

8.4 The Kronecker Product

So far the unknown in a matrix equation has been a vector: $\mathbf{Ax} = \mathbf{b}$. Many natural questions have a matrix as the unknown instead. Which matrices commute with a given $\mathbf{A}$? Which $\mathbf{X}$ solve $\mathbf{AX} + \mathbf{XB} = \mathbf{C}$? Such an equation is still linear in the entries of $\mathbf{X}$, so in principle it is an ordinary linear system. The trouble is bookkeeping: we need to know what the coefficient matrix of that system is. This section builds it from $\mathbf{A}$ and $\mathbf{B}$ with one operation, the Kronecker product, and turns matrix equations into systems we already know how to solve.

8.4.1 From a matrix equation to a linear system

Take $\mathbf{A}, \mathbf{B} \in M_2(F)$ and an unknown $\mathbf{X} = (x_{ij}) \in M_2(F)$, and look at the equation $\mathbf{AXB} = \mathbf{C}$. By Definition 0.9.9, applied twice,

$$(\mathbf{AXB})_{rs} = \sum_{i=1}^2 \sum_{j=1}^2 a_{ri} x_{ij} b_{js}.$$

So each entry of $\mathbf{AXB}$ is a linear combination of the four unknowns $x_{11}, x_{21}, x_{12}, x_{22}$, and the coefficient of x_{ij} in entry (r, s) is the product $b_{js}a_{ri}$: **an entry of $\mathbf{B}$ times an entry of $\mathbf{A}$** . Writing the four equations as a 4×4 system, every coefficient is such a product. For instance, if we list the unknowns column by column as $(x_{11}, x_{21}, x_{12}, x_{22})$ and the equations in the same order, the first row of the coefficient matrix is

$$(b_{11}a_{11} \quad b_{11}a_{12} \quad b_{21}a_{11} \quad b_{21}a_{12}).$$

The left half of this row is b_{11} times row 1 of $\mathbf{A}$, and the right half is b_{21} times row 1 of $\mathbf{A}$.

That is a recurring pattern: a big matrix whose entries are all the products of the entries of two small matrices, arranged in blocks. It deserves a name.

8.4.2 The Kronecker product

The Kronecker product of $\mathbf{A}$ and $\mathbf{B}$ is the block matrix obtained by replacing each entry a_{ij} of $\mathbf{A}$ by the block $a_{ij}\mathbf{B}$.

Definition 8.4.1. Kronecker product

Let $\mathbf{A} = (a_{ij}) \in M_{m \times n}(F)$ and $\mathbf{B} = (b_{kl}) \in M_{p \times q}(F)$, of **any** sizes. The **Kronecker product** of $\mathbf{A}$ and $\mathbf{B}$ is the $mp \times nq$ block matrix

$$\mathbf{A} \otimes \mathbf{B} := \begin{pmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \cdots & a_{1n}\mathbf{B} \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & \cdots & a_{2n}\mathbf{B} \\ \vdots & \vdots & & \vdots \\ a_{m1}\mathbf{B} & a_{m2}\mathbf{B} & \cdots & a_{mn}\mathbf{B} \end{pmatrix} \in M_{mp \times nq}(F),$$

with m rows and n columns of blocks, each block of size $p \times q$ (Definition 8.1.1).

In words: the outer layout is the layout of $\mathbf{A}$, and the inner pattern of every block is $\mathbf{B}$. The sizes multiply, rows with rows and columns with columns. No condition on the sizes is needed, unlike for an ordinary product.

To work with single entries we need to know where each product $a_{ij}b_{kl}$ sits. Block (i, j) occupies rows $(i-1)p+1, \dots, ip$ and columns $(j-1)q+1, \dots, jq$. Inside it, entry (k, l) is $a_{ij}b_{kl}$. So

$$(\mathbf{A} \otimes \mathbf{B})_{(i-1)p+k, (j-1)q+l} = a_{ij}b_{kl} \quad (1 \leq i \leq m, 1 \leq j \leq n, 1 \leq k \leq p, 1 \leq l \leq q). \quad (*)$$

Every entry is reached exactly once. Division with remainder writes each row index $r \in \{1, \dots, mp\}$ uniquely as $r = (i-1)p + k$ with $1 \leq i \leq m$ and $1 \leq k \leq p$ (divide $r-1$ by p), and likewise for columns. So $(*)$ is a complete description of $\mathbf{A} \otimes \mathbf{B}$, and it is the formula we use in proofs. In words, the rows of $\mathbf{A} \otimes \mathbf{B}$ are labeled by pairs (row of $\mathbf{A}$, row of $\mathbf{B}$) in dictionary order, and the columns by pairs (column of $\mathbf{A}$, column of $\mathbf{B}$) in dictionary order.

Examples.

1. **A 2×2 example.** With $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$,

$$\mathbf{A} \otimes \mathbf{B} = \begin{pmatrix} 1\mathbf{B} & 2\mathbf{B} \\ 3\mathbf{B} & 4\mathbf{B} \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 2 \\ 1 & 0 & 2 & 0 \\ 0 & 3 & 0 & 4 \\ 3 & 0 & 4 & 0 \end{pmatrix}.$$

Check one entry against $(*)$: row 3 = $(2-1) \cdot 2 + 1$ and column 2 = $(1-1) \cdot 2 + 2$ give $a_{21}b_{12} = 3 \cdot 1 = 3$, as displayed.

2. **Vectors.** For column vectors $\mathbf{u} \in F^m$ and $\mathbf{v} \in F^p$ (so $n = q = 1$), $\mathbf{u} \otimes \mathbf{v} \in F^{mp}$ lists all products $u_i v_k$ in dictionary order. In particular, for standard basis vectors, $\mathbf{e}_i \otimes \mathbf{e}_k = \mathbf{e}_{(i-1)p+k} \in F^{mp}$: the only nonzero product is $1 \cdot 1$ in position $(i-1)p+k$. For a column $\mathbf{u} \in F^m$ and a row $\mathbf{v}^\top$, $\mathbf{u} \otimes \mathbf{v}^\top$ is the $m \times p$ matrix with entries $u_i v_k$, which is the outer product $\mathbf{u}\mathbf{v}^\top$ of Corollary 8.1.3.
3. **Identity blocks.** $\mathbf{I}_n \otimes \mathbf{B}$ has $\mathbf{B}$ in each diagonal block and zero blocks elsewhere, so it is the block diagonal matrix $\mathbf{B} \oplus \dots \oplus \mathbf{B}$ with n copies of $\mathbf{B}$ (Definition 8.1.2). By contrast, $\mathbf{A} \otimes \mathbf{I}_m$ replaces each entry a_{ij} by the $m \times m$ block $a_{ij}\mathbf{I}_m$. Taking both, $\mathbf{I}_n \otimes \mathbf{I}_m = \mathbf{I}_{nm}$: by $(*)$ its entries are $\delta_{ij}\delta_{kl}$, which is 1 exactly when the row and column labels (i, k) and (j, l) agree.
4. **The degenerate case 1×1 .** If $\mathbf{A} = (a)$ is 1×1 , then $\mathbf{A} \otimes \mathbf{B} = a\mathbf{B}$; if $\mathbf{B} = (b)$, then $\mathbf{A} \otimes \mathbf{B} = b\mathbf{A}$. So a 1×1 factor behaves like a scalar, and $(1) \otimes \mathbf{B} = \mathbf{B}$. This case is worth keeping in mind: it shows that the Kronecker product extends scalar multiplication, which is why it is bilinear below.

Nonexample by minimal change. Swap the two factors of Example 1:

$$\mathbf{B} \otimes \mathbf{A} = \begin{pmatrix} 0\mathbf{A} & 1\mathbf{A} \\ 1\mathbf{A} & 0\mathbf{A} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 & 2 \\ 0 & 0 & 3 & 4 \\ 1 & 2 & 0 & 0 \\ 3 & 4 & 0 & 0 \end{pmatrix}.$$

It has the same size and the same multiset of entries, since both list all products $a_{ij}b_{kl}$. What fails is the position formula $(*)$: in $\mathbf{B} \otimes \mathbf{A}$ the **outer** layout comes from $\mathbf{B}$, so the products sit in different places. For instance, the $(1, 2)$ -entry is $a_{11}b_{12} = 1$ in $\mathbf{A} \otimes \mathbf{B}$ but $b_{11}a_{12} = 0$ in $\mathbf{B} \otimes \mathbf{A}$.

Why this definition. The dictionary order in $(*)$ is a convention, and the other order would simply give $\mathbf{B} \otimes \mathbf{A}$. We fix it because it makes the block picture true: the Kronecker product is literally “ $\mathbf{A}$ with entries replaced by blocks”, so block multiplication (Theorem 8.1.1) applies to it directly. The symbol $\otimes$ anticipates the tensor product of Chapter 15, of which this is the matrix version.

