Throughout this homework, the notation "$x \mid y$" will be used to mean "x divides y".

Exercise A: Minimal polynomial and Smith normal form

A1

Let us first prove the following lemma.

Lemma. *Let $A(\lambda), B(\lambda) \in \mathbb{C}^{n \times n}[\lambda]$. For $k = 1, \ldots, n$,*

- $\delta_k\big(A(\lambda)\big) \mid \delta_k\big(A(\lambda)B(\lambda)\big)$;

- $\delta_k\big(B(\lambda)\big) \mid \delta_k\big(A(\lambda)B(\lambda)\big)$.

Proof. For clarity, we write A, B instead of $A(\lambda), B(\lambda)$. Let $\boldsymbol{i}, \boldsymbol{j}$ be index k-tuples. If one denotes by $X_{\boldsymbol{ij}}$ the $k \times k$ submatrix of a matrix X taking the rows whose indices are in $\boldsymbol{i}$ and the columns whose indices are in $\boldsymbol{j}$, one can write and develop the associated k-minor, which is its determinant (using bold notation for tuples):

$$\det(AB)_{\boldsymbol{ij}} = \det(A_{\boldsymbol{in}}B_{\boldsymbol{nj}})$$
$$= \sum_{\boldsymbol{\ell}_k} A\binom{\boldsymbol{i}}{\boldsymbol{\ell}_k} B\binom{\boldsymbol{\ell}_k}{\boldsymbol{j}}$$
$$= \sum_{\boldsymbol{\ell}_k} \det A_{\boldsymbol{i\ell}_k} \det B_{\boldsymbol{\ell}_k \boldsymbol{j}},$$

where the second equality is a consequence of the Cauchy–Binet theorem. One observes that the k-minors of AB can be written as linear combinations of the k-minors of A. This implies that every common divisor of the k-minors of A must also be a common divisor of the k-minors of AB. When applied to the greatest common divisor, one finds that $\delta_k(A) \mid \delta_k(AB)$. The same reasoning is valid when taking B instead of A and therefore one has proven that $\delta_k(B) \mid \delta_k(AB)$. $\square$

One can now show that $\delta_k\big(P(\lambda)\big) = \delta_k\big(Q(\lambda)\big)$, where $Q(\lambda) = M(\lambda)P(\lambda)N(\lambda)$.

Proof. Using the lemma above, with $A = M(\lambda)P(\lambda)$ and $B = N(\lambda)$, one finds that $\delta_k\big(M(\lambda)P(\lambda)\big) \mid \delta_k\big(Q(\lambda)\big)$. Similarly, with $A = \delta_k\big(M(\lambda)\big)$ and $B = \delta_k\big(P(\lambda)\big)$, one obtains that $\delta_k(P(\lambda)) \mid \delta_k\big(M(\lambda)P(\lambda)\big)$. One can deduce that $\delta_k\big(P(\lambda)\big)$ divides $\delta_k\big(Q(\lambda)\big)$.

Using the fact that unimodular matrices are invertible as a consequence of their definition, one can write $P(\lambda) = M^{-1}(\lambda)Q(\lambda)N^{-1}(\lambda)$. Applying a similar reasoning as above, one finds that $\delta_k\big(Q(\lambda)\big)$ divides $\delta_k\big(P(\lambda)\big)$. Consequently, we conclude that $\delta_k\big(Q(\lambda)\big) = \delta_k\big(P(\lambda)\big)$. $\square$

A2

One can show that for all $k \in \{1, \ldots, n\}$, $\delta_k\big(P(\lambda)\big) = \prod_{i=1}^{k} d_i(\lambda)$, where $d_i(\lambda)$ are the diagonal entries of $D(\lambda)$ and $M(\lambda)D(\lambda)N(\lambda)$ is the Smith normal decomposition of $P(\lambda)$.

