S^\Lambda(x) := |\{t \mid t \in S, t \leq x\}| \quad \text{and} \quad \rho_S^\Lambda(x) := |\{t \mid t \in S, t < x\}|\,.
$$

π_S^Λ is called the *position function* for Λ relative to S and ρ_S^Λ the *rank function* for Λ relative to S. If $S = \Lambda$, we use π^Λ instead of π_Λ^Λ, and, similarly, we use ρ^Λ instead of ρ_Λ^Λ.

Definition 7.5 (Order relation, partially ordered set, poset). A subset $R \subseteq S \times S$ is called a *binary relation on S*. The statement $(x,y) \in R$ is also denoted by $x \; R \; y$. Likewise, $(x, y) \notin R$ is denoted by $x \not R \; y$. A binary relation on a set S is called an *order relation* if it satisfies the following three conditions and (usually written $x \preceq y$ instead of $x \; R \; y$ in this case).

- (i) (Reflexive) For all $s \in S$ we have $s \preceq s$.
- (ii) (Antisymmetric) For all $s, t \in S$ such that $s \neq t$, if $s \preceq t$ then $t \npreceq s$.
- (iii) (Transitive) For all $r, s, t \in S$, $r \preceq s$ and $s \preceq t$ implies $r \preceq t$.

A set S together with an order relation $\preceq$ is called a *partially ordered set* or *poset*, $(S, \preceq)$. If for all $(x, y) \in S \times S$, either $x \preceq y$ or $y \preceq x$ (both if $x = y$) then $(S, \preceq)$ is a *linearly ordered set* (e.g., lists 7.1, lexicographic order 7.2).

Definition 7.6 (Function). Let A and B be sets. A function f from A to B is a rule that assigns to each element $x \in A$ a unique element $y \in B$. We write $y = f(x)$. Two functions f and g from A to B are equal if $f(x) = g(x)$ for all $x \in A$.

Given a function f from A to B, we can define a set $F \subseteq A \times B$ by

$$F = \{(x, f(x)) \,|\, x \in A\} \tag{7.7}$$

We call F the *graph* of f, denoted by Graph(f). A subset $F \subseteq A \times B$ is the graph of a function from A to B if and only if it satisfies the following two conditions:

$$\text{G1}: \quad (x, y) \in F \quad \text{and} \quad (x, y') \in F \implies y = y' \tag{7.8}$$

$$\text{G2}: \quad \{x \,|\, (x, y) \in F\} = A. \tag{7.9}$$

Two functions, f and g, are equal if and only if their graphs are equal as sets: Graph(f) = Graph(g). The set A is called the *domain* of f (written A = domain(f)), and B is called the *range* of f (written B = range(f)). The notation $f : A \to B$ is used to denote that f is a function with domain A and range B. For $S \subseteq A$, define $f(S)$ (image of S under f) by $f(S) \equiv \{f(x) \,|\, x \in S\}$. In particular, $f(A)$ is called the *image* of f (written $f(A)$ = image(f)). The set of *all* functions with domain A and range B can be written $\{f \,|\, f : A \to B\}$ or simply as B^A. If A and B are finite then $\left|B^A\right|$ is $|B|^{|A|}$. The *characteristic or indicator function* of a set $S \subseteq A$, $\mathcal{X}_S : A \to \{0, 1\}$, is defined by

$$\mathcal{X}_S(x) = 1 \text{ if and only if } x \in S. \tag{7.10}$$

The *restriction* f_S of $f : A \to B$ to a subset $S \subseteq A$ is defined by

$$f_S : S \to B \text{ where } f_S(x) = f(x) \text{ for all } x \in S. \tag{7.11}$$

If $f : A \to B$ and $g : B \to C$ then the *composition* of g and f, denoted by $gf : A \to C$, is defined by

$$gf(x) = g(f(x)) \text{ for } x \in A. \tag{7.12}$$

There are many ways to describe a function. Any such description must specify the domain, the range, and the rule for assigning some range element to each domain element. You could specify a function using set notation: the domain is the set $\{1,2,3,4\}$, the range to be the set $\{a,b,c,d,e\}$, and the function f is the set: $\text{Graph}(f) = \{(1,a)\,(2,c)\,(3,a)\,(4,d)\}$. Alternatively, you could describe the same function by giving the range as $\{a,b,c,d,e\}$ and using *two line notation*

$$f = \begin{pmatrix} 1\ 2\ 3\ 4 \\ a\ c\ a\ d \end{pmatrix}. \tag{7.13}$$

If we assume the domain, in order, is 1 2 3 4, then 7.13 can be abbreviated to *one line*: $a\,c\,a\,d$. Sometimes it is convenient to describe a function $f \in B^A$ by the "maps to" notation. For example, $2 \mapsto c$, $4 \mapsto d$, otherwise $x \mapsto a$.

