Binomial Expansion for the Kronecker Powers of Vector Sums

METİN DEMİRALP
İstanbul Technical University, Informatics Institute,
Maslak, 34469, İstanbul, TÜRKİYE (TURKEY)
metin.demiralp@gmail.com

Abstract: We have developed a concise representation instead of the multivariate Taylor series expansion. The new representation is constructed in ascending Kronecker powers of a system vector. The expansion is not unique because of the Kronecker powers’ specific structures. It is possible to use this nonuniqueness to give certain desired properties to the expansion coefficients. These Kronecker powers can be treated in a specific way if the system vector is the sum of two separate vectors of same dimension. This brings the issue of handling binomial expansion type formulae, to mind. This work is devoted to the development of the binomial expansion for the Kronecker powers of vector sums. It also aim at the construction of a recursion between binomial coefficients. The basic goal is formulation and constructing theoretical background. We consider no need for any implementation.

Key–Words: Kronecker products, Kronecker powers, Direct producst and powers, Permutations, Binomial Representation

1 Introduction

There have been quite important progresses recently for the theory of ODEs in our group studies. We have developed a new approach to the solution of first order set of ODEs in such a way that the use of certain basis functions to span somehow a Kronecker power space spanned by the unknowns of the ODEs to get a first order infinite and homogeneous set of linear ODEs with constant infinite coefficient matrix has been successful for explicit first order autonomous ODE sets. Autonomy is not a restriction since a space extension provides it. The explicitness seems to be an important restriction however many practical problems are modeled through this type ODEs. Even in the case of implicitness, explicitness may be availed by solving an algebraic equation if possible. We have called this procedure “Probabilistic Evolution Approach (PEA)” because of the possibility of having initial conditions which are given through probabilistic structures. Even though the cases where initial value problem of ODEs is under consideration correspond to the strictly sharp probabilistic distributions, the dynamical systems where the unknowns become the expectation values of certain operators need to use the probability concept. We do not intend to give all details of the recent status of the theory of PEA. The curious reader can look at related references [1–5] which are quite recent publications. However this theory finds its roots in earlier reports [6–9]. Certain other works, which may be appearing as if secondary despite their important contributions, can give the insight about the state of the art of the theory [10–15].

The infinite set of ODEs can be formally solved because of the coefficient matrix constancy. The solution is the image of an initially given vector under the exponential matrix whose argument is the product of the coefficient matrix with the time variable. Hence, this exponential matrix is responsible for the evolution of the system characterized by the given set of ODEs. We call this matrix “Evolution Matrix”. The theory of PEA dictates us that this matrix is in upper block Hessenberg form and the blocks of each diagonal is generated by a rectangular matrix related to the right hand side (descriptive) functions of the original ODEs. If the descriptive functions vanish at the expansion point then the Evolution Matrix becomes upper block triangular which is a pleasant structure because of the rather simplicity of its spectral property determination. We call these situations “Triangularity Case” and prefer them to use. If the block diagonal generators except the one for main diagonal are all vanishing then the evolution matrix turns out to be partitioned to finite square blocks and obviously it is the quite pleasant situation. This is called “Block Diagonality”. The next simplest case corresponds the evolution matrices having main block diagonal and its nearest upper block diagonal neighbor as nonvanishing structures. We call this case “Conicality” due to geometrical considerations. The case of conicality enables us to construct analytical Kronecker power series as the solutions. If for these cases, the initial vector is considered as core expectation value plus its fluctuations then we need to deal with direct power se-
eries of two vector sums even though we have not yet attempted to focus on such cases.

PEA is fundamentally based on the utilization of the direct products and direct powers which are also known as Kronecker products and powers. However, until now, we needed to focus on only the direct power series of a single vector. On the other hand, as above-mentioned, our most recent studies revealed that we may need to focus on the direct power series of two vector sums. The vectors in these sums may play quite different roles in the theory depending on the modeled physical system. This necessity, however, requires the rearrangement of the direct power series in such a way that it becomes a direct power series with respect to one of the vector summands while the expansion coefficients depend on the other vector summand. The algebraic rearrangement to this end is generally based on the use of binomial expansions in the scalar cases. Hence, we need to develop a binomial expansion for nonnegative integer direct powers of a binary vector sum. This paper is devoted to this aim and the remaining sections are organized to present the construction of the binomial expansion to this end. Our main focus is conceptual derivation basically, the implementations are left to the future works.

