Exercise A: The Kronecker product

A1

The Kronecker product of two matrices $A \in \mathbb{F}^{m\times n}$ and $B \in \mathbb{F}^{p\times q}$ is the matrix of size $mp \times nq$ whose elements are all possible products between the elements of A and B arranged in the following way:

$$A \otimes B := \begin{bmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{bmatrix}$$

With $A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}, \quad B = \begin{bmatrix} b_{11} & \cdots & b_{1q} \\ \vdots & \ddots & \vdots \\ b_{p1} & \cdots & b_{pq} \end{bmatrix}$

A2

The Kronecker product is associative. Let $C \in \mathbb{F}^{s\times t}$ be a third matrix. We show that $(A \otimes B) \otimes C = A \otimes (B \otimes C)$.

With $C = \begin{bmatrix} c_{11} & \cdots & c_{1t} \\ \vdots & \ddots & \vdots \\ c_{s1} & \cdots & c_{st} \end{bmatrix}$

Proof.

$$\begin{aligned}(A \otimes B) \otimes C &= \begin{bmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{bmatrix} \otimes C \\ &= \begin{bmatrix} a_{11}b_{11}C & \cdots & a_{11}b_{1q}C & \cdots & a_{1n}b_{11}C & \cdots & a_{1n}b_{1q}C \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{11}b_{p1}C & \cdots & a_{11}b_{pq}C & \cdots & a_{1n}b_{p1}C & \cdots & a_{1n}b_{pq}C \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{m1}b_{11}C & \cdots & a_{m1}b_{1q}C & \cdots & a_{mn}b_{11}C & \cdots & a_{mn}b_{1q}C \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{m1}b_{p1}C & \cdots & a_{m1}b_{pq}C & \cdots & a_{mn}b_{p1}C & \cdots & a_{mn}b_{pq}C \end{bmatrix} \\ &= A \otimes (B \otimes C).\end{aligned}$$

□

The Kronecker product is non-commutative; we show that $A \otimes B \neq B \otimes A$

Proof. We show a counterexample to the claim of commutativity. Let

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

In that case, we have

$$A \otimes B = \begin{bmatrix} 0 & -2 & 0 & -3 \\ -2 & 2 & -3 & 3 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 1 \end{bmatrix}, \quad B \otimes A = \begin{bmatrix} 0 & 0 & -2 & -3 \\ 0 & 0 & 0 & -1 \\ -2 & -3 & 2 & 3 \\ 0 & -1 & 0 & 1 \end{bmatrix}.$$

We see that $A \otimes B \neq B \otimes A$. □

Finally, the set $\mathbb{F}^{n\times n}$ equipped with the Kronecker product is not a group as it does not satisfy the closure property. Indeed, one can easily see that the Kronecker product of two matrices in $\mathbb{F}^{n\times n}$ is an element of $\mathbb{F}^{n^2\times n^2}$.

A3

Let $A \in \mathbb{F}^{m\times n}$, $B \in \mathbb{F}^{p\times q}$, $C \in \mathbb{F}^{n\times r}$, and $D \in \mathbb{F}^{q\times s}$,

$$(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$$

Proof. We simply verify that

$$\begin{aligned}
(A\otimes B)(C\otimes D) &= \begin{bmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{bmatrix} \begin{bmatrix} c_{11}D & \cdots & c_{1r}D \\ \vdots & \ddots & \vdots \\ c_{n1}D & \cdots & c_{nr}D \end{bmatrix} \\
&= \begin{bmatrix} \sum_{k=1}^n a_{1k}c_{k1}BD & \cdots & \sum_{k=1}^n a_{1k}c_{kr}BD \\ \vdots & \ddots & \vdots \\ \sum_{k=1}^n a_{mk}c_{k1}BD & \cdots & \sum_{k=1}^n a_{mk}c_{kr}BD \end{bmatrix} \\
&= AC \otimes BD.
\end{aligned}$$

