Linear System Theory

Jun Moon

Lecture 4
Linear Algebra

March 12, 2018

Matrix, Function, Linear Algebra

Linear vector space	$\Leftrightarrow$	state space, input space, output space
Linear operators	$\Leftrightarrow$	reachability (controllability), observability operator
Normed spaces	$\Leftrightarrow$	stability
Inner product	$\Leftrightarrow$	controllability and observability grammians

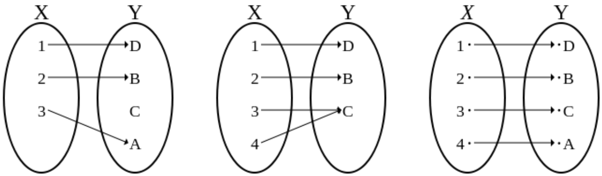
I will use more advanced materials for this lecture. I will upload it on the course website

Preliminaries

- Sets: $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{N}$, $\mathbb{R}$, $\mathbb{R}_+$, $\mathbb{C}$, $\mathbb{C}_+$
- Cartesian product: two sets X and Y, the cartesian product is $X \times Y$
- Function: $f : X \to Y \Rightarrow$ for any $x \in X$, f assigns a unique $f(x) = y \in Y$
- X: domain, Y: codomain
- $\{f(x) : x \in X\}$: range
- f: function, operator, map, transformation

Preliminaries

- Injective (one to one): f is injective if and only if $f(x_1) = f(x_2) \Rightarrow x_1 = x_2$, conversely, $x_1 \neq x_2$ implies $f(x_1) \neq f(x_2)$
- Surjective (onto): f is surjective if and only if for all $y \in Y$, there exists $x \in X$ such that $y = f(x)$
- Bijective (one to one correspondnce): f is bijective if and only if for all $y \in Y$, there exists a unique $x \in X$ such that $y = f(x)$
- equivalently, f is bijective if and only if f is injective and surjective



Fields

- A field $\mathbb{F}$ is an object that consisting of a set of elements, and two binary operations:
  addition $(+)$, multiplication $(\cdot)$ such that

Addition

(i) associative: $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$ for all $\alpha, \beta, \gamma \in \mathbb{F}$
(ii) commutative: $\alpha + \beta = \beta + \alpha$ for all $\alpha, \beta \in \mathbb{F}$
(iii) there exists an identity element 0: $\alpha + 0 = \alpha$ for all $\alpha \in \mathbb{F}$
(iv) for all $\alpha \in \mathbb{F}$, there exists an inverse $-\alpha$ such that $\alpha + (-\alpha) = 0$

Multiplication

(i) associative: $\alpha \cdot (\beta \cdot \gamma) = (\alpha \cdot \beta) \cdot \gamma$ for all $\alpha, \beta, \gamma \in \mathbb{F}$
(ii) commutative: $\alpha \cdot \beta = \beta \cdot \alpha$ for all $\alpha, \beta \in \mathbb{F}$
(iii) there exists an identity element 1: $\alpha \cdot 1 = \alpha$ for all $\alpha \in \mathbb{F}$
(iv) for all $\alpha \in \mathbb{F}$, $\alpha \neq 0$, there exists an inverse α^{-1} such that
$\alpha \cdot \alpha^{-1} = \alpha^{-1} \cdot \alpha = 1$

Fields

Examples: $\mathbb{R}$, $\mathbb{C}$, $\mathbb{Q}$

Example: Consider the set of all 2×2 matrices for the form
Matrix addition and multiplication

$$\begin{pmatrix} x & -y \\ y & x \end{pmatrix}, \; x, y \in \mathbb{R}, \; 0 \Rightarrow \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \; 1 \Rightarrow \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

The inverse always exists

How about the set of all 2×2 matrices?