2 Kronecker Power Series of Binary Vector Sums

The Kronecker power series [16–19] of an analytical multivariate function can be written as follows

\[ f(s) = \sum_{j=0}^{\infty} F_j s \otimes^j \]  

where \( s \) is defined through the following equalities

\[ s \equiv x - x^{(e)} \]

\[ x \equiv \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad x^{(e)} \equiv \begin{bmatrix} x_1^{(e)} \\ \vdots \\ x_n^{(e)} \end{bmatrix} \]

These dictates us that the number of the independent variables is \( n \) and the independent variables are denoted by \( x_1, \ldots, x_n \). The entities having superscript \( (e) \) are the components of the expansion point vector in \( n \) dimensional Euclidean space. The vector \( s \) is called "system vector" because of its frequent appearances in the theory of dynamical systems.

The Kronecker power of the state vector can be recursively defined through the following equality

\[ s \otimes^j = \begin{bmatrix} s_1 \otimes^{j-1} \\ \vdots \\ s_n \otimes^{j-1} \end{bmatrix}, \quad j = 1, 2, \ldots \]  

whose starting point is the zeroth Kronecker power which is defined as follows

\[ s \otimes^0 \equiv 1 \]  

by following the general convention for the algebraic structures. This definition implies that the \( j \)th Kronecker power of the state vector is composed of \( n^j \) elements. (1) has a scalar entity at its left hand side. Hence the right hand side’s infinite linear combination summands should also be scalar for compatibility. This can be provided by giving specific types to each summand’s coefficient matrix. This means that \( F_j \) must be of \( 1 \times n^j \) type. In other words, the rectangular matrix \( F_j \) is in fact the transpose of an \( n^j \) element vector.

Despite \( s \) is independent of the target function, \( F_j \) coefficients involve all informations about the target function since they are composed of the derivative values of the target function, evaluated at the expansion point.

In certain specific cases the state vector becomes the sum of two separate vectors like in the case of the expectations and fluctuations. We can then rewrite (1) as follows

\[ f(x + y) = \sum_{j=0}^{\infty} F_j (x + y) \otimes^j \]  

Now our purpose becomes to reexpress the right hand side of this formula in a different but more efficient way to separately cluster the \( x \) and \( y \) vector containing factors. We need to arrive at the following formula

\[ f(x + y) = \sum_{j=0}^{\infty} G_j (x) y \otimes^j \]

where the \( j \)th coefficient matrix is again the transpose of an \( n^j \) element vector which depends on the elements of the vector \( x \). To get this result what we need is the reexpression of the power \( (x + y)^j \) in a canonical form for any given \( j \) integer value in such a way that the last formula above can be obtained. This work focuses on this issue.
3 Binomial expansion for the non-negative integer power of a vector sum

Let us focus on the $j$th Kronecker power of the sum $x + y$. The case where $j = 0$ produces just 1 because of the usual convention of algebra

$$(x + y)^{\otimes 0} = 1$$

(8)

while the case where $j = 1$ produces the base sum $x + y$. There is nothing unusual in these two cases. Whereas the situation changes when $j = 2$ where we can write the following explicit formula

$$(x + y)^{\otimes 2} = x^{\otimes 2} + x \otimes y + y \otimes x + y^{\otimes 2}$$

