System Theory, KOM5108
Lecture #1

Instructor: Dr. Yavuz Eren

Lecture Book: **"Linear System Theory, W.J. Rugh''**

Spring 2024

Linear Vector Spaces(X, $\mathbb{F}$; linear vector space is associated with $\mathbb{F}$)

Definition: A set X is said to be linear vector space(LVS) if operations addition and scalar multiplication over the scalar field $\mathbb{F}$ are defined as follows :

$$(a)\ x + y \in \mathcal{X}\ for\ all\ x, y \in \mathcal{X}$$

$$(b)\ \alpha x \in \mathcal{X}\ for\ all\ x \in \mathcal{X}\ and\ \alpha \in \mathbb{F}$$

and for all $x. y, z \in X$ and $\alpha, \beta \in \mathbb{F}$, the following conditions hold:

(i) $x + y = y + x$

(ii) $(x + y) + z = x + (y + z)$

(iii) *There is a (unique) null vector* $\theta \in \mathcal{X}$ *such that* $x + \theta = x$

(iv) *For each* $x \in \mathcal{X}$, *there exists a (unique)* $-x \in \mathcal{X}$ *such that* $x + (-x) = \theta$.

(v) $\alpha(x + y) = \alpha x + \alpha y$

(vi) $(\alpha + \beta)x = \alpha x + \beta x$

(vii) $(\alpha\beta)x = \alpha(\beta x)$

(viii) $0x = \theta$

(ix) $1x = x$

Examples:

Definition(Linear independence): A set of vectors $\{x_1, \dots, x_m\} \subseteq \mathcal{X}$ is said to be linearly dependent if

$$\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_m x_m = 0 \quad \textit{implies} \quad \alpha_i = 0 \quad \forall i = 1 : m.$$

Definition(Span): Given a set of $\mathcal{V} = \{x_1, \dots, x_m\} \subseteq \mathcal{X}$, the span of $\mathcal{V}$ is the set of all linear combinations of $x_1, \dots, x_m$,

$$\operatorname{span}(\mathcal{V}) := \left\{ \sum_{i=1}^{m} \alpha_i x_i : \alpha_i \in \mathbb{F} \right\}.$$

Definitions(Basis and Dimension): A linearly independent a set of vectors, $\mathcal{V}$, is said to form a basis of $\mathcal{X}$ is span($\mathcal{V}$)= $\mathcal{X}$. The dimension of $\mathcal{X}$, denoted dim($\mathcal{X}$), is the number of elements N a basis of $\mathcal{X}$.

Examples:

Linear Subspaces

Definitions(Linear Susbspace): Assume that we have $(\mathcal{X}, \mathbb{F})$, a set $\mathcal{S} \subseteq \mathcal{X}$ is said to be a linear subspace of $\mathcal{X}$ if

$$\text{(i)}\quad x + y \in \mathcal{S} \ for\ all\ x, y \in \mathcal{S}$$

$$\text{(ii)}\quad \alpha x \in \mathcal{S} \ for\ all\ x \in \mathcal{S},\ \alpha \in \mathbb{F}.$$

Example:

Definition: Let $\mathcal{X}_1$and $\mathcal{X}_2$ be two linear vector spaces defined over the same field. The direct sum of spaces($\mathcal{X}_1 \oplus \mathcal{X}_2$ or $\mathcal{X}_1$x$\mathcal{X}_2$) is defined as the collection of ordered pairs (x_1, x_2) where $x_1 \subseteq \mathcal{X}_1$ and $x_2 \subseteq \mathcal{X}_2$.

