# AN INDEX NOTATION FOR TENSOR PRODUCTS AND KRONECKER PRODUCTS

## 1. Bases for Vector Spaces

Consider an identity matrix of order $N$, which can be written as follows:

$$[\,e_1 \quad e_2 \quad \cdots \quad e_N\,] = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = \begin{bmatrix} e^1 \\ e^2 \\ \vdots \\ e_N \end{bmatrix}. \tag{1}$$

On the LHS, the matrix is expressed as a collection of column vectors, denoted by $e_i; i = 1, 2, \ldots, N$, which form the basis of an ordinary $N$-dimensional Euclidean space. On the RHS, the matrix is expressed as a collection of row vectors $e^j; j = 1, 2, \ldots, N$, which form the basis of the conjugate dual space.

The basis vectors can be used in specifying arbitrary vectors in both spaces. In the primary space, there is the column vector

$$a = \sum_i a_i e_i = (a_i e_i), \tag{2}$$

and in the dual space, there is the row vector

$$b' = \sum_j b_j e^j = (b_j e^j). \tag{3}$$

Here, on the RHS, there is a notation that replaces the summation signs by parentheses. Its advantage will become apparent at a later stage, when the summations are over several indices.

A vector in the primary space can be converted to a vector in the conjugate dual space and vice versa by the operation of transposition. Thus $a' = (a_i e^i)$ is formed via the conversion $e_i \rightarrow e^i$ whereas $b = (b_j e_j)$ is formed via the conversion $e^j \rightarrow e_j$.

## 2. Elementary Tensor Products

A tensor product of two vectors is an outer product that entails the pairwise products of the elements of both vector. Consider two vectors

$$\begin{aligned} a &= [a_t; t = 1, \ldots T] = [a_1, a_2, \ldots, b_T]' \quad \text{and} \\ b &= [b_j; j = 1, \ldots, M] = [b_1, b_2, \ldots, b_M]', \end{aligned} \tag{4}$$

which need not be of the same order. Then, two kinds of tensor products can be defined. First, there are covariant tensor products. The covariant product of $a$ and $b$ is a column vector in a primary space:

$$a \otimes b = \sum_t \sum_j a_t b_j (e_t \otimes e_j) = (a_t b_j e_{tj}). \tag{5}$$

Here, the elements are arrayed in a long column in an order that is determined by the lexicographic variation of the indices $t$ and $j$. Thus, the index $j$ undergoes a complete cycles from $j = 1$ to $j = M$ with each increment of the index $t$ in the manner that is familiar from dictionary classifications. Thus

$$a \otimes b = \begin{bmatrix} a_1 b \\ a_2 b \\ \vdots \\ a_T b \end{bmatrix} = [a_1 b_1, \ldots, a_1 b_M, a_2 b_1, \ldots, a_2 b_M, \cdots, a_T b_1, \ldots, a_T b_M]'. \tag{6}$$

A covariant tensor product can also be formed from the row vectors $a'$ and $b'$. Thus, there is

$$a' \otimes b' = \sum_t \sum_j a_t b_j (e^t \otimes e^j) = (a_t b_j e^{tj}). \tag{7}$$

It will be observed that this is just the transpose of $a \otimes b$. That is to say

$$(a \otimes b)' = a' \otimes b' \quad \text{or, equivalently} \quad (a_t b_j e_{tj})' = (a_t b_j e^{tj}). \tag{8}$$

The order of the vectors in a covariant tensor product is crucial, since, as once can easily verify, it is the case that

$$a \otimes b \neq b \otimes a \qquad \text{and} \qquad a' \otimes b' \neq b' \otimes a'. \tag{9}$$

The second kind of tensor product of the two vectors is a so-called contravariant tensor product:

$$a \otimes b' = b' \otimes a = \sum_t \sum_j a_t b_j (e_t \otimes e^j) = (a_t b_j e_t^j). \tag{10}$$

This is just the familiar matrix product $ab'$, which can be written variously as

$$\begin{bmatrix} a_1 b \\ a_2 b \\ \vdots \\ a_T b \end{bmatrix} = [\, b_1 a \quad b_2 a \quad \cdots \quad b_M a \,] = \begin{bmatrix} a_1 b_1 & a_1 b_2 & \ldots & a_1 b_M \\ a_2 b_1 & a_2 b_2 & \ldots & a_2 b_M \\ \vdots & \vdots & & \vdots \\ a_T b_1 & a_T b_2 & \ldots & a_T b_M \end{bmatrix}. \tag{11}$$

Observe that

$$(a \otimes b')' = a' \otimes b \quad \text{or, equivalently,} \quad (a_t b_j e_t^j)' = (a_t b_j e_j^t). \tag{12}$$

We now propose to dispense with the summation signs and to write the various vectors as follows:

$$a = (a_t e_t), \quad a' = (a_t e^t) \quad \text{and} \quad b = (a_j e_j), \quad b' = (b_j e^j). \tag{13}$$

The convention here is that, when the products are surrounded by parentheses, summations are to be take in respect of each of the indices therein.

