Exercise A: Minimal polynomial and Smith normal form

A1

First we prove a lemma.

Lemma. *If $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{n\times n}$ then for $k = 1, ..., n$:*

- *$\delta_k(A)$ divides $\delta_k(AB)$*
- *$\delta_k(B)$ divides $\delta_k(AB)$*

Proof. If we denote by $(AB)_{\mathbf{IJ}}$ the k-by-k sub-matrix of AB taking the rows $\mathbf{I}$ and the columns $\mathbf{J}$ (hence $|\mathbf{I}| = |\mathbf{J}| = k$). We can write and develop the associated k-minor:

$$\begin{aligned}\det(AB)_{\mathbf{IJ}} &= \det(A_{\mathbf{I}*}B_{*\mathbf{J}})\\ &= \sum_{\mathbf{j}_k} A_{\mathbf{I}*}\binom{k}{\mathbf{j}_k} B_{*\mathbf{J}}\binom{\mathbf{j}_k}{k}\\ &= \sum_{\mathbf{j}_k} \det A_{\mathbf{I}\mathbf{j}_k} \det B_{\mathbf{j}_k\mathbf{J}}\end{aligned}$$

The second equality is derived from the Binet-Cauchy theorem. We observe that the k-minors of AB can be written as linear combinations of the k-minors of A. Thus, every common divisor of the k-minors of A must also be a common divisor of the k-minors of AB. As it is true for the greatest common divisor, we have that $\delta_k(A)$ divides $\delta_k(AB)$. The same reasoning is valid taking B instead of A and therefore we prove that $\delta_k(B)$ divides $\delta_k(AB)$. □

We now prove 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 with $A = M(\lambda)P(\lambda)$ and $B = N(\lambda)$, we have that $\delta_k(M(\lambda)P(\lambda))$ divides $\delta_k(Q(\lambda))$. Then applying once more the lemma with $A = \delta_k(M(\lambda))$ and $B = \delta_k(P(\lambda))$, we obtain that $\delta_k(P(\lambda))$ divides $\delta_k(M(\lambda)P(\lambda))$. We deduce that $\delta_k(P(\lambda))$ divides $\delta_k(Q(\lambda))$.

We can apply the exact same reasoning to $P(\lambda) = N^{-1}(\lambda)Q(\lambda)M^{-1}(\lambda)$. Doing so, we have that $\delta_k(Q(\lambda))$ divides $\delta_k(P(\lambda))$. Consequently, we conclude that $\delta_k(Q(\lambda))$ is equal to $\delta_k(P(\lambda))$. □

A2

. We now prove that for all $k \in \{1, ..., n\}$, $\delta_k(P(\lambda)) = \Pi_{i=1}^{k} d_i(\lambda)$ where $d_i(\lambda)$ are the diagonal entries of $D(\lambda)$ and $M(\lambda)D(\lambda)N(\lambda)$ is a Smith decomposition of $P(\lambda)$.

Proof. From A1, we know that $\delta_k(P(\lambda)) = \delta_k(D(\lambda))$ and therefore we will consider $\delta_k(D(\lambda))$.
We first note that all sub-matrices of $D(\lambda)$ are either triangular inferior, triangular superior or null. The determinants of these sub-matrices are hence given by the product of the diagonal elements. The only sub-matrices for which the determinant is non-zero are these where the indices of the removed rows are identical to the indices of the removed columns. Because 0 has infinitely many divisors, we should therefore only consider the case where this specific case.
We denote by $\mathbf{j} = (j_1, ..., j_k)$, where $j_1 < ... < j_k$, the indices of the rows and columns that are kept. The determinant associated to $\mathbf{j}$ can be written as:

$$\det D_{\mathbf{j}}(\lambda) = d_{j_1}(\lambda)...d_{j_k}(\lambda)$$

We now use the property that $d_i(\lambda)$ divides $d_{i+1}(\lambda)$ for all $i \in \{1, ..., n-1\}$.
If we consider the ith factor $d_{j_i}(\lambda)$, we observe that $d_i(\lambda)$ divides $d_{j_i}(\lambda)$ as j_i is either equal to i or larger than i (due to the strict inequality $j_1 < ... < j_k$). Hence $d_{j_i}(\lambda)$ can be rewritten as $a_{j_i}(\lambda)d_i(\lambda)$ for a certain $a_{j_i}(\lambda) \in \mathbb{C}[\lambda]$. The determinant can thus be developed as:

$$\det D_{\mathbf{j}}(\lambda) = a_{j_1}(\lambda)d_1(\lambda)...a_{j_k}(\lambda)d_k(\lambda) \tag{1}$$

We also observe that when $\mathbf{j} = (1, ..., k)$, we have:

$$\det D_{\mathbf{j}}(\lambda) = \Pi_{i=1}^{k} d_i(\lambda) \tag{2}$$

which is a monic polynomial as it is the product of monic polynomials.
Because $\Pi_{i=1}^{k} d_i(\lambda)$ is a divisor of (1) for any $\mathbf{j}$ and that $\delta_k(D(\lambda))$ cannot contain any additional factor due to (2), we conclude that the monic greatest common divisor of all k-minors of $D(\lambda)$ is $\Pi_{i=1}^{k} d_i(\lambda)$. Therefore, we have:

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

□

The diagonal entries $d_1(\lambda), ..., d_n(\lambda)$ of $D(\lambda)$ are unique and depend only on the sequence $\delta_1(P(\lambda))$,...,$\delta_n(P(\lambda))$.

Proof. We first consider the first entry $d_1(\lambda)$. From (3), we know that $d_1(\lambda) = \delta_1(\lambda)$. The notion of monic greatest common divisor leads to the unicity of $\delta_1(\lambda)$ and hence we deduce $d_1(\lambda)$ is unique as well. This constitutes the base case of the induction proof.
Next, we prove that if $d_k(\lambda)$ is unique then $d_{k+1}(\lambda)$ is also unique. Again from (3), we have that $d_{k+1}(\lambda) = \delta_{k+1}(P(\lambda))/d_k(\lambda)$. By the unicity of the notion of monic greatest common divisor and the unicity $d_k(\lambda)$, we conclude that $d_{k+1}(P(\lambda))$ is unique.
Clearly the diagonal entries $d_1(\lambda), ..., d_n(\lambda)$ can be computed using the sequence $\delta_1(P(\lambda)), ..., \delta_n(P(\lambda))$. First, we have $d_1(\lambda) = \delta_1(P(\lambda))$ and then we use $d_{k+1}(\lambda) = \delta_{k+1}(P(\lambda))/d_k(\lambda)$ for the next entries. □

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 $\delta_k(P(\lambda)) = \delta_k(Q(\lambda))$ for all $k \in \{1, ..., n\}$.

Proof. $\Rightarrow$ This was shown in A1.
$\Leftarrow$ From the previous proved statement we know that the diagonal entries of $D(\lambda)$ of the Smith decomposition of $P(\lambda)$ are unique and depend only on the sequence $\delta_1(P(\lambda)), ..., \delta_n(P(\lambda))$. As this sequence is the same for $P(\lambda)$ and $Q(\lambda)$, we deduce $D(\lambda)$ is the same for their Smith decompositions:

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

where $M_1(\lambda), M_1(\lambda), M_1(\lambda), M_1(\lambda) \in \mathbb{C}^{n\times n}[\lambda]$ are unimodular and $D(\lambda) \in \mathbb{C}^{n\times n}[\lambda]$ diagonal.
From the first equation, we can 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. Next, we inject this expression of $D(\lambda)$ in the second equation and we obtain:

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

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

□

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 equal to: $d_i(\lambda) = 1$ for $i = 1, ..., n-1$ and $d_n(\lambda) = (\lambda - \lambda_1)^n$.

Proof. We first write the matrix $\lambda I - A$ that we call $P(\lambda)$:

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

As we have seen the sequence $d_1(\lambda), ..., d_n(\lambda)$ can be derived by the sequence $\delta_1(P(\lambda)), ..., \delta_n(P(\lambda))$, we first derive the latter sequence. When computing each $\delta_k(P(\lambda))$, we observe that for $k < n$, a valid sub-matrix to take into account is the one which has all the -1 entries on its diagonal. We illustrate here the case where $k = n-1$:

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

As this sub-matrix is triangular inferior, its determinant is equal to the product of its diagonal entries. The determinant of such matrices is therefore 1 or -1. As 1 and -1 have only one monic divisor which is 1, we conclude that $\delta_k(P(\lambda)) = 1$ for $k \in \{1, ..., n-1\}$. From the recursive formula proposed in A2, we deduce that $d_k(\lambda) = 1$ for $k \in \{1, ..., n-1\}$.
When $k = n$, the only possible sub-matrix to consider is the whole matrix which determinant is equal to the product of the diagonal entries, hence $(\lambda - \lambda_1)^n$. We deduce that $\delta_n(P(\lambda)) = (\lambda - \lambda_1)^n$ and therefore using the recursive formula proposed in A2 we obtain $d_n(\lambda) = (\lambda - \lambda_1)^n$. □

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 size 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

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

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

Proof. From A3, we have already proved that $\lambda I - J_1$ can be written as

$$\begin{aligned} P_1(\lambda) &:= \lambda I_{n_1} - J_1 = M_1(\lambda) D_1(\lambda) N_1(\lambda) \\ \Leftrightarrow D_1 &= M_1^{-1}(\lambda) P_1(\lambda) N_1^{-1}(\lambda) \\ &= \begin{bmatrix} 1 & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & \\ & & & & (\lambda - \lambda_1)^{n_1} \end{bmatrix} \end{aligned}$$

Similarly,

$$\begin{aligned} P_2(\lambda) &:= \lambda I_{n_2} - J_2 = M_2(\lambda) D_2(\lambda) N_2(\lambda) \\ \Leftrightarrow D_2 &= M_2^{-1}(\lambda) P_2(\lambda) N_2^{-1}(\lambda) \\ &= \begin{bmatrix} 1 & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & \\ & & & & (\lambda - \lambda_2)^{n_2} \end{bmatrix} \end{aligned}$$

Note that $M_1^{-1}(\lambda)$, $N_1^{-1}(\lambda)$, $M_2^{-1}(\lambda)$ and $N_2^{-1}(\lambda)$ are unimodular as their inverse are unimodular.

Now, if we consider the diagonal matrix formed by D_1 and D_2, we have :

$$
\begin{aligned}
D(\lambda) := \begin{bmatrix} D_1 & \\ & D_2 \end{bmatrix} &= \begin{bmatrix} 1 &&&&&&&&& \\ & 1 &&&&&&&& \\ && \ddots &&&&&&& \\ &&& 1 &&&&&& \\ &&&& (\lambda-\lambda_1)^{n_1} &&&&& \\ &&&&& 1 &&&& \\ &&&&&& 1 &&& \\ &&&&&&& \ddots && \\ &&&&&&&& 1 & \\ &&&&&&&&& (\lambda-\lambda_2)^{n_2} \end{bmatrix} \\
&= \begin{bmatrix} M_1^{-1}(\lambda)P_1(\lambda)N_1^{-1}(\lambda) & \\ & M_2^{-1}(\lambda)P_2(\lambda)N_2^{-1}(\lambda) \end{bmatrix} \\
&= \underbrace{\begin{bmatrix} M_1^{-1}(\lambda) & \\ & M_2^{-1}(\lambda) \end{bmatrix}}_{:=M(\lambda)} \underbrace{\begin{bmatrix} P_1\lambda) & \\ & P_2(\lambda) \end{bmatrix}}_{:=P(\lambda)=(\lambda I-A)} \underbrace{\begin{bmatrix} N_1^{-1}(\lambda) & \\ & N_2^{-1}(\lambda) \end{bmatrix}}_{:=N(\lambda)}
\end{aligned}
$$

$M(\lambda)$ is indeed unimodular as its determinant is the product of the determinants of $M_1^{-1}(\lambda)$ and $M_2^{-1}(\lambda)$ (both unimodular). Thus, the determinant of $M_1(\lambda)$ is equal to the product of 2 nonzero constants, which is a nonzero constant.

Similarly, $N(\lambda)$ is also unimodular.

Remark : this result can be easily extended to the case of A being composed by more than 2 Jordan blocks with the same reasoning.

$$
\underbrace{\begin{bmatrix} D_1 &&& \\ & D_2 && \\ && D_3 & \\ &&& \ddots \end{bmatrix}}_{:=D(\lambda)} = \underbrace{\begin{bmatrix} M_1^{-1}(\lambda) &&& \\ & M_2^{-1}(\lambda) && \\ && M_3^{-1}(\lambda) & \\ &&& \ddots \end{bmatrix}}_{:=B^{-1}(\lambda)} \underbrace{\begin{bmatrix} P_1\lambda) &&& \\ & P_2(\lambda) && \\ && P_3(\lambda) & \\ &&& \ddots \end{bmatrix}}_{:=P(\lambda)=(\lambda I-A)} \underbrace{\begin{bmatrix} N_1^{-1}(\lambda) &&& \\ & N_2^{-1}(\lambda) && \\ && N_3^{-1}(\lambda) & \\ &&& \ddots \end{bmatrix}}_{:=C^{-1}(\lambda)} \tag{4}
$$

□

Now, we want to find the elementary polynomials of $P(\lambda) = (\lambda I - A)$. For that, we will use the property proved in A1 s.t. $\delta_k(P(\lambda)) = \delta_k(D(\lambda))$.

Indeed, it is clear that for $k = 1, 2, \ldots, (n_1 + n_2 - 2)$, we can always choose a submatrix containing only ones on the diagonal, which means that $\delta_k(P(\lambda)) = 1$. And by A2, we have that the monic $d_k(\lambda)$ are all equal to 1 for $k = 1, 2, \ldots, (n_1 + n_2 - 2)$.

Then, we need to distinguish 2 cases :

- When $\lambda_1 \neq \lambda_2$,

$$
\begin{aligned}
&\delta_{n_1+n_2-1}(P(\lambda)) = 1 && \Leftrightarrow d_{n_1+n_2-1}(\lambda) = 1 \\
&\delta_{n_1+n_2}(P(\lambda)) = (\lambda-\lambda_1)^{n_1}(\lambda-\lambda_2)^{n_2} && \Leftrightarrow d_{n_1+n_2}(\lambda) = (\lambda-\lambda_1)^{n_1}(\lambda-\lambda_2)^{n_2}
\end{aligned}
$$

- When $\lambda_1 = \lambda_2$,

$$
\begin{aligned}
&\delta_{n_1+n_2-1}(P(\lambda)) = (\lambda-\lambda_1)^{\min(n_1,n_2)} && \Leftrightarrow d_{n_1+n_2-1}(\lambda) = (\lambda-\lambda_1)^{\min(n_1,n_2)} \\
&\delta_{n_1+n_2}(P(\lambda)) = (\lambda-\lambda_1)^{n_1+n_2} && \Leftrightarrow d_{n_1+n_2}(\lambda) = \frac{(\lambda-\lambda_1)^{n_1+n_2}}{(\lambda-\lambda_1)^{\min(n_1,n_2)}} = (\lambda-\lambda_1)^{\max(n_1,n_2)}
\end{aligned}
$$

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

A5

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

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

where $E(\lambda)$ is diagonal with 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.

$$T^{-1}AT = J \qquad \Leftrightarrow \qquad A = TJT^{-1}$$

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

We can now derive :

$$\begin{aligned} \lambda I - A &= \lambda I - TJT^{-1} \\ &= T(\lambda I - J)T^{-1} \\ &= \underbrace{TB(\lambda)}_{:=M(\lambda)} \underbrace{D(\lambda)}_{:=E(\lambda)} \underbrace{C(\lambda)T^{-1}}_{:=N(\lambda)} \end{aligned}$$

The last equality is obtained with the results in A4 (cfr Equation). $M(\lambda)$ and $N(\lambda)$ are indeed unimodular as they are products of 2 unimodular matrices. □

The matrices A and J have the same eigenvalues which means that they have the same minimal polynomial as well. In A4, we have already observed that a Jordan matrix with 2 Jordan blocks has a minimal polynomial equal to its last elementary polynomial. It is trivial to see that this result can be extended to a Jordan matrix with more Jordan blocks. Hence the minimal polynomial of A is equal to the last elementary polynomial of J (obtained with $D(\lambda)$) which is also the last elementary polynomial of A (obtained with $E(\lambda) = D(\lambda)$).

Exercise B: Implementation

B1

Using the Jordan normal form is not numerically stable because taking limits does not commute with forming the Jordan canonical form. A simple example is the matrix $A = I_2$, approximated by $A_\varepsilon = \left[\begin{smallmatrix} 1 & \varepsilon \\ 0 & 1 \end{smallmatrix}\right]$, the latter having Jordan canonical form $J_\varepsilon = \left[\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right]$. However, the Jordan form of A is simply $J = I_2$. We thus have

$$\lim_{\varepsilon\to 0} A_\varepsilon = A, \quad \text{but} \quad \lim_{\varepsilon\to 0} J_\varepsilon \neq J.$$

Similarly, computing the minimal polynomials yields $p_{A_\varepsilon}(\lambda) = (\lambda - 1)^2 \neq \lambda - 1 = p_A(\lambda)$.

B2