Vector Norms and Normed Vector Spaces

Definition(Vector Norm): Let $\mathcal{X}$ be an linear vector space associated with FF. A vector norm on $\mathcal{X}$ is a function $\|.\|: \mathcal{X} \rightarrow \mathbb{R}$ which satisfies the following conditions:

$$\text{(i)}\quad \|x\| \geq 0 \ for\ all\ x \in \mathcal{X}$$

$$\text{(i')}\quad \|x\| = 0 \ if\ and\ only\ if\ x = 0$$

$$\text{(ii)}\quad \|ax\| = |a|\,\|x\| \ for\ all\ a \in \mathbb{F},\ x \in \mathcal{X}$$

$$\text{(iii)}\quad \|x + y\| \leq \|x\| + \|y\| \ for\ all\ x, y \in \mathcal{X}$$

Examples:

Convergency, Closedness and Completeness

An infinite sequence of $\mathrm{n}{\times}1$ vectors is written as $\{x_k\}_{k=0}^{\infty}$, where the subscript notation denotes different vectors rather than etries of a vector. A vector $\hat{x}$ is called as the limit of the sequence if for any given $\varepsilon > 0$ there exists a positive integer, written $K(\varepsilon)$ to indicate the integer depends on ε, such that

$$\|\hat{x} - x_k\| < \varepsilon, k > K(\varepsilon).$$

Often, we are interested in sequences of vector functions of time, denoted $\{x_k(t)\}_{k=0}^{\infty}$, and defİned on some interval ($\{t_0, t_1\}$). Such a sequence is said to converge(pointwise) on the interval if there exists a vector function $\hat{x}(t)$ such that for every $t_a \in \{t_0, t_1\}$ the sequence of vectors converge $\{x_k(t_a)\}_{k=0}^{\infty}$ converge to the vector $\hat{x}(t_a)$.

In this case, given an ε, the K can depend on both ε and t_a. The sequence of function converge uniformly on $[t_0, t_1]$ if there exists a function $\hat{x}(t)$ such that given $\varepsilon > 0$ there exists a positive integer $K(\varepsilon)$ such that for every t_a in the interval

$$\|\hat{x}(t_a) - x_k(t_a)\| < \varepsilon, k > K(\varepsilon).$$

For an infinite series of vector functions

$$\sum_{j=0}^{\infty} x_j(t) \qquad (*)$$

with each a $x_j(t)$ defined on $[t_0, t_1]$, convergence is defined in terms of sequence of partial sums

$$s_k(t) = \sum_{j=0}^{\infty} x_j(t)$$

The series converges to the function $\hat{x}_t$ if for each $t_a \in [t_0, t_1]$,

$$\lim_{k\to\infty} \|\hat{x}(t_a) - s_k(t_a)\| = 0$$

The series (*) is said to converge uniformly to $\hat{x}(t)$ on $[t_0, t_1]$. Namely, given $\varepsilon > 0$ there exists a positive integer $K(\varepsilon)$ such that for every t $\in [t_0, t_1]$,

$$\left\|\hat{x}(t) - \sum_{j=0}^{k} x_j(t)\right\| < \varepsilon$$

Suppose (*) is an infinite series of cont.-diff functions on $[t_0, t_1]$ that converges uniformly to $\hat{x}(t)$ on $[t_0, t_1]$, if the series

$$\sum_{J=0}^{\infty} \frac{d}{dt} x_j(t)$$

Converges uniformly on $[t_0, t_1]$, it converges to $\mathrm{d}\hat{x}(t)/dt$.

The infinite series (*) is said to be converge absolutely if the series of real functions

$$\sum_{J=0}^{\infty} \left|\left|x_j(t)\right|\right|$$

Converges on the interval. The key property of an absolutely convergent series is that the terms in the series can be reordered without changing the fact of convergence.

Definition:(Cauchy Sequence) A sequence $\{x_n\}$ in a normed linear space $(\mathcal{X}, \|.\|)$ is said to be Cauchy if for any $\varepsilon > 0$, there exists an N such that $n > N$ implies $\|x_m - x_n\| < \varepsilon$.

Definition:(Convergence) A sequence $\{x_k\}$ is said to be converge to x, denoted by $\{x_k\} \rightarrow x$, if for any $\varepsilon > 0$, there exists an N such that $n > N$ implies $\|x_m - x\| < \varepsilon$.

