Kronecker Product and Vectorization

Frank R. Kschischang
Department of Electrical & Computer Engineering
University of Toronto

January 16, 2022

1 Notation

The field of complex numbers is denoted as $\mathbb{C}$. The complex conjugate of $z \in \mathbb{C}$ is denoted as z^*. The set of natural numbers is $\mathbb{N} = \{0, 1, 2, \ldots\}$. For any positive integer a, let $[a] = \{1, 2, \ldots, a\}$.

Throughout this note, matrices are, for some positive integers a and b and some field F, elements of $F^{a\times b}$, the set of matrices with a rows and b columns having entries from F. For any $i \in [a]$ and any $j \in [b]$, the entry in row i and column j of $M \in F^{a\times b}$ is denoted as $M_{i,j}$, i.e., $M = [M_{i,j} \colon i \in [a], j \in [b]]$.

The transpose of M, denoted M^T, is the matrix $M^T = [M_{j,i} \colon j \in [b], i \in [a]]$ obtained by writing the rows of M as the columns of M^T. When $F = \mathbb{C}$, the conjugate transpose (or Hermitian transpose) of M, denote M^H, is the matrix $M^T = [M^*_{j,i} \colon j \in [b], i \in [a]]$ obtained by writing the complex conjugates of the rows of M as the columns of M^H. If $M \in F^{a\times b}$, then $M^T \in F^{b\times a}$; likewise if $M \in \mathbb{C}^{a\times b}$, then $M^H \in \mathbb{C}^{b\times a}$.

The ith row of $M \in F^{a\times b}$, i.e., the matrix $[M_{i,1}, \ldots, M_{i,b}]$, is denoted as $M_{i,:}$. Clearly $M_{i,:} \in F^{1\times b}$. The jth column of M, i.e., the matrix $[M_{1,j}, \ldots, M_{a,j}]^T$, is denoted as $M_{:,j}$. Clearly $M_{:,j} \in F^{a\times 1}$. For any $i \in [a]$ we have that $(M_{i,:})^T$, the transpose of the ith row of M, is equal to $(M^T)_{:,i}$, the ith column of M^T. Similarly $(M_{:,j})^T = (M^T)_{j,:}$ for any $j \in [b]$. The notation $M^T_{i,:}$ must be avoided, since it is not clear whether this denotes the ith row of M^T or the transpose of the ith row of M.

The $a \times a$ identity matrix is denoted as I_a.

For any positive integer n, denote by F^n the vector space of n-tuples over the field F. We may identify F^n either with $F^{1\times n}$ (the "vectors are rows" convention) or with $F^{n\times 1}$ (the "vectors are columns" convention). The former convention is common in coding theory, while the latter convention is common in most other disciplines. Under the vectors are rows convention, we will assume that the vector $(v_1, \ldots, v_n) \in F^n$ is equal to (or is another notation for) the matrix $[v_1 \ \cdots \ v_n] \in F^{1\times n}$. Under the vectors are columns convention, we will assume that the vector $(v_1, \ldots, v_n) \in F^n$ is equal to (or is another notation for) the matrix $[v_1 \ \cdots \ v_n]^T \in F^{n\times 1}$. The two conventions agree when $n = 1$; i.e., we may identify elements of F^1 (scalars) with 1×1 matrices.

2 Kronecker Product

The *Kronecker* product $M \otimes N$ of matrices $M = [M_{i,j}] \in F^{a\times b}$ and $N \in F^{c\times d}$ (in that order) is the $ac \times bd$ matrix given (in block form) as

$$M \otimes N = \begin{bmatrix} M_{1,1}N & M_{1,2}N & \cdots & M_{1,b}N \\ M_{2,1}N & M_{2,2}N & \cdots & M_{2,b}N \\ \vdots & \vdots & \ddots & \vdots \\ M_{a,1}N & M_{a,2}N & \cdots & M_{a,b}N \end{bmatrix}.$$

For $a \in F$, for $[a] \otimes M$ we write $a \otimes M$ and for $M \otimes [a]$ we write $M \otimes a$. The following proposition follows immediately from the definition of Kronecker product.