Warning. The Kronecker product is not commutative, and it is not the ordinary product. The two matrices in Example 1 and the nonexample differ, so $\mathbf{A} \otimes \mathbf{B} \neq \mathbf{B} \otimes \mathbf{A}$ in general, even for square matrices of the same size. Also do not confuse $\mathbf{A} \otimes \mathbf{B}$ with $\mathbf{AB}$: for $\mathbf{A}, \mathbf{B} \in M_2(F)$, $\mathbf{AB}$ is 2×2 while $\mathbf{A} \otimes \mathbf{B}$ is 4×4 , and $\mathbf{A} \otimes \mathbf{B}$ is defined for all sizes. We will see below that $\mathbf{A} \otimes \mathbf{B}$ and $\mathbf{B} \otimes \mathbf{A}$ are always **similar** when $\mathbf{A}$ and $\mathbf{B}$ are square, so they share rank, trace, and determinant, but they are not equal.

8.4.3 Algebraic properties

The Kronecker product interacts well with every operation we have. The most useful rule is the one connecting it to ordinary multiplication: multiply the two left factors together and the two right factors together.

Theorem 8.4.1. Properties of the Kronecker Product

Let $c \in F$. Whenever the sizes make each side defined:

- (a) **(Bilinearity)** $(\mathbf{A} + \mathbf{A}') \otimes \mathbf{B} = \mathbf{A} \otimes \mathbf{B} + \mathbf{A}' \otimes \mathbf{B}$, $\mathbf{A} \otimes (\mathbf{B} + \mathbf{B}') = \mathbf{A} \otimes \mathbf{B} + \mathbf{A} \otimes \mathbf{B}'$, and $(c\mathbf{A}) \otimes \mathbf{B} = c(\mathbf{A} \otimes \mathbf{B}) = \mathbf{A} \otimes (c\mathbf{B})$.
- (b) **(Associativity)** $(\mathbf{A} \otimes \mathbf{B}) \otimes \mathbf{C} = \mathbf{A} \otimes (\mathbf{B} \otimes \mathbf{C})$ for all $\mathbf{A} \in M_{m \times n}(F)$, $\mathbf{B} \in M_{p \times q}(F)$, $\mathbf{C} \in M_{r \times s}(F)$.
- (c) **(Mixed product rule)** For $\mathbf{A} \in M_{m \times n}(F)$, $\mathbf{C} \in M_{n \times t}(F)$, $\mathbf{B} \in M_{p \times q}(F)$, and $\mathbf{D} \in M_{q \times u}(F)$,

$$(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = (\mathbf{AC}) \otimes (\mathbf{BD}).$$
- (d) **(Transpose)** $(\mathbf{A} \otimes \mathbf{B})^\top = \mathbf{A}^\top \otimes \mathbf{B}^\top$.
- (e) **(Inverse)** If $\mathbf{A} \in M_n(F)$ and $\mathbf{B} \in M_m(F)$ are invertible, then $\mathbf{A} \otimes \mathbf{B}$ is invertible and $(\mathbf{A} \otimes \mathbf{B})^{-1} = \mathbf{A}^{-1} \otimes \mathbf{B}^{-1}$.

Idea. Parts (a), (b), and (d) are entry checks with (*). For (c), both factors on the left are block matrices, and the columns of blocks of $\mathbf{A} \otimes \mathbf{B}$ (n of them, each q wide) match the rows of blocks of $\mathbf{C} \otimes \mathbf{D}$ (n of them, each q tall). So we may multiply blockwise, and each block of the product is a sum $\sum_j (a_{ij}\mathbf{B})(c_{jk}\mathbf{D})$ in which the scalars come out. Part (e) is (c) with $\mathbf{C} = \mathbf{A}^{-1}$, $\mathbf{D} = \mathbf{B}^{-1}$.

Proof.

- (a) By (*), the entry of $(\mathbf{A} + \mathbf{A}') \otimes \mathbf{B}$ in position $((i-1)p+k, (j-1)q+l)$ is $(a_{ij} + a'_{ij})b_{kl} = a_{ij}b_{kl} + a'_{ij}b_{kl}$, which is the entry of $\mathbf{A} \otimes \mathbf{B} + \mathbf{A}' \otimes \mathbf{B}$ in the same position. The other identities are checked the same way, using $(ca_{ij})b_{kl} = c(a_{ij}b_{kl}) = a_{ij}(cb_{kl})$.
- (b) Both sides are $mpr \times nqs$. Let $1 \leq i \leq m$, $1 \leq k \leq p$, $1 \leq t \leq r$, and similarly j, l, u for columns. Applying (*) twice, the entry of $(\mathbf{A} \otimes \mathbf{B}) \otimes \mathbf{C}$ in row $((i-1)p+k-1)r+t$ and column $((j-1)q+l-1)s+u$ is $(\mathbf{A} \otimes \mathbf{B})_{(i-1)p+k, (j-1)q+l} c_{tu} = a_{ij}b_{kl}c_{tu}$. Likewise the entry of $\mathbf{A} \otimes (\mathbf{B} \otimes \mathbf{C})$ in row $(i-1)pr + (k-1)r+t$ and column $(j-1)qs + (l-1)s+u$ is $a_{ij}(\mathbf{B} \otimes \mathbf{C})_{(k-1)r+t, (l-1)s+u} = a_{ij}b_{kl}c_{tu}$. Since $((i-1)p+k-1)r+t = (i-1)pr + (k-1)r+t$, and similarly for columns, the two matrices have the same entry in every position; every position arises from exactly one choice of indices, by division with remainder. Hence

they are equal.

- (c) Partition $\mathbf{A} \otimes \mathbf{B}$ into m rows and n columns of blocks as in Definition 8.4.1, each block $p \times q$, and $\mathbf{C} \otimes \mathbf{D}$ into n rows and t columns of blocks, each $q \times u$. The column partition of the first matrix (n parts of size q) equals the row partition of the second, so the partitions are conformable, and by Theorem 8.1.1 the (i, k) block of the product is

$$\sum_{j=1}^n (a_{ij} \mathbf{B})(c_{jk} \mathbf{D}) = \sum_{j=1}^n a_{ij} c_{jk} \mathbf{B} \mathbf{D} = (\mathbf{A} \mathbf{C})_{ik} \mathbf{B} \mathbf{D},$$

where the first equality uses Theorem 0.9.4 (4) and the second uses Definition 0.9.9. This is the (i, k) block of $(\mathbf{A} \mathbf{C}) \otimes (\mathbf{B} \mathbf{D})$, a matrix partitioned into m rows and t columns of $p \times u$ blocks. So the two matrices agree block by block, hence entry by entry.

- (d) $(\mathbf{A} \otimes \mathbf{B})^\top$ is $nq \times mp$. By Definition 0.9.10 and (*), its entry in row $(j-1)q + l$ and column $(i-1)p + k$ is $(\mathbf{A} \otimes \mathbf{B})_{(i-1)p+k, (j-1)q+l} = a_{ij} b_{kl} = (\mathbf{A}^\top)_{ji} (\mathbf{B}^\top)_{lk}$, which by (*) is the entry of $\mathbf{A}^\top \otimes \mathbf{B}^\top$ in the same position.
- (e) By (c) and Example 3, $(\mathbf{A} \otimes \mathbf{B})(\mathbf{A}^{-1} \otimes \mathbf{B}^{-1}) = (\mathbf{A} \mathbf{A}^{-1}) \otimes (\mathbf{B} \mathbf{B}^{-1}) = \mathbf{I}_n \otimes \mathbf{I}_m = \mathbf{I}_{nm}$, and similarly $(\mathbf{A}^{-1} \otimes \mathbf{B}^{-1})(\mathbf{A} \otimes \mathbf{B}) = \mathbf{I}_{nm}$. By Definition 0.9.13, $\mathbf{A} \otimes \mathbf{B}$ is invertible with inverse $\mathbf{A}^{-1} \otimes \mathbf{B}^{-1}$. ■

The mixed product rule is the engine of everything that follows. One immediate consequence splits a Kronecker product of square matrices into two simpler factors: for $\mathbf{A} \in M_n(F)$ and $\mathbf{B} \in M_m(F)$,

$$\mathbf{A} \otimes \mathbf{B} = (\mathbf{A} \mathbf{I}_n) \otimes (\mathbf{I}_m \mathbf{B}) = (\mathbf{A} \otimes \mathbf{I}_m)(\mathbf{I}_n \otimes \mathbf{B}), \quad (**)$$

and in the same way $\mathbf{A} \otimes \mathbf{B} = (\mathbf{I}_n \otimes \mathbf{B})(\mathbf{A} \otimes \mathbf{I}_m)$. So $\mathbf{A} \otimes \mathbf{I}_m$ and $\mathbf{I}_n \otimes \mathbf{B}$ commute, even though $\mathbf{A}$ and $\mathbf{B}$ have nothing to do with each other.