Proof. From applying A1 on $P(\lambda)$ and its Smith normal form, one knows that $\delta_k\big(P(\lambda)\big) = \delta_k\big(D(\lambda)\big)$. Therefore, it suffices to consider only $\delta_k\big(D(\lambda)\big)$, without loss of generality. One can first note that all submatrices of $D(\lambda)$ are either triangular or zero. The determinants of these submatrices are hence given by the product of their diagonal elements. One can also note that only submatrices for which the determinant is nonzero are relevant when finding the gcd (as every number divides 0); these are matrices where the indices of the removed rows are identical to the indices of the removed columns. If one denotes by $\boldsymbol{j}_k = (j_1, \ldots, j_k)$, where $1 \leqslant j_1 < \cdots < j_k \leqslant n$, the indices of the rows and columns that are kept, the determinant of the minor associated to $\boldsymbol{j}_k$ can be written as:

$$\det D_{\boldsymbol{j}_k}(\lambda) = d_{j_1}(\lambda) \ldots d_{j_k}(\lambda) = \prod_{i=1}^{k} d_{j_i}(\lambda).$$

One can now use the property that $d_i(\lambda) \mid d_{i+1}(\lambda)$ for all $i \in \{1, \ldots, n-1\}$. If one considers the i-th factor $d_{j_i}(\lambda)$, one can observe that $d_i(\lambda) \mid d_{j_i}(\lambda)$ as $j_i \geqslant i$ (due to the strictness of the inequalities in

$j_1 < \cdots < j_k$). Hence $d_{j_i}(\lambda)$ can be rewritten as $a_i(\lambda)d_i(\lambda)$ for a certain $a_i(\lambda) \in \mathbb{C}[\lambda]$. The determinant can thus be expanded as

$$\det D_{\boldsymbol{j}_k}(\lambda) = a_1(\lambda)d_1(\lambda)\ldots a_k(\lambda)d_k(\lambda) = \prod_{i=1}^{k} a_i(\lambda)d_i(\lambda). \tag{1}$$

One can also observe that when $\boldsymbol{j}_k = \boldsymbol{k} = (1,\ldots,k)$, we have

$$\det D_{\boldsymbol{k}}(\lambda) = \prod_{i=1}^{k} d_i(\lambda), \tag{2}$$

as in that case $j_i = i$ and thus $a_i = 1$. This determinant is a monic polynomial, as it is a product of monic polynomials.

Because $\prod_{i=1}^{k} d_i(\lambda)$ is a divisor of (1) for any $\boldsymbol{j}_k$ and as $\delta_k\big(D(\lambda)\big)$ must divide the result of (2), one can conclude that the monic greatest common divisor of all k-minors of $D(\lambda)$ is $\prod_{i=1}^{k} d_i(\lambda)$. One can finally write

$$\delta_k\big(P(\lambda)\big) = \delta_k\big(D(\lambda)\big) = \prod_{i=1}^{k} d_i(\lambda). \tag{3}$$

$\square$

Next, one can show that the diagonal entries $d_1(\lambda),\ldots,d_n(\lambda)$ of $D(\lambda)$ are unique and depend only on the sequence $\delta_1\big(P(\lambda)\big),\ldots,\delta_n\big(P(\lambda)\big)$.

Proof. By induction. One should first consider $d_1(\lambda)$. From (3), one can infer that $d_1(\lambda) = \delta_1\big(P(\lambda)\big)$. As $\delta_k(\cdot)$ denotes a monic greatest common divisor, it must be unique. The uniqueness of $\delta_1\big(P(\lambda)\big)$ thus implies that $d_1(\lambda)$ is unique as well. This constitutes the base case of the induction proof. Next, one can show that if $d_k(\lambda)$ is unique then $d_{k+1}(\lambda)$ is also unique. Again, from (3), one has that

$$d_{k+1}(\lambda) = \frac{\delta_{k+1}\big(P(\lambda)\big)}{d_k(\lambda)}. \tag{4}$$

By uniqueness of the monic greatest common divisor and assumed uniqueness of $d_k(\lambda)$, one can conclude that $d_{k+1}\big(P(\lambda)\big)$ is also unique.