Definition 7.14 (Sets of functions). Let $\underline{n} = \{1,2,\ldots,n\}$ and let $\underline{p}^{\underline{n}}$ be all functions with domain $\underline{n}$, range $\underline{p}$. Define

$$\text{SNC}(n,p) = \{f \mid f \in \underline{p}^{\underline{n}}, i < j \implies f(i) < f(j)\} \ \ \textbf{(strictly increasing)}$$

$$\text{WNC}(n,p) = \{f \mid f \in \underline{p}^{\underline{n}}, i < j \implies f(i) \leq f(j) \ \ \textbf{(weakly increasing)}$$

$$\text{INJ}(n,p) = \{f \mid f \in \underline{p}^{\underline{n}}, i \neq j \implies f(i) \neq f(j)\} \ \ \textbf{(injective)}$$

$$\text{PER}(n) = \text{INJ}(n,n) \ \ \textbf{(permutations of } \underline{n}\textbf{)}.$$

From combinatorics, $|\text{INJ}(n,p)| = (p)_n = p(p-1)\cdots(p-n+1)$, $|\text{PER(n)}| = \text{n!}$,

$$|\text{SNC}(n,p)| = \binom{p}{n} \ \text{ and } \ |\text{WNC}(n,p)| = \binom{p+n-1}{n}.$$

More generally, if $X \subseteq \underline{n}$ and $Y \subseteq \underline{p}$, then $\text{SNC}(X,Y)$ denotes the strictly increasing functions from X to Y. We define $\text{WNC}(X,Y)$ and $\text{INJ}(X,Y)$ similarly. Sometimes "increasing" is used instead of "strictly increasing" or "nondecreasing" instead of "weakly increasing"

Definition 7.15 (Matrix). Let m,n be positive integers. An m by n matrix with entries in a set S is a function $f : \underline{m} \times \underline{n} \to S$. The sets $\underline{m}$ and $\underline{n}$ are the *row indices* and *column indices* respectively. The set of all such f is denoted by $\mathbf{M}_{m,n}(S)$. If $S = F$, a field, the set $\{x \mid f(x) \neq 0\}$ is the *support* of f, and the matrices $E_{ij} \in \mathbf{M}_{m,n}(F)$ with a 1 in position i,j and 0 elsewhere are the *standard basis elements*. Instead of E_{ij}, we use $E_{i,j}$ (with comma) when needed for clarity.

7.16 (**Additional matrix notational conventions**). More generally, a matrix is a function $f : \Phi \times \Lambda \to S$ where Φ and Λ are linearly ordered sets (row and column indices respectively). We use $A[X|Y]$ to denote the submatrix of A gotten by retaining rows indexed by the set $X \subseteq \Phi$ and columns indexed by the set $Y \subseteq \Lambda$. We use $A(X|Y)$ to denote the submatrix of A gotten by retaining rows indexed by the set $\Phi \setminus X$ (the complement of X in Φ) and columns indexed by the set $\Lambda \setminus Y$. We also use the mixed notation $A[X|Y)$ and $A(X|Y]$ with obvious meaning. We use Θ to denote the zero matrix of the appropriate size and I to denote the identity matrix.

Definition 7.17 (Partition of a set). A partition of a set Q is a collection, $\mathcal{B}(Q)$, of nonempty subsets, X, of Q such that each element of Q is contained in *exactly one* set $X \in \mathcal{B}(Q)$. The sets $X \in \mathcal{B}(Q)$ are called the *blocks* of the partition $\mathcal{B}(Q)$. A set $D \subseteq Q$ consisting of exactly one element from each block is called a *system of distinct representatives* (or "SDR") for the partition. If each $X \in \mathcal{B}(Q)$ has $|X| = 1$, then we call $\mathcal{B}(Q)$ the *discrete* partition. If the partition $\mathcal{B}(Q) = \{Q\}$ is called the *unit* partition. The set of all partitions of Q is $\Pi(Q)$. The set of all partitions of Q with k blocks is $\Pi_k(Q)$.

If $|Q| = n$, the numbers $S(n,k) := |\Pi_k(Q)|$ are called the *Stirling numbers of the second kind* and the numbers, $B(n) := |\Pi(Q)|$ are called the Bell numbers. If $Q = \{1,2,3,4,5\}$ then $\mathcal{B}(Q) = \{\{1,3,5\},\{2,4\}\} \in \Pi_2(Q)$ is a partition of Q with two blocks: $\{1,3,5\} \in \mathcal{B}(Q)$ and $\{2,4\} \in \mathcal{B}(Q)$. The set $D = \{1,2\}$ is an SDR for $\mathcal{B}(Q)$. $S(5,2) = 15$ and $B(5) = 52$.