It is obvious that every convergent sequence is Cauchy. However, not every Cauchy sequence converges to a vector in the vector space considered.

Definition:(Complete Linear Vector Space) A normed $(\mathcal{X}, \|.\|)$ is said to be complete if every Cauchy sequence in $\mathcal{X}$ converges to an element in $\mathcal{X}$.

For elementary algebra , $\mathbb{R}$ is complete.

Definition:(Banach Space) A complete normed linear space is called a Banach space.

Examples:

Matrices

Represantation :$A \subseteq \mathbb{C}^{m \times n}$; $A \subseteq \mathbb{R}^{m \times n}$
Arithmetic Operations:

(i) *Matrix summation:* Given $A, B \in \mathbb{C}^{m \times n}$,

$$A + B = [a_{ij} + b_{ij}].$$

(ii) *Scalar multiplication:* Given $A \in \mathbb{C}^{m \times n}$ and $\alpha \in \mathbb{C}$,

$$\alpha A = [\alpha\, a_{ij}].$$

(iii) *Matrix multiplication:* Given $A \in \mathbb{C}^{m \times k}$ and $B \in \mathbb{C}^{k \times n}$,

$$AB = C \in \mathbb{C}^{m \times n}, \quad \text{where} \quad c_{ij} = \sum_{q=1}^{k} a_{iq} b_{qj}.$$

Generally, $AB \neq BA$ for two matrices $A, B \in \mathbb{C}^{n \times n}$. If $AB = BA$ they are said to be commute.
Adjoint of a matrix $A = [a_{ij}] \in \mathbb{C}^{m \times n}$ is $A^T = [a_{ji}] \in \mathbb{C}^{n \times m}$.

For real matrices, the adjoint and transpose are equal to each other. Two facts about adjoints are

$$(A + B)^* = A^* + B^*$$

$$(AB)^* = B^* A^*$$

Definitions: (Range and Null Space)

(i) *The* **range** *of A is the set*

$$\mathcal{R}(A) := \{y \in \mathbb{C}^m : Ax = y \text{ for some } x \in \mathbb{C}^n\}.$$

(ii) *The* **null space** *of A is the set*

$$\mathcal{N}(A) := \{x \in \mathbb{C}^n : Ax = 0\}.$$

The range of A is also called as the image of A(Im(A)) and null space of A is also called as the kernel of A(Ker(A)).
Both of the spaces are linear subspaces of A and are never empty.(Zero vector is always a member of both.)

Theorem: For $A \subseteq \mathbb{C}^{m \times n}$, n=dim(R(A))+dim(N(A)). Proof: HW

Theorem: For $A \subseteq \mathbb{C}^{m \times n}$, R(A)⊥N($A^*$)Proof: HW

Definition: The rank of a matrix $A \subseteq \mathbb{C}^{m \times n}(\mathrm{rank(A)})$, linearly independent columns is (or equivalently, rows) of A. It is said from the definitions:

$$\mathrm{rank}(A) = \dim(\mathcal{R}(A)) \qquad \mathrm{rank}(A) \leq \min\{m, n\}$$

Lemma(Sylvester's Rank Ineq.): For any $A \subseteq \mathbb{C}^{m\times n}$ and B $\subseteq \mathbb{C}^{k\times n}$

$$\mathrm{rank}(A) + \mathrm{rank}(B) - k \leq \mathrm{rank}(AB) \leq \min\{\mathrm{rank}(A), \mathrm{rank}(B)\}.$$

Definition: (Determinant): Given $A \subseteq \mathbb{C}^{n\times n}$, the determinant of A is :

$$\begin{aligned} \det(A) &:= \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(A_{[i,j]}) \quad \textit{for any } j = 1:n \\ &= \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A_{[i,j]}) \quad \textit{for any } i = 1:n, \end{aligned}$$

where $A_{[i,j]}$ denotes the submatrix of A obtained by deleting the ith row and jth column. Determinant function has the following features:

(i) $\det(A^T) = \det(A)$ *for any* $A \in \mathbb{C}^{n\times n}$.