Clearly the diagonal entries $d_1(\lambda),\ldots,d_n(\lambda)$ can be computed using the sequence $\delta_1\big(P(\lambda)\big),\ldots,\delta_n\big(P(\lambda)\big)$. As shown previously, one has $d_1(\lambda) = \delta_1\big(P(\lambda)\big)$, and by recursively applying (4), one can obtain $d_k(\lambda)$ for $k > 1$. $\square$

Finally, one can show that there are unimodular matrices $M(\lambda), N(\lambda) \in \mathbb{C}^{n \times n}[\lambda]$ such that $Q(\lambda) = M(\lambda)P(\lambda)N(\lambda)$ if and only if $\delta_k\big(P(\lambda)\big) = \delta_k\big(Q(\lambda)\big)$ for all $k \in \{1,\ldots,n\}$.

Proof. The necessary condition was proven in A1. One thus only needs to show the sufficient condition.

One knows that the diagonal entries of $D(\lambda)$ in the Smith normal decomposition of $P(\lambda)$ are unique and depend only on the sequence $\delta_1\big(P(\lambda)\big),\ldots,\delta_n\big(P(\lambda)\big)$. As this sequence is the same for $P(\lambda)$ and $Q(\lambda)$, one can deduce that the diagonal matrix in their Smith normal forms must be the same:

$$P(\lambda) = M_1(\lambda)D(\lambda)N_1(\lambda),$$
$$Q(\lambda) = M_2(\lambda)D(\lambda)N_2(\lambda),$$

where $M_1(\lambda), N_1(\lambda), M_2(\lambda), N_2(\lambda) \in \mathbb{C}^{n \times n}[\lambda]$ are unimodular and $D(\lambda) \in \mathbb{C}^{n \times n}[\lambda]$ is diagonal. One can then write $D(\lambda) = M_1(\lambda)^{-1}P(\lambda)N_1(\lambda)^{-1}$ as $M_1(\lambda)$ and $N_1(\lambda)$ are unimodular and hence invertible. Substituting this in the expression for $Q(\lambda)$, one obtains

$$Q(\lambda) = M_2(\lambda)M_1(\lambda)^{-1}P(\lambda)N_1(\lambda)^{-1}N_2(\lambda).$$

Finally, as the inverse of a unimodular matrix is unimodular and a product of unimodular matrices is also unimodular, we have $Q(\lambda) = M(\lambda)P(\lambda)N(\lambda)$, where $M(\lambda) = M_2(\lambda)M_1(\lambda)^{-1}$ and $N(\lambda) = N_1(\lambda)^{-1}N_2(\lambda)$ are unimodular matrices. $\square$

A3

Let $A \in \mathbb{C}^{n \times n}$ be a Jordan block with eigenvalue λ_1. The elementary polynomials of $\lambda I - A$ are $d_i(\lambda) = 1$ for $i = 1, \ldots, n-1$ and $d_n(\lambda) = (\lambda - \lambda_1)^n$.

Proof. One starts by sketching the matrix $P \triangleq \lambda I - A$:

$$P(\lambda) \triangleq \lambda I - A = \begin{bmatrix} \lambda - \lambda_1 & -1 & & \\ & \ddots & \ddots & \\ & & \ddots & -1 \\ & & & \lambda - \lambda_1 \end{bmatrix}.$$

As discussed previously, the sequence $d_1(\lambda), \ldots, d_n(\lambda)$ depends only on $\delta_1(P(\lambda)), \ldots, \delta_n(P(\lambda))$. Let us start by first deriving the latter sequence. When computing each $\delta_k(P(\lambda))$, we observe that for $k < n$, a valid submatrix which needs to be taken into account is one which has all -1 entries on its diagonal. We illustrate here the case where $k = n - 1$:

$$P(\lambda) = \begin{bmatrix} \lambda - \lambda_1 & -1 & & \\ & \ddots & \ddots & \\ & & \ddots & -1 \\ & & & \lambda - \lambda_1 \end{bmatrix}.$$

As this submatrix is lower triangular, its determinant is equal to the product of its diagonal entries. The determinant of such matrices is therefore ± 1. As 1 and -1 have only one monic common divisor, which is 1, one concludes that $\delta_k(P(\lambda)) = 1$ for $k \in \{1, \ldots, n-1\}$. From (4), one can deduce that $d_k(\lambda) = 1$ for $k \in \{1, \ldots, n-1\}$.