For example, with $\mathbf{A} = \begin{pmatrix} 1 & -1 \\ 0 & 2 \end{pmatrix}$, $\mathbf{B} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$, $\mathbf{C} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$, and $\mathbf{D} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, multiplying two 4×4 matrices directly gives

$$(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = \begin{pmatrix} 2 & 1 & -2 & -1 \\ 1 & 1 & -1 & -1 \\ 0 & 0 & 4 & 2 \\ 0 & 0 & 2 & 2 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & -1 & -2 \\ 0 & 0 & -1 & -1 \\ 2 & 4 & 2 & 4 \\ 2 & 2 & 2 & 2 \end{pmatrix},$$

while the rule needs only two 2×2 products, $\mathbf{A} \mathbf{C} = \begin{pmatrix} 0 & -1 \\ 2 & 2 \end{pmatrix}$ and $\mathbf{B} \mathbf{D} = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}$, whose Kronecker product is the same matrix.

8.4.4 Swapping the factors

The warning said that $\mathbf{A} \otimes \mathbf{B}$ and $\mathbf{B} \otimes \mathbf{A}$ differ. They differ only in the order in which rows and columns are labeled: $\mathbf{A} \otimes \mathbf{B}$ uses (index of $\mathbf{A}$, index of $\mathbf{B}$) in dictionary order, $\mathbf{B} \otimes \mathbf{A}$ the reverse order. Relabeling rows and columns by the same permutation is conjugation by a permutation matrix, and that is what the next result makes precise.

First, a fact about permutation matrices. Recall that for $\sigma \in S_N$, the permutation matrix $\mathbf{P}_\sigma$ has j -th column $\mathbf{e}_{\sigma(j)}$, and $\mathbf{P}_\sigma^{-1} = \mathbf{P}_\sigma^\top$ (Definition 3.4.2, Lemma 3.4.7). For every $\mathbf{M} \in M_N(F)$,

$$(\mathbf{P}_\sigma^\top \mathbf{M} \mathbf{P}_\sigma)_{rs} = \mathbf{M}_{\sigma(r) \sigma(s)} \quad (1 \leq r, s \leq N). \quad (\dagger)$$

Indeed, by Theorem 0.9.2 column s of $\mathbf{M}\mathbf{P}_\sigma$ is $\mathbf{M}\mathbf{e}_{\sigma(s)}$, which is column $\sigma(s)$ of $\mathbf{M}$. Row r of $\mathbf{P}_\sigma^\top$ is $\mathbf{e}_{\sigma(r)}^\top$, so by the same theorem row r of $\mathbf{P}_\sigma^\top \mathbf{L}$ is row $\sigma(r)$ of $\mathbf{L}$, for any $\mathbf{L} \in M_N(F)$. Combining the two gives $(\dagger)$.

Proposition 8.4.2. Swapping the Factors Is a Permutation Similarity

Let $n, m \geq 1$. There is a permutation matrix $\mathbf{P} \in M_{nm}(F)$, depending only on n and m , such that

$$\mathbf{B} \otimes \mathbf{A} = \mathbf{P}^{-1}(\mathbf{A} \otimes \mathbf{B})\mathbf{P} \quad \text{for all } \mathbf{A} \in M_n(F), \mathbf{B} \in M_m(F).$$

In particular $\mathbf{A} \otimes \mathbf{B} \sim \mathbf{B} \otimes \mathbf{A}$, and $\mathbf{I}_m \otimes \mathbf{A} = \mathbf{P}^{-1}(\mathbf{A} \otimes \mathbf{I}_m)\mathbf{P}$.

Proof. Every $r \in \{1, \dots, nm\}$ can be written uniquely as $r = (k-1)n + i$ with $1 \leq k \leq m$ and $1 \leq i \leq n$. Define $\sigma(r) := (i-1)m + k$. As (k, i) runs over all pairs, so does (i, k) , and each value $(i-1)m + k$ is taken by exactly one pair; hence σ is a bijection of $\{1, \dots, nm\}$, that is, $\sigma \in S_{nm}$. Put $\mathbf{P} = \mathbf{P}_\sigma$.

Let $r = (k-1)n + i$ and $s = (l-1)n + j$. By $(*)$ applied to $\mathbf{B} \otimes \mathbf{A}$, $(\mathbf{B} \otimes \mathbf{A})_{rs} = b_{kl}a_{ij}$. By $(*)$ applied to $\mathbf{A} \otimes \mathbf{B}$,

$$(\mathbf{A} \otimes \mathbf{B})_{\sigma(r)\sigma(s)} = (\mathbf{A} \otimes \mathbf{B})_{(i-1)m+k, (j-1)m+l} = a_{ij}b_{kl}.$$

So by $(\dagger)$, $\mathbf{P}_\sigma^\top(\mathbf{A} \otimes \mathbf{B})\mathbf{P}_\sigma$ and $\mathbf{B} \otimes \mathbf{A}$ agree in every entry, and $\mathbf{P}_\sigma^\top = \mathbf{P}_\sigma^{-1}$ by Lemma 3.4.7 (b). The last claim is the case $\mathbf{B} = \mathbf{I}_m$. ■

For example, when $n = m = 2$, σ fixes 1 and 4 and swaps 2 and 3, so $\mathbf{P}$ is the elementary matrix $\mathbf{P}_{23}$: conjugating by it swaps rows 2, 3 and columns 2, 3. Doing this to $\mathbf{A} \otimes \mathbf{B}$ in Example 1 does give $\mathbf{B} \otimes \mathbf{A}$, as a direct check shows. For rectangular $\mathbf{A} \in M_{m \times n}(F)$ and $\mathbf{B} \in M_{p \times q}(F)$, the same argument gives $\mathbf{B} \otimes \mathbf{A} = \mathbf{P}_\sigma^\top(\mathbf{A} \otimes \mathbf{B})\mathbf{P}_\tau$ with one permutation for the rows and another for the columns.

8.4.5 Rank, trace, and determinant

Each of the three basic numbers attached to a matrix behaves multiplicatively. The formulas for trace and rank are what one would guess. The determinant is subtler, because the sizes enter as exponents.

Theorem 8.4.3. Rank, Trace, and Determinant of a Kronecker Product

- (a) For $\mathbf{A} \in M_{m \times n}(F)$ and $\mathbf{B} \in M_{p \times q}(F)$, $\text{rank}(\mathbf{A} \otimes \mathbf{B}) = \text{rank } \mathbf{A} \cdot \text{rank } \mathbf{B}$.
- (b) For $\mathbf{A} \in M_n(F)$ and $\mathbf{B} \in M_m(F)$, $\text{tr}(\mathbf{A} \otimes \mathbf{B}) = \text{tr } \mathbf{A} \cdot \text{tr } \mathbf{B}$.
- (c) For $\mathbf{A} \in M_n(F)$ and $\mathbf{B} \in M_m(F)$, $\det(\mathbf{A} \otimes \mathbf{B}) = (\det \mathbf{A})^m (\det \mathbf{B})^n$.

Idea. For rank, bring both factors to rank normal form, $\mathbf{A} = \mathbf{Q}\mathbf{N}_r\mathbf{P}$ and $\mathbf{B} = \mathbf{Q}'\mathbf{N}_s\mathbf{P}'$. The mixed product rule moves the invertible factors outside, where they do not change rank, and $\mathbf{N}_r \otimes \mathbf{N}_s$ visibly has rank rs . Trace is a sum over diagonal entries, and the diagonal entries of $\mathbf{A} \otimes \mathbf{B}$ are the products $a_{ii}b_{kk}$. For the determinant, split $\mathbf{A} \otimes \mathbf{B} = (\mathbf{A} \otimes \mathbf{I}_m)(\mathbf{I}_n \otimes \mathbf{B})$

by (**). The second factor is block diagonal with n copies of $\mathbf{B}$. The first is not, but by [Proposition 8.4.2](#) it is similar to $\mathbf{I}_m \otimes \mathbf{A}$, which is block diagonal with m copies of $\mathbf{A}$.

Proof.