(9)

which contains the $y$ factors at the right hand side of the $x$ factors in all summands except the third one. However, the Kronecker products $x \otimes y$ and $y \otimes x$ are composed of exactly same terms but maybe in different locations. In other words, the only difference between these two Kronecker products is just the ordering of the same set of elements. This implies that there should be a mapping characterized by an $n^2 \times n^2$ type matrix, between these two products. Hence we can write

$$y \otimes x = \Pi (x \otimes y)$$

(10)

where $\Pi$ stands for a permutation matrix changing element ordering in its operand to produce the left hand side Kronecker product. The following equality is given to exemplify this relation for the easiest case

$$y \otimes x = \begin{bmatrix} y_1 x_1 \\ y_1 x_2 \\ y_2 x_1 \\ y_2 x_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 y_1 \\ x_1 y_2 \\ x_2 y_1 \\ x_2 y_2 \end{bmatrix} = \Pi (x \otimes y)$$

(11)

and the fact that a Kronecker power can be distributed over a matrix-vector product and vice versa. The use of (10) in (9) produces

$$(x + y)^{\otimes 2} = x^{\otimes 2} + (I_{n^2} + \Pi) (x \otimes y) + y^{\otimes 2}$$

(12)

where $I_{n^2}$ stands for the $n^2 \times n^2$ type identity matrix. This equality’s right hand side is now in the form where no $x$ including factor appears at the right hand side of $y$ containing factors. This form is somehow canonical for our purposes.

Now we can extend our investigation to the following equality

$$(x + y)^{\otimes 3} = x^{\otimes 3} + x^{\otimes 2} y + x \otimes y \otimes x + x \otimes y^{\otimes 2} + y \otimes x \otimes y + y \otimes x \otimes y + y^{\otimes 2} x + y^{\otimes 3}$$

(13)

where the third, fifth, sixth and seventh additive terms of the right hand side destroy the canonical form and therefore they should be replaced by their canonical counterparts. We can now write the following equalities

$$x \otimes y \otimes x = x \otimes (y \otimes x) = x \otimes (\Pi (x \otimes y)) = (I_n x) \otimes (\Pi (x \otimes y)) = (I_n \otimes \Pi) (x \otimes x \otimes y) = (I_n \otimes \Pi) (x^{\otimes 2} \otimes y)$$

(14)

where we have used the fact that the direct product and the matrix product is distributive on each other. That is, $(A_1 B_1) \otimes (A_2 B_2) = (A_1 A_2) \otimes (B_1 B_2)$ where the multiplicative compatibility amongst the arbitrary matrices $A_1, B_1, A_2, B_2$ should exist. This last set of equalities show the details about how the vector Kronecker factor $y$ can be pushed to rightmost position.

A similar detailed chain of manipulations can take us to the following results by skipping the presentations of intermediate details

$$y \otimes x \otimes y = (\Pi \otimes I_n) (I_n \otimes \Pi) (x \otimes x \otimes y)$$

(15)

$$y \otimes x \otimes y = (\Pi \otimes I_n) (I_n \otimes \Pi) (x \otimes y \otimes 2)$$

(16)

$$y \otimes x \otimes x = (I_n \otimes \Pi) (I_n \otimes I_n) (x \otimes y \otimes 2)$$

(17)

Now we can rewrite (13) as follows

$$(x + y)^{\otimes 3} = B_{3,0} x^{\otimes 3} + B_{3,1} (x \otimes ^2 y) + B_{3,2} (x \otimes x \otimes y) + B_{3,3} (y \otimes 3)$$

(18)

where $B_{3,0}, B_{3,1}, B_{3,2}, B_{3,3}$ denote $n^3 \times n^3$ type matrices which depend on neither $x$ nor $y$. We call these entities “Direct Power Binomial Coefficient Matrices” or, for brevity, just “Binomial Coefficients”. Binomial coefficients may depend on the permutation matrix $\Pi$ which is a universal entity independent of the vectors $x$ and $y$ ($\Pi$ acts as identity matrix on the direct products where $x$ and $y$ are proportional. This is because
of the $\Pi$’s spectrum. It involves eigenvalues identical 1. The explicit expressions of binomial coefficients are given below

\[
\begin{align*}
B_{3,0} &= I_n^3 \\
B_{3,1} &= I_n^3 + I_n \otimes \Pi + (\Pi \otimes I_n) (I_n \otimes \Pi) \\
B_{3,2} &= I_n^3 + \Pi \otimes I_n + (I_n \otimes \Pi) (\Pi \otimes I_n) \\
B_{3,3} &= I_n^3
\end{align*}
\]

There is an apparent symmetry amongst the elements as we have in usual expansion over scalars. As easily seen, all terms except the leading and ending ones are depending on $\Pi$.