When $k = n$, the only possible submatrix to consider is the whole matrix, whose determinant is equal to the product of the diagonal entries, $(\lambda - \lambda_1)^n$. One can deduce that $\delta_n(P(\lambda)) = (\lambda - \lambda_1)^n$. Applying (4) then yields $d_n(\lambda) = (\lambda - \lambda_1)^n$. $\qquad\square$

A4

Let $J_1 \in \mathbb{C}^{n_1 \times n_1}$ and $J_2 \in \mathbb{C}^{n_2 \times n_2}$ be two Jordan blocks with respective eigenvalues λ_i and sizes n_i, $i = 1, 2$. Let $A \in \mathbb{C}^{n \times n}$ with $n = n_1 + n_2$ be the Jordan matrix consisting of the two Jordan blocks J_1 and J_2. Then $\lambda I - A$ can be reduced to the form

$$M(\lambda)(\lambda I - A)N(\lambda) = \mathrm{diag}\{ \underbrace{1, \ldots, 1}_{n_1 - 1 \text{ times}}, (\lambda - \lambda_1)^{n_1}, \underbrace{1, \ldots, 1}_{n_2 - 1 \text{ times}}, (\lambda - \lambda_2)^{n_2} \},$$

with unimodular matrices $M(\lambda), N(\lambda) \in \mathbb{C}^{n \times n}[\lambda]$.

Proof. From A3, one knows that $P_i(\lambda) \triangleq \lambda I - J_i$ (for $i = 1, 2$) can be written as

$$\begin{aligned} P_i(\lambda) &\triangleq \lambda I_{n_i} - J_i \\ &= M_i(\lambda) D_i(\lambda) N_i(\lambda) \\ \Longleftrightarrow D_i &= M_i^{-1}(\lambda) P_i(\lambda) N_i^{-1}(\lambda) \\ &= \begin{bmatrix} 1 & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & \\ & & & & (\lambda - \lambda_i)^{n_i} \end{bmatrix}. \end{aligned}$$

Note that $M_i^{-1}(\lambda), N_i^{-1}(\lambda)$ are unimodular, as they are inverses of unimodular matrices.

Now, if one considers the diagonal matrix $D(\lambda)$ formed by combining $D_1(\lambda)$ and $D_2(\lambda)$, one finds

$$
D(\lambda) \triangleq \begin{bmatrix} D_1(\lambda) & \\ & D_2(\lambda) \end{bmatrix} = \begin{bmatrix} 1 & & & & & & & & \\ & 1 & & & & & & & \\ & & \ddots & & & & & & \\ & & & 1 & & & & & \\ & & & & (\lambda-\lambda_1)^{n_1} & & & & \\ & & & & & 1 & & & \\ & & & & & & 1 & & \\ & & & & & & & \ddots & \\ & & & & & & & & 1 \\ & & & & & & & & & (\lambda-\lambda_2)^{n_2} \end{bmatrix}
$$

$$
= \begin{bmatrix} M_1^{-1}(\lambda)P_1(\lambda)N_1^{-1}(\lambda) & \\ & M_2^{-1}(\lambda)P_2(\lambda)N_2^{-1}(\lambda) \end{bmatrix}
$$

$$
= \underbrace{\begin{bmatrix} M_1^{-1}(\lambda) & \\ & M_2^{-1}(\lambda) \end{bmatrix}}_{\triangleq M(\lambda)} \underbrace{\begin{bmatrix} P_1(\lambda) & \\ & P_2(\lambda) \end{bmatrix}}_{\triangleq P(\lambda) = (\lambda I - A)} \underbrace{\begin{bmatrix} N_1^{-1}(\lambda) & \\ & N_2^{-1}(\lambda) \end{bmatrix}}_{\triangleq N(\lambda)}.
$$

$M(\lambda)$ is unimodular as its determinant is the product of the determinants of $M_1^{-1}(\lambda)$ and $M_2^{-1}(\lambda)$, both of which are unimodular; the determinant of $M(\lambda)$ is equal to the product of two nonzero constants, which is also a nonzero constant. Similarly, $N(\lambda)$ is also unimodular. $\qquad\square$