Vector Space

- vector space = linear space = linear vector space
- A linear space over a field $\mathbb{F}$, $(\mathbb{V}, \mathbb{F})$, consists of a set $\mathbb{V}$ of vectors, a field $\mathbb{F}$, and two operations, vector addition and scalar multiplication
- The two operations satisfy

Addition
(a) addition: for all $x, y \in \mathbb{V}$, $x + y \in \mathbb{V}$
(b) associative: for all $x, y, z \in \mathbb{V}$, $(x + y) + z = x + (y + z)$
(c) commutative: for all $x, y \in \mathbb{V}$, $x + y = y + x$
(d) there exists a unique zero vector $0 \in \mathbb{V}$ such that $x + 0 = 0 + x = x$ for all $x \in \mathbb{V}$
(e) there exists a unique inverse $-x \in \mathbb{V}$ such that $x + (-x) = 0$ for all $x \in \mathbb{V}$

Vector Space

- vector space = linear space = linear vector space
- A linear space over a field $\mathbb{F}$, $(\mathbb{V}, \mathbb{F})$, consists of a set $\mathbb{V}$ of vectors, a field $\mathbb{F}$, and two operations, vector addition and scalar multiplication
- The two operations satisfy

Multiplication
(a) multiplication: for any $\alpha \in \mathbb{F}$ and $x \in \mathbb{V}$, $\alpha x \in \mathbb{V}$
(b) associative: for any $\alpha, \beta \in \mathbb{F}$ and $x \in \mathbb{V}$, $\alpha(\beta x) = (\alpha\beta)x$
(c) distributive w.r.t. scalar addition:

for any $\alpha \in \mathbb{F}$ and $x, y \in \mathbb{V}$, $\alpha(x + y) = \alpha x + \alpha y$

(d) distributive w.r.t. scalar multiplication

for any $\alpha, \beta \in \mathbb{F}$ and $x \in \mathbb{V}$, $(\alpha + \beta)x = \alpha x + \beta x$

(e) there exists a unique $1 \in \mathbb{F}$ such that for any $x \in \mathbb{V}$, $1x = x$
(f) there exists a unique $0 \in \mathbb{F}$ such that for any $x \in \mathbb{V}$, $0x = 0$

Vector Space

- vector space = linear space = linear vector space
- A linear space over a field $\mathbb{F}$, $(\mathbb{V}, \mathbb{F})$, consists of a set $\mathbb{V}$ of vectors, a field $\mathbb{F}$, and two operations, vector addition and scalar multiplication

Example: $(\mathbb{F}^n, \mathbb{F})$ where $\mathbb{F}^n = \mathbb{F} \times \cdots \times \mathbb{F}$

Example: $(\mathbb{R}^n, \mathbb{R})$, $(\mathbb{C}^n, \mathbb{C})$, $(\mathbb{C}^n, \mathbb{R})$

Example: $(\mathbb{R}, \mathbb{C})$ is not a vector space! (why?) $(1 + i)1 = 1 + i \notin \mathbb{R}$

Example: a continuous function $f : [t_0, t_1] \to \mathbb{R}^n$, the set of such functions, $(C([t_0, t_1], \mathbb{R}^n), \mathbb{R})$, is a linear space

Example: You will see other examples in HW

Vector Subspaces (Linear Subspaces)

Let $(\mathbb{V}, \mathbb{F})$ be a linear space, and $\mathbb{W} \subset \mathbb{V}$, i.e., $\mathbb{W}$ is a subset of $\mathbb{V}$.

Then $(\mathbb{W}, \mathbb{F})$ is called a subspace of $(\mathbb{V}, \mathbb{F})$ if $(\mathbb{W}, \mathbb{F})$ is itself a vector space

We can show that $(\mathbb{W}, \mathbb{F})$ is a subspace if

- $\mathbb{W} \subset \mathbb{V}$
- for any $w_1, w_1 \in \mathbb{W}$ and $\alpha, \beta \in \mathbb{F}$, $\alpha w_1 + \beta w_2 \in \mathbb{W}$