The canonical formula given by (18) is not peculiar to the cubic direct power of a binary vector sum. We could write the following equalities for the cases where the Kronecker power is 2, 1, or, 0.

\[
\begin{align*}
(x + y)^{\otimes 2} &= B_{2,0} (x^{\otimes 2}) + B_{2,1} (x \otimes y) \\
&\quad + B_{2,2} (y^{\otimes 2}) \\
(x + y)^{\otimes 1} &= B_{1,0} (x^{\otimes 1}) + B_{1,1} (y^{\otimes 1}) \\
(x + y)^{\otimes 0} &= B_{0,0}
\end{align*}
\]

where

\[
\begin{align*}
B_{0,0} &= I_n^0 = 1 \\
B_{1,0} &= I_n^1 \\
B_{1,1} &= I_n^1 \\
B_{2,0} &= I_n^2 \\
B_{2,1} &= I_n^2 + \Pi \\
B_{2,2} &= I_n^2
\end{align*}
\]

An interesting and expected issue in all these formula is that the case where $x$ matches $y$ (or becomes parallel to $y$) converts the present binomial formula to the well-known scalar binomial formula.

All these results urge us to write the following formula for the canonical form of the binomial expansion of the sum of two vectors under direct (Kronecker) powering

\[
(x + y)^{\otimes j} = \sum_{k=0}^{j} B_{j,k} (x^{\otimes j-k} \otimes y^{\otimes k})
\]  

where the binomial coefficient matrices can be determined as functions of the permutation matrix $\Pi$, by tracing exactly the same route used above. Now this formula matches its scalar counterpart when $y$ becomes propotional to $x$. Indeed, in the case of $y = \alpha x$

the binomial coefficients turn out to have the following structure

\[
B_{j,k} = \alpha^k \binom{j}{k} I_n^j
\]

and (24) becomes

\[
(x + y)^{\otimes j} = (1 + \alpha) x^{\otimes j}.
\]

4 Recursion amongst the binomial coefficients

The determination of the binomial coefficients appearing in the above canonical expansion formula where the second vector containing factors come after the first vector involving terms can be realized for each given direct power separately as we have done above. However the cost of this operation rapidly increases as $j$ grows and brings the possibility of making mistakes in the derivations even if a symbolic programming or scripting tool is used via the computer. This operation involves the manipulations of the products of certain $x$ and $y$ direct powers. It is possible at least to by-pass these manipulations and construct recursions only amongst the binomial coefficients. The main focus on which the construction of the recursion is based is the following equality

\[
(x + y)^{\otimes j+1} = (x + y) \otimes (x + y)^{\otimes j}
\]

which can be used to construct the recursion we want. We can start to proceed by writing

\[
x \otimes (x + y)^{\otimes j} = \sum_{k=0}^{j} B_{j,k} (x^{\otimes j-k} \otimes y^{\otimes k})
\]

\[
= \sum_{k=0}^{j} (I_n \otimes B_{j,k}) (x^{\otimes j+1-k} \otimes y^{\otimes k}),
\]

where we have assumed vanishing $B_{j,j+1}$ because of its nonexistence, and continue through

\[
y \otimes (x + y)^{\otimes j} = \sum_{k=0}^{j} B_{j,k} (y^{\otimes j-k} \otimes x^{\otimes k})
\]

\[
= \sum_{k=0}^{j} (I_n \otimes B_{j,k}) (y^{\otimes j+1-k} \otimes x^{\otimes k}).
\]
which urges us to focus on the direct product $y \otimes x^{\otimes j-k} \otimes y^{\otimes k}$ whose leftmost vector factor $y$ should not be standing at that position not to destroy the canonicality. It is not hard to see that the only difference between the direct powers, $y \otimes x^{\otimes j}$ and $x^{\otimes j} \otimes y$ is the different positionings of the elements of same set. Hence, by passing from one to the other, the elements are in fact permuted. This means that these two direct product must be related to each other through the following equality