□

This allows us to say that (if $A \in \mathbb{F}^{n\times n}$ and $B \in \mathbb{F}^{m\times m}$ are nonsingular)

$$(A\otimes B)(A^{-1}\otimes B^{-1}) = AA^{-1} \otimes BB^{-1} = I_n \otimes I_m = I_{nm},$$

and hence that

$$(A\otimes B)^{-1} = A^{-1} \otimes B^{-1}.$$

A4

We first show the first property, P1:

$$A^{\otimes k}B^{\otimes k} = (AB)^{\otimes k}.$$

Proof. By induction. The base case is trivial:

$$A^{\otimes 1}B^{\otimes 1} = AB = (AB)^{\otimes 1}.$$

Next, we assume the property holds for k, and we prove it for $k+1$:

$$\begin{aligned}
A^{\otimes k+1}B^{\otimes k+1} &= \left(A^{\otimes k}\otimes A\right)\left(B^{\otimes k}\otimes B\right) \\
&\overset{\text{A3}}{=} \left(A^{\otimes k}B^{\otimes k}\right)\otimes AB \\
&= (AB)^{\otimes k} \otimes AB \\
&= (AB)^{\otimes k+1}.
\end{aligned}$$

□

Next, we show the second property, P2:

$$(A^{\otimes k})^T = (A^T)^{\otimes k}$$

Proof. We start by proving an auxiliary lemma, L1.

$$(A\otimes B)^\top = \begin{bmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{bmatrix}^\top = \begin{bmatrix} a_{11}B^\top & \cdots & a_{m1}B^\top \\ \vdots & \ddots & \vdots \\ a_{1n}B^\top & \cdots & a_{mn}B^\top \end{bmatrix} = A^\top \otimes B^\top.$$

We then proceed by induction. The base case is trivial as before:

$$\left(A^{\otimes 1}\right)^\top = A^\top = \left(A^\top\right)^{\otimes 1}.$$

Next, we assume the property holds for k, and we prove it for $k+1$:

$$\begin{aligned}
\left(A^{\otimes k+1}\right)^\top &= \left(A^{\otimes k}\otimes A\right)^\top \\
&\overset{\text{L1}}{=} \left(A^{\otimes k}\right)^\top \otimes A^\top \\
&= \left(A^\top\right)^{\otimes k} \otimes A^\top \\
&= \left(A^\top\right)^{\otimes k+1}.
\end{aligned}$$

□

Finally, we show the following: for any vector $v \in \mathbb{R}^n$,

$$\|v^{\otimes k}\| = \|v\|^k,$$

where $\|v\| = \sqrt{v^\top v}$.

Proof.

$$\begin{aligned}
\|v^{\otimes k}\| &= \sqrt{\left(v^{\otimes k}\right)^\top v^{\otimes k}} \\
&\overset{\text{P2}}{=} \sqrt{\left(v^\top\right)^{\otimes k} v^{\otimes k}} \\
&\overset{\text{P1}}{=} \sqrt{(v^\top v)^{\otimes k}} \\
&= \sqrt{\left(v^\top v\right)^k} \\
&= \left(\sqrt{v^\top v}\right)^k \\
&= \|v\|^k,
\end{aligned}$$

where the fourth equality follows from a simplification of the Kronecker product for scalars, and the fifth equality is a property of the square root. □

A5

The determinant of a square matrix $A \in \mathbb{F}^{n\times n}$ is given by

$$\det(A) = \sum_{\boldsymbol{j}} (-1)^{t(\boldsymbol{j})} a_{1j_1} \cdot a_{2j_2} \cdots a_{nj_n},$$

where the index vector $\boldsymbol{j}$ constitutes a permutation of $\{1, 2, \ldots, n\}$, and $t(\boldsymbol{j})$ denotes the parity of each quasi-diagonal.

First, we show that $\det(A \otimes I_m) = \det(A)^m$.