- (a) Let $r = \text{rank } \mathbf{A}$ and $s = \text{rank } \mathbf{B}$. By Theorem 4.3.3 (b) and Definition 4.3.1, there are invertible $\mathbf{Q}_1, \mathbf{P}_1, \mathbf{Q}_2, \mathbf{P}_2$ with $\mathbf{N}_r = \mathbf{Q}_1 \mathbf{A} \mathbf{P}_1$ and $\mathbf{N}_s = \mathbf{Q}_2 \mathbf{B} \mathbf{P}_2$, where $\mathbf{N}_r \in M_{m \times n}(F)$ and $\mathbf{N}_s \in M_{p \times q}(F)$ have $\mathbf{I}_r, \mathbf{I}_s$ in the top left corner and zeros elsewhere. Hence $\mathbf{A} = \mathbf{Q}_1^{-1} \mathbf{N}_r \mathbf{P}_1^{-1}$ and $\mathbf{B} = \mathbf{Q}_2^{-1} \mathbf{N}_s \mathbf{P}_2^{-1}$, and by [Theorem 8.4.1](#) (c), used twice,

$$\mathbf{A} \otimes \mathbf{B} = (\mathbf{Q}_1^{-1} \otimes \mathbf{Q}_2^{-1})(\mathbf{N}_r \otimes \mathbf{N}_s)(\mathbf{P}_1^{-1} \otimes \mathbf{P}_2^{-1}).$$

The outer factors are invertible by [Theorem 8.4.1](#) (e), so $\text{rank}(\mathbf{A} \otimes \mathbf{B}) = \text{rank}(\mathbf{N}_r \otimes \mathbf{N}_s)$ by Theorem 3.5.9.

By (*), the entry of $\mathbf{N}_r \otimes \mathbf{N}_s$ in position $((i-1)p+k, (j-1)q+l)$ is 1 if $i = j \leq r$ and $k = l \leq s$, and 0 otherwise. So for each of the rs pairs (i, k) with $i \leq r, k \leq s$, column $(i-1)q+k$ equals $\mathbf{e}_{(i-1)p+k}$, and every other column is zero. Distinct pairs give distinct standard basis vectors, so these rs columns are linearly independent and span the column space. Hence $\text{rank}(\mathbf{N}_r \otimes \mathbf{N}_s) = rs$ by Definition 3.5.4, which proves (a). (If $r = 0$ or $s = 0$, the matrix is zero and both sides are 0.)

- (b) By (*) with $p = q = m$, the diagonal entries of $\mathbf{A} \otimes \mathbf{B}$ are those with $(i-1)m+k = (j-1)m+l$, that is $i = j$ and $k = l$ by uniqueness of division with remainder, and they equal $a_{ii}b_{kk}$. By Definition 0.9.12,

$$\text{tr}(\mathbf{A} \otimes \mathbf{B}) = \sum_{i=1}^n \sum_{k=1}^m a_{ii}b_{kk} = \left(\sum_{i=1}^n a_{ii} \right) \left(\sum_{k=1}^m b_{kk} \right) = \text{tr } \mathbf{A} \cdot \text{tr } \mathbf{B}.$$

- (c) First, a block diagonal matrix $\mathbf{M}_1 \oplus \cdots \oplus \mathbf{M}_t$ with square blocks $\mathbf{M}_1, \dots, \mathbf{M}_t$ has determinant $\det \mathbf{M}_1 \cdots \det \mathbf{M}_t$. For $t = 1$ there is nothing to prove, and for $t \geq 2$ the matrix is $\mathbf{M}_1 \oplus (\mathbf{M}_2 \oplus \cdots \oplus \mathbf{M}_t)$, so Theorem 8.1.4 (d) (or Theorem 7.8.4) and induction on t give the claim.

By Example 3, $\mathbf{I}_n \otimes \mathbf{B} = \mathbf{B} \oplus \cdots \oplus \mathbf{B}$ with n blocks, so $\det(\mathbf{I}_n \otimes \mathbf{B}) = (\det \mathbf{B})^n$. Likewise $\mathbf{I}_m \otimes \mathbf{A} = \mathbf{A} \oplus \cdots \oplus \mathbf{A}$ with m blocks, so $\det(\mathbf{I}_m \otimes \mathbf{A}) = (\det \mathbf{A})^m$. By [Proposition 8.4.2](#), $\mathbf{A} \otimes \mathbf{I}_m$ is similar to $\mathbf{I}_m \otimes \mathbf{A}$, so $\det(\mathbf{A} \otimes \mathbf{I}_m) = (\det \mathbf{A})^m$ by Corollary 7.5.7. Finally, by (**) and Theorem 7.5.4,

$$\det(\mathbf{A} \otimes \mathbf{B}) = \det(\mathbf{A} \otimes \mathbf{I}_m) \det(\mathbf{I}_n \otimes \mathbf{B}) = (\det \mathbf{A})^m (\det \mathbf{B})^n.$$

■

The exponents are easy to get backwards, and a size count fixes them. In $\mathbf{A} \otimes \mathbf{B}$ with $\mathbf{A} \in M_n(F)$, $\mathbf{B} \in M_m(F)$, the matrix $\mathbf{A}$ is “used” m times (once for each index of $\mathbf{B}$), so $\det \mathbf{A}$ appears to the power m , the size of the **other** factor. For $\mathbf{A} = c\mathbf{I}_n$ and $\mathbf{B} = \mathbf{I}_m$, $\mathbf{A} \otimes \mathbf{B} = c\mathbf{I}_{nm}$ has determinant $c^{nm} = (c^n)^m = (\det \mathbf{A})^m$, which confirms the exponent.

For the matrices of the example after [Theorem 8.4.1](#), $\det \mathbf{A} = 2$, $\det \mathbf{B} = 1$, $\text{tr } \mathbf{A} = 3$, and $\text{tr } \mathbf{B} = 3$, so $\det(\mathbf{A} \otimes \mathbf{B}) = 2^2 \cdot 1^2 = 4$, $\text{tr}(\mathbf{A} \otimes \mathbf{B}) = 9$, and $\text{rank}(\mathbf{A} \otimes \mathbf{B}) = 2 \cdot 2 = 4$. Directly, the diagonal of $\mathbf{A} \otimes \mathbf{B}$ is 2, 1, 4, 2, with sum 9.

Quick check. Let $\mathbf{A} \in M_2(\mathbb{R})$ have $\text{tr } \mathbf{A} = 4$, $\det \mathbf{A} = 3$, and let $\mathbf{B} \in M_3(\mathbb{R})$ have $\text{tr } \mathbf{B} = 2$, $\det \mathbf{B} = -1$. Find $\text{tr}(\mathbf{A} \otimes \mathbf{B})$, $\det(\mathbf{A} \otimes \mathbf{B})$, and $\det(\mathbf{B} \otimes \mathbf{A})$.

Solution. By [Theorem 8.4.3](#), $\text{tr}(\mathbf{A} \otimes \mathbf{B}) = 4 \cdot 2 = 8$ and $\det(\mathbf{A} \otimes \mathbf{B}) = (\det \mathbf{A})^3 (\det \mathbf{B})^2 = 27 \cdot 1 = 27$. With the roles swapped, $\det(\mathbf{B} \otimes \mathbf{A}) = (\det \mathbf{B})^2 (\det \mathbf{A})^3 = 27$, as it must be since $\mathbf{B} \otimes \mathbf{A} \sim \mathbf{A} \otimes \mathbf{B}$ ([Proposition 8.4.2](#)). The common wrong answer $(\det \mathbf{A})^2 (\det \mathbf{B})^3 = -9$ puts each matrix's own size in its exponent.

8.4.6 Vectorization

We return to the question of the opening: turning a matrix equation into a linear system. For that we need to flatten the unknown matrix into a column, in an order that matches (*).

To vectorize a matrix is to stack its columns into one long column.

Definition 8.4.2. Vectorization

Let $\mathbf{X} = (x_{ij}) \in M_{m \times n}(F)$ have columns $\mathbf{x}_1, \dots, \mathbf{x}_n \in F^m$. Its **vectorization** is the column vector

$$\text{vec } \mathbf{X} := \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_n \end{pmatrix} \in F^{mn},$$

whose entry in position $(j-1)m + i$ is x_{ij} , for $1 \leq i \leq m$, $1 \leq j \leq n$.

For example, $\text{vec} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = (1, 3, 2, 4)$, read down the first column and then down the second.

For a column $\mathbf{x} \in F^m$ (the case $n = 1$), $\text{vec } \mathbf{x} = \mathbf{x}$. The map $\text{vec}: M_{m \times n}(F) \rightarrow F^{mn}$ is linear, since each entry of $\text{vec } \mathbf{X}$ is an entry of $\mathbf{X}$, and bijective, since it only rearranges entries; so it is an isomorphism. It sends the matrix unit $\mathbf{E}_{ij}$ to $\mathbf{e}_{(j-1)m+i} = \mathbf{e}_j \otimes \mathbf{e}_i$ (Example 2, with the index of the column first). More generally, $\text{vec}(\mathbf{u}\mathbf{v}^\top) = \mathbf{v} \otimes \mathbf{u}$ for $\mathbf{u} \in F^m$ and $\mathbf{v} \in F^n$: column j of $\mathbf{u}\mathbf{v}^\top$ is $v_j \mathbf{u}$.