Vector Subspaces

Example: In $(\mathbb{R}^2, \mathbb{R})$, every straight line passing through the origin is a subspace of $(\mathbb{R}^2, \mathbb{R})$. That is,

$$\begin{pmatrix} x_1 \\ cx_2 \end{pmatrix}$$

for any fixed $c \in \mathbb{R}$ is a subspace of $(\mathbb{R}^2, \mathbb{R})$

Example: The vector space $(\mathbb{R}^n, \mathbb{R})$ is a subspace of $(\mathbb{C}^n, \mathbb{R})$

Vector Subspaces

Question: Prove that if $\mathbb{W}_1$ and $\mathbb{W}_2$ are subspaces of $\mathbb{V}$, then

- $\mathbb{W}_1 \cap \mathbb{W}_2$ is a subspace
- $\mathbb{W}_1 \cup \mathbb{W}_2$ is not necessarily a subspace
  Proof of (ii): We provide a counter example

$$\mathbb{W}_1 = \text{span}\left\{ \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}, \ \mathbb{W}_2 = \text{span}\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\}$$

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} \notin \mathbb{W}_1 \cup \mathbb{W}_2$$

Linear Independence and Dependence

Suppose $(\mathbb{V}, \mathbb{F})$ is a linear space (vector space). The set of vectors $\{v_1, ..., v_p\}$, $v_i \in \mathbb{V}$ is said to be *linearly independent* if and only if

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_p v_p = 0$$
$$\Rightarrow \alpha_1 = \alpha_2 = \cdots = \alpha_p = 0$$

where $\alpha_i \in \mathbb{F}$

The set of vectors is said to be linearly dependent if and only if it is not linearly independent

If the set of vectors is linearly dependent, then at least one (not everyone!!!) of them can be written as a linear combination of others

Linear Independence and Dependence

Example: $x_1, ..., x_n \in \mathbb{R}^n$ and $x_1 = 0$. This is linearly dependent, since $\alpha_1 = 1$, and $\alpha_2, ..., \alpha_n = 0$

Example:

$$\alpha_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} -3 \\ 2 \end{pmatrix} + \alpha_3 \begin{pmatrix} 2 \\ 4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} 1 & 0 & 16/5 \\ 0 & 1 & 2/5 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow \begin{pmatrix} \alpha_1 \\ \alpha_1 \end{pmatrix} = -\alpha_3 \begin{pmatrix} 16/5 \\ 2/5 \end{pmatrix}$$

Linear Independence and Dependence

Example: $\mathbb{F} = \mathbb{R}$, $k \in \{0, 1, ...\}$, and define

$$f_k : [-1, 1] \to \mathbb{R} \text{ such that } f_k(t) = t^k$$

The set $\{f_0, ..., f_n\}$ is linearly independent in $(\mathbb{F}([-1, 1], \mathbb{R}), \mathbb{R})$

We need to show $\alpha_0 + \alpha_1 t \alpha_2 t^2 + \cdots + \alpha_n t^n = 0$, where $t \in [-1, 1]$ and $\alpha_i \in \mathbb{R}$, implies $\alpha_i = 0$ for all i

Dimension and Basis

Dimension: The maximal number of linear independent vectors in a linear space $(\mathbb{V}, \mathbb{F})$ is called the dimension of the linear space $(\mathbb{V}, \mathbb{F})$

Example: $(\mathbb{F}^n, \mathbb{F})$

Dimension and Basis

Basis: A set of linearly independent vectors of a linear space $(\mathbb{V}, \mathbb{F})$ is said to be a basis of $\mathbb{V}$ if every vector in $\mathbb{V}$ can be expressed as a unique linear combination of these vectors

5mm Basis (equivalent definition): For a linear space $(\mathbb{V}, \mathbb{F})$, a set of vectors $\{v_1, ..., v_n\}$ is called basis of $\mathbb{V}$ if

- $\{v_1, ..., v_n\}$ spans $\mathbb{V}$, $\mathbb{V} = \{\sum_{i=1}^{n} \alpha_i v_i : \ \alpha_i \in \mathbb{F}, \forall i\}$
- $\{v_1, ..., v_n\}$ is linearly independent

Example: Suppose that $\{v_1, .., v_n\}$ is a basis of $\mathbb{V}$. Then for any $v \in \mathbb{V}$, there exists a unique $\alpha_1, ..., \alpha_n \in \mathbb{F}$ such that

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$$

Dimension and Basis

Example: $(\mathbb{R}^n, \mathbb{R})$, $(\mathbb{C}, \mathbb{C})$, $(\mathbb{C}^n, \mathbb{C})$, $(\mathbb{C}, \mathbb{R})$, $(\mathbb{C}^n, \mathbb{R})$
Example: $(\mathbb{R}^{2\times 2}, \mathbb{R})$, $(\mathbb{R}^{n\times n}, \mathbb{R})$

Example: $\mathbb{F} = \mathbb{R}$, $k \in \{0, 1, ...\}$, and define

$$f_k : (-\infty, \infty) \to \mathbb{R} \text{ such that } f_k(t) = t^k$$

The set $\{f_0, f_1 ...\}$ is linearly independent in $(\mathbb{F}((-\infty, \infty), \mathbb{R}), \mathbb{R})$
Equivalently, there exist no real constants α_i's not all zero, such that

$$\sum_{k=1}^{\infty} \alpha_i t^k = 0$$

$\Rightarrow$ The dimension of $(\mathbb{F}((-\infty, \infty), \mathbb{R}), \mathbb{R})$ is not finite (infinite-dimensional space)

Dimension and Basis

Question: What is the relationship between dimension and basis in a vector space?

Theorem

In an n-dimensional vector space, $(\mathbb{V}, \mathbb{F})$, any set of n linearly independent vectors forms a basis

Dimension and Basis

By definition of the n-dimensional vector space, there exists a set of n linearly independent vectors, $\{v_1, ..., v_n\}$ (note that n is the maximum number of linearly independent vectors that we can have).

Let $x \in \mathbb{V}$, and consider $\{x, v_1, ..., v_n\}$ (note that they are linearly dependent!). There exist $\alpha_0, \alpha_1, .., \alpha_n$ in $\mathbb{F}$ not all zero such that

$$\alpha_0 x + \alpha_1 v_1 + \cdots + \alpha_n v_n = 0$$

Claim: $\alpha_0 \neq 0$ (why??) $\Rightarrow$ by contradiction, suppose that $\alpha_0 = 0$. Then

$$\alpha_1 v_1 + \cdots + \alpha_n v_n = 0$$

implies $\alpha_1, ..., \alpha_n = 0$ contradicts the fact that $\alpha_i \neq 0$ for all i

Dimension and Basis

Hence,

$$\begin{aligned} x &= (-\alpha_1/\alpha_0)v_1 + \cdots + (-\alpha_n/\alpha_0)v_n \\ &= \beta_1 v_1 + \cdots + \beta_n v_n \end{aligned}$$

This implies that every vector $x \in \mathbb{V}$ can be expressed as a linear combination of $\{v_1, ..., v_n\}$

$$\begin{aligned} x &= (-\alpha_1/\alpha_0)v_1 + \cdots + (-\alpha_n/\alpha_0)v_n \\ &= \beta_1 v_1 + \cdots + \beta_n v_n \end{aligned}$$

What is the next? $\Rightarrow$ uniqueness
Suppose that there is another linear combination of x, namely

$$x = \beta_1 v_1 + \cdots + \beta_n v_n = \gamma_1 v_1 + \cdots + \gamma_n v_n$$

We have show that $\beta_i = \gamma_i$ for all i. Consider,

$$0 = (\beta_1 - \gamma_1)v_1 + \cdots + (\beta_n - \gamma_n)v_n$$

Dimension and Basis

Dimension and Basis: If a linear space $(\mathbb{V}, \mathbb{F})$ has a basis of n elements, then $(\mathbb{V}, \mathbb{F})$ is said to be of dimension n. In other words, $(\mathbb{V}, \mathbb{F})$ is a n-dimensional linear space, and we write

$$\dim(\mathbb{V}) = n$$

- Basis is not unique!!
- Given a basis, every vector $x \in \mathbb{V}$ can be uniquely represented by a set of scalars in $\mathbb{F}$. Namely, for any $v \in \mathbb{V}$, there exists a unique α_is in $\mathbb{F}$ such that