Remark. *This result can easily be extended to an arbitrary number of Jordan blocks m.*

$$
\underbrace{\begin{bmatrix} D_1(\lambda) & & \\ & \ddots & \\ & & D_m(\lambda) \end{bmatrix}}_{\triangleq D(\lambda)} = \underbrace{\begin{bmatrix} M_1^{-1}(\lambda) & & \\ & \ddots & \\ & & M_m^{-1}(\lambda) \end{bmatrix}}_{\triangleq M^{-1}(\lambda)} \underbrace{\begin{bmatrix} P_1(\lambda) & & \\ & \ddots & \\ & & P_m(\lambda) \end{bmatrix}}_{\triangleq P(\lambda)=(\lambda I - A)} \underbrace{\begin{bmatrix} N_1^{-1}(\lambda) & & \\ & \ddots & \\ & & N_m^{-1}(\lambda) \end{bmatrix}}_{\triangleq N^{-1}(\lambda)}. \quad (5)
$$

Proof. By induction, let us consider $A = \begin{bmatrix} B & \\ & J_p \end{bmatrix}$, with $B \in \mathbb{C}^{n\times n}$ a block diagonal matrix with $(p-1)$ Jordan blocks and $J_p \in \mathbb{C}^{n_p \times n_p}$ the p-th Jordan block. The previous proof verifies the base case $p = 2$.

For $p > 2$, one assumes that

$$
\underbrace{\begin{bmatrix} D_1(\lambda) & & \\ & \ddots & \\ & & D_{p-1} \end{bmatrix}}_{\triangleq D_B(\lambda)} = \underbrace{\begin{bmatrix} M_1^{-1}(\lambda) & & \\ & \ddots & \\ & & M_{p-1}^{-1}(\lambda) \end{bmatrix}}_{\triangleq M_B^{-1}(\lambda)} \underbrace{\begin{bmatrix} P_1(\lambda) & & \\ & \ddots & \\ & & P_{p-1}(\lambda) \end{bmatrix}}_{\triangleq P_B(\lambda)=(\lambda I - B)} \underbrace{\begin{bmatrix} N_1^{-1}(\lambda) & & \\ & \ddots & \\ & & N_{p-1}^{-1}(\lambda) \end{bmatrix}}_{\triangleq N_B^{-1}(\lambda)}.
$$

Then,

$$
P \triangleq \lambda I - A = \begin{bmatrix} \lambda I - B & \\ & \lambda I - J_p \end{bmatrix}
$$

$$
= \begin{bmatrix} M_B(\lambda)D_B(\lambda)N_B(\lambda) & \\ & \lambda I - J_p \end{bmatrix}
$$

$$
= \begin{bmatrix} M_B(\lambda)D_B(\lambda)N_B(\lambda) & \\ & M_p(\lambda)D_p(\lambda)N_p(\lambda) \end{bmatrix}
$$

$$
= \underbrace{\begin{bmatrix} M_B(\lambda) & \\ & M_p(\lambda) \end{bmatrix}}_{\triangleq M(\lambda)} \underbrace{\begin{bmatrix} D_B(\lambda) & \\ & D_p(\lambda) \end{bmatrix}}_{\triangleq D(\lambda)} \underbrace{\begin{bmatrix} N_B(\lambda) & \\ & N_p(\lambda) \end{bmatrix}}_{\triangleq N(\lambda)}.
$$

From the first equation to the second, the inductive hypothesis was used, and from the second to the third, one simply needs to apply the results of A3. $\qquad\square$

4

Now, one can find the elementary polynomials of $P(\lambda) \triangleq (\lambda I - A)$. For that, the property proved in A1 applied to the Smith decomposition, that is, $\delta_k\big(P(\lambda)\big) = \delta_k\big(D(\lambda)\big)$, will be used.

We wish to bring the reader's attention to the fact that the diagonal matrix $D(\lambda)$ that one obtains here is not the same matrix as the diagonal matrix in the Smith normal form, containing the elementary polynomials $d_i(\lambda)$, despite potentially confusing notations.

Using A1, it is clear that for $k = 1, 2, \ldots, (n_1 + n_2 - 2)$, one can always choose a submatrix having only ones on its diagonal, which means that $\delta_k\big(P(\lambda)\big) = 1$.

Note that for this proof, one need not consider zero minors as they are irrelevant in the search for the (monic) greatest common divisor.