Proof. Laplace's theorem states that for a matrix $B \in \mathbb{F}^{n\times n}$ and a p-tuple of rows $\boldsymbol{i}_p$, we have:

$$\det(B) = \sum_{\boldsymbol{j}_p} B\begin{pmatrix}\boldsymbol{i}_p \\ \boldsymbol{j}_p\end{pmatrix} B^c \begin{pmatrix}\boldsymbol{i}_p \\ \boldsymbol{j}_p\end{pmatrix}.$$

We apply this theorem to the matrix $B = A \otimes I_m$ with the n-tuple $\boldsymbol{i}_n = (1, m+1, 2m+1, \ldots, (n-1)m+1)$ (if $n = 1$ then the tuple is just (1)). For every n-tuple $\boldsymbol{j}_n$ that contains another index than those present in $\boldsymbol{i}_n$, the minor $B\binom{\boldsymbol{i}_n}{\boldsymbol{j}_n}$ is zero. Indeed, if we consider a new matrix B' only containing the rows whose indices are in $\boldsymbol{i}_n$, we have:

$$B' = \begin{bmatrix}
\underbrace{\begin{matrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & 0 & \cdots & 0 \end{matrix}}_{m \text{ columns}} & \begin{matrix} a_{12} & 0 & \cdots & 0 \\ a_{22} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n2} & 0 & \cdots & 0 \end{matrix} & \begin{matrix} \cdots \\ \cdots \\ \ddots \\ \cdots \end{matrix} & \begin{matrix} a_{1n} & 0 & \cdots & 0 \\ a_{2n} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{nn} & 0 & \cdots & 0 \end{matrix}
\end{bmatrix}.$$

Thus if $\boldsymbol{j}_n$ contains any index not present in $\boldsymbol{i}_n$, a column of zeros is included leading to a zero minor. We are left with

$$\begin{aligned}
\det(A \otimes I_m) = \det(B) &= B\begin{pmatrix}\boldsymbol{i}_n \\ \boldsymbol{i}_n\end{pmatrix} B^c \begin{pmatrix}\boldsymbol{i}_n \\ \boldsymbol{i}_n\end{pmatrix} & (1) \\
&= \det(A)\det(A \otimes I_{m-1}). & (2)
\end{aligned}$$

The first term of (2) can be easily found by inspecting B' and taking only columns whose indices are in $\boldsymbol{i}_n$. The second term is derived by removing the rows and columns of $A \otimes I_m$ whose indices are in $\boldsymbol{i}_n$, which gives the following matrix:

$$\begin{bmatrix} a_{11}I_{m-1} & \cdots & a_{1n}I_{m-1} \\ \vdots & \ddots & \vdots \\ a_{n1}I_{m-1} & \cdots & a_{nn}I_{m-1} \end{bmatrix}.$$

From (2), we then conclude that $\det(A \otimes I_m) = \det(A)^m$. □

Similarly, we can also prove that $\det(I_m \otimes A) = \det(A)^m$.

Proof. Indeed, if we take $B = I_m \otimes A$ and $\boldsymbol{i_n} = (1, 2, \dots n)$, then we have :

$$B' = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & \underbrace{0 \quad \cdots \quad 0}_{(m-1)n \text{ columns}} \end{bmatrix}$$

With the same argument as before, if $\boldsymbol{j}_n$ contains any index not present in $\boldsymbol{i}_n$, a column of zeros is included leading to a zero minor.

We are left with

$$\begin{aligned}\det(I_m \otimes A) = \det(B) &= B\binom{\boldsymbol{i}_n}{\boldsymbol{i}_n} B^c \binom{\boldsymbol{i}_n}{\boldsymbol{i}_n} \\ &= \det(A)\det(I_{m-1} \otimes A).\end{aligned}$$

Which leads us to $\det(I_m \otimes A) = \det(A)^m$.

□

From this, we can deduce that for $A \in \mathbb{F}^{n\times n}$ and $B \in \mathbb{F}^{m\times m}$, $\det(A \otimes B) = \det(A)^m \det(B)^n$.

Proof. We can write

$$\begin{aligned} A \otimes B &= (AI_n) \otimes (I_m B) \\ &\overset{A3}{=} (A \otimes I_m)(I_n \otimes B). \end{aligned}$$

Taking the determinant on both sides, and using the fact that $\det(AB) = \det(A)\det(B)$ (Exercise 1.18 in the lecture notes), we then get

$$\begin{aligned} \det(A \otimes B) &= \det(A \otimes I_m)\det(I_n \otimes B) \\ &= \det(A)^m \det(B)^n. \end{aligned}$$

□

A6

The rank of a matrix $A \in \mathbb{F}^{m\times n}$ is equal to the largest size of its nonzero minors. From this, we prove the following property: $\operatorname{rank}(A \otimes B) = \operatorname{rank}(A)\operatorname{rank}(B) = \operatorname{rank}(B \otimes A)$.

