Lecture notes 12 developed properties of the LTI representation $R = [A, B, C, D]$ for the case in which $A$ had distinct eigenvalues (so that $R$'s state space has a basis which consists of the eigenvectors of $A$). In lecture notes 13, we develop properties of $R$ for the case in which $A$ is a general n×n matrix.

**GOALS OF THIS LECTURE:**

- Introduce: $A$-invariant subspaces
- Direct sum of subspaces
- Functions of a matrix $A$
  - Case 2: repeated eigenvalues
  - minimal polynomial
  - geometric structure of eigenspaces
  - generalized eigenvectors

**REFS:** Callier and Desoer, Chapter 4.
A-invariant Subspace and the Second Rep^n Theorem

Consider a vector space \((V, \mathbb{F})\) and a linear map \(A: V \rightarrow V\). A subspace \(M \subseteq V\) is said to be \(A\)-invariant, or \(A\)-invariant under \(A\), if, given \(x \in M\), \(Ax \in M\).

[often written as \(A[M] \subseteq M\).]

Examples of \(A\)-invariant subspace:

(i) \(N(A)\) is \(A\)-invariant

(ii) \(R(A)\)

(iii) \(N(A - \lambda_i I)\) where \(\lambda_i \in \lambda(A)\)

if \(p(A) = A^k + \lambda_1 A^{k-1} + \ldots + \lambda_{k-1} A + \lambda_k I\)

(iv) Then \(N(p(A))\) is \(A\)-invariant

(v) Let the subspaces \(M_1\) and \(M_2\) be \(A\)-invariant.

let \(M_1 + M_2 := \{x \in V : x = x_1 + x_2, x_i \in M_i\ \text{for} \ i = 1, 2\}\)

\(M_1 \cap M_2\) and \(M_1 + M_2\) are \(A\)-invariant.

Direct sum of subspaces:

We say that \(V\) is the direct sum of \(M_1, M_2, \ldots, M_k\), denoted as \(V = M_1 \oplus M_2 \oplus \ldots \oplus M_k\), if \(V = \bigoplus_{i=1}^{k} M_i\) and \(\forall x \in V, x = x_1 + x_2 + \ldots + x_k\) for 
\(x_i \in M_i, i = 1, \ldots, k\) such that 
\(x = x_1 + x_2 + \ldots + x_k\).
Direct sum is the generalization of linear independence. Check for example that if
\[ V = M_1 \oplus M_2 \oplus \cdots \oplus M_k \] then \[ M_i \cap M_j = \mathbb{Z}^0 \] sub-vector.

**Example:** Let \( A \in \mathbb{R}^{n \times n} \) have \( n \) distinct eigenvalues \( \lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{C} \). Then:
\[ \mathbb{C}^n = N(A - \lambda_1 I) \oplus N(A - \lambda_2 I) \oplus \cdots \oplus N(A - \lambda_n I) \]

**Theorem (2nd Representation Theorem).**
\[ \dim(A) = \dim(M) + \dim(K) + \dim(N) \]
Let \( V = M_1 \oplus M_2 \) be a finite dimensional vector space. If \( M_1 \) is \( A \)-invariant, then \( A \) has a representation of the form:
\[ A = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} \]

**Proof:** Let \( \{ b_1, b_2, \ldots, b_k \} \) be a basis for \( M_1 \), and let \( \{ b_{k+1}, \ldots, b_n \} \) be a basis for \( M_2 \).
If \( x \in M_1 \), then \( x = \sum_{i=1}^{k} x_i b_i \)
Since \( Ax \in M_1 \), \( Ax = \sum_{i=1}^{k} \eta_i b_i \)
Thus the first \( k \) columns of \( A \) have zeros in the last \( (n-k) \) rows.
\[ \square \]
If both $M_1$ and $M_2$ are $A$-invariant then $A$ has the representation:

\[
\begin{pmatrix}
\begin{array}{c}
\lambda_1 \\
\vdots \\
\lambda_k
\end{array}
& 0
\end{array}
\begin{array}{c}
0 \\
\vdots \\
0
\end{array}
\end{pmatrix} \in \mathbb{F}^{n \times n}
\]

Function of a Matrix with repeated eigenvalues

We know that:

\[
\det(sI - A) = \hat{x}_A(s) \quad \text{(characteristic polynomial of } A)\]

We can write:

\[
\hat{x}_A(s) = (s - \lambda_1)^{d_1} (s - \lambda_2)^{d_2} \ldots (s - \lambda_k)^{d_k}
\]

where $d_1, d_2, \ldots, d_k$ are the multiplicities of $\lambda_1, \lambda_2, \ldots, \lambda_k \in \mathbb{C}$, where $d_1 + d_2 + \ldots + d_k = n$.