(i') $\det(A^*) = \overline{\det(A)}$ *for any* $A \in \mathbb{C}^{n\times n}$.

(ii) $\det(AB) = \det(A)\det(B)$ *for any* $A, B \in \mathbb{C}^{n\times n}$.

(iii) $\det(I_n + AB) = \det(I_m + BA)$ *for any* $A \in \mathbb{C}^{n\times m}$ *and* $B \in \mathbb{C}^{m\times n}$.

(iv) $\det(\alpha A) = \alpha^n \det A$ *for any* $\alpha \in \mathbb{C}$, $A \in \mathbb{C}^{n\times n}$.

Definition(Invertibility): A matrix $A \subseteq \mathbb{C}^{n\times n}$ is said to be invertible if there exists a unique matrix in $\mathbb{C}^{n\times n}$, denoted by A^{-1}, such that

$$AA^{-1} = A^{-1}A = I.$$

The matrix A^{-1} is then called the inverse of A.

Some ways to check invertibility:

(i) *A is invertible*

(ii) $\mathrm{rank}(A) = n$

(iii) *The rows of A are linearly independent*

(iv) *The columns of A are linearly independent*

(v) $\det(A) \neq 0$

(vi) $\mathcal{R}(A) = \mathbb{C}^n$

(vii) $\mathcal{N}(A) = \{0\}$

(viii) $Ax = 0$ *implies* $x = 0$.

(ix) *For any* $b \in \mathbb{C}^n$, *there exists a unique* x *such that* $Ax = b$

(x) *0 is not an eigenvalue of A (see the next section).*

Examples:

Remark: Given an invertible matrix $A \in \mathbb{C}^{n\times n}$, the inverse of A can be computed as

$$A^{-1} = \frac{1}{\det A} C^T$$

chere C is the cofactor of A which is $c_{ij} = (-1)^{i+j} det A_{[i,j]}$.

For numerical solution(which requires much less operations)of inverse of such a function, the problem should be casted as solution of system equations. That is, solve for $x_i \in \mathbb{C}^{n\times 1}$ such that $Ax_i = e_i(\forall i = 1...n)$ where e_i denotes the ith column of the nth dimensioned identity matrix(I_n). After solving the problem, the inverse of A can be given as:

Remark: Some useful formulas for inverse operations:

$$A^{-1} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}$$

(i) $$(A + BCD)^{-1} = A^{-1} - A^{-1}B(C^{-1} + DA^{-1}B)^{-1}DA^{-1}$$

(ii) $$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} = \begin{bmatrix} \Delta^{-1} & -\Delta^{-1}BD^{-1} \\ -D^{-1}C\Delta^{-1} & D^{-1} + D^{-1}C\Delta^{-1}BD^{-1} \end{bmatrix}$$

$$= \begin{bmatrix} A^{-1} + A^{-1}B\nabla^{-1}CA^{-1} & -A^{-1}B\nabla^{-1} \\ -\nabla^{-1}CA^{-1} & \nabla^{-1} \end{bmatrix}.$$

where $\Delta \triangleq A - BD^{-1}C$ and $\nabla \triangleq A^{-1} - A^{-1}B(C^{-1} + DA^{-1}B)^{-1}DA^{-1}$ (if there exists)

Theorem(Schur's determinant identity): Given matrices $A \in \mathbb{C}^{n\times n}$, $B \in \mathbb{C}^{n\times m}$, $C \in \mathbb{C}^{m\times n}$ and $D \in \mathbb{C}^{m\times m}$ and also A is invertible, the following holds:

$$\det \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \det A \cdot \det(D - CA^{-1}B).$$

Proof: (HW)