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n$$

- $\alpha = (\alpha_1, ..., \alpha_n) \in \mathbb{F}^n$ is called the representation of v or the coordination vector of v
- α_i is called the coordinate of v w.r.t. v_i

Dimension and Basis

Example: $(\mathbb{R}^n, \mathbb{R})$, $(\mathbb{C}, \mathbb{C})$, $(\mathbb{C}^n, \mathbb{C})$, $(\mathbb{C}, \mathbb{R})$, $(\mathbb{C}^n, \mathbb{R})$
Example: $(\mathbb{R}^{2\times 2}, \mathbb{R})$, $(\mathbb{R}^{n\times n}, \mathbb{R})$

Example: $\mathbb{F} = \mathbb{R}$, $k \in \{0, 1, ...\}$, and define

$$f_k : (-\infty, \infty) \to \mathbb{R} \text{ such that } f_k(t) = t^k$$

The set $\{f_0, f_1 ...\}$ is linearly independent in $(\mathbb{F}((-\infty, \infty), \mathbb{R}), \mathbb{R})$
Equivalently, there exist no real constants α_i's not all zero, such that

$$\sum_{k=1}^{\infty} \alpha_i t^k = 0$$

$\Rightarrow$ The dimension of $(\mathbb{F}((-\infty, \infty), \mathbb{R}), \mathbb{R})$ is not finite (infinite-dimensional space)
$\Rightarrow$ Basis $\{1, t, t^2, ...\}$

Change of Basis

- Basis is not unique
- For each basis, there are different representations of a vector x
- Suppose that $\{v_1, ..., v_n\}$ and $\{e_1, ..., e_n\}$ are bases (plural of basis) of the linear space $(\mathbb{V}, \mathbb{F})$. Then

$$x = \sum_{i=1}^{n} \alpha_i v_i = \sum_{i=1}^{n} \beta_i e_i, \ (\alpha_i \neq \beta_i, \ \forall i)$$

- We need to figure out the relationship between the different representations of the same vector

Change of Basis

Suppose that $\{v_1, ..., v_n\}$ and $\{e_1, ..., e_n\}$ are bases (plural of basis) of the linear space $(\mathbb{V}, \mathbb{F})$. Then for any $x \in \mathbb{V}$, we have a unique representation of x for each basis:

$$x = \sum_{i=1}^{n} \alpha_i v_i = \sum_{i=1}^{n} \beta_i e_i, \ (\alpha_i \neq \beta_i, \ \forall i)$$

$$= \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} e_1 & \cdots & e_n \end{pmatrix} \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} = [v]\alpha = [e]\beta$$

Now, for each e_i, there is a unique representation w.r.t. v

$$e_i = \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix} \begin{pmatrix} \gamma_{1i} \\ \vdots \\ \gamma_{ni} \end{pmatrix}$$

Change of Basis

For each e_i, there is a unique representation w.r.t. v

$$e_i = \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix} \begin{pmatrix} \gamma_{1i} \\ \vdots \\ \gamma_{ni} \end{pmatrix}$$

Hence,

$$\begin{pmatrix} e_1 & e_2 & \cdots e_n \end{pmatrix} = \begin{pmatrix} v_1 & v_2 & \cdots v_n \end{pmatrix} \begin{pmatrix} \gamma_{11} & \gamma_{12} & \cdots & \gamma_{1n} \\ \gamma_{21} & \gamma_{22} & \cdots & \gamma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{n1} & \gamma_{n2} & \cdots & \gamma_{nn} \end{pmatrix}$$

$$\Rightarrow \quad [e] = [v][\gamma]$$

Now,

$$x = [v]\alpha = [e]\beta = [v][\gamma]\beta$$

Since for each $x \in \mathbb{V}$ there is a unique representation of x w.r.t. $(v_1, ..., v_n)$, we have