Proof. Let $A \in \mathbb{F}^{m\times n}, B \in \mathbb{F}^{p\times q}$.

First, we note that $B \otimes A$ can be obtained by permuting rows and columns of $A \otimes B$. As elementary operations do not affect the rank of a matrix, we deduce that $\operatorname{rank}(A \otimes B) = \operatorname{rank}(B \otimes A)$.

Let R_1 and Q_1 be products of elementary transformations such that

$$R_1 B Q_1 = \begin{bmatrix} I_r & 0_{r\times(q-r)} \\ 0_{(p-r)\times r} & 0_{(p-r)\times(q-r)} \end{bmatrix}.$$

By Theorem 1.8 of the lecture notes, we know such matrices exist. The scalar r is the rank of B.

Next, we multiply on both sides the matrix $A \otimes B$ by matrices with R_1 and Q_1 on the diagonal:

$$\begin{aligned}
&\begin{bmatrix} R_1 & & \\ & \ddots & \\ & & R_1 \end{bmatrix} \begin{bmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{bmatrix} \begin{bmatrix} Q_1 & & \\ & \ddots & \\ & & Q_1 \end{bmatrix} \\
&= \begin{bmatrix} a_{11}R_1BQ_1 & \cdots & a_{1n}R_1BQ_1 \\ \vdots & & \vdots \\ a_{m1}R_1BQ_1 & \cdots & a_{mn}R_1BQ_1 \end{bmatrix} \\
&= \begin{bmatrix} a_{11}I_r & 0_{r\times(q-r)} & \cdots & a_{1n}I_r & 0_{r\times(q-r)} \\ 0_{(p-r)\times r} & 0_{(p-r)\times(q-r)} & \cdots & 0_{(p-r)\times r} & 0_{(p-r)\times(q-r)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1}I_r & 0_{r\times(q-r)} & \cdots & a_{mn}I_r & 0_{r\times(q-r)} \\ 0_{(p-r)\times r} & 0_{(p-r)\times(q-r)} & \cdots & 0_{(p-r)\times r} & 0_{(p-r)\times(q-r)} \end{bmatrix}.
\end{aligned}$$

By manipulating the last matrix with permutation matrices, the following matrix is obtained:

$$\begin{bmatrix} \begin{matrix} A & & \\ & \ddots & \\ & & A \end{matrix} & 0_{mr\times n(q-r)} \\ 0_{m(p-r)\times nr} & 0_{m(p-r)\times n(q-r)} \end{bmatrix},$$

where the matrix A appears r times on the diagonal and the rest of the matrix is only zeros. Its rank is still equal to the rank of $A \otimes B$ as only elementary operations have been applied.

Let R_2 and Q_2 be products of elementary transformations such that:

$$R_2AQ_2 = \begin{bmatrix} I_s & 0_{s\times(n-s)} \\ 0_{(m-s)\times s} & 0_{(m-s)\times(n-s)} \end{bmatrix}.$$

We multiply on both sides the matrix obtained previously by matrices with R_2 and Q_2 on the diagonal :

$$\begin{bmatrix} R_2 & & & \\ & \ddots & & 0_{mr\times m(p-r)} \\ & & R_2 & \\ 0_{m(p-r)\times nr} & & & 0_{m(p-r)\times m(p-r)} \end{bmatrix} \begin{bmatrix} A & & & \\ & \ddots & & 0_{nr\times n(q-r)} \\ & & A & \\ 0_{n(q-r)\times nr} & & & 0_{n(q-r)\times n(q-r)} \end{bmatrix} \begin{bmatrix} Q_2 & & & \\ & \ddots & & 0_{nr\times n(q-r)} \\ & & Q_2 & \\ 0_{n(q-r)\times nr} & & & 0_{n(q-r)\times n(q-r)} \end{bmatrix}$$