By A2, one has that the monic $d_i(\lambda)$ are all equal to 1 for $i = 1, 2, \ldots, (n_1 + n_2 - 2)$.

One then needs to distinguish two cases:

- When $\lambda_1 \neq \lambda_2$, the $(n_1 + n_2 - 1)$-minors are $(\lambda - \lambda_1)^{n_1}, (\lambda - \lambda_2)^{n_2}$, and $(\lambda - \lambda_1)^{n_1}(\lambda - \lambda_2)^{n_2}$. Their gcd is $\delta_{n_1+n_2-1}\big(P(\lambda)\big) = 1$. Using (4), one finds that $d_{n_1+n_2-1}(\lambda) = 1$. Similarly, the only $(n_1 + n_2)$-minor is $(\lambda - \lambda_1)^{n_1}(\lambda - \lambda_2)^{n_2}$, hence $\delta_{n_1+n_2}\big(P(\lambda)\big) = (\lambda - \lambda_1)^{n_1}(\lambda - \lambda_2)^{n_2}$. Using the recursive formula of (4), one thus finds $d_{n_1+n_2}(\lambda) = (\lambda - \lambda_1)^{n_1}(\lambda - \lambda_2)^{n_2}$.

- When $\lambda_1 = \lambda_2$, the $(n_1 + n_2 - 1)$-minors are $(\lambda - \lambda_1)^{n_1}, (\lambda - \lambda_1)^{n_2}$, and $(\lambda - \lambda_1)^{n_1+n_2}$, hence their gcd is $\delta_{n_1+n_2-1}\big(P(\lambda)\big) = (\lambda - \lambda_1)^{\min\{n_1, n_2\}}$. Applying (4) then yields that $d_{n_1+n_2-1}(\lambda) = (\lambda - \lambda_1)^{\min\{n_1, n_2\}}$. Similarly, the only $(n_1 + n_2)$-minor is $(\lambda - \lambda_1)^{n_1+n_2}$, hence $\delta_{n_1+n_2}\big(P(\lambda)\big) = (\lambda - \lambda_1)^{n_1+n_2}$. Using (4), one can deduce that $d_{n_1+n_2}(\lambda) = (\lambda - \lambda_1)^{\max\{n_1, n_2\}}$.

We can observe that the minimal polynomial of A is equal to $d_{n_1+n_2}(\lambda)$.

Remark. *This result can also be extended to the general case where A has an arbitrary number of Jordan blocks.*

Proof. By induction, let us consider $A = \begin{bmatrix} B & \\ & J_p \end{bmatrix}$, with $B \in \mathbb{C}^{n \times n}$ a block diagonal matrix with $(p-1)$ Jordan blocks and $J_p \in \mathbb{C}^{n_p \times n_p}$ the p-th Jordan block. The base case $p = 2$ has already been proven.

For $p > 2$, assume that the minimal polynomial of B is equal to the last elementary polynomial of $(\lambda I - B)$: $d_n^B(\lambda) = \prod_{i=1}^{k}(\lambda - \lambda_i)^{n_i}$, with λ_i, $i = 1, \ldots, k$, the distinct eigenvalues of A and n_i the size of the largest Jordan block associated to the corresponding eigenvalue.

We know that

$$P(\lambda) \triangleq \lambda I - A = \underbrace{\begin{bmatrix} M_B(\lambda) & \\ & M_p(\lambda) \end{bmatrix}}_{\triangleq M(\lambda)} \underbrace{\begin{bmatrix} D_B(\lambda) & \\ & D_p(\lambda) \end{bmatrix}}_{\triangleq D(\lambda)} \underbrace{\begin{bmatrix} N_B(\lambda) & \\ & N_p(\lambda) \end{bmatrix}}_{\triangleq N(\lambda)}.$$

This means that $\delta_{n+n_p}\big(P(\lambda)\big) = \delta_{n+n_p}\big(D(\lambda)\big) = \prod_{i=1}^{p}(\lambda - \lambda_i)^{n_i'}$ with p the total number of Jordan blocks and n_i' the size of each Jordan block.

As before, one need not consider zero minors, as they are irrelevant in the search for the (monic) greatest common divisor.

We must then consider two cases:

- If $\lambda_i \neq \lambda_p$, for all $i \neq p$, then $(\lambda - \lambda_p)^{n_p'}$ is not a divisor of any i-minor with $i < (n + n_p)$. This means that $d_{n+n_p}(\lambda) = d_n^B(\lambda)(\lambda - \lambda_p)^{n_p'}$. This corresponds to the minimal polynomial of A.

- If there exists $i \neq p$ such that $\lambda_p = \lambda_i$, then let us denote y as the number of Jordan blocks with an eigenvalue equal to λ_p. It is clear that if one chooses any submatrix of dimension $(n + n_p - y + 1)$, there will be at least one occurrence of $(\lambda - \lambda_p)^z$ in the diagonal terms used to compute the minor (with $z \in \mathcal{Z}_p$, $\mathcal{Z}_p$ being the set of all sizes of Jordan blocks with eigenvalue equal to λ_p). This then means that $(\lambda - \lambda_p)^{\min(\mathcal{Z}_p)}$ must divide $\delta_{n+n_p-y+1}\big(P(\lambda)\big)$ and $d_{n+n_p-y+1}^A(\lambda)$. This reasoning can be continued by considering $\mathcal{Z}_p \setminus \min(\mathcal{Z}_p)$, and submatrices of dimension $(n + n_p - y + 2)$, and so forth, continually removing the minimum in the set until one arrives at matrices of dimension $(n + n_p)$, where

$$d_{n+n_p}^A(\lambda) = \frac{\delta_{n+n_p}\big(P(\lambda)\big)}{\prod_{z \in \mathcal{Z}_p \setminus \max \mathcal{Z}_p}(\lambda - \lambda_p)^z},$$

which means that the only contribution of λ_p to $d_{n+n_p}^A(\lambda)$ is $(\lambda - \lambda_p)^{\max(\mathcal{Z}_p)}$. One can then, from $d_{n+n_p}^B(\lambda)$, take out the factor with λ_p in it, and replace it with the new $(\lambda - \lambda_p)^{\max(\mathcal{Z}_p)}$. With this, one obtains the minimal polynomial of A.

The previous conclusion for $p = 2$ thus still holds in the general case. $\square$

A5

Let $A \in \mathbb{C}^{n \times n}$. There are unimodular matrices $M(\lambda), N(\lambda) \in \mathbb{C}^{n \times n}[\lambda]$ such that

$$\lambda I - A = M(\lambda) E(\lambda) N(\lambda),$$

where $E(\lambda)$ is diagonal with its elements being either 1's or polynomials of the form $(\lambda - \lambda_i)^{n_i}$ where λ_i are the eigenvalues of A and n_i the size of the corresponding Jordan block in the Jordan decomposition of A.

Proof. First, let us write the Jordan decomposition of A.

$$T^{-1}AT = J \iff A = TJT^{-1},$$

with J the Jordan matrix and T a similarity transformation (this implies that T is invertible, its determinant is nonzero, which means that it is a unimodular matrix).

One can then expand as follows:

$$
\begin{aligned}
\lambda I - A &= \lambda I - TJT^{-1} \\
&= T(\lambda I - J)T^{-1} \\
&= \underbrace{TM'(\lambda)}_{\triangleq M(\lambda)} \underbrace{D(\lambda)}_{\triangleq E(\lambda)} \underbrace{N'(\lambda)T^{-1}}_{\triangleq N(\lambda)},
\end{aligned}
$$

where the last equality is obtained with the results of A4, (5) in particular. $M(\lambda)$ and $N(\lambda)$ are indeed unimodular as they are products of unimodular matrices. $\square$

The matrices A and J have the same eigenvalues which means that they have the same minimal polynomial as well. In A4, one could already observe that a Jordan matrix has a minimal polynomial equal to its last elementary polynomial.

The minimal polynomial of A is thus equal to the last elementary polynomial of $\lambda I - J$ (obtained by studying $D(\lambda)$) which is also the last elementary polynomial of $\lambda I - A$ (obtained by studying $E(\lambda) = D(\lambda)$).