$$y \otimes x^{\otimes J} = \Pi_J (x^{\otimes J} \otimes y), \quad J = 0, 1, 2, \ldots$$  (30)

where $\Pi_J$ stands for a permutation matrix of $n^{J+1} \times n^{J+1}$ type. It is quite natural to expect that this matrix depends on the former $n \times n$ permutation matrix $\Pi$. To find the explicit structure of the permutation matrix $\Pi_J$ it is possible to use the elementwise comparison of the direct products, $y \otimes x^{\otimes J}$ and $x^{\otimes J} \otimes y$. However, this action is quite tedious and hard to be generalized.

One of the other but efficient ways to determine the permutation matrices, $\Pi_J$, is to construct a recursion between two consecutive members of this sequence, that is, $\Pi_J$ and $\Pi_{J+1}$. To this end we can write first

$$y \otimes x^{\otimes J+1} = (y \otimes x^{\otimes J}) \otimes x$$  (31)

and then use (30) in the left hand side factor of the right hand side between the parentheses. This enables us to write the following equalities

$$\left( y \otimes x^{\otimes J} \right) \otimes x = (\Pi_J (x^{\otimes J} \otimes y)) \otimes x$$
$$= (\Pi_J (x^{\otimes J} \otimes y)) \otimes (I_n x)$$
$$= (\Pi_J \otimes I_n) (x^{\otimes J} \otimes y \otimes x)$$  (32)

$$x^{\otimes J} \otimes y \otimes x = x^{\otimes J} \otimes (\Pi (x \otimes y))$$
$$= (I_n, x^{\otimes J}) \otimes (\Pi (x \otimes y))$$
$$= (I_n, x^{\otimes J}) \otimes (x^{\otimes J+1} \otimes y)$$  (33)

which can be combined and used in (31) to give

$$y \otimes x^{\otimes J+1} = (\Pi_J \otimes I_n) (I_n, x^{\otimes J+1} \otimes y)$$  (34)

whose comparison with (30) takes us to the following recursion

$$\Pi_{J+1} = (\Pi_J \otimes I_n) (I_n, \Pi), \quad J = 0, 1, 2, \ldots$$  (35)

which can be initiated by $\Pi_0 = I_n$. This formula replaces (29) by

$$y \otimes (x + y)^{\otimes j} = \sum_{k=0}^{j+1} (I_n \otimes B_{j,k-1})$$
$$\times (\Pi_{j+1-k} \otimes I_n^{k-1}) \left( x^{\otimes j+1-k} \otimes y^{\otimes k} \right)$$  (36)

where we have changed all appearances of $k$ by $k - 1$ and used the indexed entity convention such that a negative indexed entity is taken as vanishing. The combination of this equality with (28) produces the following recursive relation between two consecutive sets of binomial coefficients

$$B_{j+1,k} = I_n \otimes B_{j,k}$$
$$+ (I_n \otimes B_{j,k-1}) (\Pi_{j+1-k} \otimes I_n^{k-1})$$
$$j = 0, 1, 2, \ldots \quad k = 0, 1, \ldots, j + 1$$  (37)

where the $B$ matrices whose second index is greater than the first, and, any negative indexed entity is conventionally assumed to be vanishing. The last formula above is the ultimate form of the recursive relation we desired to construct between the set of $(j + 1)$th direct power binomial coefficient matrices and the set of $j$th direct power binomial coefficient matrices.

5 Concluding Remarks

The basic goal has been solely the construction of a formula which extends the very well-known binomial expansion for the nonnegative integer powers of two scalar sums to the case where the Kronecker power of two vector sums is under consideration. The goal is achieved and the resulting formula will be used in our future applications especially in the framework of the Probabilistic Evolution Approach (PEA).

References:


