This assignment checks your ability to write down a matrix representation of a linear function in a given basis, but mostly explores the benefits of a coordinate-free approach to calculating eigenvalues of a linear transformation.

For a positive integer \( n \) let \( \mathcal{P}_n \) denote the complex vector space of all polynomials \( p = p(x) = p_0 + p_1 x + \cdots + p_{n-1} x^{n-1} \) of degree less than \( n \) of a scalar variable \( x \), with complex coefficients \( p_i \). Let \( A_n : \mathcal{P}_n \mapsto \mathcal{P}_n \) be the function mapping a polynomial \( p \in \mathcal{P}_n \) to the polynomial \( q = A_n p \in \mathcal{P}_n \) defined by

\[
q(x) = \frac{x + 1}{x - 1} [p(x) - p(1)] \quad (x \neq 1).
\]

It is easy to see that \( A_n \) is a linear operator on \( \mathcal{P}_n \), i.e. a linear function \( \mathcal{P}_n \mapsto \mathcal{P}_n \).

(a) Find the matrix \( M_n \) of \( A_n \) in the basis \( \{v_1, \ldots, v_n\} \), where \( v_k(x) \equiv x^{k-1} \). Give the answer specifically for \( n = 4 \), and describe the entries of \( M_n \) for a general \( n \).

Using the polynomial identity

\[
x^m - 1 = (x - 1)(x^{m-1} + x^{m-2} + \cdots + x + 1)
\]

we get

\[
A_n(x^i) = x^{i-1} + 2x^{i-2} + \cdots + 2x + 1
\]

for \( 1 < i \leq n \), and \( A_n(1) = 0 \). Therefore, the matrix \( M_n \) for general \( n \) is given by

\[
M_n = \begin{pmatrix}
0 & 1 & 1 & \cdots & 1 & 1 & 1 \\
0 & 1 & 2 & \cdots & 2 & 2 & 2 \\
0 & 0 & 1 & \cdots & 2 & 2 & 2 \\
& & & \ddots & & & \\
0 & 0 & 0 & \cdots & 1 & 2 & 2 \\
0 & 0 & 0 & \cdots & 0 & 1 & 2 \\
0 & 0 & 0 & \cdots & 0 & 0 & 1
\end{pmatrix}
\]

and for \( n = 4 \) by

\[
M_4 = \begin{pmatrix}
0 & 1 & 1 & 1 \\
0 & 1 & 2 & 2 \\
0 & 0 & 1 & 2 \\
0 & 0 & 0 & 1
\end{pmatrix}
\]
(b) Find all eigenvalues of $A_n$ (Hint: describe all polynomials $p = p(x)$ satisfying the identity

$$\lambda p(x)(x - 1) \equiv (x + 1)[p(x) - p(1)]$$

for some $\lambda \in \mathbb{C}$.)

All eigenvectors $p(x)$ of $A_n$ satisfy the identity $\lambda p(x) = M_n p(x) = g(x)$, where $\lambda$ are scalars equal to the eigenvalues of $M_n$. Consequently they satisfy

$$\lambda p(x)(x - 1) = (x + 1)[p(x) - p(1)] \Leftrightarrow p(x) = -\frac{(x + 1)}{x(\lambda - 1) - (\lambda + 1)} p(1)$$

But $p(x)$ is a polynomial so the fraction in the above expression should also be a polynomial which can only happen for $\lambda = 1$ and $\lambda = 0$. Thus $\lambda = 1$ and $\lambda = 0$ are the eigenvalues of $A_n$.

(c) The set $V_n$ consisting of all polynomials $p \in P_n$ such that $p(-1) = 0$ is a linear subspace of $P_n$ which is invariant with respect to $A_n$. Find all eigenvalues of the restriction of $A_n$ onto $V_n$.

In question (b) we found that the eigenvalues of $A_n$ on $P_n$ are $\lambda = 0$ and $\lambda = 1$. Thus the eigenvalues of the restriction of $A_n$ on $V_n$ will be a nonempty subset of $\{0, 1\}$. They should also satisfy the identity $\lambda p(x)(x - 1) \equiv (x + 1)[p(x) - p(1)]$. But for $\lambda = 0$ the above identity gives us $(x + 1)[p(x) - p(1)] = 0 \Leftrightarrow p(x) = p(1) = \text{constant} \ \forall x$, which is a contradiction. So $\lambda = 1$ has to hold. Also, given that the constraint $p(-1) = 0$ removes one degree of freedom from the vector space, we get $\dim \{V_n\} = n - 1$ and thus the algebraic multiplicity of $\lambda = 1$ is $n - 1$.

(d) Use the results from (b) and (c) to find the characteristic polynomial of $A_n$.

$A_n$ is applied on the vector space $P_n$ of all polynomials of degree less than $n$, so it has $n$ eigenvalues because $\dim \{P_n\} = n$. From (c) we know that $A_n$ has $n - 1$ eigenvalues $\lambda = 1$ and combining with (b) we can conclude that it has exactly one eigenvalue $\lambda = 0$. Consequently the characteristic polynomial of $A_n$ can be expressed as $c(\lambda) = \lambda(\lambda - 1)^{n-1}$.

(e) Write MATLAB function $\text{ps1 lc(n)}$ checking the result of (d) for a given positive integer $n$. Provide the code with your solution.

See attached solution at the end of the file.
**Problem 1.2**

This assignment lets you derive an important linear algebra statement used extensively in systems applications. Your understanding of the relation between eigenvalues and the characteristic polynomial, as well as knowledge of standard conditions of feasibility of linear equations are tested.

For positive integers $n, m$ let $\mathbb{C}^{m,n}$ denote the complex vector space of all complex $m$-by-$n$ matrices. Let $A \in \mathbb{C}^{m,m}$ and $B \in \mathbb{C}^{n,n}$ be fixed matrices with eigenvalues $a_1, \ldots, a_m$ and $b_1, \ldots, b_n$ respectively, in the sense that

$$\det(sI_m - A) = (s - a_1) \cdots (s - a_m), \quad \det(sI_n - B) = (s - b_1) \cdots (s - b_n).$$

Let $L_{A,B} : \mathbb{C}^{m,n} \mapsto \mathbb{C}^{m,n}$ and $H_{A,B} : \mathbb{C}^{m,n} \mapsto \mathbb{C}^{m,n}$ be the linear operators mapping $X \in \mathbb{C}^{m,n}$ to $AX + XB$ and $AXB$ respectively.

(a) Assuming that $a_i \neq a_j$ and $b_i \neq b_j$ for $i \neq j$, show that matrices $X = v_i f_j$, where $v_i \neq 0$ are column vectors such that $Av_i = a_i v_i$, and $f_j$ are row vectors such that $f_j B = b_j f_j$, are eigenvectors of $L_{A,B}$ and $H_{A,B}$. Find the corresponding eigenvalues of $L_{A,B}$ and $H_{A,B}$ in terms of the numbers $a_i, b_j$. Use this observation to describe the characteristic polynomial of $L_{A,B}$ and $H_{A,B}$ in terms of the numbers $a_i, b_j$. Explain how the assumption that $a_i \neq a_j$ and $b_i \neq b_j$ for $i \neq j$ was used in the derivation (if it was).

In order for matrices $X_{ij} = v_i f_j$ to be eigenvectors of $L_{A,B}$ and $H_{A,B}$ we only need to show that $L_{A,B}(X_{ij}) = \lambda_L X_{ij}$ and $H_{A,B}(X_{ij}) = \lambda_H X_{ij}$, where $\lambda_L$ and $\lambda_H$ scalars equal to the eigenvalues of $L_{A,B}$ and $H_{A,B}$ respectively. Indeed,

$$L_{A,B}(X_{ij}) = AX_{ij} + X_{ij}B$$

$$= Av_i f_j + v_i f_j B$$

$$= a_i v_i f_j + b_j v_i f_j$$

$$= (a_i + b_j) v_i f_j$$

$$= \lambda_L v_i f_j$$

$$H_{A,B}(X_{ij}) = AX_{ij}B$$

$$= Av_i f_j B$$

$$= a_i v_i b_j f_j$$

$$= \lambda_H v_i f_j$$

Thus matrices $X_{ij}$ are eigenvectors of $L_{A,B}$ and $H_{A,B}$ with eigenvalues $\lambda_L = a_i + b_j$ and $\lambda_H = a_i b_j$ respectively. Given that $i = 1, \ldots, m$ and $j = 1, \ldots, n$ we expect
to have $mn$ values for $\lambda_{L_{ij}}$ and $mn$ values for $\lambda_{H_{ij}}$. Given also that both linear operators are applied on a vector space of dimension $m \times n$ they will have exactly $mn$ eigenvalues. Consequently the characteristic polynomials of the two linear operators can be conjectured to be

$$c_L(\lambda) = \prod_{i,j} (\lambda - a_i - b_j), \quad c_H(\lambda) = \prod_{i,j} (\lambda - a_i b_j)$$

respectively. When $a_i + b_j \neq a_k + b_r$ unless $i = k$ and $j = r$ the conjecture is proven by the calculations given above. Otherwise, the multiplicity of some eigenvalues $a_i + b_j$ needs to be figured out, as shown in the solution of (b).

(b) According to the Shur decomposition theorem, there exists a basis $\{u_1, \ldots, u_m\}$ of column vectors $u_i \in \mathbb{C}^m$ such that $Au_k - a_k u_k$ is a linear combination of $u_i$ with $i < k$ for all $k$ (the matrix of $A$, as a linear operator $A : u \in \mathbb{C}^m \mapsto Au \in \mathbb{C}^m$, in this basis is upper triangular, with the $k$-th diagonal element being $a_k$). Similarly, there exists a basis $\{g_1, \ldots, g_n\}$ of row vectors $g_i \in \mathbb{C}^{1,n}$ such that $g_k B - b_k g_k$ is a linear combination of $g_i$ with $i < k$ for all $k$. Arrange the $m \cdot n$ matrices $X = u_i g_j$ in such an order $(X_1, \ldots, X_{mn})$ that $L_{A,B} X_k$ and $H_{A,B} X_k$ are linear combinations of $X_i$ with $i \leq k$ for all $k \in \{1, \ldots, mn\}$. Use this to generalize the result of (a) by describing the characteristic polynomials of $L_{A,B}$ and $H_{A,B}$ in terms of the numbers $a_i, b_j$, without relying on the assumptions made in (a).

Let’s try to express $L_{A,B} X_{k(i,j)}$ and $H_{A,B} X_{k(i,j)}$ as linear combinations of $X_{q(i,j)}$ with $k(s,t) \leq k(i,j)$, $\forall k(i,j) \in \{1, \ldots, mn\}$ and $k(i,j) = (i - 1) * n + j$:

$$L_{A,B} (X_{k(i,j)}) = AX_{k(i,j)} + X_{k(i,j)}B$$

$$= Au_i g_j + u_i g_j B$$

$$= (a_i u_i + \sum_{s=1}^{i-1} r_s u_s) g_j + u_i (b_j g_j + \sum_{t=1}^{j-1} p_t g_t)$$

$$= (a_i + b_j) u_i g_j + \sum_{s=1}^{i-1} r_s u_s g_j + u_i \sum_{t=1}^{j-1} p_t g_t$$

$$= (a_i + b_j) X_{k(i,j)} + \sum_{s=1}^{i-1} r_s X_{k(s,j)} + \sum_{t=1}^{j-1} p_t X_{k(i,t)}.$$
\[ H_{A,B}(X_{k(i,j)}) = AX_{k(i,j)}B = Au_{i}g_{j}B \]
\[ = (a_{i}u_{i} + \sum_{s=1}^{i-1} r_{s}u_{s})(b_{j}g_{j} + \sum_{t=1}^{j-1} p_{t}g_{t}) \]
\[ = a_{i}b_{j}X_{k(i,j)} + b_{j}\sum_{s=1}^{i-1} r_{s}X_{k(s,j)} + a_{i}\sum_{t=1}^{j-1} p_{t}X_{k(i,t)} + \sum_{s=1}^{i-1} r_{s}\sum_{t=1}^{j-1} p_{t}X_{k(s,t)}. \]

So for both \( L_{A,B} \) and \( H_{A,B} \) an order of \( X_{k(i,j)} \) that satisfies the Schur decomposition is \( X_{k(1,1)}, \ldots, X_{k(1,2)}, \ldots, X_{k(1,n)} \), \( \ldots, X_{k(n,2)}, \ldots, X_{k(n,n)} \), where \( X_{k(i,j)} = u_{i}g_{j} \). Easily one can observe that this is not the only order of \( X_{k(i,j)} \) that satisfies the Schur decomposition!

With respect to the above basis the matrix representations of \( L_{A,B} \) and \( H_{A,B} \) can be written in the forms:

\[
M_{L} = \begin{pmatrix}
  a_{1} + b_{1} & * & * & \ldots & * & * & * \\
  0 & a_{1} + b_{2} & * & \ldots & * & * & * \\
  0 & 0 & a_{1} + b_{3} & \ldots & * & * & * \\
  \vdots & & & & & & \\
  0 & 0 & 0 & \ldots & a_{m} + b_{n-2} & * & * \\
  0 & 0 & 0 & \ldots & 0 & a_{m} + b_{n-1} & * \\
  0 & 0 & 0 & \ldots & 0 & 0 & a_{m} + b_{n}
\end{pmatrix}
\]

and

\[
M_{H} = \begin{pmatrix}
  a_{1}b_{1} & * & * & \ldots & * & * & * \\
  0 & a_{1}b_{2} & * & \ldots & * & * & * \\
  0 & 0 & a_{1}b_{3} & \ldots & * & * & * \\
  \vdots & & & & & & \\
  0 & 0 & 0 & \ldots & a_{m}b_{n-2} & 0 & * \\
  0 & 0 & 0 & \ldots & 0 & a_{m}b_{n-1} & * \\
  0 & 0 & 0 & \ldots & 0 & 0 & a_{m}b_{n}
\end{pmatrix}
\]

where with '*' we denote values which are irrelevant to the task at hand. Easily we can derive the characteristic polynomials of the above matrices by multiplying the diagonal elements of the determinants \( |\lambda I - M_{L}| \) and \( |\lambda I - M_{H}| \). So,

\[ c_{L}(\lambda) = \prod_{i,j}(\lambda - a_{i} - b_{j}), \quad c_{H}(\lambda) = \prod_{i,j}(\lambda - a_{i}b_{j}) \]
which generalizes the result of (a).