Definition: The trace of $A \in \mathbb{C}^{n\times n}$ denoted as $Trace(A)$ is the sum of diagonal elements:

$$\mathbf{Trace}(A) = \sum_{i=1}^{n} a_{ii}$$

Note that trace function is linear. So that, for any $A, B \in \mathbb{C}^{n\times n}$

$$(i)\ \ \mathbf{Trace}(\alpha\, A) = \alpha\ \mathbf{Trace}(A) \text{ for any } \alpha \in \mathbb{C}$$

$$(ii)\ \ \mathbf{Trace}(A + B) = \mathbf{Trace}(A) + \mathbf{Trace}(B).$$

Proposition: For $A \in \mathbb{C}^{m\times n}$ and $B \in \mathbb{C}^{n\times m}$ Trace(AB)=Trace(BA).

Proof: (HW)

Eigenvalues and Eigenvectors

For a matrix $A \in \mathbb{C}^{n\times n}$, the equation $Ax = \lambda x$ where $x \in \mathbb{C}^n$(not equal to zero) is called the eigenvalue-eigenvector equations of A.

The scalar λ and the corresponding non-zero vectors x that satisfy $Ax = \lambda x$ are called the eigenvalue and eigenvectors of A, respectively.

The set of eigenvalues of A(also including the repetitions) is called as spectrum of A, as denoted as $\sigma(A)$.

Consider the rearranged form of above equations

$$(\lambda I - A)x = 0, \quad x \neq 0.$$

There exists a nontrival solution of to this equation iff the matrix $\lambda I - A$ is rank deficient as $p_A(\lambda) \triangleq \det(\lambda I - A) = 0$ where $p_A(\lambda)$ is the nth-order polynomial in λ and called as characteristic polynomial of A.

Note that, if A is real the spectrum of A is symmetric about the real axis(If λ_* is an eigenvalue of A, then so is λ_*, in other words $p_A(\lambda_*) = 0 \Leftrightarrow \overline{p_A(\lambda_*)} = 0 \Leftrightarrow p_A(\lambda_*) = 0$.)

Charactersitic polynomial: $p_A(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2) \dots (\lambda - \lambda_n)$.

Definition(Minimal Polynomial): Minimal polynomial of square matrix A is the monic polynomial $\phi(\lambda)$ of least degree such that $\phi(\lambda) = 0$.(Monic: The coeff. of highest term is one.)

Proposition: Two important facts regarding the eigenvalues of a matrix, $A \in \mathbb{C}^{n\times n}$ (A may have real, complex, distinct or repeted eigenvalues):

$$\det(A) = \prod_{i=1}^{n} \lambda_i$$

$$\mathbf{Trace}(A) = \sum_{i=1}^{n} \lambda_i$$

Proof: HW

Spectral Decomposition

Definition(Similarity): The matrices $\{A, B\} \in \mathbb{R}^{n\times n}$ are said to be similar is there exists a nonsingular $S \in \mathbb{R}^{n\times n}$ such that $B = S^{-1}AS$, the transformation $A \to S^{-1}AS$ is called a similarity transformation of A under S.

Definition(Diagonalizability): A square matrix is said to be diagonalizable if it is similar to a diagonal matrix.

Theorem: A matrix $A \in \mathbb{R}^{n\times n}$ is diagonalizable iff it has n linearly independent eigenvectors.

Remark: Any square matrix A can be decomposed as A= MJM^{-1} for some $J \in \mathbb{C}^{n\times n}$ as so-called Jordan canonical form with some invertable $M \in \mathbb{C}^{n\times n}$, as seen below

$$\begin{bmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_k \end{bmatrix}$$

where $k \leq n$ and each $J_i \in \mathbb{C}^{n_i\times n_i}$ is in the following form :

$$J_i = \begin{bmatrix} \lambda_i & 1 & 0 & \cdots & 0 \\ 0 & \lambda_i & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_i & 1 \\ 0 & 0 & \cdots & 0 & \lambda_i \end{bmatrix} \in \mathbb{C}^{n_i \times n_i}$$

If each Jordan block is scalar($n_i = 1$ for all i=1,..,k), A is said to be diagonalizable. So, it is said that A is diagonalizable iff it has n linearly independent eigenvectors.