Two conventions are in use, and they differ. Some books stack the **rows** instead, which exchanges the roles of the two factors in the next theorem. We always stack columns.

Theorem 8.4.4. The vec Identity

Let $\mathbf{A} \in M_{k \times m}(F)$, $\mathbf{X} \in M_{m \times n}(F)$, and $\mathbf{B} \in M_{n \times l}(F)$. Then

$$\text{vec}(\mathbf{A}\mathbf{X}\mathbf{B}) = (\mathbf{B}^\top \otimes \mathbf{A}) \text{vec } \mathbf{X}.$$

Idea. Both sides are columns in F^{kl} . We compare one entry, using the formula for a triple product on the left and (*) on the right. The factor $\mathbf{A}$ acts on the row index of $\mathbf{X}$, which

is the inner index in vec ; the factor $\mathbf{B}$ acts on the column index, which is the outer index. That is why $\mathbf{B}$ comes first, and it enters transposed because it multiplies $\mathbf{X}$ from the right.

Proof. Both sides lie in F^{kl} : $\mathbf{AXB} \in M_{k \times l}(F)$, and $\mathbf{B}^\top \otimes \mathbf{A} \in M_{lk \times nm}(F)$ acts on $\text{vec } \mathbf{X} \in F^{mn}$. Fix $1 \leq r \leq k$ and $1 \leq s \leq l$, and consider position $(s-1)k + r$. By Definition 8.4.2 and Definition 0.9.9 applied twice, the left side has entry

$$(\mathbf{AXB})_{rs} = \sum_{i=1}^m \sum_{j=1}^n a_{ri} x_{ij} b_{js}.$$

On the right, write the column index $c \in \{1, \dots, nm\}$ of $\mathbf{B}^\top \otimes \mathbf{A}$ uniquely as $c = (j-1)m + i$ with $1 \leq j \leq n$, $1 \leq i \leq m$. By (*), $(\mathbf{B}^\top \otimes \mathbf{A})_{(s-1)k+r, (j-1)m+i} = (\mathbf{B}^\top)_{sj} a_{ri} = b_{js} a_{ri}$, and by Definition 8.4.2 the c -th entry of $\text{vec } \mathbf{X}$ is x_{ij} . Hence the entry of the right side in position $(s-1)k + r$ is

$$\sum_{c=1}^{nm} (\mathbf{B}^\top \otimes \mathbf{A})_{(s-1)k+r, c} (\text{vec } \mathbf{X})_c = \sum_{j=1}^n \sum_{i=1}^m b_{js} a_{ri} x_{ij},$$

which equals the left side. As (r, s) ranges over all pairs, $(s-1)k + r$ ranges over all positions, so the two vectors are equal. ■

Taking $\mathbf{B} = \mathbf{I}_n$ or $\mathbf{A} = \mathbf{I}_m$ gives the two one-sided forms, which are used most often:

$$\text{vec}(\mathbf{AX}) = (\mathbf{I}_n \otimes \mathbf{A}) \text{vec } \mathbf{X}, \quad \text{vec}(\mathbf{XB}) = (\mathbf{B}^\top \otimes \mathbf{I}_m) \text{vec } \mathbf{X}.$$

The first says that multiplying $\mathbf{X}$ by $\mathbf{A}$ on the left multiplies each column of $\mathbf{X}$ by $\mathbf{A}$: $\mathbf{I}_n \otimes \mathbf{A} = \mathbf{A} \oplus \dots \oplus \mathbf{A}$ acts on the stacked columns one at a time.

Now every linear matrix equation built from products becomes a linear system. The most important one is the **Sylvester equation**

$$\mathbf{AX} + \mathbf{XB} = \mathbf{C}, \quad \mathbf{A} \in M_m(F), \mathbf{B} \in M_n(F), \mathbf{C} \in M_{m \times n}(F),$$

for an unknown $\mathbf{X} \in M_{m \times n}(F)$. Applying vec , which is linear, and the two one-sided forms,

$$(\mathbf{I}_n \otimes \mathbf{A} + \mathbf{B}^\top \otimes \mathbf{I}_m) \text{vec } \mathbf{X} = \text{vec } \mathbf{C}.$$

So the Sylvester equation has **exactly one** solution for every $\mathbf{C}$ if and only if the $mn \times mn$ matrix $\mathbf{I}_n \otimes \mathbf{A} + \mathbf{B}^\top \otimes \mathbf{I}_m$ is invertible (Theorem 3.4.4, transported through the isomorphism vec). The equation $\mathbf{AX} = \mathbf{XA}$ of the opening, which describes the matrices commuting with $\mathbf{A}$, is the case $\mathbf{B} = -\mathbf{A}$, $\mathbf{C} = \mathbf{O}$: its solutions form the null space of $\mathbf{I}_m \otimes \mathbf{A} - \mathbf{A}^\top \otimes \mathbf{I}_m$.

Example. A Sylvester Equation as a Linear System

Solve $\mathbf{AX} + \mathbf{XB} = \mathbf{C}$ over $\mathbb{Q}$, where

$$\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \quad \mathbf{C} = \begin{pmatrix} 1 & -1 \\ 1 & 3 \end{pmatrix}.$$

Solution. Here $m = n = 2$. Compute the two Kronecker products from [Definition 8.4.1](#):

$$\mathbf{I}_2 \otimes \mathbf{A} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}, \quad \mathbf{B}^\top \otimes \mathbf{I}_2 = \begin{pmatrix} 1\mathbf{I}_2 & 1\mathbf{I}_2 \\ 0\mathbf{I}_2 & 1\mathbf{I}_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

With $\text{vec } \mathbf{X} = (x_{11}, x_{21}, x_{12}, x_{22})$ and $\text{vec } \mathbf{C} = (1, 1, -1, 3)$, the system $(\mathbf{I}_2 \otimes \mathbf{A} + \mathbf{B}^\top \otimes \mathbf{I}_2) \text{vec } \mathbf{X} = \text{vec } \mathbf{C}$ is

$$\begin{pmatrix} 2 & 1 & 1 & 0 \\ 0 & 3 & 0 & 1 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} x_{11} \\ x_{21} \\ x_{12} \\ x_{22} \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ -1 \\ 3 \end{pmatrix}.$$

The coefficient matrix is upper triangular with nonzero diagonal entries 2, 3, 2, 3, so its determinant is $36 \neq 0$ (Theorem 7.4.6) and the solution is unique. Back substitution gives $x_{22} = 1$; then $2x_{12} + 1 = -1$, so $x_{12} = -1$; then $3x_{21} + 1 = 1$, so $x_{21} = 0$; then $2x_{11} + 0 - 1 = 1$, so $x_{11} = 1$. Hence

$$\mathbf{X} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}.$$

Check: $\mathbf{A}\mathbf{X} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$ and $\mathbf{X}\mathbf{B} = \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix}$, whose sum is $\mathbf{C}$. As a check on the bookkeeping, row 1 of the system is the $(1, 1)$ entry of $\mathbf{A}\mathbf{X} + \mathbf{X}\mathbf{B}$, namely $(x_{11} + x_{21}) + (x_{11} + x_{12})$.

In this example the coefficient matrix came out triangular, with diagonal entries $a_{ii} + b_{jj}$. That is no accident when $\mathbf{A}$ is upper triangular and $\mathbf{B}$ lower triangular, as [Exercise \(C2: Kronecker sums\)](#) below shows. In general, whether $\mathbf{I}_n \otimes \mathbf{A} + \mathbf{B}^\top \otimes \mathbf{I}_m$ is invertible depends on the eigenvalues of $\mathbf{A}$ and $\mathbf{B}$: over $\mathbb{C}$, the equation $\mathbf{A}\mathbf{X} - \mathbf{X}\mathbf{B} = \mathbf{C}$ is uniquely solvable for every $\mathbf{C}$ exactly when $\mathbf{A}$ and $\mathbf{B}$ have no common eigenvalue. That criterion needs eigenvalues and is proved in Chapter 12, and its quantitative form appears in Chapter 20.