By Cayley-Hamilton, we know that: $\hat{x}_A(A) = \Theta_{n \times n}$

Let $\tilde{x}_A(s)$ be the polynomial of least degree such that $\hat{x}_A(A) = \Theta_{n \times n}$.

Claim $\tilde{x}_A(s)$ divides $\hat{x}_A(s)$.

Proof if not, then let

\[
\frac{\hat{x}_A(s)}{\tilde{x}_A(s)} = \hat{q}(s) + \frac{\hat{r}(s)}{\tilde{x}_A(s)} \quad \text{with degree } \hat{r} < \text{degree } \hat{x}_A
\]

Thus $\hat{x}_A(A) = \hat{q}(A) \tilde{x}_A(A) + \hat{r}(A) = \Theta_{n \times n}$

\[
\Rightarrow \hat{r}(A) = \Theta_{n \times n}
\]

which contradicts our hypothesis. 

Hence, we have that:

\[ \hat{\chi}_A(s) = (s - \lambda_1)^{m_1} \cdots (s - \lambda_\Delta)^{m_\Delta} \]

with \( m_1 \leq d_1, m_2 \leq d_2, \ldots, m_\Delta \leq d_\Delta \).

\( \hat{\chi}_A(s) \) is called the minimal polynomial of \( A \).

**Examples:**

1) Consider \( A = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{bmatrix} \), \( \hat{\chi}_A(s) = (s - \lambda_1)^2(s - \lambda_2) \)

2) Consider \( A = \begin{bmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_1 \end{bmatrix} \), \( \hat{\chi}_A(s) = (s - \lambda_1)^3 \)

**Theorem**

\[ C^n = N(A - \lambda_1 I)^{m_1} \oplus N(A - \lambda_2 I)^{m_2} \oplus \cdots \oplus N(A - \lambda_\Delta I)^{m_\Delta}. \]

**Proof**

\[ \frac{1}{\hat{\chi}_A(s)} = \frac{1}{(s - \lambda_1)^{m_1} \cdots (s - \lambda_\Delta)^{m_\Delta}} = \frac{\hat{\chi}_1(s)}{(s - \lambda_1)^{m_1}} + \cdots + \frac{\hat{\chi}_\Delta(s)}{(s - \lambda_\Delta)^{m_\Delta}} \]

\[ 1 = \hat{\chi}_1(s) \hat{p}_1(s) + \cdots + \hat{\chi}_\Delta(s) \hat{p}_\Delta(s) \]

where \( \hat{p}_i(s) = \frac{\hat{\chi}_A(s)}{(s - \lambda_i)^{m_i}} \).

Thus:

\[ I = \hat{\chi}_1(A) \hat{p}_1(A) + \cdots + \hat{\chi}_\Delta(A) \hat{p}_\Delta(A) \]

\[ x = \underbrace{\hat{\chi}_1(A) \hat{p}_1(A) x + \cdots + \hat{\chi}_\Delta(A) \hat{p}_\Delta(A) x}_{x, x} \]
\[ x_i = \hat{\eta}_i(A) \hat{p}_i(A) = \hat{\eta}_i(A) \cdot \frac{\hat{\gamma}_A(A)}{(A - \lambda_i I)^{m_i}} \]

\[ \Rightarrow (A - \lambda_i I)^{m_i} x_i = \Theta_n \]
\[ \Rightarrow x_i \in N (A - \lambda_i I)^{m_i} \]

To show that the decomposition is unique, we proceed by contradiction. Let \( x_i \in N (A - \lambda_i I)^{m_i} \) so that \( x_1 + x_2 + \cdots + x_k = \Theta_n \), and without loss of generality, assume \( x_1 \neq \Theta_n \).

Thus \( x_1 = -x_2 - x_3 \cdots - x_k \)

\[ \therefore (A - \lambda_2 I)^{m_2} \cdots (A - \lambda_k I)^{m_k} x_1 = \Theta_n \]

ie. \( \hat{p}_1(A) x_1 = \Theta_n \)

But \( \hat{p}_1(s) \) and \( (s - \lambda_i)^{m_i} \) are coprime, meaning that \( \hat{h}_1(s) \hat{p}_1(s) + \hat{h}_2(s) (s - \lambda_i)^{m_i} = 1 \).

