Stretching and Condensing Tensors.
Application to tensor and matrix equations

Juan Ramón Ruiz-Tolosa

Department of Applied Mathematics and Computational Sciences,
University of Cantabria,
Santander
Spain

Enrique Castillo

Department of Applied Mathematics and Computational Sciences,
University of Cantabria,
Santander
Spain

Abstract

Since tensors are vectors, no matter their order and valencies, they can be represented in matrix form and in particular as column matrices of components with respect to their corresponding bases. This consideration allows using all the available tools for vectors and can be extremely advantageous to deal with some interesting problems in linear algebra, as solving tensor linear equations. Once, the tensors have been operated as vectors, they can be returned to their initial notation as tensors. This technique, that is specially useful for implementing computer programs, is illustrated by several examples of applications. In particular, some interesting tensor and matrix equations are solved. Several numerical examples are used to illustrate the proposed methodology.

1   Introduction and Motivation

The paper describes a technique that permits to treat tensors as vectors, that is, to extend (stretch) the tensor components, as column matrices that repre-

sent their vector components, solve as vectors the desired problems, and recover again the initial format of the tensor (condensation) (see (21)). This technique is described and used to solve a wide set of tensor and matrix equations.

The paper is structured as follows. In Section 2 we introduce some necessary background for the reader to understand the rest of the paper. In particular we deal with basis ordering and matrix representation of tensors. Since the technique used in this paper is based on representing tensors as vectors to take full advantage of vector theory and tools, and we stretch and condense tensors to move from vectors to tensors and vice versa, we present the corresponding extension and condensation operators, which permit moving from one of these representation to the other, and vice versa. As an initial application of these operators, Section 3 deals with some important tensor equations that are solved by the stretching-condensing technique. In Section 4 some interesting applications to matrix equations, including the generalized Sylvester equation, the tensor product commuters equation, the Schur tensor product equation, the square equation and some applications to probability and Statistics, are presented. Finally, Section 5 gives some conclusions.

2 Background

In this section we include some important material that is required to understand the following sections.

2.1 Basis ordering

Consider tensors of the form

$$t \in V_1^{n_1} \otimes V_2^{n_2} \otimes \cdots \otimes V_k^{n_k}.$$

To work with these tensors and avoid important errors, it is very important that the different basic vectors

$$\{\vec{e}_{i_1} \otimes \vec{e}_{i_2} \otimes \cdots \otimes \vec{e}_{i_k};\ i_1 = 1, 2, \ldots, n_1;\ i_2 = 1, 2, \ldots, n_2;\ i_k = 1, 2, \ldots, n_k\}$$

be "ordered" by means of an agreed upon criteria. In this paper we use the following criterion: The vector $\vec{e}_{i_1} \otimes \vec{e}_{i_2} \otimes \cdots \otimes \vec{e}_{i_k}$ in this ordering will be ahead of vector $\vec{e}_{j_1} \otimes \vec{e}_{j_2} \otimes \cdots \otimes \vec{e}_{j_k}$ iff at least one of the following condition holds:

1. $i_1 < j_1$

2. Being $i_1 = j_1$ it is $i_2 < j_2$.

3. Being $i_s = j_s$ for $s = 1, 2, \ldots, t - 1$, it is $i_t < j_t$, for any t.

Then, if the indices of the ordered sequence $i_1 i_2 \cdots i_r$ of basic vectors were considered as digits of a natural number, they would appear as ordered with respect to the standard ordering.

This order will play an important role in this paper.

2.2 Matrix representation of a tensor

The components of a tensor can be organized in matrix form. However, there are many ways of representing a tensor in matrix form. In this paper we consider only representations such that the order of its components are kept by rows. For example, the third order tensor $\mathbf{A} \in V^2 \otimes V^3 \otimes V^4$ can be written as

$$
M_{(\{1,2\},\{3\})}(\mathbf{A}) = \left(
\begin{array}{cccc}
a^{111}_{\circ\circ\circ} & a^{112}_{\circ\circ\circ} & a^{113}_{\circ\circ\circ} & a^{114}_{\circ\circ\circ} \\
a^{121}_{\circ\circ\circ} & a^{122}_{\circ\circ\circ} & a^{123}_{\circ\circ\circ} & a^{124}_{\circ\circ\circ} \\
a^{131}_{\circ\circ\circ} & a^{132}_{\circ\circ\circ} & a^{133}_{\circ\circ\circ} & a^{134}_{\circ\circ\circ} \\
\hline
a^{211}_{\circ\circ\circ} & a^{212}_{\circ\circ\circ} & a^{213}_{\circ\circ\circ} & a^{214}_{\circ\circ\circ} \\
a^{221}_{\circ\circ\circ} & a^{222}_{\circ\circ\circ} & a^{223}_{\circ\circ\circ} & a^{224}_{\circ\circ\circ} \\
a^{231}_{\circ\circ\circ} & a^{232}_{\circ\circ\circ} & a^{233}_{\circ\circ\circ} & a^{234}_{\circ\circ\circ}
\end{array}
\right),
$$

where the indices in the sets $\{1,2\}$ and $\{3\}$ appear ordered in each column and row, respectively; it can also be written as

$$
M_{(\{1\},\{2,3\})}(\mathbf{A}) = \left(
\begin{array}{cccccccccccc}
a^{111}_{\circ\circ\circ} a^{112}_{\circ\circ\circ} a^{113}_{\circ\circ\circ} a^{114}_{\circ\circ\circ} a^{121}_{\circ\circ\circ} a^{122}_{\circ\circ\circ} a^{123}_{\circ\circ\circ} a^{124}_{\circ\circ\circ} a^{131}_{\circ\circ\circ} a^{132}_{\circ\circ\circ} a^{133}_{\circ\circ\circ} a^{134}_{\circ\circ\circ} \\
\hline
a^{211}_{\circ\circ\circ} a^{212}_{\circ\circ\circ} a^{213}_{\circ\circ\circ} a^{214}_{\circ\circ\circ} a^{221}_{\circ\circ\circ} a^{222}_{\circ\circ\circ} a^{223}_{\circ\circ\circ} a^{224}_{\circ\circ\circ} a^{231}_{\circ\circ\circ} a^{232}_{\circ\circ\circ} a^{233}_{\circ\circ\circ} a^{234}_{\circ\circ\circ}
\end{array}
\right),
$$

where the indices in the sets $\{1\}$ and $\{2,3\}$ appear ordered in each column and row, respectively; finally, it can be written with all its components in a single row $(M_{(\emptyset,\{1,2,3\})}(\mathbf{A}))$ or with all its components in a single column $(M_{(\{1,2,3\},\emptyset)}(\mathbf{A}))$.

Note that the subindices of M refer to the indices that are represented as rows or columns, respectively.

The convenience of each representation depends on the application we are interested in. In the following sections we shall use different representations according to different criteria (saving printing space, convenience, simplicity of calculations, etc.).

2.2.1 Stretching and condensing matrices

Consider the linear space $K^{m \times n}(+, \circ)$ of matrices $\{\mathbf{A}_{m,n}\}$. We refer vectors in this space to its "canonical basis" $\mathcal{B} = \{\mathbf{E}_{ij}\}$, which consists of the simplest

matrices of $K^{m \times n}$, that is,

$$\mathcal{B} = \{(\mathbf{E}_{ij})_{mn}\} \equiv \{\mathbf{E}_{11}, \mathbf{E}_{12}, \ldots, \mathbf{E}_{1n}, \mathbf{E}_{21}, \mathbf{E}_{22}, \ldots, \mathbf{E}_{mn}\}$$

where $(\mathbf{E}_{ij})_{mn} = [k_{\alpha\beta}]$ and

$$k_{\alpha\beta} = \begin{cases} 1 & \text{if } \alpha = i \text{ and } \beta = j \\ 0 & \text{otherwise.} \end{cases}$$

When choosing the basis $\mathcal{B}$, matrix $\mathbf{A}_{m,n}$ is expressed as a linear manifold spanned by $\mathcal{B}$. If a "matrix form" is adopted to notify the manifold, we get:

$$\mathbf{A}_{m,n} = (\mathbf{E}_{11}|\mathbf{E}_{12}|\ldots|\mathbf{E}_{ij}|\ldots|\mathbf{E}_{mn}) \begin{pmatrix} a_{11} \\ \hline a_{12} \\ \hline \vdots \\ \hline a_{1n} \\ \hline a_{21} \\ \hline a_{22} \\ \hline \vdots \\ \hline a_{2n} \\ \hline \vdots \\ \hline a_{m1} \\ \hline \vdots \\ \hline a_{mn} \end{pmatrix} \equiv ||\mathbf{E}_{ij}||\mathbf{X}_{mn \times 1},$$

where all the elements of matrix $\mathbf{A}_{mn}$ appear "stacked" in a column matrix $\mathbf{X}$ according to the ordering criteria imposed by the given basis "$\mathcal{B}$", and the matrix product must be understood in symbolic form and as products of blocks. When one desires a given matrix $\mathbf{A}_{m,n}$ in this form, the English speaking texts write: "obtained by stacking the elements of the rows of $\mathbf{A}_{m,n}$ in sequence."

However, we want to note that there is no "need" of expressing this result by "words" of a given language, but using the universal language of Linear Algebra. Given a matrix $\mathbf{T}_{m,n}$ of elements of a field K, and calling $\sigma = m \cdot n$ (not a prime number) the dimension of the linear space $\mathbf{T}_{m,n}(K^\sigma)$ of matrices, we define by "extension" the mapping: $E : K^{m \times n} \to K^\sigma$ such that $\forall \mathbf{T}_{m,n} \in K^{m \times n} : E(\mathbf{T}_{m,n}) = \mathbf{T}_{\sigma,1}$ with $\mathbf{T}_{\sigma,1} \in K^\sigma$, that is, the "staked" optic is replaced by "staked and extend the given matrix and write it in column form".

In other words, the extension of a matrix $\mathbf{T}$ is a transformation that converts its tensor representation $M_{(\{1\},\{2\})}(\mathbf{T})$ to the column representation $M_{(\{1,2\},\emptyset)}(\mathbf{T})$.

If $\mathcal{B}'_m = \{E_i\}_{1 \leq i \leq m}$ is the canonical basis of matrices in $\mathbb{R}^{m \times 1}$, the "stacked" column matrix $\mathbf{T}_{\sigma,1}$, associated with matrix $\mathbf{T}_{m,n}$ can be obtained by:

$$(\mathbf{T}_{\sigma,1})^t = \begin{pmatrix} \mathbf{E}_1^t & |\mathbf{E}_2^t & | & \cdots & | & \mathbf{E}_m^t \end{pmatrix}_{1,m^2} \bullet (\mathbf{I}_m \otimes \mathbf{T}_{m,n}) = \mathbf{J}_{1,m^2} \bullet (\mathbf{I}_m \otimes \mathbf{T}_{m,n}), \quad (1)$$

where we have denoted $\mathbf{J}_{1,m^2}$ to $(\ \mathbf{E}_1^t\quad |\mathbf{E}_2^t\quad |\quad \cdots\quad |\quad \mathbf{E}_m^t\ )_{1,m^2}$, and $\bullet$ is the standard product of matrices. Then, we have:

<u>Extension formula</u>:

$$\mathbf{T}_{\sigma,1} = (\mathbf{I}_m \otimes \mathbf{T}_{m,n}^t) \bullet \mathbf{J}_{1,m^2}^t. \tag{2}$$

Example 2.1 (Extension of a matrix) *Consider the matrix*

$$\mathbf{A}_{3,4} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{pmatrix};$$

where $m = 3$ and $n = 4$, then

$$\mathbf{I}_3 \otimes \mathbf{A}_{3,4} = \begin{pmatrix} \mathbf{A}_{3,4} & | & \mathbf{\Omega}_{3,4} & | & \mathbf{\Omega}_{3,4} \\ \hline \mathbf{\Omega}_{3,4} & | & \mathbf{A}_{3,4} & | & \mathbf{\Omega}_{3,4} \\ \hline \mathbf{\Omega}_{3,4} & | & \mathbf{\Omega}_{3,4} & | & \mathbf{A}_{3,4} \end{pmatrix}_{9,12}.$$

where $\mathbf{\Omega}$ is the null matrix. Applying (1) one obtains

$$\mathbf{X}^t = (100|010|001)_{1,9} \bullet (\mathbf{I}_3 \otimes \mathbf{A}_{3,4}) \equiv (a_{11}a_{12}\ldots a_{14}a_{21}\ldots a_{33}a_{34}).\blacksquare$$

Similarly, given a "staked" matrix, $\mathbf{T}_{\sigma,1}$ we can be interested in its "condensation", that is, recover its original format $\mathbf{T}_{m,n}$ as a matrix.