Definition 7.18 (Coimage partition). Let $f : A \to B$ be a function with domain A and range B. Let $\text{image}(f) = \{f(x) \,|\, x \in A\}$ (7.6). The inverse image of an element $y \in B$ is the set $f^{-1}(y) \equiv \{x \,|\, f(x) = y\}$. The *coimage* of f is the set of subsets of A :

$$\text{coimage}(f) = \{f^{-1}(y) \,|\, y \in \text{image}(f)\}. \tag{7.19}$$

The coimage(f) is a partition of A (7.17) called the *coimage partition of A induced by f*.

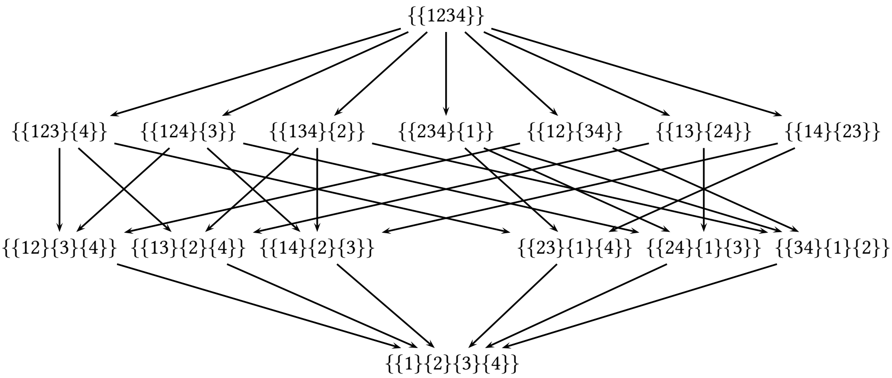
For the function f of 7.13, we have $\text{image}(f) = \{a,c,d\}$. Thus, the coimage of f is

$$\text{coimage}(f) = \{f^{-1}(a), f^{-1}(c), f^{-1}(d)\} = \{\{1,3\},\{2\},\{4\}\}. \tag{7.20}$$

7.21 (**Posets of subsets and partitions**). If $S = \mathbb{P}(Q)$ denotes all subsets of Q, then $(S, \preceq)$ is a poset (7.5) if $\preceq := \subseteq$ (set inclusion). In this case, $(S, \subseteq)$ is called the *poset of subsets* or *lattice of subsets* of Q.

Suppose $S = \Pi(Q)$ is the collection of all partitions, $\mathcal{B}(Q)$, of Q. Let $\preceq$ denote *refinement* of partitions where $\mathcal{B}_1(Q) \preceq \mathcal{B}_2(Q)$ means the blocks of $\mathcal{B}_1(Q)$ are obtained by further subdividing the blocks of $\mathcal{B}_2(Q)$ (e.g., if $\mathcal{B}_1(Q) = \{\{1,3\},\{5\},\{2\},\{4\}\}$ and $\mathcal{B}_2(Q) = \{\{1,3,5\},\{2,4\}\}$ then $\mathcal{B}_1(Q) \preceq \mathcal{B}_2(Q)$). Using the general notation for posets, if $x = \mathcal{B}_1$ and $y = \mathcal{B}_2$, $x, y \in \mathcal{B}(Q)$, $x \prec y$, then *x is covered by y*, written $x \prec_c y$, if $\{z \,|\, x \prec z \prec y\} = \emptyset$. Diagram 7.22 represents the poset $(\Pi(\underline{4}), \preceq)$ by its *covering relation* where $x \prec_c y$ is represented by $y \to x$. Such a diagram for a poset is called a *Hasse diagram*.

7.22 (**Diagram of covering relation for** $\Pi(\underline{4})$).



8. Discussion and acknowledgements

This article is a rewrite of notes I originally prepared for first year graduate students in combinatorics seminars in which multilinear algebra was applied to combinatorics (UCSD Mathematics and CSE). Additional references in multilinear algebra for these seminars were provided by Professor Marvin Marcus, my thesis advisor, friend and mentor in this subject.

I would like to thank Professor Mike Sharpe, UCSD Department of Mathematics, for considerable LaTeX typesetting assistance and for his Linux Libertine font options to the newtxmath package.

S. Gill Williamson, 2015
http://cseweb.ucsd.edu/~gill

References

[Mar73] M Marcus. *Finite Dimensional Multilinear Algebra, Part I,*. Marcel Dekker, Inc., New York, 1973.

[Mar75] M. Marcus. *Finite Dimensional Multilinear Algebra, Part II,*. Marcel Dekker, Inc., New York, 1975.