This implies that \( \hat{h}_1(A) \hat{p}_1(A) x_1 + \hat{h}_2(A) (A - \lambda_i I)^{m_i} x_1 = x_1 \)

\[ \Rightarrow x_1 = \Theta_n, \text{ establishing the contradiction} \]

Fact: \( \text{Dim} \ N (A - \lambda_i I)^{m_i} = d_i \)

(* In proof, see Callier: Desoer, p 115).
Geometric structure of Eigenspaced: EXAMPLE.

Consider

\[ X_A(s) = (s - \lambda)^{d_1} \quad d_1 = n \]
\[ Y_A(s) = (s - \lambda)^{m_1} \quad m_1 \geq 1 \]

(let us say, for example, that \( m_1 = 3 \) and \( n = 6 \))

Let \( N(A-\lambda I) = \text{Sp } \{e_1, e_2, e_3\} \)
\[ N(A-\lambda I)^2 = \text{Sp } \{e_1, e_2, e_3, v_1, v_2, v_3\} \supset N(A-\lambda I) \]
\[ \text{i.e. } N(A-\lambda I)^2 = N(A-\lambda I) \oplus \text{Sp } \{v_1, v_2\} \]

Let \( v_1 \) and \( v_2 \) be independent solutions of

\[ (A-\lambda I)x = e_1 \]
\[ (A-\lambda I)x = e_2 \]
\[ (A-\lambda I)x = e_3 \]

Since \( R(A-\lambda I) \subseteq \mathbb{C}^n \), the equations may not have solutions for all \( e_i, i=1,2,3 \). Let us say there are solutions for \( e_1 \) and \( e_2 \); Then:

\[ (A-\lambda I)v_1 = e_1 \]
\[ (A-\lambda I)v_2 = e_2 \]

Since \( R(A-\lambda I)^3 \subseteq \mathbb{C}^n \), the equations may not have solutions for all \( e_i, i=1,2,3 \). Let us say there are solutions for \( e_1 \) and \( e_2 \); Then:

\[ (A-\lambda I)v_1 = e_1 \]
\[ (A-\lambda I)v_2 = e_2 \]
Eigenvector and generalized eigenvector chains

$$\text{Sp} \{ e_1, v_1, w_1 \} \oplus \text{Sp} \{ e_2, v_2 \} \oplus \text{Sp} \{ e_3 \} = \mathbb{R}^6$$

Thus the representation of $A$ with respect to:
$$\{ e_1, v_1, w_1, e_2, v_2, e_3 \}$$
is

$$\begin{bmatrix}
\lambda & 1 & 0 \\
0 & \lambda & 1 \\
0 & 0 & \lambda
\end{bmatrix} = T A T^{-1}$$

where
$$T^{-1} = \begin{bmatrix}
e_1, v_1, w_1, e_2, v_2, e_3
\end{bmatrix}$$

and
$$A e_1 = \lambda e_1 \quad A v_1 = \lambda v_1 + e_1 \quad A w_1 = \lambda w_1 + v_1$$
$$A e_2 = \lambda e_2 \quad A v_2 = \lambda v_2 + e_2$$
$$A e_3 = \lambda e_3$$

Functions of a Matrix

**Defn.** Let $\hat{f}(s)$ be any function of $s$ analytic on the spectrum of $A$ and $\hat{p}(s)$ be a polynomial such that:
$$\hat{f}^k(de) = \hat{p}^k(de) \quad 0 \leq k \leq m - 1$$
$$1 \leq e \leq \delta$$

Then
$$\hat{f}(A) = \hat{p}(A)$$
In fact, if \( m := \sum_{i=1}^{\infty} m_i \) then

\[
\hat{p}(s) = a_1 s^{m-1} + a_2 s^{m-2} + \cdots + a_m s^0
\]

where \( a_1, a_2, \ldots \) are functions of

\[
(\hat{f}(\lambda_1), \hat{f}'(\lambda_1), \hat{f}''(\lambda_1), \ldots, \hat{f}^{m_1}(\lambda_1), \hat{f}(\lambda_2), \ldots)
\]

and hence

\[
\hat{f}(A) = a_1 A^{m-1} + \cdots + a_m A^0
\]

\[
= \sum_{k=1}^{\infty} \sum_{k=0}^{m_k-1} p_k e(A) \cdot f^k(\lambda_k)
\]
Functions of a Matrix (repeated eigenvalues)