The convention can be applied to provide summary representations of the products under (5), (7) and (10):

$$a \otimes b' = (a_t e_t) \otimes (b_j e^j) = (a_t b_j e_t^j), \tag{14}$$

$$a' \otimes b' = (a_t e^t) \otimes (b_j e^j) = (a_t b_j e^{tj}), \tag{15}$$

$$a \otimes b = (a_t e_t) \otimes (b_j e_j) = (a_t b_j e_{tj}). \tag{16}$$

Such products are described as decomposable tensors.

## 3. Non-decomposable Tensor Products

Non-decomposable tensors are the result of taking weighted sums of decomposable tensors. Consider an arbitrary matrix $X = [x_{tj}]$ of order $T \times M$. This can be expressed as the following weighted sum of the contravariant tensor products formed from the basis vectors:

$$X = (x_{tj} e_t^j) = \sum_t \sum_j x_{tj} (e_t \otimes e^j) \tag{17}$$

The indecomposability lies in the fact that the elements $x_{tj}$ cannot be written as the products of an element indexed by $t$ and an element indexed by $j$.

From $X = (x_{tj} e_t^j)$, the following associated tensors products may be derived:

$$X' = (x_{tj} e_j^t), \tag{18}$$

$$X^r = (x_{tj} e^{tj}), \tag{19}$$

$$X^c = (x_{tj} e_{jt}). \tag{20}$$

Here, $X'$ is the transposed matrix, whereas $X^c$ is a long column vector and $X^r$ is a long row vector. Notice that, in forming $X^c$ and $X^r$ from $X$, the index that moves assumes a position at the head of the string of indices to which it is joined.

It is evident that

$$X^r = X'^{c'} \qquad \text{and} \qquad X^c = X'^{r'} \tag{21}$$

Thus, it can be seen that $X^c$ and $X^r$ are not related to each other by simple transpositions. A consequence of this is that the indices of the elements in $X^c$ follow the reverse of a lexicographic ordering.

**Example.** Consider the equation

$$y_{tj} = \mu + \gamma_t + \delta_j + \varepsilon_{tj} \tag{22}$$

wherein $t = 1, \ldots, T$ and $j = 1, \ldots, M$. This relates to a two-way analysis of variance. For a concrete interpretation, we may imagine that $y_{tj}$ is an observation taken at time $t$ in the $j$th region. Then the parameter $\gamma_t$ represents an effect that is common to all observations taken at time $t$, whereas the parameter $\delta_j$ represents a characteristic of the $j$th region that prevails through time.

In ordinary matrix notation, the set of $TM$ equations becomes

$$Y = \mu\iota_T\iota_M' + \gamma\iota_M' + \iota_T\delta' + \mathcal{E}, \tag{23}$$

where $Y = [y_{tj}]$ and $\mathcal{E} = [\varepsilon_{tj}]$ are matrices of order $T \times M$, $\gamma = [\gamma_1, \ldots, \gamma_T]'$ and $\delta = [\delta_1, \ldots, \delta_M]'$ are vectors of orders $T$ and $M$ respectively, and $\iota_T$ and $\iota_M$ are vectors of units whose orders are indicated by their subscripts. In terms of the index notation, the $TM$ equations are represented by

$$(y_{tj}e_t^j) = \mu(e_t^j) + (\gamma_t e_t^j) + (\delta_j e_t^j) + (\varepsilon_{tj}e_t^j). \tag{24}$$