In this context, we have two cases as A does or does not have linearly independent eigenvectors:

Case 1(A has linearly independent eigenvectors): In this case $Ax = \lambda x$ can be given a matrix equations as follows:

$$AU = U\Lambda$$

where

$$U := \begin{bmatrix} x^1 & x^2 & \cdots & x^n \end{bmatrix} \qquad \Lambda := \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

and x^i is the eigenvector corresponding to the eigenvalue λ_i. Since U is now invertible by the above theorem , we get

$$A = U\Lambda U^{-1}$$

and k=n which means each Jordan block is a scalar.(Spectral Decomposition with U=M and $\Lambda = J$.)

A sufficient condition for A to have linearly independent eigenvectors is that it have distinct eigenvalues as given lemma below

Lemma: For $A \in \mathbb{R}^{n \times n}$, let $\lambda_1, \ldots, \lambda_k$ with $k \leq n$ be distinct eigenvalues of A and let $x^1, .., x^k$ be the associated eigenvalues. Then, the set of $(x^1, .., x^k)$ is linearly independent.
Proof: HW

Note that, the reverse implication is not necessarily true. Because, there may be more than one independent eigenvector associated with a repeated eigenvalue.
Ex.: I_n, the identity matrix of dimension n. All n of its igenvalues are equal to 1. However it has n linearly independent eigenvectors(taken by its colunms).

Note that: When we talk about the multiplicity of an eigenvalue, we usually mean its algebraic multiplicity, meaning its multiplicity as a root of the characteristic polynomial. The number of linearly independent eigenvectors associated with an eigenvalue is defined as geometric multiplicity. Hence, in the case of I_n, algebraic multiplicity and geometric multiplicity of the eigenvalue ,1, are both n.

Note that, Assuming the matrix A is real, but it has complex eigenvalues(so complex eigenvectors). In this case, we may prefer to work with real matrices M and J. Let us to accomplish this case as follows:

If A is real, the eigenvalues of A are symmetric about the real axis and the corresponding eigenvectors are complex conjugates of each other. For simplicity, suppose n=2 and the two eigenvalues of A are complex congugate to each other. As denoting as the eigenvalues as $\alpha \pm \beta j \in \mathbb{C}$ and the corresponding eigenvectors as $x \pm yj \in \mathbb{C}^n$. Then, we have the following transformation

$$
\begin{aligned}
A = T\Lambda T^{-1} &= \begin{bmatrix} x+yj & x-yj \end{bmatrix} \begin{bmatrix} \alpha+\beta j & 0 \\ 0 & \alpha-\beta j \end{bmatrix} \begin{bmatrix} x+yj & x-yj \end{bmatrix}^{-1} \\
&= \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 1 & 1 \\ j & -j \end{bmatrix} \begin{bmatrix} \alpha+j\beta & 0 \\ 0 & \alpha-j\beta \end{bmatrix} \begin{bmatrix} 1 & 1 \\ j & -j \end{bmatrix}^{-1} \begin{bmatrix} x & y \end{bmatrix}^{-1} \\
&= \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 1 & 1 \\ j & -j \end{bmatrix} \begin{bmatrix} \alpha+j\beta & 0 \\ 0 & \alpha-j\beta \end{bmatrix} \begin{bmatrix} 1/2 & -j/2 \\ 1/2 & j/2 \end{bmatrix} \begin{bmatrix} x & y \end{bmatrix}^{-1} \\
&= \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix} \begin{bmatrix} x & y \end{bmatrix}^{-1}.
\end{aligned}
$$

Case 2(A does not have linearly independent eigenvectors): "A" does have repeated eigenvalues and the eigenvectors associated with, are not all linearly independent. For this case, we have to resort to generalized eigenvectors and the Jordan form of A.