Consider \( J = \begin{bmatrix} \lambda & 1 & \cdots & 0 \\ 0 & \lambda & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda \end{bmatrix}^n \)

Claim: \( f(J) = \begin{bmatrix} f(\lambda) & f'(\lambda) & \frac{f''(\lambda)}{2!} & \cdots & \frac{f^{(n-1)}(\lambda)}{(n-1)!} \\ 0 & f'(\lambda) & \frac{f''(\lambda)}{2!} & \cdots & \frac{f^{(n-2)}(\lambda)}{(n-2)!} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & f'(\lambda) & \frac{f''(\lambda)}{2!} \\ 0 & 0 & \cdots & 0 & f(\lambda) \end{bmatrix} \)

Proof: min polynomial = \((s-\lambda)^n\)

Thus \( f(J) = \sum_{k=0}^{n-1} f^{(k)}(\lambda) p_k(J) \)

Choose \( f_1(s) = 1 \Rightarrow f_1(J) = I = f_1^{(0)}(\lambda) p_0(J) \)

\( \Rightarrow p_0(J) = I \)

\( f_2(s) = s-\lambda \Rightarrow f_2(J) = J - \lambda I = f_2^{(1)}(\lambda) p_1(J) \)

\( \Rightarrow p_1(J) = J - \lambda I \)

\( f_3(s) = (s-\lambda)^2 \Rightarrow f_3(J) = (J - \lambda I)^2 = f_3^{(2)}(\lambda) p_2(J) \)

\( \Rightarrow 2 p_2(J) = (J - \lambda I)^2 \)

Thus \( f(J) = \begin{bmatrix} f(\lambda) & f'(\lambda) & f''(\lambda)/2 & \cdots & f^{(n-1)}(\lambda)/(n-1)! \\ 0 & f'(\lambda) & f''(\lambda)/2 & \cdots & f^{(n-2)}(\lambda)/(n-2)! \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & f'(\lambda) & f''(\lambda)/2 \\ 0 & 0 & \cdots & 0 & f(\lambda) \end{bmatrix} \)
Hence we have:

**Spectral Mapping Theorem**

\[ \Delta(f(J)) = f(\Delta(J)) \]
\[ = \{f(\lambda_1), f(\lambda_2), \ldots, f(\lambda_n)\} \]

and, more generally,

\[ \Delta(f(A)) = f(\Delta(A)) \]

More generally, if we had

\[
J = \begin{bmatrix}
\lambda_1 & 1 & 0 & 0 & 0 \\
0 & \lambda_1 & 1 & 0 & 0 \\
0 & 0 & \lambda_1 & 1 & 0 \\
0 & 0 & 0 & \lambda_2 & 1 \\
0 & 0 & 0 & 0 & \lambda_2
\end{bmatrix}
\]

(recall that this Jordan form may be obtained from \( A \) by the similarity transform:

\[ J = T A T^{-1} \]

Where \( T^{-1} = \begin{bmatrix}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
e_1 v_1 w_1 e_2 v_2 e_3 v_3 e_4 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1
\end{bmatrix} \] where \( e_1, e_2, e_3, e_4 \) are eigenvectors and \( v_1, v_2, v_3, w_1 \) are generalized eigenvectors.)
Then \( f(A) = f(T^{-1}JT) \)
\[ = T^{-1}f(J)T \]

where \( f(J) = \begin{bmatrix} f(\lambda_1) & f'(\lambda_1) & f''(\lambda_1) \\ 0 & f(\lambda_1) & f'(\lambda_1) \\ 0 & 0 & f(\lambda_1) \end{bmatrix} \)

Example (from old prelim question — to be worked in class):

(a) A matrix \( A \) has minimal polynomial \((s-\lambda_1)^2 (s-\lambda_2)^3\). Find \( \cos(e^A) \) as a polynomial in \( A \).

(b) Now, assume further that \( A \) has characteristic polynomial \((s-\lambda_1)^5(s-\lambda_2)^3\) and that it has four linearly independent eigenvectors. Write down the Jordan form \( J \) of this matrix, and write down \( \cos(e^A) \) explicitly.