Proposition 1. *Let A, B, and C be matrices with entries from the field F, let v and w be vectors with components in F, and let $a, b \in F$ be scalars. In the following, expressions involving a matrix sum, a matrix product, or a matrix inverse are well defined only when the matrices are conformable for addition or multiplication, or are invertible, respectively. The following properties then hold:*

1. *(associativity)* $A \otimes (B \otimes C) = (A \otimes B) \otimes C$,
2. *(the distributive laws)* $A \otimes (B + C) = (A \otimes B) + (A \otimes C)$ *and* $(A + B) \otimes C = (A \otimes C) + (B \otimes C)$,
3. *(interaction with scalars)* $a \otimes A = A \otimes a = aA$, $aA \otimes bB = ab(A \otimes B)$,
4. *(mixed product) for conforming matrices,* $(A \otimes B)(C \otimes D) = AC \otimes BD$,
5. *(transposition)* $(A \otimes B)^T = A^T \otimes B^T$, $(A \otimes B)^H = A^H \otimes B^H$,
6. *(outer products) when vectors are rows,* $v^T w = v^T \otimes w = w \otimes v^T$*; when vectors are columns,* $vw^T = v \otimes w^T = w^T \otimes v$,
7. *(partitioned matrices)* $[A \mid B] \otimes C = [A \otimes C \mid B \otimes C]$*, but* $A \otimes [B \mid C] \neq [A \otimes B \mid A \otimes C]$,
8. *(invertible matrices)* $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$,
9. *(determinants) if* $A \in F^{m\times m}$ *and* $B \in F^{n\times n}$ *then* $\det(A \otimes B) = (\det(A))^n(\det(B))^m$,
10. *(trace)* $\operatorname{tr}(A \otimes B) = \operatorname{tr}(A) \operatorname{tr}(B)$,
11. *(rank)* $\operatorname{rank}(A \otimes B) = \operatorname{rank}(A) \operatorname{rank}(B)$.

A consequence of the mixed product property is obtained by taking B and C to be identity matrices. Let matrix A have a columns, let $B = I_a$, let matrix D have d rows, and let C $= I_d$. Then

$$A \otimes D = AI_a \otimes I_d D = (A \otimes I_a)(I_d \otimes D).$$

Similarly, if B has b rows and C has c columns,

$$B \otimes C = I_b B \otimes C I_c = (I_b \otimes B)(C \otimes I_c).$$

3 Vectorization

The vectorization map $\mathrm{vec}\colon F^{a \times b} \to F^{ab}$ takes a matrix $M = [\, M_{i,j} \colon i \in [a],\ j \in [b]\,]$ to a vector. Depending on the convention followed—vectors are rows or vectors are columns—it is natural to define two different maps.

When vectors are rows:

$$\begin{aligned}\mathrm{vec}(M) &= (M_{1,:}, M_{2,:}, \dots, M_{a,:}) \\ &= (\, M_{1,1}, \dots, M_{1,b}, \\ &\qquad M_{2,1}, \dots, M_{2,b}, \\ &\qquad \dots, M_{a,1}, \dots, M_{a,b}\,),\end{aligned}$$

i.e., the rows of the matrix are appended (in order) to form the vector. For example,

$$\mathrm{vec}\begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} = (a, b, c, d, e, f).$$

When vectors are columns:

$$\begin{aligned}\mathrm{vec}(M) &= (M_{:,1}, M_{:,2}, \dots, M_{:,a}) \\ &= (\, M_{1,1}, \dots, M_{a,1}, \\ &\qquad M_{1,2}, \dots, M_{a,2}, \\ &\qquad \dots, M_{1,b}, \dots, M_{a,b}\,),\end{aligned}$$

i.e., the columns of the matrix are appended (in order) to form the vector. For example,

$$\mathrm{vec}\begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix} = (a, d, b, e, c, f).$$

Note that vectorization is a linear transformation, so in particular $\mathrm{vec}(M + N) = \mathrm{vec}(M) + \mathrm{vec}(N)$ for any matrices M and N commensurate for addition.