Warning. Keep the order and the transpose in $\text{vec}(\mathbf{AXB}) = (\mathbf{B}^\top \otimes \mathbf{A}) \text{vec } \mathbf{X}$. Both $\mathbf{A} \otimes \mathbf{B}^\top$ and $\mathbf{B} \otimes \mathbf{A}$ look plausible, and both are wrong. With $\mathbf{X} = \mathbf{E}_{11} \in M_2(F)$, $\mathbf{A} = \mathbf{I}_2$, and $\mathbf{B} = \mathbf{E}_{21}$, we get $\mathbf{AXB} = \mathbf{E}_{11}\mathbf{E}_{21} = \mathbf{O}$, and indeed $(\mathbf{B}^\top \otimes \mathbf{I}_2)\mathbf{e}_1 = (\mathbf{E}_{12} \otimes \mathbf{I}_2)\mathbf{e}_1 = \mathbf{O}$, because the first column of $\mathbf{E}_{12} \otimes \mathbf{I}_2$ is zero. But $(\mathbf{B} \otimes \mathbf{I}_2)\mathbf{e}_1 = (\mathbf{E}_{21} \otimes \mathbf{I}_2)\mathbf{e}_1 = \mathbf{e}_3 \neq \mathbf{O}$. So $\mathbf{B} \otimes \mathbf{A}$ fails. A similar test with $\mathbf{A} = \mathbf{E}_{12}$, $\mathbf{X} = \mathbf{E}_{21}$, $\mathbf{B} = \mathbf{I}_2$ refutes $\mathbf{A} \otimes \mathbf{B}^\top$. If a formula is in doubt, test it on matrix units.

The Kronecker product is the coordinate shadow of a basis-free construction: the tensor product of linear maps, in Chapter 15, whose matrix in suitable bases is exactly $\mathbf{A} \otimes \mathbf{B}$. The dictionary order in $(*)$ will reappear there as an ordering of a basis $v_i \otimes w_k$.

8.4.7 Exercises

8.4.7.1 A. Check your understanding

Exercise. A1

- (a) Define $\mathbf{A} \otimes \mathbf{B}$ for $\mathbf{A} \in M_{m \times n}(F)$ and $\mathbf{B} \in M_{p \times q}(F)$, and state its size.
- (b) True or false: for all $\mathbf{A}, \mathbf{B} \in M_2(F)$, $\mathbf{A} \otimes \mathbf{B} = \mathbf{B} \otimes \mathbf{A}$. Justify your answer.
- (c) State the mixed product rule, including the conditions on the sizes.
- (d) True or false: $\det(\mathbf{A} \otimes \mathbf{B}) = \det \mathbf{A} \det \mathbf{B}$ for all $\mathbf{A}, \mathbf{B} \in M_2(\mathbb{R})$. Justify your answer.
- (e) State the vec identity, and name the matrix $\mathbf{M}$ with $\text{vec}(\mathbf{XB}) = \mathbf{M} \text{vec} \mathbf{X}$ for $\mathbf{X} \in M_{m \times n}(F)$, $\mathbf{B} \in M_n(F)$.

Solution.

- (a) $\mathbf{A} \otimes \mathbf{B}$ is the $mp \times nq$ block matrix with (i, j) block $a_{ij}\mathbf{B}$ (Definition 8.4.1); equivalently, its entry in position $((i-1)p+k, (j-1)q+l)$ is $a_{ij}b_{kl}$.
- (b) False. For $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, the $(1, 2)$ entry of $\mathbf{A} \otimes \mathbf{B}$ is $a_{11}b_{12} = 1$ and that of $\mathbf{B} \otimes \mathbf{A}$ is $b_{11}a_{12} = 0$. The two are similar (Proposition 8.4.2), not equal.
- (c) If $\mathbf{A} \in M_{m \times n}(F)$, $\mathbf{C} \in M_{n \times t}(F)$, $\mathbf{B} \in M_{p \times q}(F)$, and $\mathbf{D} \in M_{q \times u}(F)$, then $(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = (\mathbf{AC}) \otimes (\mathbf{BD})$ (Theorem 8.4.1 (c)). The conditions are exactly that $\mathbf{AC}$ and $\mathbf{BD}$ are defined.
- (d) False. By Theorem 8.4.3, $\det(\mathbf{A} \otimes \mathbf{B}) = (\det \mathbf{A})^2(\det \mathbf{B})^2$. For $\mathbf{A} = \mathbf{B} = 2\mathbf{I}_2$, $\mathbf{A} \otimes \mathbf{B} = 4\mathbf{I}_4$ has determinant 256, while $\det \mathbf{A} \det \mathbf{B} = 16$.
- (e) $\text{vec}(\mathbf{AXB}) = (\mathbf{B}^\top \otimes \mathbf{A}) \text{vec} \mathbf{X}$ (Theorem 8.4.4). With $\mathbf{A} = \mathbf{I}_m$, $\mathbf{M} = \mathbf{B}^\top \otimes \mathbf{I}_m$.

8.4.7.2 B. Practice**Exercise. B1: Computing with Kronecker products**

Let $\mathbf{A} = \begin{pmatrix} 2 & 0 \\ 1 & -1 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} 1 & 1 \\ 0 & 3 \end{pmatrix}$ over $\mathbb{Q}$.

- (a) Write out $\mathbf{A} \otimes \mathbf{B}$ and $\mathbf{B} \otimes \mathbf{A}$.
- (b) Compute tr , $\det$, and rank of $\mathbf{A} \otimes \mathbf{B}$ from Theorem 8.4.3, and check the trace directly.
- (c) Compute $\mathbf{A}^2 \otimes \mathbf{B}^2$ and use it to write down $(\mathbf{A} \otimes \mathbf{B})^2$ without multiplying 4×4 matrices. Justify the step.
- (d) Hence find $(\mathbf{A} \otimes \mathbf{B})^{-1}$.

Solution.

(a) By [Definition 8.4.1](#),

$$\mathbf{A} \otimes \mathbf{B} = \begin{pmatrix} 2\mathbf{B} & 0\mathbf{B} \\ 1\mathbf{B} & -1\mathbf{B} \end{pmatrix} = \begin{pmatrix} 2 & 2 & 0 & 0 \\ 0 & 6 & 0 & 0 \\ 1 & 1 & -1 & -1 \\ 0 & 3 & 0 & -3 \end{pmatrix},$$

$$\mathbf{B} \otimes \mathbf{A} = \begin{pmatrix} 1\mathbf{A} & 1\mathbf{A} \\ 0\mathbf{A} & 3\mathbf{A} \end{pmatrix} = \begin{pmatrix} 2 & 0 & 2 & 0 \\ 1 & -1 & 1 & -1 \\ 0 & 0 & 6 & 0 \\ 0 & 0 & 3 & -3 \end{pmatrix}.$$

(b) $\text{tr } \mathbf{A} = 1$, $\text{tr } \mathbf{B} = 4$, $\det \mathbf{A} = -2$, $\det \mathbf{B} = 3$, and both have rank 2. By [Theorem 8.4.3](#), $\text{tr}(\mathbf{A} \otimes \mathbf{B}) = 4$, $\det(\mathbf{A} \otimes \mathbf{B}) = (-2)^2 \cdot 3^2 = 36$, and $\text{rank}(\mathbf{A} \otimes \mathbf{B}) = 4$. Directly, the diagonal of $\mathbf{A} \otimes \mathbf{B}$ is 2, 6, -1, -3, with sum 4.

(c) By the mixed product rule ([Theorem 8.4.1 \(c\)](#)) with $\mathbf{C} = \mathbf{A}$ and $\mathbf{D} = \mathbf{B}$, $(\mathbf{A} \otimes \mathbf{B})^2 = \mathbf{A}^2 \otimes \mathbf{B}^2$. Here $\mathbf{A}^2 = \begin{pmatrix} 4 & 0 \\ 1 & 1 \end{pmatrix}$ and $\mathbf{B}^2 = \begin{pmatrix} 1 & 4 \\ 0 & 9 \end{pmatrix}$, so

$$(\mathbf{A} \otimes \mathbf{B})^2 = \begin{pmatrix} 4\mathbf{B}^2 & \mathbf{O} \\ \mathbf{B}^2 & \mathbf{B}^2 \end{pmatrix} = \begin{pmatrix} 4 & 16 & 0 & 0 \\ 0 & 36 & 0 & 0 \\ 1 & 4 & 1 & 4 \\ 0 & 9 & 0 & 9 \end{pmatrix}.$$

(d) $\mathbf{A}^{-1} = \frac{1}{-2} \begin{pmatrix} -1 & 0 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & 0 \\ \frac{1}{2} & -1 \end{pmatrix}$ and $\mathbf{B}^{-1} = \frac{1}{3} \begin{pmatrix} 3 & -1 \\ 0 & 1 \end{pmatrix}$ by [Theorem 0.9.8](#). By [Theorem 8.4.1 \(e\)](#),

$$(\mathbf{A} \otimes \mathbf{B})^{-1} = \mathbf{A}^{-1} \otimes \mathbf{B}^{-1} = \begin{pmatrix} \frac{1}{2}\mathbf{B}^{-1} & \mathbf{O} \\ \frac{1}{2}\mathbf{B}^{-1} & -\mathbf{B}^{-1} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & -\frac{1}{6} & 0 & 0 \\ 0 & \frac{1}{6} & 0 & 0 \\ \frac{1}{2} & -\frac{1}{6} & -1 & \frac{1}{3} \\ 0 & \frac{1}{6} & 0 & -\frac{1}{3} \end{pmatrix}.$$

Multiplying by $\mathbf{A} \otimes \mathbf{B}$ gives $\mathbf{I}_4$.