Since we know that $\sigma = m \cdot n$, we define as "condensation", the mapping:

$$\mathcal{C} : K^\sigma \to K^{m \times n}$$

such that $\forall \mathbf{T}_{\sigma,1} \in K^\sigma, \mathcal{C}(\mathbf{T}_{\sigma,1}) = \mathbf{T}_{m,n}$ with $\mathbf{T}_{m,n} \in K^{m \times n}$.

<u>Condensation formula</u>:

$$\begin{aligned}
\mathbf{T}_{m,n} &= \left(\mathbf{I}_m \otimes \mathbf{T}_{\sigma,1}^t\right)_{(m,m^2n)} \bullet \left(\left(\begin{array}{c} \mathbf{E}_1 \\ \hline \mathbf{E}_2 \\ \hline \vdots \\ \hline \mathbf{E}_m \end{array}\right) \otimes \mathbf{I}_n\right)_{(m^2n,n)} \\
&= \left(\mathbf{I}_m \otimes \mathbf{T}_{\sigma,1}^t\right) \bullet \left(\mathbf{J}_{1,m^2}^t \otimes \mathbf{I}_n\right) \tag{3}
\end{aligned}$$

2.2.2   Stretching and condensing Tensors

Let $V_1^{n_1} \otimes V_2^{n_2} \otimes \cdots \otimes V_k^{n_k}(K)$ be the linear space "tensor product" of the linear spaces $V_1^{n_1}, V_2^{n_2} \ldots V_k^{n_k}$, which vectors $\vec{t}$ are called tensors, and are referred to a basis β of the given space, that is denoted

$$\beta = \{\vec{e}_{1i_1} \otimes \vec{e}_2^{*i_2} \otimes \cdots \otimes \vec{e}_{k-1}^{*i_{k-1}} \otimes \vec{e}_{ki_k}\}, \tag{4}$$

with the ranges: $1 \leq i_1 \leq n_1; 1 \leq i_2 \leq n_2; \cdots, 1 \leq i_k \leq n_k$ and dimension $\prod_{h=1}^{k} n_h$. Any tensor in this space can be written as:

$$\vec{t} = t^{i_1 \circ i_3}_{\circ i_2 \circ} \cdots {}^{\circ \ i_k}_{i_{k-1} \circ} \vec{e}_{1i_1} \otimes \vec{e}_2^{*i_2} \otimes \cdots \otimes \vec{e}_{k-1}^{*i_{k-1}} \otimes \vec{e}_{ki_k}, \tag{5}$$

where all scalars $t^{i_1 \circ i_3}_{\circ i_2 \circ} \cdots {}^{\circ \ i_k}_{i_{k-1} \circ} \in K$.

The stretching-condensation formulas can be extended to tensors. It is interesting to know the existing relation between the position n that occupies a component in the stretched vector representation of a tensor (ordered set of all of its components), and the indices $i_1, i_2, \cdots, i_k$ of the given component in the initial representation of the tensor.

This leads to two problems:

Direct problem :   Assuming known the indices $i_1, i_2, \cdots, i_k$ of a given component and the ranges $n_1, n_2, \cdots, n_k$ of each index, determine the position n (position that occupies such component in the ordered set), that is given by

$$\begin{aligned}
n &= i_k + n_k((i_{k-1} - 1) + n_{k-1}((i_{k-2} - 1) + \cdots (\cdots + n_2((i_1 - 1))))) \\
&= i_k + \sum_{j=1}^{k-1} (i_j - 1) \left(\prod_{s=j+1}^{k} n_s \right),
\end{aligned}$$

where the first equality shows the most convenient formula for computer implementation and the second equality gives the shortest expression.

Reverse problem :   The reverse problem consists of assuming known the position n (position of a given component) and the ranges $r_1, r_2, \ldots, r_k$ of all indices, determine the indices $i_1, i_2, \cdots, i_k$ of the given component.

The indices $i_1, i_2, \cdots, i_k$ can be obtained by means of the following algorithm:

1. $q = n - 1$

2. repeat with j from k to 2, by -1: $i_j = \mod[q, r_j] + 1;\ \ q = \text{quotient}[q, r_j]$

3. $i_1 = q + 1,$

where quotient$[q, r]$ and mod$[q, r]$ are the quotient and the rest, respectively, when dividing q by r.

These two procedures define the stretching-condensation transformations for general tensors.

This extended vector can be condensated as a two dimensional matrix by rows, independently of the order of the tensor, to have a matrix representation of the tensor. The following example shows how this matrix representation has an stretching transformation that is independent on the tensor orders and ranges. However, to recover the tensor, we need to know its order and ranges.

Example 2.2 (stretching-condensing tensors) *Consider the fourth order tensor $a^{ijk\ell}_{\circ\circ\circ\circ} \in V^2 \otimes V^3 \otimes V^2 \otimes V^4$ that can be represented in matrix form as:*

$$M_{(\{1,2\},\{3,4\})}(\mathbf{A}) = \left(\begin{array}{cccc|cccc}
a^{1111}_{\circ\circ\circ\circ} & a^{1112}_{\circ\circ\circ\circ} & a^{1113}_{\circ\circ\circ\circ} & a^{1114}_{\circ\circ\circ\circ} & a^{1121}_{\circ\circ\circ\circ} & a^{1122}_{\circ\circ\circ\circ} & a^{1123}_{\circ\circ\circ\circ} & a^{1124}_{\circ\circ\circ\circ} \\
a^{1211}_{\circ\circ\circ\circ} & a^{1212}_{\circ\circ\circ\circ} & a^{1213}_{\circ\circ\circ\circ} & a^{1214}_{\circ\circ\circ\circ} & a^{1221}_{\circ\circ\circ\circ} & a^{1222}_{\circ\circ\circ\circ} & a^{1223}_{\circ\circ\circ\circ} & a^{1224}_{\circ\circ\circ\circ} \\
a^{1311}_{\circ\circ\circ\circ} & a^{1312}_{\circ\circ\circ\circ} & a^{1313}_{\circ\circ\circ\circ} & a^{1314}_{\circ\circ\circ\circ} & a^{1321}_{\circ\circ\circ\circ} & a^{1322}_{\circ\circ\circ\circ} & a^{1323}_{\circ\circ\circ\circ} & a^{1324}_{\circ\circ\circ\circ} \\
\hline
a^{2111}_{\circ\circ\circ\circ} & a^{2112}_{\circ\circ\circ\circ} & a^{2113}_{\circ\circ\circ\circ} & a^{2114}_{\circ\circ\circ\circ} & a^{2121}_{\circ\circ\circ\circ} & a^{2122}_{\circ\circ\circ\circ} & a^{2123}_{\circ\circ\circ\circ} & a^{2124}_{\circ\circ\circ\circ} \\
a^{2211}_{\circ\circ\circ\circ} & a^{2212}_{\circ\circ\circ\circ} & a^{2213}_{\circ\circ\circ\circ} & a^{2214}_{\circ\circ\circ\circ} & a^{2221}_{\circ\circ\circ\circ} & a^{2222}_{\circ\circ\circ\circ} & a^{2223}_{\circ\circ\circ\circ} & a^{2224}_{\circ\circ\circ\circ} \\
a^{2311}_{\circ\circ\circ\circ} & a^{2312}_{\circ\circ\circ\circ} & a^{2313}_{\circ\circ\circ\circ} & a^{2314}_{\circ\circ\circ\circ} & a^{2321}_{\circ\circ\circ\circ} & a^{2322}_{\circ\circ\circ\circ} & a^{2323}_{\circ\circ\circ\circ} & a^{2324}_{\circ\circ\circ\circ}
\end{array}\right),$$

the third order tensor $a^{ijk}_{\circ\circ\circ} \in V^2 \otimes V^3 \otimes V^8$ that can be represented in matrix form as:

$$M_{(\{1,2\},\{3\})}(\mathbf{A}) = \left(\begin{array}{cccccccc}
a^{111}_{\circ\circ\circ} & a^{112}_{\circ\circ\circ} & a^{113}_{\circ\circ\circ} & a^{114}_{\circ\circ\circ} & a^{115}_{\circ\circ\circ} & a^{116}_{\circ\circ\circ} & a^{117}_{\circ\circ\circ} & a^{118}_{\circ\circ\circ} \\
a^{121}_{\circ\circ\circ} & a^{122}_{\circ\circ\circ} & a^{123}_{\circ\circ\circ} & a^{124}_{\circ\circ\circ} & a^{125}_{\circ\circ\circ} & a^{126}_{\circ\circ\circ} & a^{127}_{\circ\circ\circ} & a^{128}_{\circ\circ\circ} \\
a^{131}_{\circ\circ\circ} & a^{132}_{\circ\circ\circ} & a^{133}_{\circ\circ\circ} & a^{134}_{\circ\circ\circ} & a^{135}_{\circ\circ\circ} & a^{136}_{\circ\circ\circ} & a^{137}_{\circ\circ\circ} & a^{138}_{\circ\circ\circ} \\
\hline
a^{211}_{\circ\circ\circ} & a^{212}_{\circ\circ\circ} & a^{213}_{\circ\circ\circ} & a^{214}_{\circ\circ\circ} & a^{215}_{\circ\circ\circ} & a^{216}_{\circ\circ\circ} & a^{217}_{\circ\circ\circ} & a^{218}_{\circ\circ\circ} \\
a^{221}_{\circ\circ\circ} & a^{222}_{\circ\circ\circ} & a^{223}_{\circ\circ\circ} & a^{224}_{\circ\circ\circ} & a^{225}_{\circ\circ\circ} & a^{226}_{\circ\circ\circ} & a^{227}_{\circ\circ\circ} & a^{228}_{\circ\circ\circ} \\
a^{231}_{\circ\circ\circ} & a^{232}_{\circ\circ\circ} & a^{233}_{\circ\circ\circ} & a^{234}_{\circ\circ\circ} & a^{235}_{\circ\circ\circ} & a^{236}_{\circ\circ\circ} & a^{237}_{\circ\circ\circ} & a^{238}_{\circ\circ\circ}
\end{array}\right),$$

and the second order tensor $a^{ij}_{\circ\circ} \in V^6 \otimes V^8$ that can be represented in matrix

form as:

$$M_{(\{1\},\{2\})}(\mathbf{A}) = \begin{pmatrix} a^{11}_{\circ\circ} & a^{12}_{\circ\circ} & a^{13}_{\circ\circ} & a^{14}_{\circ\circ} & a^{15}_{\circ\circ} & a^{16}_{\circ\circ} & a^{17}_{\circ\circ} & a^{18}_{\circ\circ} \\ a^{21}_{\circ\circ} & a^{22}_{\circ\circ} & a^{23}_{\circ\circ} & a^{24}_{\circ\circ} & a^{25}_{\circ\circ} & a^{26}_{\circ\circ} & a^{27}_{\circ\circ} & a^{28}_{\circ\circ} \\ a^{31}_{\circ\circ} & a^{32}_{\circ\circ} & a^{33}_{\circ\circ} & a^{34}_{\circ\circ} & a^{35}_{\circ\circ} & a^{36}_{\circ\circ} & a^{37}_{\circ\circ} & a^{38}_{\circ\circ} \\ a^{41}_{\circ\circ} & a^{42}_{\circ\circ} & a^{43}_{\circ\circ} & a^{44}_{\circ\circ} & a^{45}_{\circ\circ} & a^{46}_{\circ\circ} & a^{47}_{\circ\circ} & a^{48}_{\circ\circ} \\ a^{51}_{\circ\circ} & a^{52}_{\circ\circ} & a^{53}_{\circ\circ} & a^{54}_{\circ\circ} & a^{55}_{\circ\circ} & a^{56}_{\circ\circ} & a^{57}_{\circ\circ} & a^{58}_{\circ\circ} \\ a^{61}_{\circ\circ} & a^{62}_{\circ\circ} & a^{63}_{\circ\circ} & a^{64}_{\circ\circ} & a^{65}_{\circ\circ} & a^{66}_{\circ\circ} & a^{67}_{\circ\circ} & a^{68}_{\circ\circ} \end{pmatrix},$$

then the stretching formula, that is, the transformation giving the column of the tensor components with respect to its ordered basis for the three tensors is the same. In other words, the order of the tensor and the dimensions of the original vector spaces are not needed. However, the condensation formula must keep record of the tensor order and the initial vector spaces dimensions. These are also required for other transformations, as basis changes, contractions and other homomorphisms, etc. ∎

Consider the following tensors: $\mathbf{A} \in V_1^{n_1} \otimes V_2^{n_2}$ and $\mathbf{B} \in V_{*3}^{n_3} \otimes V_{*4}^{n_4}$, then we have

$$M_{(\{1\},\{2\})}\mathbf{A} \otimes M_{(\{3\},\{4\})}\mathbf{B} = M_{(\{1,3\},\{2,4\})}\mathbf{A} \otimes \mathbf{B}. \tag{6}$$

3   Tensor equations

In this section we deal with tensor equations, i.e., equations whose unknowns are tensors. They will be the basic material for the following section on matrix equations.