$$= \begin{bmatrix} R_2AQ_2 & & & \\ & \ddots & & 0_{mr\times n(q-r)} \\ & & R_2AQ_2 & \\ 0_{m(p-r)\times nr} & & & 0_{m(p-r)\times n(q-r)} \end{bmatrix}$$

$$= \begin{bmatrix} \begin{matrix} I_s & 0_{s\times(n-s)} \\ 0_{(m-s)\times s} & 0_{(m-s)\times(n-s)} \end{matrix} & & & \\ & \ddots & & 0_{mr\times n(q-r)} \\ & & \begin{matrix} I_s & 0_{s\times(n-s)} \\ 0_{(m-s)\times s} & 0_{(m-s)\times(n-s)} \end{matrix} & \\ & 0_{m(p-r)\times nr} & & 0_{m(p-r)\times n(q-r)} \end{bmatrix}.$$

As the identity matrix I_s appears r times on the diagonal, we deduce that

$$\operatorname{rank}(A\otimes B) = sr = \operatorname{rank}(A)\operatorname{rank}(B). \qquad \square$$

A7

We show that : $\operatorname{vec}(AXB) = \left(B^\top \otimes A\right)\operatorname{vec}(X)$.

Proof. Let $A \in \mathbb{F}^{m\times n}, B \in \mathbb{F}^{p\times q}$, and $X \in \mathbb{F}^{n\times p}$.

We develop the right-hand side of the equality we want to prove:

$$\begin{aligned} \left(B^\top \otimes A\right)\operatorname{vec}(X) &= \begin{bmatrix} b_{11}A & b_{21}A & \cdots & b_{p1}A \\ b_{12}A & b_{22}A & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ b_{1q}A & \cdots & \cdots & b_{pq}A \end{bmatrix} \begin{bmatrix} X_{:,1} \\ X_{:,2} \\ \vdots \\ X_{:,p} \end{bmatrix} \\ &= \begin{bmatrix} b_{11}AX_{:,1} + b_{21}AX_{:,2} + \cdots + b_{p1}AX_{:,p} \\ b_{12}AX_{:,1} + \cdots + b_{p2}AX_{:,p} \\ \vdots \\ b_{1q}AX_{:,1} + \cdots + b_{pq}AX_{:,p} \end{bmatrix}. \end{aligned}$$

We recognize the elements of a product D of three matrices in vectorized form:

$$d_i = \sum_{r=1}^{p} b_{ri} \sum_{k=1}^{n} a_{rk}x_{kr},$$

which shows D is simply AXB, and hence proves $\operatorname{vec}(AXB) = \left(B^\top \otimes A\right)\operatorname{vec}(X)$. $\square$

The proven equality can be used to solve the Sylvester equation: $AX + XA^\top = B$ where X is the unknown. Indeed, we can vectorize both sides of the equation:

$$\operatorname{vec}(AX) + \operatorname{vec}\left(XA^\top\right) = \operatorname{vec}(B).$$

Then, we use some identity matrices to be able to apply the proven relation:

$$\begin{aligned}
\operatorname{vec}(B) = \operatorname{vec}(AX) + \operatorname{vec}\left(XA^\top\right) &= \operatorname{vec}(AXI) + \operatorname{vec}\left(IXA^\top\right) \\
&= (I \otimes A)\operatorname{vec}(X) + (A \otimes I)\operatorname{vec}(X) \\
&= (I \otimes A + A \otimes I)\operatorname{vec}(X).
\end{aligned}$$

The term $\operatorname{vec}(X)$ can then be isolated:

$$\operatorname{vec}(X) = (I \otimes A + A \otimes I)^{-1}\operatorname{vec}(B).$$

Finally, the matrix X can be simply reconstructed from $\operatorname{vec}(X)$.

1 Exercise B: The matrix exponential

B1

If $\lambda \in \mathbb{C}$ is an eigenvalue of A then e^λ is an eigenvalue of e^A.