Exercise. B2: One-sided equations

Let $\mathbf{A} \in M_m(F)$ and $\mathbf{C} \in M_{m \times n}(F)$, and consider $\mathbf{A}\mathbf{X} = \mathbf{C}$ for $\mathbf{X} \in M_{m \times n}(F)$.

- Write the equation in the form $\mathbf{M} \text{vec } \mathbf{X} = \text{vec } \mathbf{C}$, and describe $\mathbf{M}$ as a block matrix.
- Explain why the system splits into n separate systems, one for each column of $\mathbf{X}$.
- Hence show that $\mathbf{A}\mathbf{X} = \mathbf{C}$ has exactly one solution for every $\mathbf{C}$ if and only if $\mathbf{A}$ is invertible.
- Solve $\mathbf{X}\mathbf{A} = \mathbf{C}$ over $\mathbb{Q}$ by the same method, where $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix}$ and $\mathbf{C} = \begin{pmatrix} 3 & 7 \\ 0 & 1 \end{pmatrix}$.

Solution.

- (a) By [Theorem 8.4.4](#) with $\mathbf{B} = \mathbf{I}_n$, $\text{vec}(\mathbf{A}\mathbf{X}) = (\mathbf{I}_n \otimes \mathbf{A}) \text{vec } \mathbf{X}$, so $\mathbf{M} = \mathbf{I}_n \otimes \mathbf{A} = \mathbf{A} \oplus \cdots \oplus \mathbf{A}$, with n blocks.
- (b) Write $\text{vec } \mathbf{X}$ and $\text{vec } \mathbf{C}$ as stacks of the columns $\mathbf{x}_j$ and $\mathbf{c}_j$. By [Theorem 8.1.1](#), $(\mathbf{A} \oplus \cdots \oplus \mathbf{A}) \text{vec } \mathbf{X}$ is the stack of the $\mathbf{A}\mathbf{x}_j$. So the system says $\mathbf{A}\mathbf{x}_j = \mathbf{c}_j$ for $j = 1, \dots, n$, and the unknowns of different columns never meet.
- (c) ($\Leftarrow$) If $\mathbf{A}$ is invertible, then $\mathbf{A}\mathbf{X} = \mathbf{C}$ forces $\mathbf{X} = \mathbf{A}^{-1}\mathbf{C}$, and this $\mathbf{X}$ is a solution. ($\Rightarrow$) Suppose $\mathbf{A}\mathbf{X} = \mathbf{C}$ is solvable for every $\mathbf{C}$. Given $\mathbf{c} \in F^m$, take $\mathbf{C}$ with every column equal to $\mathbf{c}$; by (b), a solution gives $\mathbf{x}$ with $\mathbf{A}\mathbf{x} = \mathbf{c}$. So $\mathbf{A}\mathbf{x} = \mathbf{c}$ is solvable for every $\mathbf{c}$, and $\mathbf{A}$ is invertible by [Theorem 3.4.4](#) ((e) $\Rightarrow$ (a)).
- (d) By [Theorem 8.4.4](#) with the left factor $\mathbf{I}_2$, $\text{vec}(\mathbf{X}\mathbf{A}) = (\mathbf{A}^\top \otimes \mathbf{I}_2) \text{vec } \mathbf{X}$. With $\mathbf{A}^\top = \begin{pmatrix} 1 & 1 \\ 2 & 3 \end{pmatrix}$, $\text{vec } \mathbf{X} = (x_{11}, x_{21}, x_{12}, x_{22})$, and $\text{vec } \mathbf{C} = (3, 0, 7, 1)$, the system is

$$\begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 2 & 0 & 3 & 0 \\ 0 & 2 & 0 & 3 \end{pmatrix} \begin{pmatrix} x_{11} \\ x_{21} \\ x_{12} \\ x_{22} \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \\ 7 \\ 1 \end{pmatrix}.$$

This time the system splits by the **rows** of $\mathbf{X}$: equations 1, 3 involve only x_{11}, x_{12} , and equations 2, 4 only x_{21}, x_{22} . From $x_{11} + x_{12} = 3$ and $2x_{11} + 3x_{12} = 7$ we get $x_{12} = 1$, $x_{11} = 2$. From $x_{21} + x_{22} = 0$ and $2x_{21} + 3x_{22} = 1$ we get $x_{22} = 1$, $x_{21} = -1$. Hence

$$\mathbf{X} = \begin{pmatrix} 2 & 1 \\ -1 & 1 \end{pmatrix}.$$

$$\text{Check: } \mathbf{X}\mathbf{A} = \begin{pmatrix} 2+1 & 4+3 \\ -1+1 & -2+3 \end{pmatrix} = \begin{pmatrix} 3 & 7 \\ 0 & 1 \end{pmatrix} = \mathbf{C}.$$

Exercise. B3: Which identities hold?

Determine which of the following statements hold for **all** $\mathbf{A}, \mathbf{B}, \mathbf{C} \in M_2(\mathbb{R})$. Justify your answer.

- (a) $(\mathbf{A} \otimes \mathbf{B})(\mathbf{A} \otimes \mathbf{B}) = \mathbf{A}^2 \otimes \mathbf{B}^2$.
- (b) $(\mathbf{A} + \mathbf{B}) \otimes \mathbf{C} = \mathbf{A} \otimes \mathbf{C} + \mathbf{B} \otimes \mathbf{C}$.
- (c) $(\mathbf{A} \otimes \mathbf{B})^\top = \mathbf{B}^\top \otimes \mathbf{A}^\top$.
- (d) $\mathbf{A} \otimes \mathbf{B} = \mathbf{O}$ implies $\mathbf{A} = \mathbf{O}$ or $\mathbf{B} = \mathbf{O}$.
- (e) $\text{tr}(\mathbf{A} \otimes \mathbf{B}) = \text{tr}(\mathbf{B} \otimes \mathbf{A})$.
- (f) $\mathbf{A} \otimes \mathbf{B} + \mathbf{B} \otimes \mathbf{A} = 2(\mathbf{A} \otimes \mathbf{B})$.

Solution.

- (a) Holds, by the mixed product rule ([Theorem 8.4.1](#) (c)).
- (b) Holds, by bilinearity ([Theorem 8.4.1](#) (a)).
- (c) Fails. By [Theorem 8.4.1](#) (d), $(\mathbf{A} \otimes \mathbf{B})^\top = \mathbf{A}^\top \otimes \mathbf{B}^\top$, and this differs from $\mathbf{B}^\top \otimes \mathbf{A}^\top$ in

general. Take $\mathbf{A} = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \mathbf{B}^\top$. By $(*)$, the $(1, 2)$ entry of $\mathbf{A}^\top \otimes \mathbf{B}^\top$ is $(\mathbf{A}^\top)_{11}b_{12} = 1$, while that of $\mathbf{B}^\top \otimes \mathbf{A}^\top$ is $b_{11}(\mathbf{A}^\top)_{12} = 0$. Unlike $(\mathbf{AB})^\top = \mathbf{B}^\top \mathbf{A}^\top$, the order of the factors does not reverse.

- (d) Holds. If $\mathbf{A} \neq \mathbf{O}$ and $\mathbf{B} \neq \mathbf{O}$, then $\text{rank } \mathbf{A} \geq 1$ and $\text{rank } \mathbf{B} \geq 1$, so $\text{rank}(\mathbf{A} \otimes \mathbf{B}) \geq 1$ by [Theorem 8.4.3](#) (a), and $\mathbf{A} \otimes \mathbf{B} \neq \mathbf{O}$. (Directly: if $a_{ij} \neq 0$ and $b_{kl} \neq 0$, the entry $a_{ij}b_{kl}$ of $\mathbf{A} \otimes \mathbf{B}$ is nonzero.)
- (e) Holds: both equal $\text{tr } \mathbf{A} \text{tr } \mathbf{B}$ by [Theorem 8.4.3](#) (b).
- (f) Fails, since it is equivalent to $\mathbf{B} \otimes \mathbf{A} = \mathbf{A} \otimes \mathbf{B}$, which is false for the matrices of Example 1 after [Definition 8.4.1](#) (see the nonexample there).

8.4.7.3 C. Going deeper

Exercise. C1: When is a Kronecker product invertible?