Consider the following tensors

$$\mathbf{A} \in V_1^{n_1} \otimes V_2^{n_2} \otimes \cdots \otimes V_r^{n_r}$$

$$\mathbf{X} \in V_{*r+1}^{n_{r+1}} \otimes V_{*r+2}^{n_{r+2}} \otimes \cdots \otimes V_{*r+p}^{n_{r+p}}$$

$$\mathbf{B} \in V_{r+p+1}^{n_{r+p+1}} \otimes V_{r+p+2}^{n_{r+p+2}} \otimes \cdots \otimes V_{r+p+q}^{n_{r+p+q}}$$

where the asterisks refer to the dual spaces, it is assumed that $p < r + q$, and for the sake of simplicity, though this condition can be relaxed, the tensors $\mathbf{A}$ and $\mathbf{B}$ are assumed to be totally contravariant, and the tensor $\mathbf{X}$ totally covariant.

Consider the following tensor equation

$$\mathbf{C} = \mathcal{C}\left(S_1, S_2\right)\left(\mathbf{A} \otimes \mathbf{X} \otimes \mathbf{B}\right), \tag{7}$$

where $\mathbf{A}$ and $\mathbf{B}$ are given tensors, $\mathbf{C}$ is a given totally contravariant tensor of order $r-p+q$, $\mathbf{X}$ is the unknown tensor, and $\mathcal{C}\,(S_1, S_2)$ refers to a contraction of the indices in the two ordered sets of cardinality p, $S_2 \equiv \{r+1, r+2, \ldots, r+p\}$ and $S_1 \subset S$, where $S \equiv \{1, 2, \ldots, r\} \cup \{r+p+1, r+p+2, \ldots, r+p+q\}$. Note that each of the p indices of tensor $\mathbf{X}$ is contracted with one of the indices of either $\mathbf{A}$ or $\mathbf{B}$; the first index $r+1$ of $\mathbf{X}$ is contracted with the first index in S_1, the second index $r+2$ of $\mathbf{X}$ is contracted with the second index in S_1, and so on.

Since Equation (7) is equivalent to the following standard system of linear equations

$$M_{(S-S_1, S_1)}\,(\mathbf{A} \otimes \mathbf{B})\,E(\mathbf{X}) = E(\mathbf{C}), \tag{8}$$

where $E(\mathbf{X}) = M_{(\emptyset, S_2)}(\mathbf{X})$ and $E(\mathbf{C}) = M_{(\emptyset, S-S_1)}(\mathbf{C})$ are the matrices of extended or stretched forms of tensors $\mathbf{X}$ and $\mathbf{C}$, respectively, solving Equation (8), the linear tensor equation (7) is solved.

In summary, the tensor equation (7) can be solved using the following algorithm:

Algorithm for solving Equation (7).

Input: Tensors $\mathbf{A}, \mathbf{B}$ and $\mathbf{C}$ and sets of contracted indices S_1 and S_2.

Output: The tensor $\mathbf{X}$ solution to Equation (7).

Step 1. Use different index numbers I_A, $I_X \equiv S_2$ and I_B, for the indices of tensors $\mathbf{A}, \mathbf{X}$ and $\mathbf{B}$, respectively.

Step 2. Determine the set $S = I_A \cup I_B$.

Step 3. Calculate the tensor $\mathbf{A} \otimes \mathbf{B}$ and write it in the matrix form $M_{(S-S_1, S_1)}\mathbf{A} \otimes \mathbf{B}$.

Step 4. Write the tensor $\mathbf{C}$ in matrix form $E(\mathbf{C}) = M_{(\emptyset, S-S_1)}(\mathbf{C})$.

Step 5. Solve the system of linear equations $M_{(S-S_1, S_1)}\,(\mathbf{A} \otimes \mathbf{B})\,E(\mathbf{X}) = E(\mathbf{C})$ to obtain the column matrix form $M_{(\emptyset, S_2)}(\mathbf{X})$ of the tensor $\mathbf{X}$.

The tensor equation (7) can be generalized to

$$\sum_{i=1}^{I} \mathcal{C}\,(S_{1i}, S_{2i})\,(\mathbf{A}_i \otimes \mathbf{X} \otimes \mathbf{B}_i) = \mathbf{C}, \tag{9}$$

where the S_{1i}, S_{2i} sets and the tensors $\mathbf{A}_i$, $\mathbf{B}_i$ and $\mathbf{C}$ are assumed to be such that Equation (9) is well defined.

Then, Equation (9) is equivalent to

$$\left(\sum_{i=1}^{I} M_{(S_i - S_{1i}, S_{1i})} \left(\mathbf{A}_i \otimes \mathbf{B}_i \right) \right) E(\mathbf{X}) = E(\mathbf{C}), \tag{10}$$

a standard system of linear equations that solves (9).

The tensor equation (7) can also be generalized to

$$\sum_{i=1}^{I} \mathcal{C} \left(S_{1i}, S_{2i} \right) \left(\mathbf{A}_i \otimes \mathbf{X}_i \otimes \mathbf{B}_i \right) = \mathbf{C}, \tag{11}$$

where the S_{1i}, S_{2i} sets and the tensors $\mathbf{A}_i$, $\mathbf{B}_i$ and $\mathbf{C}$ are assumed to be such that Equation (11) is well defined, and we have I unknown tensors.

Then, Equation (11) is equivalent to

$$\left(M_{(S_1 - S_{11}, S_{11})} \left(\mathbf{A}_1 \otimes \mathbf{B}_1 \right) \middle| \cdots \middle| M_{(S_I - S_{1I}, S_{1I})} \left(\mathbf{A}_I \otimes \mathbf{B}_I \right) \right) \begin{pmatrix} E(\mathbf{X}_1) \\ - - - \\ E(\mathbf{X}_2) \\ - - - \\ \vdots \\ - - - \\ E(\mathbf{X}_I) \end{pmatrix}$$

$$= E(\mathbf{C}), \tag{12}$$

a standard system of linear equations that solves (11).

The extension of this procedure to systems of equations of the form (11) is obvious. The same is true if one has inequalities instead of equalities.

Example 3.1 (Tensor equation) *Consider the following tensor equation*

$$c^{i_2\, i_7\, i_9\, i_{10}}_{\circ\ \circ\ \circ\ \circ} = a^{\lambda\, i_2\, \mu}_{\circ\ \circ\ \circ} x^{\circ\ \circ\ \circ}_{\theta\, \lambda\, \mu} b^{i_7\, \theta\, i_9\, i_{10}}_{\circ\ \circ\ \circ\ \circ}; \quad \forall i_2, i_7, i_9, i_{10} \in \{1, 2\}, \tag{13}$$

which can be written as

$$\mathbf{C} = \mathcal{C} \left(\{8, 1, 3\}, \{4, 5, 6\} \right) \left(\mathbf{A} \otimes \mathbf{X} \otimes \mathbf{B} \right), \tag{14}$$

where

$$\mathbf{A} \in V_1^2 \otimes V_2^2 \otimes V_3^2$$
$$\mathbf{B} \in V_7^2 \otimes V_8^2 \otimes V_9^2 \otimes V_{10}^2$$
$$\mathbf{C} \in V_2^2 \otimes V_7^2 \otimes V_9^2 \otimes V_{10}^2$$

are given totally contravariant tensors and

$$\mathbf{X} \in V_{*4}^2 \otimes V_{*5}^2 \otimes V_{*6}^2$$

is the unknown totally covariant tensor.

  Here $S_1 = \{8, 1, 3\}$, $S_2 = \{4, 5, 6\}$ *and* $S = \{1, 2, 3, 7, 8, 9, 10\}$. *Thus,* $S - S_1 \equiv \{2, 7, 9, 10\}$.

  Suppose that the tensors in matrix form are:

$$M_{(\{1,2\},\{3\})}(\mathbf{A}) = \begin{pmatrix} 1 & 2 \\ 0 & 0 \\ 1 & 0 \\ 2 & -1 \end{pmatrix}; \quad M_{(\{7,8\},\{9,10\})}(\mathbf{B}) = \begin{pmatrix} 1 & 1 & 1 & 0 \\ -1 & 1 & 0 & -1 \\ 1 & 0 & 0 & 2 \\ 1 & -1 & 0 & 1 \end{pmatrix};$$

$$M_{(\{2,7\},\{9,10\})}(\mathbf{C}) = \begin{pmatrix} 1 & 3 & 2 & -1 \\ 3 & -1 & 0 & 5 \\ 1 & -1 & 0 & 1 \\ -1 & 1 & 0 & -1 \end{pmatrix}.$$

  Then, the tensor $\mathbf{A} \otimes \mathbf{B}$ *in matrix form* $M_{(\{2,7,9,10\},\{8,1,3\})}(\mathbf{A} \otimes \mathbf{B})$ *is*

$$M_{(\{2,7,9,10\},\{8,1,3\})}(\mathbf{A} \otimes \mathbf{B}) = \begin{pmatrix} 1 & 2 & 1 & 0 & -1 & -2 & -1 & 0 \\ 1 & 2 & 1 & 0 & 1 & 2 & 1 & 0 \\ 1 & 2 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & -2 & -1 & 0 \\ 1 & 2 & 1 & 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & -1 & -2 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 4 & 2 & 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 2 & -1 & 0 & 0 & -2 & 1 \\ 0 & 0 & 2 & -1 & 0 & 0 & 2 & -1 \\ 0 & 0 & 2 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -2 & 1 \\ 0 & 0 & 2 & -1 & 0 & 0 & 2 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & -2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 4 & -2 & 0 & 0 & 2 & -1 \end{pmatrix},$$

so that one can write Equation (13) in its equivalent form (see (8)):

$$M_{(\{2,7,9,10\},\{8,1,3\})}(\mathbf{A} \otimes \mathbf{B}) \begin{pmatrix} x^{\circ\circ\circ}_{111} \\ x^{\circ\circ\circ}_{112} \\ x^{\circ\circ\circ}_{121} \\ x^{\circ\circ\circ}_{122} \\ x^{\circ\circ\circ}_{211} \\ x^{\circ\circ\circ}_{212} \\ x^{\circ\circ\circ}_{221} \\ x^{\circ\circ\circ}_{222} \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 2 & 1 & 0 & -1 & -2 & -1 & 0 \\ 1 & 2 & 1 & 0 & 1 & 2 & 1 & 0 \\ 1 & 2 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & -2 & -1 & 0 \\ 1 & 2 & 1 & 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & -1 & -2 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 4 & 2 & 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 2 & -1 & 0 & 0 & -2 & 1 \\ 0 & 0 & 2 & -1 & 0 & 0 & 2 & -1 \\ 0 & 0 & 2 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -2 & 1 \\ 0 & 0 & 2 & -1 & 0 & 0 & 2 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & -2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 4 & -2 & 0 & 0 & 2 & -1 \end{pmatrix} \begin{pmatrix} x_{111}^{\circ\circ\circ} \\ x_{112}^{\circ\circ\circ} \\ x_{121}^{\circ\circ\circ} \\ x_{122}^{\circ\circ\circ} \\ x_{211}^{\circ\circ\circ} \\ x_{212}^{\circ\circ\circ} \\ x_{221}^{\circ\circ\circ} \\ x_{222}^{\circ\circ\circ} \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \\ 2 \\ -1 \\ 3 \\ -1 \\ 0 \\ 5 \\ 1 \\ -1 \\ 0 \\ 1 \\ -1 \\ 1 \\ 0 \\ -1 \end{pmatrix} ,$$

where the coefficient matrix $M_{(\{2,7,9,10\},\{8,1,3\})}(\mathbf{A} \otimes \mathbf{B})$ is the convenient matrix representation of the tensor $\mathbf{A} \otimes \mathbf{B}$ and the independent term is $E(\mathbf{C})$. Note that this is a simple linear system of equations that allows us solving the tensor equation (14) or (13).