Proposition 2. *Let $L \in F^{a \times b}$, $M \in F^{b \times c}$, and $N \in F^{c \times d}$ be three matrices commensurate for multiplication in that order. Then,*

when vectors are rows:

$$\mathrm{vec}(LMN) = \mathrm{vec}(M)(L^T \otimes N).$$

Consequently,

$$\begin{aligned}\mathrm{vec}(MN) &= \mathrm{vec}(M)(I_b \otimes N), \\ \mathrm{vec}(LM) &= \mathrm{vec}(M)(L^T \otimes I_c).\end{aligned}$$

when vectors are columns:

$$\mathrm{vec}(LMN) = (N^T \otimes L)\,\mathrm{vec}(M).$$

Consequently,

$$\begin{aligned}\mathrm{vec}(MN) &= (N^T \otimes I_b)\,\mathrm{vec}(M), \\ \mathrm{vec}(LM) &= (I_c \otimes L)\,\mathrm{vec}(M).\end{aligned}$$

Proof. For $i \in [a]$, the ith row of LMN is given as $(LMN)_{i,:} = L_{i,1}M_{1,:}N + \cdots + L_{i,b}M_{b,:}N$, while for $j \in [d]$, the jth column of LMN is given as $(LMN)_{:,j} = LM_{:,1}N_{1,j} + \cdots + LM_{:,c}N_{c,j}$.

Thus,

when vectors are rows:

$$\begin{aligned}
\operatorname{vec}(LMN) &= (L_{1,1}M_{1,:}N + \cdots + L_{1,b}M_{b,:}N, \\
&\quad \ldots, L_{a,1}M_{1,:}N + \cdots + L_{a,b}M_{b,:}N) \\
&= [M_{1,:}, \ldots, M_{b,:}] \begin{bmatrix} L_{1,1}N & \cdots & L_{a,1}N \\ \vdots & \ddots & \vdots \\ L_{1,b}N & \cdots & L_{a,b}N \end{bmatrix} \\
&= \operatorname{vec}(M)(L^T \otimes N).
\end{aligned}$$

when vectors are columns:

$$\begin{aligned}
\operatorname{vec}(LMN) &= (LM_{:,1}N_{1,1} + \cdots + LM_{:,c}N_{c,1}, \\
&\quad \ldots, LM_{:,1}N_{1,d} + \cdots + LM_{:,c}N_{c,d}) \\
&= \begin{bmatrix} LN_{1,1} & \cdots & LN_{c,1} \\ \vdots & \ddots & \vdots \\ LN_{1,d} & \cdots & LN_{c,d} \end{bmatrix} \begin{bmatrix} M_{:,1} \\ \vdots \\ M_{:,c} \end{bmatrix} \\
&= (N^T \otimes L) \operatorname{vec}(M).
\end{aligned}$$

The two consequences follow by setting $L = I_a$ and $N = I_c$, respectively. □

The vectorization of a matrix and that of its transpose are related by a permutation. Let $K^{(a,b)}$ denote the $ab \times ab$ *commutation matrix*, that satisfies,

when vectors are rows:

$$\operatorname{vec}(M)K^{(a,b)} = \operatorname{vec}(M^T).$$

when vectors are columns:

$$K^{(a,b)} \operatorname{vec}(M) = \operatorname{vec}(M^T).$$

Note that $(K^{(a,b)})^T = K^{(b,a)}$, and furthermore, $K^{(a,b)}K^{(b,a)} = I_{ab}$, i.e., the inverse of $K^{(a,b)}$ is its own transpose.

For example,

$$K^{(2,3)} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

and indeed $[a, b, c, d, e, f]K^{(2,3)} = [a, d, b, e, c, f]$ while $K^{(2,3)}[a, d, b, e, c, f]^T = [a, b, c, d, e, f]^T$.

The terminology "commutation matrix" arises from the following proposition.

Proposition 3. *For any matrices $L \in F^{a\times b}$ and $N \in F^{c\times d}$,*

$$K^{(c,a)}(L \otimes N)K^{(b,d)} = N \otimes L$$

Proof. Choose an arbitrary $M \in F^{c\times a}$, and note by Proposition 2—using the vectors are rows convention—that

$$\begin{aligned}
\operatorname{vec}(M)K^{(c,a)}(L \otimes N) = \operatorname{vec}(M^T)(L \otimes N) = \operatorname{vec}(L^T M^T N) &= \operatorname{vec}(N^T ML)K^{(d,b)} \\
= \operatorname{vec}(M)(N \otimes L)K^{(d,b)},
\end{aligned}$$

which shows that $K^{(c,a)}(L\otimes N) = (N\otimes L)K^{(d,b)}$ since M was chosen arbitrarily. Multiplying on the right by $K^{(b,d)}$ recovers the statement of the proposition. □

A special case arises when $b = d = 1$, in which case $L \in F^{a\times 1}$ and $N \in F^{c\times 1}$ are column vectors and $K^{(c,a)}(L \otimes N) = N \otimes L$.