- (a) Let $\mathbf{A} \in M_n(F)$ and $\mathbf{B} \in M_m(F)$. Prove that $\mathbf{A} \otimes \mathbf{B}$ is invertible if and only if both $\mathbf{A}$ and $\mathbf{B}$ are invertible.
- (b) Let $\mathbf{A} \in M_{m \times n}(F)$ and $\mathbf{B} \in M_{p \times q}(F)$ with $mp = nq$, so that $\mathbf{A} \otimes \mathbf{B}$ is square. Prove that if $m \neq n$, then $\mathbf{A} \otimes \mathbf{B}$ is **not** invertible.
- (c) Give an example of a 2×1 matrix $\mathbf{A}$ and a 1×2 matrix $\mathbf{B}$ for which $\mathbf{A} \otimes \mathbf{B}$ is a nonzero 2×2 matrix, and compute its rank.

Solution.

- (a) By [Theorem 8.4.3](#) (c), $\det(\mathbf{A} \otimes \mathbf{B}) = (\det \mathbf{A})^m (\det \mathbf{B})^n$. In a field a product is nonzero if and only if each factor is, so this is nonzero if and only if $\det \mathbf{A} \neq 0$ and $\det \mathbf{B} \neq 0$ (here $n, m \geq 1$). By [Theorem 7.5.2](#), applied to all three matrices, $\mathbf{A} \otimes \mathbf{B}$ is invertible if and only if $\mathbf{A}$ and $\mathbf{B}$ are.
- (b) Suppose $m < n$. Since $mp = nq$ and $m < n$, we get $p > q$. By [Theorem 8.4.3](#) (a) and the bound $\text{rank } \mathbf{M} \leq (\text{number of rows})$ and $\leq (\text{number of columns})$,

$$\text{rank}(\mathbf{A} \otimes \mathbf{B}) = \text{rank } \mathbf{A} \text{rank } \mathbf{B} \leq m \cdot q < n \cdot q = nq,$$

so the $nq \times nq$ matrix $\mathbf{A} \otimes \mathbf{B}$ is not invertible ([Theorem 7.5.3](#) (d)). If $m > n$, then $p < q$, and $\text{rank } \mathbf{A} \text{rank } \mathbf{B} \leq n \cdot p < m \cdot p = mp = nq$ in the same way.

- (c) Take $\mathbf{A} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} 3 & 1 \end{pmatrix}$. Then $\mathbf{A} \otimes \mathbf{B} = \begin{pmatrix} 1\mathbf{B} \\ 2\mathbf{B} \end{pmatrix} = \begin{pmatrix} 3 & 1 \\ 6 & 2 \end{pmatrix}$, the outer product $\mathbf{AB}$, of rank $1 \cdot 1 = 1$: its second row is twice its first. So it is square and nonzero but singular, as (b) predicts.

Exercise. C2: Kronecker sums

Let $\mathbf{A} \in M_n(F)$ and $\mathbf{B} \in M_m(F)$, and put $\mathbf{K} := \mathbf{A} \otimes \mathbf{I}_m + \mathbf{I}_n \otimes \mathbf{B} \in M_{nm}(F)$.

- (a) Prove that $(\mathbf{A} \otimes \mathbf{I}_m)(\mathbf{I}_n \otimes \mathbf{B}) = (\mathbf{I}_n \otimes \mathbf{B})(\mathbf{A} \otimes \mathbf{I}_m)$, and deduce that $\mathbf{K}^2 = \mathbf{A}^2 \otimes \mathbf{I}_m + 2(\mathbf{A} \otimes \mathbf{B}) + \mathbf{I}_n \otimes \mathbf{B}^2$.
- (b) Show that for $\mathbf{X} \in M_{m \times n}(F)$, $\text{vec}(\mathbf{BX} + \mathbf{XA}^\top) = \mathbf{K} \text{vec } \mathbf{X}$.

- (c) Suppose $\mathbf{A}$ and $\mathbf{B}$ are upper triangular. Prove that $\mathbf{K}$ is upper triangular with diagonal entries $a_{ii} + b_{kk}$, and deduce that $\mathbf{B}\mathbf{X} + \mathbf{X}\mathbf{A}^\top = \mathbf{C}$ has exactly one solution for every $\mathbf{C} \in M_{m \times n}(F)$ if and only if $a_{ii} + b_{kk} \neq 0$ for all i, k .
- (d) For $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}$ over $\mathbb{Q}$, decide whether $\mathbf{B}\mathbf{X} + \mathbf{X}\mathbf{A}^\top = \mathbf{C}$ is solvable for every $\mathbf{C}$. If not, find a $\mathbf{C}$ with no solution.

Solution.

- (a) By the mixed product rule, both products equal $(\mathbf{A}\mathbf{I}_n) \otimes (\mathbf{I}_m\mathbf{B}) = (\mathbf{I}_n\mathbf{A}) \otimes (\mathbf{B}\mathbf{I}_m) = \mathbf{A} \otimes \mathbf{B}$, as in (**). Expanding $\mathbf{K}^2$ by distributivity (Theorem 0.9.4) gives $(\mathbf{A} \otimes \mathbf{I}_m)^2 + (\mathbf{A} \otimes \mathbf{I}_m)(\mathbf{I}_n \otimes \mathbf{B}) + (\mathbf{I}_n \otimes \mathbf{B})(\mathbf{A} \otimes \mathbf{I}_m) + (\mathbf{I}_n \otimes \mathbf{B})^2$. The middle terms are both $\mathbf{A} \otimes \mathbf{B}$, and the outer ones are $\mathbf{A}^2 \otimes \mathbf{I}_m$ and $\mathbf{I}_n \otimes \mathbf{B}^2$ by the mixed product rule.
- (b) By Theorem 8.4.4, $\text{vec}(\mathbf{B}\mathbf{X}\mathbf{I}_n) = (\mathbf{I}_n \otimes \mathbf{B})\text{vec } \mathbf{X}$ and $\text{vec}(\mathbf{I}_m\mathbf{X}\mathbf{A}^\top) = ((\mathbf{A}^\top)^\top \otimes \mathbf{I}_m)\text{vec } \mathbf{X} = (\mathbf{A} \otimes \mathbf{I}_m)\text{vec } \mathbf{X}$. Add, using linearity of vec .
- (c) By (*), the entry of $\mathbf{A} \otimes \mathbf{I}_m$ in position $((i-1)m+k, (j-1)m+l)$ is $a_{ij}\delta_{kl}$, which is zero unless $i \leq j$ and $k = l$; in that case $(i-1)m+k \leq (j-1)m+l$. The entry of $\mathbf{I}_n \otimes \mathbf{B}$ there is $\delta_{ij}b_{kl}$, zero unless $i = j$ and $k \leq l$, and again the row index is at most the column index. So both are upper triangular, hence so is $\mathbf{K}$. On the diagonal, $i = j$ and $k = l$, and the entry is $a_{ii} + b_{kk}$. By Theorem 7.4.6, $\det \mathbf{K} = \prod_{i,k} (a_{ii} + b_{kk})$. By (b), the equation is $\mathbf{K} \text{vec } \mathbf{X} = \text{vec } \mathbf{C}$, and since vec is an isomorphism, it has exactly one solution for every $\mathbf{C}$ if and only if $\mathbf{K}$ is invertible (Theorem 3.4.4), that is $\det \mathbf{K} \neq 0$ (Theorem 7.5.2), that is every $a_{ii} + b_{kk} \neq 0$.
- (d) The diagonal entries $a_{ii} + b_{kk}$ are $1 + 1 = 2$, $1 + 3 = 4$, $-1 + 1 = 0$, $-1 + 3 = 2$, so by (c) the equation is not solvable for every $\mathbf{C}$. Explicitly,

$$\mathbf{K} = \begin{pmatrix} 1\mathbf{I}_2 & 2\mathbf{I}_2 \\ 0\mathbf{I}_2 & -1\mathbf{I}_2 \end{pmatrix} + \begin{pmatrix} \mathbf{B} & \mathbf{O} \\ \mathbf{O} & \mathbf{B} \end{pmatrix} = \begin{pmatrix} 2 & 0 & 2 & 0 \\ 0 & 4 & 0 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix},$$

whose third row is zero. So $\mathbf{K} \text{vec } \mathbf{X}$ always has third entry 0, and any $\mathbf{C}$ with $(\text{vec } \mathbf{C})_3 = c_{12} \neq 0$ has no solution, for example $\mathbf{C} = \mathbf{E}_{12}$. (Directly: the $(1, 2)$ entry of $\mathbf{B}\mathbf{X} + \mathbf{X}\mathbf{A}^\top$ is $x_{12} + (x_{11}a_{21} + x_{12}a_{22}) = x_{12} - x_{12} = 0$.)