Proof. We know λ is an eigenvalue of A. Hence $Av = \lambda v$ for some eigenvector v.

$$\begin{aligned}
e^A v &= \left(I + \sum_{k=1}^{\infty} \frac{1}{k!} A^k\right) v \\
&= v + \sum_{k=1}^{\infty} \frac{1}{k!} A^k v \\
&= v + \sum_{k=1}^{\infty} \frac{1}{k!} \lambda^k v \\
&= \left(\sum_{k=0}^{\infty} \frac{1}{k!} \lambda^k\right) v \\
&= e^\lambda v.
\end{aligned}$$

The third line is derived using the equality $A^k v = \lambda^k v$ and the last line using the Taylor series of the exponential function.

This proves that e^λ is an eigenvalue of e^A, with eigenvector v. $\square$

B2

For any matrix $A \in \mathbb{C}^{n\times n}$, $\operatorname{rank}(e^A) = n$.

Proof. The matrix e^A is invertible and its inverse is e^{-A}. In fact,

$$e^A e^{-A} = e^{-A} e^A = e^{A-A} = I.$$

The second equality is valid since A and $-A$ commute. Since e^A is non-singular, its determinant is nonzero. Theorem 1.9 of the course notes states the rank of e^A is equal to the largest size of its nonzero minors which in this case is the whole matrix. We deduce $\operatorname{rank}(e^A) = n$. $\square$

B3

If $A \in \mathbb{C}^{n\times n}$ is skew-Hermitian, i.e. $A = -A^*$, then e^A is unitary, i.e. $e^A(e^A)^* = I$.

Proof. We have:

$$\begin{aligned}
(e^A)^* &= \left(I + \sum_{k=1}^{\infty} \frac{1}{k!} A^k\right)^* \\
&= \left(I^* + \sum_{k=1}^{\infty} \frac{1}{k!} (A^k)^*\right) \\
&= \left(I + \sum_{k=1}^{\infty} \frac{1}{k!} (A^*)^k\right) \\
&= e^{A^*} \\
&= e^{-A},
\end{aligned}$$

where the second and third lines are derived using the following properties:

- $(A+B)^* = A^* + B^*$;
- $\left(A^k\right)^* = \left(A^*\right)^k$.

The fourth line comes from the definition of the exponential matrix and the last line from the fact that A is skew-Hermitian.

Finally can write :

$$\begin{aligned}
e^A(e^A)^* &= e^A(e^{-A}) \\
&= e^{A-A} \\
&= I.
\end{aligned}$$

The second line is valid since A and $-A$ commute. □

B4

We start by showing that $\left(J_n(0)\right)^n = 0$.

Proof. One trivially notes that $A := J_n(0) \in \mathbb{C}^{n\times n}$ is strictly upper triangular, for any $n \geqslant 1$. We want to prove by induction on $k \geqslant 1$ that $a_{ij}^k = 0$ whenever $i \leqslant j + k - 1$.

- When $k = 1$, this is trivially true, as this precisely states that A is strictly upper triangular.
- For the induction step, we fix $k \geqslant 1$ and we suppose that $a_{ij}^k = 0$ when $i \leqslant j + k - 1$.

  Fix i, j and suppose that $i \leqslant j + (k+1) - 1 = j + k$. In that case,

  $$a_{ij}^{k+1} = \sum_{\ell=0}^{n} a_{i\ell}^k a_{\ell j}.$$

  However, regarding the base case, we know that $a_{\ell j} = 0$ whenever $\ell \leqslant j$, so that

  $$a_{ij}^{k+1} = \sum_{\ell=j+1}^{n} a_{i\ell}^k a_{\ell j}.$$

  Moreover, keeping in mind $i \leqslant j + k$, we observe that if $j + 1 \leqslant \ell \leqslant n$, then

  $$i \leqslant j + k = (j+1) + k - 1 \leqslant \ell + k - 1$$

  so that $a_{i\ell}^k = 0$ by the induction hypothesis. This allows us to conclude that $a_{ij}^{k+1} = 0$ when $i \leqslant j + (k+1) - 1$, as desired.