(c) Write MATLAB function `ps1_2c(m,n)` checking the result of (b) for given positive integers $m, n$ while generating $A, B$ randomly. Provide the code with your solution. See attached solution at the end of the file.

(d) Use the result of (b) to formulate (and prove) necessary and sufficient conditions, to be imposed on the numbers $a_i, b_i$, for the equation $AX + XB = Y$ to have a solution $X \in \mathbb{C}^{m,n}$ for every $Y \in \mathbb{C}^{m,n}$. Similarly, derive necessary and sufficient conditions for the equation $X - AXB = Y$ to have a solution $X$ for every $Y$.

A necessary and sufficient condition for the equation $CX = Y$ with $X, Y \in V$ to have a solution is $\ker(C) = \{0\}$. So,

- for $AX + XB = Y$: $\lambda_{L_{ij}} \neq 0 \iff a_i + b_j \neq 0$
- for $X - AXB = Y$: $\lambda_{I - H_{ij}} \neq 0 \iff a_i b_j \neq 1$
Function ps1_le(n)

function ps1_lc(n)

%Construct matrix Mn
Mn=zeros(n,n);
c=[1 zeros(1,n-1)];
r=[1 repmat(2,1,n-1)];
Mn=toeplitz(c,r);
Mn(1,:)=1;
Mn(:,1)=0;

%Compute eigenvalues and characteristic polynomial of Mn
n
display('The eigenvalues vector of Mn is:')
lamda=eig(Mn)
display('The characteristic polynomial of Mn is:')
p=poly(Mn)

Output

n = 4

The eigenvalues vector of Mn is:

lamda =

0
1
1
1

The characteristic polynomial of Mn is:

p =

1  -3  3  -1  0
Function ps1_2c(m,n)

function ps1_2c(m,n)
    n = m
    %Generate two random matrices A and B
    A=rand(m);
    B=rand(n);
    %Compute eigenvalues of A and B
    lamdaA=eig(A);
    lamdaB=eig(B);
    %Construct Ml and Mh
    Ml=kron(B',eye(m))+kron(eye(n),A);
    Mh=kron(B',A);
    %Compute characteristic polynomials of Ml and Mh according to
    %the result of b)
    L=repmat(lamdaA(:),1,n)+repmat(lamdaB(:,').',m,1)
    H=repmat(lamdaA(:,1,n).*repmat(lamdaB(:,').',m,1)
    display(['Characteristic polynomials of Ml and Mh computed'...'
    ' according to the result of b)']);
    Pl1=poly(L(:))
    Ph1=poly(H(:))
    %Compute characteristic polynomials of Ml and Mh directly
    %from the matrices
    display(['Characteristic polynomials of Ml and Mh computed'...
    ' directly from the matrices']);
    Pl2=poly(Ml)
    Ph2=poly(Mh)
    DPl=Pl1-Pl2
    DPh=Ph1-Ph2
    if sum(abs(DPl))<10^-10 && sum(abs(DPh))<10^-10
        display(['The two methods give the same characteristic'...
    ' polynomials']);
    end
end

Output

n = 3
m = 2

Characteristic polynomials of Ml and Mh computed according to the result of b)

Pl1 = 1.0000  -7.1773  18.8106  -21.9563  10.6507  -1.0168  -0.4026
\[ \text{Ph}_1 = \begin{bmatrix} 1.0000 & -2.0521 & 0.1491 & 0.5611 & -0.1101 & -0.0125 \\ 0.0014 \end{bmatrix} \]

Characteristic polynomials of \( M_l \) and \( M_h \) computed directly from the matrices

\[ \text{P}_{l2} = \begin{bmatrix} 1.0000 & -7.1773 & 18.8106 & -21.9563 & 10.6507 & -1.0168 & -0.4026 \end{bmatrix} \]

\[ \text{Ph}_2 = \begin{bmatrix} 1.0000 & -2.0521 & 0.1491 & 0.5611 & -0.1101 & -0.0125 \\ 0.0014 \end{bmatrix} \]

The two methods give the same characteristic polynomials!