Exercise B: Implementation

B1

The Jordan normal form relies on an eigenvalue computation, which is not numerically stable. This instability also means that any technique for studying the boundedness of trajectories of a linear dynamical system based on explicit computation of the eigenvalues of the transition matrix is necessarily numerically unstable.

Where this differs from techniques based on the Smith normal form is then that the Smith normal form can be computed exactly in a numerically stable way, and that the minimal polynomial can be deduced from the Smith normal form without introducing instabilities. Finally, as mentioned in the exercise statement, one then need not explicitly compute the eigenvalues (i.e., the roots of the minimal polynomial), as the Routh–Hurwitz criterion provides an exact way to determine boundedness of trajectories.

B2

We want to compute the Smith normal form of $P_i(\lambda) \triangleq \lambda I - A_i$ where A_i is one of the three matrices A_1, A_2, and A_3 defined in the homework statement. The Smith normal forms are the following:

$$P_1(\lambda) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \lambda^4 + 4\lambda^3 + \frac{21\lambda^2}{4} + \frac{11\lambda}{4} + \frac{1}{2} \end{bmatrix},$$

$$P_2(\lambda) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \lambda^4 + \frac{7\lambda^3}{2} + \frac{7\lambda^2}{2} + \lambda \end{bmatrix},$$

$$P_3(\lambda) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \lambda^4 + 3\lambda^3 + 2\lambda^2 \end{bmatrix}.$$

The minimal polynomial $p_{A_i}(\lambda)$ of A_i can be deduced from the last elementary polynomial of those Smith normal forms.

$$\left(P_1(\lambda)\right)_{nn} = \lambda^4 + 4\lambda^3 + \frac{21\lambda^2}{4} + \frac{11\lambda}{4} + \frac{1}{2} \implies p_{A_1}(\lambda) = (\lambda + 1/2)^2(\lambda + 1)(\lambda + 2),$$

$$\left(P_2(\lambda)\right)_{nn} = \lambda^4 + \frac{7\lambda^3}{2} + \frac{7\lambda^2}{2} + \lambda \implies p_{A_2}(\lambda) = \lambda(\lambda + 1/2)(\lambda + 1)(\lambda + 2),$$

$$\left(P_3(\lambda)\right)_{nn} = \lambda^4 + 3\lambda^3 + 2\lambda^2 \implies p_{A_3}(\lambda) = \lambda^2(\lambda + 1)(\lambda + 2).$$

If one wants to verify the boundednesss of the trajectories of the associated linear dynamical systems with transition matrices A_1, A_2 and A_3 defined in the homework statement, Statement 2 in the introduction of Exercise A of Homework 3 is useful: "All the eigenvalues of A_i are in the open left-hand plane, or are on the imaginary axis and are simple." This statement has been shown to be equivalent to boundedness of trajectories.

To check whether this statement is satisfied or not, one must look at the vectors of eigenvalues Λ_i of the matrices, which one can compute from the factorization of the minimal polynomial.

$$\Lambda_1 = \begin{bmatrix} -1/2 \\ -1/2 \\ -1 \\ -2 \end{bmatrix}, \quad \Lambda_2 = \begin{bmatrix} 0 \\ -1/2 \\ -1 \\ -2 \end{bmatrix}, \quad \Lambda_3 = \begin{bmatrix} 0 \\ 0 \\ -1 \\ -2 \end{bmatrix}.$$

One can observe that these eigenvalues are identical to the eigenvalues found for the third homework. The conclusion is thus similar: all eigenvalues of A_1 (for all i) are such that $\Re(\lambda_{1,i}) < 0$. A_1 thus satisfies Statement 2, and has bounded trajectories. The matrix A_2 has eigenvalues such that $\Re(\lambda_{2,i}) < 0$ for all i except for $(\lambda_{2,1} = 0)$, which is simple (the multiplicity of this eigenvalue in the factorization of the minimal polynomial is one), so the statement is again verified, and the trajectories are bounded. The matrix A_3 has eigenvalues $(\lambda_{3,1} = \lambda_{3,2} = 0)$ with zero real part and with a multiplicity of two in the factorization of the minimal polynomial, which is not simple. Therefore, it does not satisfy the statement and does not have bounded trajectories. We conclude that only the first two systems have bounded trajectories, which is the same conclusion we arrived at in the previous homework.