Suppose λ is an eigenvalue of algebraic multiplicity q, but with geometric multiplicity 1(λ is a q time repeated eigenvalue of A, but there is only one eigenvectors associated with it). Then we first solve for eigenvector associated with "$Ax_1 = \lambda x_1$" and solve $x_2, x_3, \ldots, x_q$ such that

$$\begin{gathered} Ax_2 = x_1 + \lambda x_2 \\ Ax_3 = x_2 + \lambda x_3 \\ \vdots \\ Ax_q = x_{q-1} + \lambda x_q \end{gathered}$$

In this formulation, $x_2, x_3, \ldots, x_q$ are called as generalized eigenvectors associated with λ. Then, we can show that the set of $\{x_1, x_2, \ldots, x_q\}$ generalized eigenvectors.

$$A \begin{bmatrix} x_1 & \cdots & x_q \end{bmatrix} = \begin{bmatrix} x_1 & \cdots & x_q \end{bmatrix} \begin{bmatrix} \lambda & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 1 \\ 0 & \cdots & \cdots & \lambda \end{bmatrix}$$

By repeating the same process for each of the eigenvalues, we get the diagonal form.

Theorem(Cayley Hamilton): Every square matrix satisfies its own charactersitic polynomial.

Proof: (HW)

Ex:

- By the conclusion of Cayley Hamilton theorem, it can be said that ''similar matrices have the same eigenvalues''.

Lemma: If matrices $\mathrm{A}, \mathrm{B} \in \mathbb{R}^{n \times n}$ are similar, then $\sigma(A) = \sigma(B)$.

Proof: Let $B = S^{-1}AS$ for the invertable S. Alsıo, we know that the eigenvalues of A and B are the roots of the polynomials $p_A(\lambda)$ and $p_B(\lambda)$.

$p_B(\lambda) = det(\lambda I - S^{-1}AS) = det(S^{-1}(\lambda I - A)S)$=$det(S^{-1})\,\det(\lambda I - A)\,\det(S) = \det(\lambda I - A)$

It is easily seen that, $\det(\lambda I - A)$ is nothing but $p_A(\lambda)$. So that, A and B have the same characteristic polynomials and the same spectra.

Remark: For similarity, having the same spectrum is necessary but not sufficient.

Ex: Consider the following matrices. The eigenvalues of both are zero with multiplicity 2 but they are not similar.

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Left Eigenvectors: Let $\mathrm{A} \in \mathbb{C}^{n\times n}$, a scalar $\lambda \in \mathbb{C}$ and a vector $\mathrm{x} \in \mathbb{C}^n$ are said to satisfy the eigenvalue, the left eigenvector equation if

$$x^*A = \lambda x^*, \qquad x \neq 0.$$

Note that, there is no distinction between left and right eigenvalue of A. Moreover, if (λ, x) is an eigenvalue, right eigenvector of A, then (λ, x) is an eigenvalue, left eigenvector pair of A^*.

To see how the right and left eigenvectors are related, assume that A has linearly independent right eigenvectors. Then, A can be decomposed as $A = S\Lambda S^{-1}$ where S is the matrix of right eigenvectors as $S \triangleq [x_1 \quad x_2 \quad \cdots \quad x_n]$.

Also, one can show that A can be decomposed as $A = T^{-1}\Lambda T$ where T is the matrix of left eigenvectors as follows

$$T = \triangleq \begin{bmatrix} z_1^* \\ z_2^* \\ \vdots \\ z_n^* \end{bmatrix}$$

where $z_1^* A = \lambda_i z_i^*$. The left eigenvector can be obtained from the rows of the inverse of the right eigenvector matrix. Therefore A can be decomposed as

$$A = S\Lambda T = \sum_{i=1}^{n} x_i z_i^*.$$