Its solution is

$$M_{(\{1,2,3\},\emptyset)}(\mathbf{X}) = \begin{pmatrix} 1 - \rho_3 - 2\rho_4 \\ \rho_4 \\ 1 + \rho_3 \\ 2 + 2\rho_3 \\ 1 - \rho_1 - 2\rho_2 \\ \rho_2 \\ \rho_1 \\ 1 + 2\rho_1 \end{pmatrix} ; \quad \rho_1, \rho_2, \rho_3, \rho_4 \in \mathbb{R},$$

which is the column matrix representation of the solution tensor $\mathbf{X}$. ∎

Example 3.2 (Inverse tensors.) *Consider the three fourth order tensors*

$$\mathbf{A} \equiv a_{ijk\ell}^{\circ\circ\circ\circ} \in V_*^n \otimes V_*^n \otimes V_*^n \otimes V_*^n,$$

$$\mathbf{B} \equiv b_{\circ\circ\circ\circ}^{\alpha\beta\gamma\delta} \in V^n \otimes V^n \otimes V^n \otimes V^n$$

and

$$\mathbf{C} \equiv c_{ij\circ\circ}^{\circ\circ\gamma\delta} \in V_*^n \otimes V_*^n \otimes V^n \otimes V^n$$

and the tensor equation

$$\mathcal{C}\left(\{k,\ell\},\{\alpha,\beta\}\right)(\mathbf{A} \otimes \mathbf{B}) = \mathbf{C}, \tag{15}$$

where $\mathcal{C}\left(\{k,\ell\},\{\alpha,\beta\}\right)$ *refers to the corresponding contraction of indices, tensors* $\mathbf{A}$ *and* $\mathbf{C}$ *are considered data, and tensor* $\mathbf{B}$ *the unknown.*

Using the Einstein's notation of repeated indices, equation (15) can be written as

$$a_{ijk\ell}^{\circ\circ\circ\circ} b_{\circ\circ\circ\circ}^{k\ell\gamma\delta} = c_{ij\circ\circ}^{\circ\circ\gamma\delta}, \tag{16}$$

which rearranged in matrix form becomes

$$M_{(\{1,2\},\{3,4\})}(\mathbf{A}) M_{(\{1,2\},\{3,4\})}(\mathbf{B}) = M_{(\{1,2\},\{3,4\})}(\mathbf{C}), \tag{17}$$

showing that any tensor equation of the form (16) is equivalent to a matrix equation of the form (17).

Note that the tensor $\mathbf{A}$ can be interpreted as a linear map that transforms vectors in $V^n \otimes V^n \otimes V^n \otimes V^n$ to vectors in $V^n_* \otimes V^n_* \otimes V^n \otimes V^n$.

Next, our problem consist of finding the inverse map of $\mathbf{A}$, that is, finding another tensor

$$\mathbf{D} \equiv d_{\circ\circ\circ\circ}^{k\ell rs} \in V^n \otimes V^n \otimes V^n \otimes V^n,$$

such that

$$b_{\circ\circ\circ\circ}^{k\ell\gamma\delta} = d_{\circ\circ\circ\circ}^{k\ell rs} c_{rs\circ\circ}^{\circ\circ\gamma\delta}. \tag{18}$$

Replacing (18) into (16) one gets

$$a_{ijk\ell}^{\circ\circ\circ\circ} d_{\circ\circ\circ\circ}^{k\ell rs} c_{rs\circ\circ}^{\circ\circ\gamma\delta} = c_{ij\circ\circ}^{\circ\circ\gamma\delta}, \tag{19}$$

which holds if

$$a_{ijk\ell}^{\circ\circ\circ\circ} d_{\circ\circ\circ\circ}^{k\ell rs} = \delta_{i\circ}^{\circ r} \delta_{j\circ}^{\circ s}. \tag{20}$$

where $\delta_{i\circ}^{\circ r}$ and $\delta_{j\circ}^{\circ s}$ are the Kronecker's deltas.

If tensor A is represented in matrix form as $M_{(\{1,2\},\{3,4\})}(\mathbf{A})$, then Expression (20) shows that

$$M_{(\{1,2\},\{3,4\})}(\mathbf{A}) M_{(\{1,2\},\{3,4\})}(\mathbf{D}) = \mathbf{I},$$

i.e.,

$$M_{(\{1,2\},\{3,4\})}(\mathbf{D}) = M_{(\{1,2\},\{3,4\})}^{-1}(\mathbf{A}),$$

where $\mathbf{I}$ is the identity matrix. This implies that using the adequate ordering, pseudo-inverses of tensors can be obtained using matrix tools, and the inverse tensor $\mathbf{D}$ of tensor $\mathbf{A}$ can be obtained by inverting the matrix $M_{(\{1,2\},\{3,4\})}(\mathbf{A})$ associated with tensor $\mathbf{A}$.

Thus, a necessary and sufficient condition for the existence of the inverse tensor is that the matrix $M_{(\{1,2\},\{3,4\})}(\mathbf{A})$ be regular.

Tensor $\mathbf{A}$ *in matrix form* $M_{(\{1,2\},\{3,4\})}(\mathbf{A})$ *is*

$$
\begin{pmatrix}
a^{1111}_{\circ\circ\circ\circ} & a^{1112}_{\circ\circ\circ\circ} & a^{1113}_{\circ\circ\circ\circ} & \vert & a^{1121}_{\circ\circ\circ\circ} & a^{1122}_{\circ\circ\circ\circ} & a^{1123}_{\circ\circ\circ\circ} & \vert & a^{1131}_{\circ\circ\circ\circ} & a^{1132}_{\circ\circ\circ\circ} & a^{1133}_{\circ\circ\circ\circ} \\
a^{1211}_{\circ\circ\circ\circ} & a^{1212}_{\circ\circ\circ\circ} & a^{1213}_{\circ\circ\circ\circ} & \vert & a^{1221}_{\circ\circ\circ\circ} & a^{1222}_{\circ\circ\circ\circ} & a^{1223}_{\circ\circ\circ\circ} & \vert & a^{1231}_{\circ\circ\circ\circ} & a^{1232}_{\circ\circ\circ\circ} & a^{1233}_{\circ\circ\circ\circ} \\
a^{1311}_{\circ\circ\circ\circ} & a^{1312}_{\circ\circ\circ\circ} & a^{1313}_{\circ\circ\circ\circ} & \vert & a^{1321}_{\circ\circ\circ\circ} & a^{1322}_{\circ\circ\circ\circ} & a^{1323}_{\circ\circ\circ\circ} & \vert & a^{1331}_{\circ\circ\circ\circ} & a^{1332}_{\circ\circ\circ\circ} & a^{1333}_{\circ\circ\circ\circ} \\
\hline
a^{2111}_{\circ\circ\circ\circ} & a^{2112}_{\circ\circ\circ\circ} & a^{2113}_{\circ\circ\circ\circ} & \vert & a^{2121}_{\circ\circ\circ\circ} & a^{2122}_{\circ\circ\circ\circ} & a^{2123}_{\circ\circ\circ\circ} & \vert & a^{2131}_{\circ\circ\circ\circ} & a^{2132}_{\circ\circ\circ\circ} & a^{2133}_{\circ\circ\circ\circ} \\
a^{2211}_{\circ\circ\circ\circ} & a^{2212}_{\circ\circ\circ\circ} & a^{2213}_{\circ\circ\circ\circ} & \vert & a^{2221}_{\circ\circ\circ\circ} & a^{2222}_{\circ\circ\circ\circ} & a^{2223}_{\circ\circ\circ\circ} & \vert & a^{2231}_{\circ\circ\circ\circ} & a^{2232}_{\circ\circ\circ\circ} & a^{2233}_{\circ\circ\circ\circ} \\
a^{2311}_{\circ\circ\circ\circ} & a^{2312}_{\circ\circ\circ\circ} & a^{2313}_{\circ\circ\circ\circ} & \vert & a^{2321}_{\circ\circ\circ\circ} & a^{2322}_{\circ\circ\circ\circ} & a^{2323}_{\circ\circ\circ\circ} & \vert & a^{2331}_{\circ\circ\circ\circ} & a^{2332}_{\circ\circ\circ\circ} & a^{2333}_{\circ\circ\circ\circ} \\
\hline
a^{3111}_{\circ\circ\circ\circ} & a^{3112}_{\circ\circ\circ\circ} & a^{3113}_{\circ\circ\circ\circ} & \vert & a^{3121}_{\circ\circ\circ\circ} & a^{3122}_{\circ\circ\circ\circ} & a^{3123}_{\circ\circ\circ\circ} & \vert & a^{3131}_{\circ\circ\circ\circ} & a^{3132}_{\circ\circ\circ\circ} & a^{3133}_{\circ\circ\circ\circ} \\
a^{3211}_{\circ\circ\circ\circ} & a^{3212}_{\circ\circ\circ\circ} & a^{3213}_{\circ\circ\circ\circ} & \vert & a^{3221}_{\circ\circ\circ\circ} & a^{3222}_{\circ\circ\circ\circ} & a^{3223}_{\circ\circ\circ\circ} & \vert & a^{3231}_{\circ\circ\circ\circ} & a^{3232}_{\circ\circ\circ\circ} & a^{3233}_{\circ\circ\circ\circ} \\
a^{3311}_{\circ\circ\circ\circ} & a^{3312}_{\circ\circ\circ\circ} & a^{3313}_{\circ\circ\circ\circ} & \vert & a^{3321}_{\circ\circ\circ\circ} & a^{3322}_{\circ\circ\circ\circ} & a^{3323}_{\circ\circ\circ\circ} & \vert & a^{3331}_{\circ\circ\circ\circ} & a^{3332}_{\circ\circ\circ\circ} & a^{3333}_{\circ\circ\circ\circ}
\end{pmatrix},
$$

Assume that the components of tensors $\mathbf{A}$ *and* $\mathbf{C}$ *in matrix form are:*

$$
M_{(\{1,2\},\{3,4\})}(\mathbf{A}) =
\begin{pmatrix}
1 & 1 & -1 & 1 & 1 & 1 & 1 & -1 & 0 \\
-1 & 1 & 1 & -1 & 0 & -1 & 1 & -1 & -1 \\
-1 & 1 & 1 & 1 & 1 & 0 & 1 & -1 & 1 \\
0 & -1 & 0 & 0 & 0 & 1 & 1 & 0 & -1 \\
-1 & 1 & -1 & 0 & 1 & 0 & 0 & 0 & 1 \\
0 & 0 & -1 & 1 & 1 & 0 & 1 & 0 & 1 \\
0 & -1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\
-1 & 0 & 0 & 1 & -1 & -1 & 1 & 0 & 1 \\
-1 & 1 & 0 & 1 & 0 & 0 & -1 & 1 & -1
\end{pmatrix}
$$

and

$$
M_{(\{1,2\},\{3,4\})}(\mathbf{C}) =
\begin{pmatrix}
2 & -3 & -2 & -3 & 1 & 1 & 1 & 3 & 4 \\
0 & 4 & -1 & 1 & 0 & -4 & -1 & -4 & -5 \\
0 & -1 & 1 & 1 & 1 & -1 & 0 & -3 & 0 \\
-1 & -2 & 1 & -1 & 0 & -2 & 0 & 4 & -2 \\
2 & -1 & -1 & 2 & -3 & 0 & 3 & -2 & 2 \\
3 & -4 & 1 & 1 & -1 & 2 & 3 & 2 & 2 \\
-1 & -2 & 3 & 0 & 2 & 0 & -1 & 2 & -2 \\
0 & -2 & 2 & 1 & 1 & 1 & 0 & -1 & -1 \\
-4 & 2 & -2 & -2 & -1 & -2 & -2 & 0 & 1
\end{pmatrix}.
$$

Then, the tensor $\mathbf{D}$ *inverse of* $\mathbf{A}$ *in matrix form* $M_{(\{1,2\},\{3,4\})}(\mathbf{B})$ *is* $M^{-1}_{(\{1,2\},\{3,4\})}(\mathbf{A})$,

that is:

$$M_{(\{1,2\},\{3,4\})}(\mathbf{D}) = \begin{pmatrix} 20 & 3 & -28 & -16 & 48 & -39 & 54 & 1 & -5 \\ 33 & 5 & -46 & -26 & 80 & -65 & 89 & 2 & -8 \\ 20 & 3 & -28 & -16 & 49 & -40 & 55 & 1 & -5 \\ -19 & -3 & 27 & 15 & -47 & 38 & -52 & -1 & 5 \\ -8 & -1 & 11 & 6 & -19 & 16 & -21 & -1 & 2 \\ 16 & 2 & -22 & -12 & 39 & -32 & 43 & 1 & -4 \\ 32 & 5 & -45 & -25 & 78 & -63 & 87 & 2 & -8 \\ 53 & 8 & -75 & -42 & 130 & -105 & 145 & 3 & -13 \\ 15 & 2 & -21 & -12 & 37 & -30 & 41 & 1 & -4 \end{pmatrix}.$$

and finally, the tensor $\mathbf{B}$ *in matrix form* $M_{(\{1,2\},\{3,4\})}(\mathbf{B})$ *is*

$$M_{(\{1,2\},\{3,4\})}^{-1}(\mathbf{D})M_{(\{1,2\},\{3,4\})}(\mathbf{C}),$$

that is:

$$M_{(\{1,2\},\{3,4\})}(\mathbf{B}) = \begin{pmatrix} 1 & 0 & 0 & -1 & 1 & 1 & 0 & 1 & 1 \\ 0 & 1 & -1 & -1 & 0 & -1 & 0 & -1 & 1 \\ -1 & 1 & 1 & 0 & 1 & -1 & -1 & -1 & -1 \\ -1 & -1 & 0 & -1 & 1 & 1 & -1 & 1 & 1 \\ 1 & 0 & 0 & 1 & -1 & 0 & 1 & 0 & 0 \\ -1 & -1 & 0 & -1 & 0 & -1 & 0 & 1 & 1 \\ 1 & -1 & 1 & 0 & 0 & -1 & 1 & 1 & -1 \\ 0 & 0 & 1 & 0 & -1 & -1 & 1 & 1 & 0 \\ 1 & -1 & 1 & 1 & 0 & 1 & 1 & -1 & 1 \end{pmatrix}.$$

∎

4 Some important matrix equations

Once tensor equations have been discussed, matrix equations arise as particular cases of them.

There are a long list of references on Matrix equations, as for example (6), (11), (17), (2), (3; 4; 5), (20), (14), (19), (15), (1), (8), (24; 23), (27), (26), (9), (10), (7), (13), (18), (28), (12), (25), (22), (16), (21), etc.

In this section we solve some important matrix equations, i.e., some equations where the unknowns are matrices, using the results of the previous section.

4.1 The generalized Sylvester matrix equations

The following matrix equation (see (15)):

$$\sum_{i=1}^{k} (\mathbf{A}_i \mathbf{X} \mathbf{B}_i) = \mathbf{C}, \tag{21}$$

where the dimensions of matrices $\mathbf{A}_i, \mathbf{X}, \mathbf{B}_i$ and $\mathbf{C}$ are $m \times n, n \times p, p \times q$ and $m \times q$, respectively, can be called the generalized Sylvester equation, because Sylvester (1884) considered only the particular case $n = m$ and $p = q$.

Equation (21) can be written as

$$\sum_{i=1}^{k} a_i{}^{i_1\,\theta}_{\circ\,\circ} x^{\circ\,\circ}_{\theta\,\lambda} b_i{}^{\lambda\,i_6}_{\circ\,\circ} = c^{i_1\,i_6}_{\circ\,\circ}, \tag{22}$$

i.e.,

$$\sum_{i=1}^{k} \mathcal{C}(\{2,5\},\{3,4\})\,(\mathbf{A}_i \otimes \mathbf{X} \otimes \mathbf{B}_i) = \mathbf{C}, \tag{23}$$

which shows it is a particular case of (9). Then, according to (10), Equation (21) is equivalent to

$$\left(\sum_{i=1}^{k} M_{(\{1,6\},\{2,5\})}\,(\mathbf{A}_i \otimes \mathbf{B}_i)\right) E(\mathbf{X}) = E(\mathbf{C}), \tag{24}$$

where $E(\mathbf{X})$ and $E(\mathbf{C})$ are the stretched (column) matrices associated with matrices $\mathbf{X}$ and $\mathbf{C}$, respectively, which taking into account (6) leads to

$$\left(\sum_{i=1}^{k} \left(M_{(\{1\},\{2\})}\mathbf{A}_i \otimes M_{(\{6\},\{5\})}\mathbf{B}_i\right)\right) E(\mathbf{X}) = \left(\sum_{i=1}^{k} \left(\mathbf{A}_i \otimes \mathbf{B}_i^t\right)\right) E(\mathbf{X}) = E(\mathbf{C}), \tag{25}$$

the system of linear equations which solves the matrix equation (21).

Another generalization of the Sylvester equation is the following matrix equation

$$\sum_{i=1}^{k} (\mathbf{A}_i \mathbf{X}_i \mathbf{B}_i) = \mathbf{C}, \tag{26}$$

where the dimensions of matrices $\mathbf{A}_i, \mathbf{X}_i, \mathbf{B}_i$ and $\mathbf{C}$ are $m \times n, n \times p, p \times q$ and $m \times q$, respectively. Note that there are k unknown matrices.

Similarly, using (12), Equation (26) can be written in stretched form as

$$(\,\mathbf{A}_1 \otimes \mathbf{B}_1^t \quad | \cdots \quad | \quad \mathbf{A}_k \otimes \mathbf{B}_k^t\,) \begin{pmatrix} E(\mathbf{X}_1) \\ -- \\ E(\mathbf{X}_2) \\ -- \\ \vdots \\ -- \\ E(\mathbf{X}_k) \end{pmatrix} = E(\mathbf{C}), \tag{27}$$

the system of linear equations (27) which solves the matrix equation (26).

Example 4.1 (Generalized Sylvester equation (Case 1)) *Consider the following matrix equation*

$$\mathbf{A}_1\mathbf{X}\mathbf{B}_1 + \mathbf{A}_2\mathbf{X}\mathbf{B}_2 + \mathbf{A}_3\mathbf{X}\mathbf{B}_3 = \mathbf{C} \tag{28}$$

where

$$\mathbf{A}_1 = \begin{pmatrix} -2 & 2 & -2 \\ -1 & 0 & 1 \\ 1 & -1 & 1 \\ 0 & 2 & 0 \end{pmatrix}; \quad \mathbf{A}_2 = \begin{pmatrix} 1 & 0 & 2 \\ 1 & 0 & 1 \\ 2 & 1 & 0 \\ 1 & 2 & 0 \end{pmatrix}; \quad \mathbf{A}_3 = \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \\ 1 & 0 & 0 \end{pmatrix}; \tag{29}$$

$$\mathbf{B}_1 = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 0 & 2 \end{pmatrix}; \quad \mathbf{B}_2 = \begin{pmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \end{pmatrix}; \quad \mathbf{B}_3 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 0 \end{pmatrix}; \tag{30}$$

$$\mathbf{C} = \begin{pmatrix} 17 & 13 & 1 \\ 9 & 7 & -3 \\ 2 & -5 & 1 \\ 22 & 8 & 8 \end{pmatrix}. \tag{31}$$

Then, Equation (25) becomes

$$\left(\sum_{i=1}^{k} \mathbf{A}_i \otimes \mathbf{B}_i^t \right) E(\mathbf{X}) = \begin{pmatrix} -2 & 0 & 5 & 3 & 0 & 2 \\ -2 & 2 & 3 & 2 & -2 & 4 \\ 0 & -4 & 3 & 4 & 2 & -4 \\ -1 & 0 & 1 & 1 & 3 & 2 \\ -2 & 0 & 1 & 2 & 0 & 0 \\ 0 & -2 & 1 & 0 & 2 & 2 \\ 4 & 3 & -2 & -1 & 3 & 2 \\ -1 & 0 & -3 & -2 & 2 & 2 \\ 3 & 2 & -1 & -2 & 2 & 2 \\ 2 & 2 & 6 & 4 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 \\ 2 & 0 & 4 & 4 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_{11} \\ x_{12} \\ x_{21} \\ x_{22} \\ x_{31} \\ x_{32} \end{pmatrix} = E(\mathbf{C}), \tag{32}$$

which is a compatible system and has a unique solution, that after condensation leads to the solution

$$\mathbf{X} = \begin{pmatrix} -2 & 4 \\ 3 & 0 \\ 2 & -1 \end{pmatrix}. \tag{33}$$

∎

Example 4.2 (Generalized Sylvester equation (Case 2)) *Consider the following matrix equation*

$$\mathbf{A}_1\mathbf{X}\mathbf{B}_1 + \mathbf{A}_2\mathbf{X}\mathbf{B}_2 + \mathbf{A}_3\mathbf{X}\mathbf{B}_3 = \mathbf{C}, \tag{34}$$

where

$$\mathbf{A}_1 = \begin{pmatrix} 1 & 1 & 1 & 1 \\ -1 & 0 & 1 & 1 \end{pmatrix}; \quad \mathbf{A}_2 = \begin{pmatrix} 1 & 0 & 2 & 0 \\ 1 & 0 & 1 & 1 \end{pmatrix}; \quad \mathbf{A}_3 = \begin{pmatrix} 1 & 1 & 2 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix}; \tag{35}$$

$$\mathbf{B}_1 = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 0 & 2 \end{pmatrix}; \quad \mathbf{B}_2 = \begin{pmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \end{pmatrix}; \quad \mathbf{B}_3 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 0 \end{pmatrix}; \qquad (36)$$

$$\mathbf{C} = \begin{pmatrix} 33 & 20 & 26 \\ 22 & 7 & 7 \end{pmatrix}. \qquad (37)$$

Then, Equation (25) becomes

$$\left(\sum_{i=1}^{k} \mathbf{A}_i \otimes \mathbf{B}_i^t \right) E(\mathbf{X}) = \begin{pmatrix} 4 & 3 & 3 & 2 & 6 & 5 & 3 & 2 \\ 1 & 2 & 2 & 2 & 1 & 4 & 2 & 2 \\ 3 & 2 & 2 & 2 & 5 & 2 & 2 & 2 \\ -1 & 0 & 1 & 1 & 3 & 2 & 3 & 2 \\ -2 & 0 & 1 & 2 & 0 & 0 & 0 & 0 \\ 0 & -2 & 1 & 0 & 2 & 2 & 2 & 2 \end{pmatrix} \begin{pmatrix} x_{11} \\ x_{12} \\ x_{21} \\ x_{22} \\ x_{31} \\ x_{32} \\ x_{41} \\ x_{42} \end{pmatrix} = E(\mathbf{C}),$$

$$(38)$$

which is a compatible system and has a unique solution, that after condensation leads to the solution

$$\mathbf{X} = \begin{pmatrix} -2 - 7\rho_1 - \rho_2 & 4 + 11\rho_1 + 5\rho_2 \\ 3 + 38\rho_1 + 26\rho_2 & -26\rho_1 - 14\rho_2 \\ 2 - 3\rho_1 - 5\rho_2 & -1 - 13\rho_1 - 11\rho_2 \\ 3 + 8\rho_1 & 2 + 8\rho_2 \end{pmatrix}; \quad \rho_1, \rho_2 \in \mathbb{R}. \qquad (39)$$

■

Next, we solve the following particular cases of matrix equations.

4.2   Linear equations in a single matrix variable

Consider the matrix equation:

$$\mathbf{A}_{n,n} \bullet \mathbf{X}_{n,m} + \mathbf{X}_{n,m} \bullet \mathbf{B}_{m,m} = \mathbf{C}_{n,m}, \qquad (40)$$

in which the unknown is the matrix $\mathbf{X}_{n,m}$.

Equation (40) is a particular case of Equation (9). To solve this equation, we proceed to write it in an equivalent form, using the "tensor product"

$$(\mathbf{A}_{n,n} \otimes \mathbf{I}_m + \mathbf{I}_n \otimes \mathbf{B}_{m,m}^t) \bullet E(\mathbf{X}) = E(\mathbf{C}) <> \mathbf{M} \bullet E(\mathbf{X}) = E(\mathbf{C}) \qquad (41)$$

where

$$\mathbf{M}_{mn,mn} = \mathbf{A}_{n,n} \otimes \mathbf{I}_m + \mathbf{I}_n \otimes \mathbf{B}_{m,m}^t.$$

The solution $\mathbf{X}$ is unique if $\mathbf{M}$ is non-singular.

Example 4.3 (Equivalent matrices) *Assume that one is interested in obtaining the most general matrix $\mathbf{C}$ such that $\mathbf{AC} = \mathbf{CB}$, where*

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 2 \\ 1 & -1 & 0 \\ -1 & 1 & -2 \end{pmatrix} \quad and \quad \mathbf{B} = \begin{pmatrix} -3 & -2 & 3 \\ 7 & 2 & -5 \\ 3 & 0 & -1 \end{pmatrix}. \qquad (42)$$

If such a matrix **C** *exists, then we have*

$$\mathbf{AC} - \mathbf{CB} = \mathbf{\Omega},$$

which is a matrix equation of the form (40); thus, after stretching the matrix **C**, *we get (see Equation (41)):*

$$(\mathbf{A} \otimes \mathbf{I}_3 - \mathbf{I}_3 \otimes \mathbf{B}^t) \bullet E(\mathbf{C}) =
\left(\begin{array}{ccc|ccc|ccc}
1 & 0 & 0 & 1 & 0 & 0 & 2 & 0 & 0 \\
0 & 1 & 0 & 0 & 1 & 0 & 0 & 2 & 0 \\
0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 2 \\
\hline
1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & -1 & 0 & 0 & 0 \\
\hline
-1 & 0 & 0 & 1 & 0 & 0 & -2 & 0 & 0 \\
0 & -1 & 0 & 0 & 1 & 0 & 0 & -2 & 0 \\
0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 & -2
\end{array}\right)
\left(\begin{array}{c}
c_{11} \\ c_{12} \\ c_{13} \\ \hline c_{21} \\ c_{22} \\ c_{23} \\ \hline c_{31} \\ c_{32} \\ c_{33}
\end{array}\right)$$

$$-\left(\begin{array}{ccc|ccc|ccc}
-3 & 7 & 3 & 0 & 0 & 0 & 0 & 0 & 0 \\
-2 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
3 & -5 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & -3 & 7 & 3 & 0 & 0 & 0 \\
0 & 0 & 0 & -2 & 2 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 3 & -5 & -1 & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & 0 & 0 & -3 & 7 & 3 \\
0 & 0 & 0 & 0 & 0 & 0 & -2 & 2 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 3 & -5 & -1
\end{array}\right)
\left(\begin{array}{c}
c_{11} \\ c_{12} \\ c_{13} \\ \hline c_{21} \\ c_{22} \\ c_{23} \\ \hline c_{31} \\ c_{32} \\ c_{33}
\end{array}\right) = \mathbf{\Omega},$$

$$(43)$$

where c_{ij} are the elements of matrix **C** *and the block representation has been used for illustrating the relation of the new matrix with matrices* **A** *and* **B**.

Whence

$$\left(\begin{array}{ccc|ccc|ccc}
4 & -7 & -3 & 1 & 0 & 0 & 2 & 0 & 0 \\
2 & -1 & 0 & 0 & 1 & 0 & 0 & 2 & 0 \\
-3 & 5 & 2 & 0 & 0 & 1 & 0 & 0 & 2 \\
\hline
1 & 0 & 0 & 2 & -7 & -3 & 0 & 0 & 0 \\
0 & 1 & 0 & 2 & -3 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & -3 & 5 & 0 & 0 & 0 & 0 \\
\hline
-1 & 0 & 0 & 1 & 0 & 0 & 1 & -7 & -3 \\
0 & -1 & 0 & 0 & 1 & 0 & 2 & -4 & 0 \\
0 & 0 & -1 & 0 & 0 & 1 & -3 & 5 & -1
\end{array}\right)
\left(\begin{array}{c}
c_{11} \\ c_{12} \\ c_{13} \\ \hline c_{21} \\ c_{22} \\ c_{23} \\ \hline c_{31} \\ c_{32} \\ c_{33}
\end{array}\right) = \mathbf{\Omega}. \qquad (44)$$

The solution of (44), is the linear subspace:

$$
\begin{pmatrix} c_{11} \\ c_{12} \\ c_{13} \\ c_{21} \\ c_{22} \\ c_{23} \\ c_{31} \\ c_{32} \\ c_{33} \end{pmatrix} = \begin{pmatrix} 0 & -3 & 1 \\ 3 & -2 & 3 \\ -6 & 2 & -5 \\ 3 & 4 & 0 \\ 3 & 2 & 1 \\ -5 & -3 & -2 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \end{pmatrix} = \begin{pmatrix} -3\rho_2 + \rho_3 \\ 3\rho_1 - 2\rho_2 + 3\rho_3 \\ -6\rho_1 + 2\rho_2 - 5\rho_3 \\ 3\rho_1 + 4\rho_2 \\ 3\rho_1 + 2\rho_2 + \rho_3 \\ -5\rho_1 - 3\rho_2 - 2\rho_3 \\ \rho_3 \\ \rho_2 \\ \rho_1 \end{pmatrix},
$$

which after condensing, implies that the most general matrix $\mathbf{C}$ *that satisfies equation* $\mathbf{AC} = \mathbf{CB}$ *is*

$$
\mathbf{C} = \begin{pmatrix} -3\rho_2 + \rho_3 & 3\rho_1 - 2\rho_2 + 3\rho_3 & -6\rho_1 + 2\rho_2 - 5\rho_3 \\ 3\rho_1 + 4\rho_2 & 3\rho_1 + 2\rho_2 + \rho_3 & -5\rho_1 - 3\rho_2 - 2\rho_3 \\ \rho_3 & \rho_2 & \rho_1 \end{pmatrix}, \tag{45}
$$

where ρ_1, ρ_2 *and* ρ_3 *are arbitrary real constants. Its determinant is:*

$$
|\mathbf{C}| = -9\rho_1^3 - 33\rho_1^2\rho_2 - 31\rho_1\rho_2^2 - \rho_2^3 - 3\rho_1^2\rho_3 - 16\rho_1\rho_2\rho_3 - 21\rho_2^2\rho_3 + \rho_1\rho_3^2 + 5\rho_2\rho_3^2 - \rho_3^3,
$$

and thus, the most general change of basis matrix transforming matrix $\mathbf{B}$ *into matrix* $\mathbf{A}$ *is that given by (45) subject to*

$$
-9\rho_1^3 - 33\rho_1^2\rho_2 - 31\rho_1\rho_2^2 - \rho_2^3 - 3\rho_1^2\rho_3 - 16\rho_1\rho_2\rho_3 - 21\rho_2^2\rho_3 + \rho_1\rho_3^2 + 5\rho_2\rho_3^2 - \rho_3^3 \neq 0. \blacksquare
$$

Similarly, if the equation is

$$
\mathbf{A}_{m,n} \bullet \mathbf{X}_{n,p} \bullet \mathbf{B}_{p,q} = \mathbf{C}_{m,q}, \tag{46}
$$

the equivalent system of linear equations becomes

$$
(\mathbf{A}_{m,n} \otimes \mathbf{B}_{m,m}^t) \bullet E(\mathbf{X})_{np,1} = E(\mathbf{C})_{mq,1}, \tag{47}
$$

which solves the problem.

4.3   Linear equations in several matrix variables

We start with the case of two unknown matrices and then we generalize to k unknown matrices.

Consider the matrix equation:

$$
\mathbf{A}_{m,p} \bullet \mathbf{X}_{p,q} + \mathbf{Y}_{m,r} \bullet \mathbf{B}_{r,q} = \mathbf{C}_{m,q}, \tag{48}
$$

where the unknown matrices are $\mathbf{X}_{p,q}$ and $\mathbf{Y}_{m,r}$.

To solve this equation we proceed to write it stretched in an equivalent form, using the "tensor product":

$$(\mathbf{A}_{m,p} \otimes \mathbf{I}_q) \bullet E(\mathbf{X})_{pq,1} + (\mathbf{I}_m \otimes \mathbf{B}_{q,r}^t) \bullet E(\mathbf{Y})_{mr,1} = E(\mathbf{C})_{mq,1} \qquad (49)$$

to get

$$\left(\mathbf{A}_{m,p} \otimes \mathbf{I}_q | (\mathbf{I}_m \otimes (\mathbf{B}^t)_{q,r}\right) \begin{pmatrix} E(\mathbf{X}_{pq,1}) \\ -\,-\,-\,- \\ E(\mathbf{Y}_{mr,1}) \end{pmatrix} = E(\mathbf{C})_{mq,1}, \qquad (50)$$

which is equivalent to Equation (48) but written in the usual form. Then, the solution $(\mathbf{X}, \mathbf{Y})$ can be obtained by solving a linear system of equations.

Example 4.4 (Matrix linear equation) *Obtain the most general matrices* $\mathbf{X}$ *and* $\mathbf{Y}$ *such that* $\mathbf{AX} + \mathbf{YB} = \mathbf{C}$, *where:*

$$\mathbf{A} = \begin{pmatrix} -2 & 4 & -4 \\ 3 & 0 & 4 \\ 2 & -1 & 3 \end{pmatrix}; \quad \mathbf{B} = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 0 & 2 \end{pmatrix}; \quad \mathbf{C} = \begin{pmatrix} 7 & -4 & 10 \\ 5 & 10 & 0 \\ -3 & 4 & -4 \end{pmatrix}. \quad (51)$$

Since $\mathbf{AX} + \mathbf{YB} = \mathbf{C}$ *is of the form (48), after stretching matrices* $\mathbf{X}$ *and* $\mathbf{Y}$, *one gets (see Equation (50)):*

$$\left(\begin{array}{ccc|ccc|ccc||cc|cc|cc}
-2 & 0 & 0 & 4 & 0 & 0 & -4 & 0 & 0 & 2 & 1 & 0 & 0 & 0 & 0 \\
0 & -2 & 0 & 0 & 4 & 0 & 0 & -4 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & -2 & 0 & 0 & 4 & 0 & 0 & -4 & 1 & 2 & 0 & 0 & 0 & 0 \\
\hline
3 & 0 & 0 & 0 & 0 & 0 & 4 & 0 & 0 & 0 & 0 & 2 & 1 & 0 & 0 \\
0 & 3 & 0 & 0 & 0 & 0 & 0 & 4 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 3 & 0 & 0 & 0 & 0 & 0 & 4 & 0 & 0 & 1 & 2 & 0 & 0 \\
\hline
2 & 0 & 0 & -1 & 0 & 0 & 3 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 1 \\
0 & 2 & 0 & 0 & -1 & 0 & 0 & 3 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 2 & 0 & 0 & -1 & 0 & 0 & 3 & 0 & 0 & 0 & 0 & 1 & 2
\end{array}\right) \begin{pmatrix} x_{11} \\ x_{12} \\ x_{13} \\ x_{21} \\ x_{22} \\ x_{23} \\ x_{31} \\ x_{32} \\ x_{33} \\ - \\ y_{11} \\ y_{12} \\ y_{21} \\ y_{22} \\ y_{31} \\ y_{32} \end{pmatrix} = \begin{pmatrix} 7 \\ -4 \\ 10 \\ -- \\ 5 \\ 10 \\ 0 \\ -- \\ -3 \\ 4 \\ -4 \end{pmatrix},$$

and solving the resulting homogeneous system and condensing $\mathbf{X}$ *and* $\mathbf{Y}$ *one finally gets:*

$$\mathbf{X} = \begin{pmatrix} -2 - \rho_3 - 2\rho_4 - 4\rho_7 & 4 - \rho_4 - 4\rho_6 & -4 - 2\rho_3 - \rho_4 - 4\rho_5 \\ 3 + \rho_1 + 2\rho_2 - 2\rho_3 - 4\rho_4 + \rho_7 & \rho_2 - 2\rho_4 + \rho_6 & 4 + 2\rho_1 + \rho_2 - 4\rho_3 - 2\rho_4 + \rho_5 \\ 2 + 3\rho_7 & -1 + 3\rho_6 & 3 + 3\rho_5 \end{pmatrix},$$

$$\mathbf{Y} = \begin{pmatrix} -4\rho_2 + 6\rho_4 & -1 - 4\rho_1 + 6\rho_3 \\ 2 + 3\rho_4 & -1 + 3\rho_3 \\ -1 + \rho_2 & \rho_1 \end{pmatrix}.$$

where $\rho_1, \rho_2, \ldots, \rho_7 \in \mathbb{R}$.   ∎

Example 4.5 (One application to probability) *Assume that $\mathbf{\Sigma}_{n,n}$ is the covariance matrix of the n-dimensional random variable $\mathbf{X}$, then, the covariance matrix of the n-dimensional random variable $\mathbf{Y}_{n,1} = \mathbf{C}_{n,n}\mathbf{X}_{n,1}$ is $\mathbf{\Sigma}^*_{n,n} = \mathbf{C}_{n,n}\mathbf{\Sigma}_{n,n}\mathbf{C}^t_{n,n}$. If we look for $\mathbf{\Sigma}_{n,n} = \mathbf{I}_n$, it must be:*

$$\mathbf{\Sigma}^* = \mathbf{C}\mathbf{C}^t \Leftrightarrow \mathbf{I}_n\mathbf{C}^t - \mathbf{C}^{-1}\mathbf{\Sigma}^* = 0.$$

To obtain all changes of basis matrices $\mathbf{C}$ leading to this result, we initially solve the equation:

$$\mathbf{I}_n\mathbf{X} - \mathbf{Y}\mathbf{\Sigma}^* = 0,$$

which is of the form (48) and then it can be written as

$$\left(\mathbf{I}_n \otimes \mathbf{I}_n | (\mathbf{I}_n \otimes (\mathbf{\Sigma}^*)^t_{n,n}\right)\begin{pmatrix} E(\mathbf{X})_{nn,1} \\ ---- \\ E(\mathbf{Y})_{nn,1} \end{pmatrix} = E(\mathbf{C})_{nn,1}, \tag{52}$$

from which matrices $\mathbf{X}$ and $\mathbf{Y}$ can be obtained. Next, it suffices to impose the condition

$$\mathbf{Y} = \left((\mathbf{X})^t\right)^{-1}.$$

As an example, if the covariance matrix is

$$\mathbf{\Sigma}^* = \begin{pmatrix} 3 & 1 & -1 \\ 1 & 4 & 0 \\ -1 & 0 & 2 \end{pmatrix}$$

then, we get

$$\mathbf{X} = \begin{pmatrix} -\rho_7 + \rho_8 + 3\rho_9 & 4\rho_8 + \rho_9 & 2\rho_7 - \rho_9 \\ -\rho_4 + \rho_5 + 3\rho_6 & 4\rho_5 + \rho_6 & 2\rho_4 - \rho_6 \\ -\rho_1 + \rho_2 + 3\rho_3 & 4\rho_2 + \rho_3 & 2\rho_1 - \rho_3 \end{pmatrix}; \quad \mathbf{Y} = \begin{pmatrix} \rho_9 & \rho_8 & \rho_7 \\ \rho_6 & \rho_5 & \rho_4 \\ \rho_3 & \rho_2 & \rho_1 \end{pmatrix}.\blacksquare$$

Finally we mention that Equations (23) and (48) can be immediately generalized to

$$\sum_{i=1}^{I}(\mathbf{A}_i)_{m,p_i} \bullet (\mathbf{X}_i)_{p_i,q} + \sum_{j=1}^{J}(\mathbf{Y}_j)_{m,r_j} \bullet (\mathbf{B}_j)_{r_j,q} = \mathbf{C}_{m,q} \tag{53}$$

leading to the following solution, that generalizes (50):

$$\left((\mathbf{A}_1)_{m,p_1} \otimes \mathbf{I}_{q,q} | \ldots | (\mathbf{A}_I)_{m,p_I} \otimes \mathbf{I}_{q,q} | \mathbf{I}_{m,m} \otimes (\mathbf{B}_1)^t_{r_1,q} | \cdots | \mathbf{I}_{m,m} \otimes (\mathbf{B}_J)^t_{r_J,q}\right) \bullet$$

$$\begin{pmatrix} (E(\mathbf{X})_1)_{p_1 q,1} \\ \vdots \\ (E(\mathbf{X})_I)_{p_I q,1} \\ --- \\ (E(\mathbf{Y})_1)_{mr_1,1} \\ \vdots \\ (E(\mathbf{Y})_J)_{mr_J,1} \end{pmatrix} = E(\mathbf{C})_{mq,1}. \tag{54}$$

4.4   Matrix tensor product commuters equation

We call *"transposition matrix"* of order *"n"*, to every matrix resulting from exchanging any two rows of the unit matrix $\mathbf{I}_n$, and leaving the remaining rows unchanged. The transposition matrices are always, regular ($|\mathbf{P}| \neq 0$, symmetric ($\mathbf{P} = \mathbf{P}^t$), involutive ($\mathbf{P} = \mathbf{P}^{-1}$) and orthogonal ($\mathbf{P}^{-1} = \mathbf{P}^t$).

We call "permutation matrix" to the classic or tensor product of several "transposition matrices" (in the second case they can be of different order ($\mathbf{P} = \mathbf{P}_{m,m} \otimes \mathbf{P}_{n,n}$).

Consider the equation:

$$\mathbf{B} \otimes \mathbf{A} = \mathbf{P}_1 \bullet (\mathbf{A} \otimes \mathbf{B}) \bullet \mathbf{P}_2, \tag{55}$$

where $\mathbf{P}_1 \in \{\text{permutations of } \mathbf{I}_{mp}\}$ and $\mathbf{P}_2 \in \{\text{permutations of } \mathbf{I}_{nq}\}$ are the unknown matrices.

Note that in general $\mathbf{A} \otimes \mathbf{B} \neq \mathbf{B} \otimes \mathbf{A}$, where $(\mathbf{A} \otimes \mathbf{B})_{mp,nq}$, i.e., the tensor product is not commutative. Thus, since direct reversing of the tensor product is not permitted, Equation (55) allows to find two correction matrices $\mathbf{P}_1$ and $\mathbf{P}_2$ for reversing the tensor product, that will be called "transposer matrices".

We shall give two different expressions for the solution matrices $\mathbf{P}_1$ and $\mathbf{P}_2$.

The first solution is as follows. The permutation matrices $\mathbf{P}_1$ and $\mathbf{P}_2$ (orthogonal matrices $\mathbf{P}_1^{-1} = \mathbf{P}_1^t; \mathbf{P}_2^{-1} = \mathbf{P}_2^t$) that solve Equation (55), for the products $\mathbf{A}_{m,n} \otimes \mathbf{B}_{p,q}$ and $\mathbf{B}_{p,q} \otimes \mathbf{A}_{m,n}$ are as follows:

$$\mathbf{P}_1(m,p) \equiv \mathbf{P}_{mp,mp} = [p_{1ij}],$$

where

$$p_{1ij} = \begin{cases} 1 & \text{if } i = \left(1 + \left\lfloor \dfrac{j-1}{p} \right\rfloor\right) + \left((j-1) - p \left\lfloor \dfrac{j-1}{p} \right\rfloor\right) m \\ 0 & \text{otherwise} \end{cases} ; \tag{56}$$

for $i, j = 1, 2, \ldots, mp$, where $\lfloor x \rfloor$ is the integer part of x, and

$$\mathbf{P}_2(n,q) \equiv \mathbf{P}_{nq,nq} = [p_{2ij}],$$

where

$$p_{2ij} = \begin{cases} 1 & \text{if } j = \left(1 + \left\lfloor \dfrac{i-1}{q} \right\rfloor\right) + \left((i-1) - q \left\lfloor \dfrac{i-1}{q} \right\rfloor\right) n \\ 0 & \text{otherwise} \end{cases} ; \tag{57}$$

for $i, j = 1, 2, \ldots, nq$, showing that

$$\mathbf{P}_2(n,q) = \mathbf{P}_1^t(n,q). \tag{58}$$

Example 4.6 (Commuting the tensor product) *Consider the partic-*
ular case $\mathbf{A}_{3,3}$ *and* $\mathbf{B}_{3,3}$ *$(m = n = p = q = 3)$, with* $\mathbf{A}_{3,3} = [a_{ij}]; \mathbf{B} = [b_{ij}]$.

Applying the indicated formulas, one obtains $\mathbf{P}_1 \equiv \mathbf{P}_{mp,mp} \equiv \mathbf{P}_{9,9}$ *with*
$p_{ij} = 0$, *with the exception of*

row	1	4	7	2	5	8	3	6	9
column	1	2	3	4	5	6	7	8	9

that is, in the positions

$$(i, j) \equiv (1, 1), (4, 2), (7, 3), (2, 4), (5, 5), (8, 6), (3, 7), (6, 8), (9, 9),$$

that take value 1, and $\mathbf{P}_2 \equiv \mathbf{P}_{nq,nq} \equiv \mathbf{P}_{9,9}$ *with* $p_{ij} = 0$, *with the exception of*

row	1	2	3	4	5	6	7	8	9
column	1	4	7	2	5	8	3	6	9

that leads to a value 1 in positions

$$(i, j) \equiv (1, 1), (2, 4), (3, 7), (4, 2), (5, 5), (6, 8), (7, 3), (8, 6), (9, 9).$$

As one can see, the results are identical, and then $\mathbf{P} = \mathbf{P}_1 = \mathbf{P}_2$, *where* $\mathbf{P}$
is symmetric, involutive and orthogonal, thus, we get

$$\begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix} \otimes \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

$$= \mathbf{P} \bullet \left(\left(\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \otimes \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix} \right) \right) \bullet \mathbf{P}$$

with

$$\mathbf{P} = \begin{pmatrix}
1\ 0\ 0 & 0\ 0\ 0 & 0\ 0\ 0 \\
0\ 0\ 0 & 1\ 0\ 0 & 0\ 0\ 0 \\
0\ 0\ 0 & 0\ 0\ 0 & 1\ 0\ 0 \\
\hline
0\ 1\ 0 & 0\ 0\ 0 & 0\ 0\ 0 \\
0\ 0\ 0 & 0\ 1\ 0 & 0\ 0\ 0 \\
0\ 0\ 0 & 0\ 0\ 0 & 0\ 1\ 0 \\
\hline
0\ 0\ 1 & 0\ 0\ 0 & 0\ 0\ 0 \\
0\ 0\ 0 & 0\ 0\ 1 & 0\ 0\ 0 \\
0\ 0\ 0 & 0\ 0\ 0 & 0\ 0\ 1
\end{pmatrix}.$$

■

For the case with $p = m, q = n$, that is, matrices $\mathbf{A}_{mn}$ and $\mathbf{B}_{mn}$ that have the same number of rows and columns, we have $\mathbf{P}_1 = \mathbf{P}_2^t$ and since they are orthogonal, the matrices $\mathbf{B} \otimes \mathbf{A}$ and $\mathbf{A} \otimes \mathbf{B}$ are "similar inside the orthogonal group"

$$\mathbf{P}^{-1} \bullet (\mathbf{A} \otimes \mathbf{B}) \bullet \mathbf{P} \equiv \mathbf{P}^t \bullet (\mathbf{A} \otimes \mathbf{B}) \bullet \mathbf{P} = \mathbf{B} \otimes \mathbf{A}$$

As a final result of the analysis of the matrices $\mathbf{P}_1$ and $\mathbf{P}_2$, that appear in Formula (55), we shall propose a second and simple general expression of such matrices,

Let $\mathcal{B}_1 = \{\mathbf{E}_{11}, \mathbf{E}_{12}, \cdots, \mathbf{E}_{ij}, \cdots, \mathbf{E}_{mp}\}$ be the canonical basis, with $m \times p$ matrices, of the $\mathbb{R}^{m \times p}$ matrix linear space.

Let $\mathcal{B}_2 = \{\mathbf{E}'_{11}, \mathbf{E}'_{12}, \cdots, \mathbf{E}'_{k\ell}, \cdots, \mathbf{E}'_{nq}\}$ be the canonical basis with $n \times q$ matrices, in the $\mathbb{R}^{n \times q}$ matrix linear space.

Matrices $\mathbf{P}_1$ and $\mathbf{P}_2$ will be represented by blocks:

$$\mathbf{P}_1 \equiv \mathbf{P}_{mp,mp} = \begin{pmatrix} \mathbf{E}_{11} & \mathbf{E}_{21} & \cdots & \mathbf{E}_{m1} \\ \mathbf{E}_{12} & \mathbf{E}_{22} & \cdots & \mathbf{E}_{m2} \\ \cdots & \cdots & \cdots & \cdots \\ \mathbf{E}_{1p} & \mathbf{E}_{2p} & \cdots & \mathbf{E}_{mp} \end{pmatrix} \tag{59}$$

$$\mathbf{P}_2 \equiv \mathbf{P}_{nq,nq} = \begin{pmatrix} \mathbf{E}'_{11} & \mathbf{E}'_{21} & \cdots & \mathbf{E}'_{n1} \\ \mathbf{E}'_{12} & \mathbf{E}'_{22} & \cdots & \mathbf{E}'_{n2} \\ \cdots & \cdots & \cdots & \cdots \\ \mathbf{E}'_{1q} & \mathbf{E}'_{2q} & \cdots & \mathbf{E}'_{nq} \end{pmatrix}^t \tag{60}$$

Special attention must be given to the "block ordering" inside matrices $\mathbf{P}_1$ and $\mathbf{P}_2$; it is not the canonical order but the transpose. As an example, an application is given below.

Example 4.7 (Commuting the tensor product) *Consider again the matrices* $\mathbf{A}_{3,3}$ *and* $\mathbf{B}_{3,3}$ *in Example 4.6. The matrices* $\mathbf{P}_1$ *and* $\mathbf{P}_2$ *that solve our application* $(m = n = p = q = 3)$ *have now a direct construction:*

$$
\mathbf{P}_1 = \mathbf{P}_2 = \mathbf{P} \;=\;
\begin{pmatrix}
\mathbf{E}_{11} & \mathbf{E}_{21} & \mathbf{E}_{31} \\
\mathbf{E}_{12} & \mathbf{E}_{22} & \mathbf{E}_{32} \\
\mathbf{E}_{13} & \mathbf{E}_{23} & \mathbf{E}_{33}
\end{pmatrix}
=
\left(
\begin{array}{ccc|ccc|ccc}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
\hline
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
\hline
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{array}
\right),
$$

$$(61)$$

that evidently coincides with the $\mathbf{P}$ *obtained in Example 4.6, that was obtained after using a certainly complicated subindex relationships.*

∎

4.5 The Schur tensor product equation

Consider the following matrix equation, which allows us replacing a tensor product by a Schur product:

$$
\mathbf{A}_{m,n} \diamond \mathbf{B}_{m,n} = \mathbf{Q}_{m,m^2} \bullet (\mathbf{A}_{m,n} \otimes \mathbf{B}_{m,n}) \bullet \mathbf{P}_{n^2,n}, \tag{62}
$$

where $\mathbf{P}$ and $\mathbf{Q}$, the unknowns, never are square matrices. Note that this is a direct relationship among the "three matrix products".

Solution matrices $\mathbf{Q}_{m,m^2}$ and $\mathbf{P}_{n^2,n}$ are, respectively, given by

$$
\mathbf{Q}_{m,m^2} = [q_{ij}]; \quad q_{ij} \in \{0,1\} \text{ with } q_{ij} = \begin{cases} 1 & \text{if } j = i(m+1) - m \\ 0 & \text{otherwise} \end{cases}
$$

and

$$
\mathbf{P}_{n^2,n} = [p_{ij}]; \quad p_{ij} \in \{0,1\} \text{ with } p_{ij} = \begin{cases} 1 & \text{if } i = j(n+1) - n \\ 0 & \text{otherwise.} \end{cases}
$$

Nevertheless, and following the previous criterion of having a faster formulation for matrices $\mathbf{Q}_{m,m^2}$ and $\mathbf{P}_{n^2,n}$ in Formula (62) we propose the following block alternative:

$$
\mathbf{Q}_{m,m^2} = \begin{pmatrix} \mathbf{E}_{11} & | & \mathbf{E}_{22} & | & \cdots & | & \mathbf{E}_{m,m} \end{pmatrix}, \tag{63}
$$

where $\mathcal{B}_{m^2} \equiv \{\mathbf{E}_{ij}\}$ is the canonical basis of the matrix linear space $\mathbb{R}^{m,m}$, and

$$
\mathbf{P}_{n^2,n} =
\begin{pmatrix}
\mathbf{E}'_{11} \\
-- \\
\mathbf{E}'_{22} \\
-- \\
\vdots \\
-- \\
\mathbf{E}'_{nn}
\end{pmatrix},
\tag{64}
$$

where $\mathcal{B}'_{n^2} = \{\mathbf{E}'_{ij}\}$ is the canonical basis of the matrix linear space $\mathbb{R}^{n,n}$.

Example 4.8 (Replacing a Schur product by a tensor product) *Returning to the case $\mathbf{A}_{3,3}$ and $\mathbf{B}_{3,3}$ of a previous example, we have*

$$
\mathbf{A}_{3,3} \diamond \mathbf{B}_{3,3} = \mathbf{Q}_{3,9} \bullet (\mathbf{A}_{3,3} \otimes \mathbf{B}_{3,3}) \bullet \mathbf{P}_{9,3}
$$

where $\mathbf{Q} = \mathbf{P}^t$ or $\mathbf{P} = \mathbf{Q}^t$ and

$$
\mathbf{Q} =
\begin{pmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix},
$$

as it can be checked easily, using the previous formulas. ∎

As a consequence of (62), a relation between dot and tensor products can be obtained for the particular case $p = m, q = n$. In fact, we know that

$$
\mathbf{A}_{m,n} \diamond \mathbf{B}_{m,n} = \mathbf{Q}_{m,m^2} \bullet (\mathbf{A}_{m,n} \otimes \mathbf{B}_{m,n}) \bullet \mathbf{P}_{n^2,n}
$$
$$
\mathbf{B}_{m,n} \diamond \mathbf{A}_{m,n} = \mathbf{Q}_{m,m^2} \bullet (\mathbf{B}_{m,n} \otimes \mathbf{A}_{m,n}) \bullet \mathbf{P}_{n^2,n}
$$

and applying the commutative property to the left hand side member and equalling the right hand side members, we get

$$
\mathbf{Q}_{m,m^2} \bullet (\mathbf{A}_{m,n} \otimes \mathbf{B}_{m,n}) \bullet \mathbf{P}_{n^2,n} = \mathbf{Q}_{m,m^2} \bullet (\mathbf{B}_{m,n} \otimes \mathbf{A}_{m,n}) \bullet \mathbf{P}_{n^2,n}.
$$

4.6 The square equation

Consider the matrix equation

$$
\mathbf{P}^2\mathbf{Z} + 2\mathbf{PZQ} + \mathbf{ZQ}^2 = \mathbf{M}
\tag{65}
$$

where the matrices $\mathbf{P}_{n,n}, \mathbf{Q}_{m,m}$ and $\mathbf{M}_{n,m}$ are data and the matrix $\mathbf{Z}_{n,m}$ is the unknown.

Since this is a particular case of the generalized Sylvester equation, we can solve it by writing it as

$$\left(\mathbf{P}^2 \otimes \mathbf{I}_m + 2\mathbf{P} \otimes \mathbf{Q}^t + \mathbf{I}_n \otimes \mathbf{Q}^2\right) E(\mathbf{Z}) = E(\mathbf{M}). \tag{66}$$

However, there is a better alternative, that consists of realizing that the linear system (66) can be written as

$$\mathbf{N}^2\mathbf{z} = \mathbf{N}\left(\mathbf{N}\mathbf{z}\right) = \mathbf{m} \tag{67}$$

where $\mathbf{N} = \mathbf{P} \otimes \mathbf{I}_m + \mathbf{I}_n \otimes \mathbf{Q}^t$, that suggests solving Equation (67) in two steps:

1. Solve Equation $\mathbf{N}\mathbf{y} = \mathbf{m}$.

2. Solve Equation $\mathbf{N}\mathbf{z} = \mathbf{y}$.

Example 4.9 (Square equation) *Consider the equation*

$$\mathbf{P}^2\mathbf{Z} + 2\mathbf{P}\mathbf{Z}\mathbf{Q} + \mathbf{Z}\mathbf{Q}^2 = \mathbf{M} \tag{68}$$

where

$$\mathbf{P} = \begin{pmatrix} 0 & 1 & -1 \\ 1 & 0 & 1 \\ -1 & 1 & 2 \end{pmatrix}; \quad \mathbf{Q} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}; \quad \mathbf{M} = \begin{pmatrix} -2 & 3 \\ 10 & 21 \\ 22 & 29 \end{pmatrix}.$$

Using equation (66) we get

$$\left(\mathbf{P}^2 \otimes \mathbf{I}_m + 2\mathbf{P} \otimes \mathbf{Q}^t + \mathbf{I}_n \otimes \mathbf{Q}^2\right)\mathbf{z} = \begin{pmatrix} 3 & 0 & 1 & 0 & -3 & 0 \\ 4 & 3 & 4 & 1 & -4 & -3 \\ 1 & 0 & 3 & 0 & 3 & 0 \\ 4 & 1 & 4 & 3 & 4 & 3 \\ -3 & 0 & 3 & 0 & 11 & 0 \\ -4 & -3 & 4 & 3 & 12 & 11 \end{pmatrix} \mathbf{z} = \begin{pmatrix} -2 \\ 3 \\ 10 \\ 21 \\ 22 \\ 29 \end{pmatrix},$$

which leads to

$$\mathbf{z} = \begin{pmatrix} 3 & 0 & 1 & 0 & -3 & 0 \\ 4 & 3 & 4 & 1 & -4 & -3 \\ 1 & 0 & 3 & 0 & 3 & 0 \\ 4 & 1 & 4 & 3 & 4 & 3 \\ -3 & 0 & 3 & 0 & 11 & 0 \\ -4 & -3 & 4 & 3 & 12 & 11 \end{pmatrix}^{-1} \begin{pmatrix} -2 \\ 3 \\ 10 \\ 21 \\ 22 \\ 29 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 1 \\ 0 \\ 2 \\ 1 \end{pmatrix}$$

Using the alternative method, we have

$$\mathbf{N} = \mathbf{P} \otimes \mathbf{I}_m + \mathbf{I}_n \otimes \mathbf{Q}^t = \begin{pmatrix} 1 & 0 & 1 & 0 & -1 & 0 \\ 2 & 1 & 0 & 1 & 0 & -1 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 1 \\ -1 & 0 & 1 & 0 & 3 & 0 \\ 0 & -1 & 0 & 1 & 2 & 3 \end{pmatrix}$$

and then

$$\mathbf{y} = \left(\mathbf{N}^{-1}\right)^2 \mathbf{m} = \begin{pmatrix} -\frac{1}{2} & 0 & 1 & 0 & -\frac{1}{2} & 0 \\ -3 & -\frac{1}{2} & \frac{5}{2} & 1 & -\frac{3}{2} & -\frac{1}{2} \\ 1 & 0 & -\frac{1}{2} & 0 & \frac{1}{2} & 0 \\ \frac{5}{2} & 1 & -3 & -\frac{1}{2} & \frac{3}{2} & \frac{1}{2} \\ -\frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 \\ -\frac{3}{2} & -\frac{1}{2} & \frac{3}{2} & \frac{1}{2} & -1 & 0 \end{pmatrix}^2 \begin{pmatrix} -2 \\ 3 \\ 10 \\ 21 \\ 22 \\ 29 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 1 \\ 0 \\ 2 \\ 1 \end{pmatrix},$$

which after condensation leads to

$$M_{(\{1\},\{2\})}\left(\mathbf{Z}\right) = \begin{pmatrix} 1 & 2 \\ 1 & 0 \\ 2 & 1 \end{pmatrix}.$$

∎

5   Conclusions

In this paper a general method for stretching-condensing tensors has been described. The technique is useful for dealing with tensors using all the already available and powerful tools for vectors. First, the tensors are written as vectors using the stretching transformation that converts the initial structure of tensors components to a column matrix of its vector components with respect to a given basis. Next, the required vector transformations are used. Finally, the vectors are transformed back to their initial tensor representation. In addition, this technique has been shown to be useful for solving important problems in algebra, as solving linear tensor and matrix equations, and to have important applications.

References

[1] J. K. Baksalary and R. Kala. The matrix equation $AXB + CYD = E$. *Linear Algebra Appl.*, 30:141–147, 1980.

[2] C. S. Ballantine. A note on the matrix equation $H = AP + PA$. *Linear Algebra and Appl.*, 2:37–47, 1969.

[3] S. Barnett. Remarks on solution of $AX + XB = C$. *Electron. Lett.*, 7:385, 1971.

[4] S. Barnett. *Matrices in control theory.* Krieger, Florida, 1984.

[5] S. Barnett. *Matrices. Methods and Applications.* Oxford University Press, Oxford, 1990.

[6] H. T. Burgess. Solution of the matrix equation $X^{-1}AX = N$. *Ann. of Math. (2)*, 19(1):30–36, 1917.

[7] M. Dobovišek. On minimal solutions of the matrix equation $AX - YB = 0$. *Linear Algebra Appl.*, 325(1-3):81–99, 2001.

[8] I. Gohberg, P. Lancaster, and L. Rodman. *Matrix polynomials.* Academic Press Inc., New York, 1982. Computer Science and Applied Mathematics.

[9] G. H. Golub and C. F. Van Loan. *Matrix computations.* Johns Hopkins University Press, Baltimore, second edition, 1989.

[10] C. N. He. The general solution of the matrix equation $AXB + CYD = F$. *Acta Sci. Natur. Univ. Norm. Hunan.*, 19(1):17–20, 1996.

[11] M. H. Ingraham and H. C. Trimble. On the matric equation $TA = BT + C$. *Amer. J. Math.*, 63:9–28, 1941.

[12] T. Jiang and M. Wei. On solutions of the matrix equations $X - AXB = C$ and $X - A\overline{X}B = C$. *Linear Algebra Appl.*, 367:225–233, 2003.

[13] Y. M. Li. The generalized quaternion matrix equation $AX - DXB = R$. *J. Math. (Wuhan)*, 22(3):281–286, 2002.

[14] C. S. Lu. Solution of the matrix equation $AX + XB = C$. *Electron. Lett.*, 7:185–186, 1971.

[15] C. C. Mac Dufee. *The Theory of matrices.* Chelsea Scientific books, New York, 1946.

[16] H. X. Niu. The matrix equation $AX + YA = C$ over an arbitrary division ring. *Qufu Shifan Daxue Xuebao Ziran Kexue Ban*, 29(3):39–40, 2003.

[17] W. V. Parker. The matrix equation $AX = XB$. *Duke Math. J.*, 17:43–51, 1950.

[18] Z. Y. Peng. The solutions of the matrix equation $AXC + BYD = E$ and their optimal approximation. *Math. Theory Appl. (Changsha)*, 22(2):99–103, 2002.

[19] C. Reischer and D. A. Simovici. On the matrix equation $\mathbf{PAP}^T = \mathbf{B}$ in the two-element Boolean algebra and some applications in graph theory. *An. Şti. Univ. "Al. I. Cuza" Iaşi Secţ. I a Mat. (N.S.)*, 17:263–274, 1971.

[20] H. H. Rosenbrock. *State-space and multivariable theory.* John Wiley & Sons, Inc., New York, 1970.

[21] J. R. Ruiz-Tolosa and E. Castillo. *From vectors to tensors.* Springer-Verlag, New York, 2005. Universitext.

[22] S. Y. Shim and Y. Chen. Least squares solution of matrix equation $AXB^* + CYD^* = E$. *SIAM J. Matrix Anal. Appl.*, 24(3):802–808 (electronic), 2003.

[23] J. J. Sylvester. Sur la résolution générale de l'équation linéaire en matrices d'un ordre quelconque. *Comptes Rendues Academie of Sciences*, 99:409–412, 432–436, 1984.

[24] J. J. Sylvester. Sur la solution du cas le plus général des équations linéaires en quantités binaires, c'est-à-dire en quaternions ou en matrices du second ordre. *Comptes Rendues Academie of Sciences*, 99:117–118, 1984.

[25] Y. Tian. Ranks of solutions of the matrix equation $AXB = C$. *Linear Multilinear Algebra*, 51(2):111–125, 2003.

[26] H. K. Wimmer. The matrix equation $X - AXB = C$ and an analogue of Roth's theorem. *Linear Algebra Appl.*, 109:145–147, 1988.

[27] Q. K. Ye. Solving Lyapunov matrix equation $A^T P + PA = -Q$ by optimization. *Acta Math. Sci. (English Ed.)*, 8(2):145–149, 1988.

[28] X. H. You, T. Yan, and S. R. Ma. Anti-symmetric solution of the matrix equation $AXB = C$. *J. Nanjing Norm. Univ. Nat. Sci. Ed.*, 26(1):6–10, 2003.