An illustration is provided by the case where $T = M = 3$. Then equations (23) and (24) represent the following structure:

$$\begin{aligned}\begin{bmatrix} y_{11} & y_{12} & y_{13} \\ y_{21} & y_{22} & y_{23} \\ y_{31} & y_{32} & y_{33} \end{bmatrix} &= \mu \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} + \begin{bmatrix} \gamma_1 & \gamma_1 & \gamma_1 \\ \gamma_2 & \gamma_2 & \gamma_2 \\ \gamma_3 & \gamma_3 & \gamma_3 \end{bmatrix} \\ &\quad + \begin{bmatrix} \delta_1 & \delta_2 & \delta_3 \\ \delta_1 & \delta_2 & \delta_3 \\ \delta_1 & \delta_2 & \delta_3 \end{bmatrix} + \begin{bmatrix} \varepsilon_{11} & \varepsilon_{12} & \varepsilon_{13} \\ \varepsilon_{21} & \varepsilon_{22} & \varepsilon_{23} \\ \varepsilon_{31} & \varepsilon_{32} & \varepsilon_{33} \end{bmatrix}.\end{aligned} \tag{25}$$

## 4. Multiple Tensor Products

The tensor product entails an associative operation that combines matrices or vectors of any order. Let $B = [b_{lj}]$ and $A = [a_{ki}]$ be arbitrary matrices of orders $t \times n$ and $s \times m$ respectively. Then, their tensor product $B \otimes A$, which is also know as a Kronecker product, is defined in terms of the index notation by writing

$$(b_{lj}e_l^j) \otimes (a_{ki}e_k^i) = (b_{lj}a_{ki}e_{lk}^{ji}). \tag{26}$$

Here, $e_{lk}^{ji}$ stands for a matrix of order $st \times mn$ with a unit in the row indexed by $lk$—the $\{(l-1)s + k\}$th row—and in the column indexed by $ji$—the $\{(j-1)m + i\}$th column—and with zeros elsewhere.

In the matrix array, the row indices $lk$ follow a lexicocographic order, as do the column indices $ji$. Also, the indices $lk$ are not ordered relative to the indices $ji$. That is to say,

$$\begin{aligned} e_{lk}^{ji} &= e_l \otimes e_k \otimes e^j \otimes e^i \\ &= e^j \otimes e^i \otimes e_l \otimes e_k \\ &= e^j \otimes e_l \otimes e_k \otimes e^i \\ &= e_l \otimes e^j \otimes e^i \otimes e_k \\ &= e_l \otimes e^j \otimes e_k \otimes e^i \\ &= e^j \otimes e_l \otimes e^i \otimes e_k. \end{aligned} \tag{27}$$

The virtue of the index notation is that it makes no distinction amongst these various products on the RHS—unless a distinction can be found between such expressions as $e^{j\ i}_{\ l\ k}$ and $e_{l\ k}^{\ j\ i}$.

For an example, consider the Kronecker of two matrices as follows:

$$\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \otimes \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} b_{11}\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} & b_{12}\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \\ b_{21}\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} & b_{22}\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \end{bmatrix}$$
$$= \begin{bmatrix} b_{11}a_{11} & b_{11}a_{12} & b_{12}a_{11} & b_{12}a_{12} \\ b_{11}a_{21} & b_{11}a_{22} & b_{12}a_{21} & b_{12}a_{22} \\ b_{21}a_{11} & b_{21}a_{12} & b_{22}a_{11} & b_{22}a_{12} \\ b_{21}a_{21} & b_{21}a_{22} & b_{22}a_{21} & b_{22}a_{22} \end{bmatrix}. \tag{28}$$

Here, it can be see that the composite row indices $lk$ follow the lexicographic sequence $\{11, 12, 21, 22\}$. The column indices follow the same sequence.

## 5. Compositions

In order to demonstrate the rules of matrix composition, let us consider the matrix equation

$$Y = AXB', \tag{29}$$

which can be construed as a mapping from $X$ to $Y$. In the index notation, this is written as

$$\begin{aligned}(y_{kl}e_k^l) &= (a_{ki}e_k^i)(x_{ij}e_i^j)(b_{lj}e_j^l) \\ &= (\{a_{ki}x_{ij}b_{lj}\}e_k^l).\end{aligned} \tag{30}$$

Here, there is

$$\{a_{ki}x_{ij}b_{lj}\} = \sum_i \sum_j a_{ki}x_{ij}b_{lj}; \tag{31}$$

which is to say that the braces surrounding the expression on the LHS are to indicate that summations are taken with respect to the repeated indices $i$ and $j$. The operation of composing two factors depends upon the cancellation of a superscript (column) index, or string of indices, in the leading factor with an equivalent subscript (row) index, or string of indices, in the following factor.

The matrix equation of (29) can be vectorised in a variety of ways. In order to represent the mapping from $X^c = (x_{ij}e_{ji})$ to $Y^c = (y_{kl}e_{lk})$, we may write

$$\begin{aligned}(y_{kl}e_{lk}) &= (\{a_{ki}x_{ij}b_{lj}\}e_{lk}) \\ &= (a_{ki}b_{lj}e_{lk}^{ji})(x_{ij}e_{ji}).\end{aligned} \tag{32}$$

Notice that the product $a_{ki}b_{lj}$ within $(a_{ki}b_{lj}e_{lk}^{ji})$ does not need to be surrounded by braces since it contains no repeated indices. Nevertheless, there would be no harm in writing $\{a_{ki}b_{lj}\}$.

The matrix $(a_{ki}b_{lj}e_{lk}^{ji})$ is decomposable. That is to say

$$\begin{aligned}(a_{ki}b_{lj}e_{lk}^{ji}) &= (b_{lj}e_l^j) \otimes (a_{ki}e_k^i) \\ &= B \otimes A;\end{aligned} \tag{33}$$

and, therefore, the vectorised form of equation (29) is

$$\begin{aligned}Y^c &= (AXB')^c \\ &= (B \otimes A)X^c.\end{aligned} \tag{34}$$

**Example.** The equation under (22), which relates to a two-way analysis of variance, can be vectorised to give

$$(y_{tj}e_{jt}) = \mu(e_{jt}) + (e_{jt}^{t})(\gamma_t e_t) + (e_{jt}^{j})(\delta_j e_j) + (\varepsilon_{tj}e_{jt}). \tag{35}$$

Using the notation of the Kronecker product, this can also be rendered as

$$\begin{aligned} Y^c &= \mu(\iota_M \otimes \iota_T) + (\iota_M \otimes I_T)\gamma + (I_M \otimes \iota_T)\delta + \mathcal{E}^c \\ &= X\beta + \mathcal{E}^c. \end{aligned} \tag{36}$$

The latter can also be obtained by applying the rule of (34) to equation (23). The various elements of (23) have been vectorised as follows:

$$\begin{aligned} (\mu\iota_T\iota_M')^c &= (\iota_T\mu\iota_M')^c = (\iota_M \otimes \iota_T)\mu, \\ (\gamma\iota_M')^c &= (I_t\gamma\iota_M')^c = (\iota_M \otimes I_T)\gamma, \\ (\iota_T\delta')^c &= (\iota_T\delta' I_M)^c = (I_M \otimes \iota_T)\delta'^c, \quad \delta'^c = \delta. \end{aligned} \tag{37}$$

In comparing (35) and (36), we see, for example, that $(e_{jt}^{t}) = (e_j) \otimes (e_t^t) = \iota_M \otimes I_T$. We recognise that $(e_t^t)$ is the sum over the index $t$ of the matrices of order $T$ which have a unit in the $t$th diagonal position and zeros elsewhere; and this sum amounts, of course, to the identity matrix of order $T$.

The vectorised form of equation (25) is

$$\begin{bmatrix} y_{11} \\ y_{21} \\ y_{31} \\ y_{12} \\ y_{22} \\ y_{32} \\ y_{13} \\ y_{23} \\ y_{33} \end{bmatrix} = \left[\begin{array}{c|ccc|ccc} 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ \hline 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 \\ \hline 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right] \left[\begin{array}{c} \mu \\ \hline \gamma_1 \\ \gamma_2 \\ \gamma_3 \\ \hline \delta_1 \\ \delta_2 \\ \delta_3 \end{array}\right] + \begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{21} \\ \varepsilon_{31} \\ \varepsilon_{12} \\ \varepsilon_{22} \\ \varepsilon_{32} \\ \varepsilon_{13} \\ \varepsilon_{23} \\ \varepsilon_{33} \end{bmatrix}. \tag{38}$$

## 6. Rules for Decomposable Tensor Products

The following rules govern the decomposable tensors product of matrices, which are commonly described as Kronecker products:

$$\begin{aligned} &\text{(i)} \quad (A \otimes B)(C \otimes D) = AC \otimes BD, \\ &\text{(ii)} \quad (A \otimes B)' = A' \otimes B', \\ &\text{(iii)} \quad A \otimes (B + C) = (A \otimes B) + (A \otimes C), \\ &\text{(iv)} \quad \lambda(A \otimes B) = \lambda A \otimes B = A \otimes \lambda B, \\ &\text{(v)} \quad (A \otimes B)^{-1} = (A^{-1} \otimes B^{-1}). \end{aligned} \tag{39}$$

The Kronecker product is non-commutative, which is to say that $A \otimes B \neq B \otimes A$. However, observe that

$$A \otimes B = (A \otimes I)(I \otimes B) = (I \otimes B)(A \otimes I). \tag{40}$$