Finally, the result is obtained by taking the case $k = n$, which tells us that $a_{ij}^n = 0$ whenever $i \leqslant j + n - 1$, but this is always true as $i \leqslant n$ and $j + n - 1 \geqslant n$ for all $1 \leqslant i, j \leqslant n$. This shows that $A = \left(J_n(0)\right)^n = 0$. □

Next, we show that

$$e^{J_n(\lambda)} = e^{\lambda}\left(I + \sum_{k=1}^{n-1} \frac{1}{k!}\left(J_n(0)\right)^k\right).$$

Proof. We note that $J_n(\lambda) = J_n(0) + \lambda I$. Furthermore, $J_n(0)$ and λI commute as the identity matrix multiplied by a scalar commutes with every matrix. Thus, we can write:

$$\begin{aligned} e^{J_n(\lambda)} &= e^{\lambda I + J_n(0)} \\ &= e^{\lambda I} e^{J_n(0)}. \end{aligned}$$

If we develop the last right-hand side and use the Taylor series expansion of the exponential, we obtain

$$\begin{aligned} e^{J_n(0)} e^{\lambda I} &= \sum_{j=0}^{\infty} \frac{1}{j!}(\lambda I)^j \left(\sum_{k=0}^{\infty} \frac{1}{k!}\left(J_n(0)\right)^k\right) \\ &= I \sum_{j=0}^{\infty} \frac{1}{j!}\lambda^j \sum_{k=0}^{\infty} \frac{1}{k!}\left(J_n(0)\right)^k \\ &= e^{\lambda}\left(\sum_{k=0}^{\infty} \frac{1}{k!}\left(J_n(0)\right)^k\right) \\ &= e^{\lambda}\left(I + \sum_{k=1}^{\infty} \frac{1}{k!}\left(J_n(0)\right)^k\right). \end{aligned}$$

Putting the left-hand side in its original form, this is precisely what we were trying to prove. □

B5

We want to show that for any matrices $A, B \in \mathbb{C}^{n \times n}$,

$$e^{A \otimes I + I \otimes B} = e^A \otimes e^B.$$

Proof. We first note that $A \otimes I$ and $I \otimes B$ commute. Indeed, by using the mixed product property, we observe that

$$(A \otimes I)(I \otimes B) = (AI) \otimes (IB) = (I \otimes B)(A \otimes I).$$

Hence we can write

$$e^{A \otimes I + I \otimes B} = e^{A \otimes I} e^{I \otimes B}.$$

Next, we develop the last equality using the definition of the matrix exponential:

$$\begin{aligned} e^{A \otimes I} e^{I \otimes B} &= \sum_{k=0}^{\infty} \frac{1}{k!}(A \otimes I)^k \sum_{j=0}^{\infty} \frac{1}{j!}(I \otimes B)^j \\ &= \sum_{k=0}^{\infty}\sum_{j=0}^{\infty} \frac{1}{k!}\frac{1}{j!}(A \otimes I)^k (I \otimes B)^j \\ &= \sum_{k=0}^{\infty}\sum_{j=0}^{\infty} \frac{(k+j)!}{(k+j)!}\frac{1}{k!}\frac{1}{j!}(A \otimes I)^k (I \otimes B)^j \\ &= \sum_{k=0}^{\infty}\sum_{j=0}^{\infty} \frac{1}{(k+j)!}\binom{k+j}{j}(A \otimes I)^k (I \otimes B)^j. \end{aligned}$$

Let $m = k + j$; we can then write this as

$$\begin{aligned}
&= \sum_{m=0}^{\infty} \sum_{j=0}^{m} \frac{1}{m!} \binom{m}{j} (A \otimes I)^{m-j} (I \otimes B)^{j} \\
&= \sum_{m=0}^{\infty} \frac{1}{m!} (A \otimes I + I \otimes B)^{m} \\
&= e^{A \otimes I + I \otimes B},
\end{aligned}$$

where the second equality comes from the binomial formula, which we can use because of the commutativity we observed earlier. Note that the bounds on the sum indexed by j have changed; this is because of the change of variables and is consistent with the definition of $m \geqslant j$. $\square$