$$\alpha = [\gamma]\beta$$

Change of Basis

Now,

$$x = [v]\alpha = [e]\beta = [v][\gamma]\beta$$

Since for each $x \in \mathbb{V}$ there is a unique representation of x w.r.t. $(v_1, ..., v_n)$, we have

$$\alpha = [\gamma]\beta \Rightarrow \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} \gamma_{11} & \gamma_{12} & \cdots & \gamma_{1n} \\ \gamma_{21} & \gamma_{22} & \cdots & \gamma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{n1} & \gamma_{n2} & \cdots & \gamma_{nn} \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix}$$

or

$$\begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix} = \begin{pmatrix} \gamma_{11} & \gamma_{12} & \cdots & \gamma_{1n} \\ \gamma_{21} & \gamma_{22} & \cdots & \gamma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{n1} & \gamma_{n2} & \cdots & \gamma_{nn} \end{pmatrix}^{-1} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}$$

Change of Basis

Similarly,

$$v_i = \begin{pmatrix} e_1 & \cdots & e_n \end{pmatrix} \begin{pmatrix} p_{1i} \\ \vdots \\ p_{ni} \end{pmatrix}$$

Hence,

$$[v] = [e] \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{pmatrix} = [e][p]$$

We have

$$x = [v]\alpha = [e]\beta = [e][p]\alpha$$

$$\Rightarrow \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix} = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}, \quad \alpha = [p]^{-1}\beta$$

Change of Basis

Now, we can see that

$$\begin{aligned} x &= [v]\alpha = [e]\beta = [e][p]\alpha \\ &= [v]\alpha = [e]\beta = [v][\gamma]\beta \end{aligned}$$

This implies

$$\begin{aligned} \beta = [p]\alpha,\ \alpha = [\gamma]\beta &\Rightarrow \alpha = [\gamma][p]\alpha \\ &\Rightarrow [\gamma][p] = I \end{aligned}$$

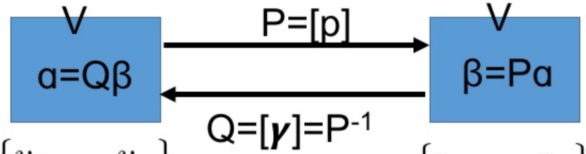
Hence,

$$[\gamma] = [p]^{-1} \quad \text{or} \quad [p] = [\gamma]^{-1}$$

Change of Basis

This is known as Change of Basis!!! For x, suppose that the two bases (plural of basis) $\{v_1, ..., v_n\}$ and $\{e_1, ..., e_n\}$ are given, and you know that $x = \sum_{i=1}^{n} \alpha_i v_i = [v]\alpha$. Then you can represent x w.r.t. $\{e_1, ..., e_n\}$ with $x = [e][p]\alpha = [e]\beta$

Change of Basis



$\{v_1, \dots, v_n\}$ $\{e_1, \dots, e_n\}$

Example

$$v = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}, \ e = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix} \right\}, \ x = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \ \alpha = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

$$[e] = [v][\gamma], \ [\gamma] = [e], \ [p] = [\gamma]^{-1} = [e]^{-1}$$

Hence,

$$x = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = [e][p]\alpha = [e][e]^{-1}\alpha = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1.5 \\ 0.5 \end{pmatrix}$$

Linear Maps (Linear Operators)

Recall that

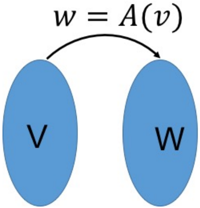
- Function: $f : X \to Y \Rightarrow$ for any $x \in X$, f assigns a unique $f(x) = y \in Y$
- X: domain, Y: codomain
- $\{f(x) : x \in X\}$: range
- f: function, operator, map, transformation
- injective, surjective, bijective

Linear Maps (Linear Operators)

Linear Mapping:

Let $(\mathbb{V}, \mathbb{F})$ and $(\mathbb{W}, \mathbb{F})$ be (finite-dimensional) linear vector spaces on the SAME FIELD!!!. Let $\mathcal{A}$ be a map from $\mathbb{V}$ to $\mathbb{W}$, i.e., $\mathcal{A} : \mathbb{V} \to \mathbb{W}$ such that $\mathcal{A}(v) = w$, $v \in \mathbb{V}$ and $w \in \mathbb{W}$. Then $\mathcal{A}$ is said to be a linear mapping (equiv. linear transformation, linear operator, linear map), if and only if

$$\mathcal{A}(\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 \mathcal{A}(v_1) + \alpha_2 \mathcal{A}(v_2), \ \forall \alpha_1, \alpha_2 \in \mathbb{F}$$



Linear Maps (Linear Operators)

We will see that if $\mathcal{A}$ is linear, then $\mathcal{A}$ is a matrix. In other words, we will see that any linear mapping between finite-dimensional linear spaces can be represented as matrix multiplication

Linear Maps (Linear Operators)

Example: Consider the following mapping on the set of polynomials of degree 2:

$$\mathcal{A} : as^2 + bs + c \to cs^2 + bs + a,$$

where $a, b, c \in \mathbb{R}$. Is this a linear map?
Let $= v_1 = a_1 s^2 + b_1 s + c_1$ and $v_2 = a_2 s^2 + b_2 s + c_2$. Then for any $\alpha_1, \alpha_2 \in \mathbb{R}$

$$\begin{aligned}
&\mathcal{A}(\alpha_1 v_1 + \alpha_2 v_2) \\
&= \mathcal{A}(\alpha_1(a_1 s^2 + b_1 s + c_1) + \alpha_2 a_2 s^2 + b_2 s + c_2) \\
&= (\alpha_1 c_1 + \alpha_2 c_2)s^2 + (\alpha_1 b_1 + \alpha_2 b_2)s + \alpha_1 a_1 + \alpha_2 a_2 \\
&= \alpha_1 \mathcal{A}(v_1) + \alpha_2 \mathcal{A}(v_2)
\end{aligned}$$

Linear Maps (Linear Operators)

Example: The mapping $\mathcal{A}(v)$ where

$$w = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 7 \\ 7 & 0 & 16 \end{pmatrix} v$$

Example: Let $\mathcal{U}$ be the set of continuous functions defined on $[0, T]$. Then $(\mathcal{U}, \mathbb{R})$ a vector space (why???). Let g be a continuous function defined on $[0, T]$. Then for $u \in \mathcal{U}$, the transform

$$y(t) = \int_0^T g(t - \tau)u(\tau)d\tau$$

is a linear transformation. In this case, y is a linear mapping from $(\mathcal{U}, \mathbb{R})$ to $(\mathcal{U}, \mathbb{R})$.

Linear Maps (Linear Operators)

Theorem

Let $(\mathbb{V}, \mathbb{F})$ have a basis $\{v_1, ..., v_n\}$, and let $(\mathbb{W}, \mathbb{F})$ have a basis $\{w_1, ..., w_m\}$. Then w.r.t. these bases, the linear map $\mathcal{A}$ is represented by $m \times n$ matrix. Namely, $y = \mathcal{A}(x) = Ax$, where A is a $m \times n$ matrix.

Linear Maps (Linear Operators)

Proof: For any $x \in \mathbb{V}$, there exists a unique $\alpha_1, ..., \alpha_n \in \mathbb{F}$ such that

$$x = \sum_{j=1}^{n} \alpha_i v_j = [v]\alpha$$

By linearity,

$$y = \mathcal{A}(x) = \alpha_1 \mathcal{A}(v_1) + \cdots + \alpha_n \mathcal{A}(v_n) = \sum_{j=1}^{n} \alpha_j \mathcal{A}(v_j)$$

Note that $\mathcal{A} : \mathbb{V} \to \mathbb{W}$; hence for each $\mathcal{A}(v_i) \in \mathbb{W}$, it has a unique representation w.r.t. $\{w_1, ..., w_m\}$:

$$\mathcal{A}(v_j) = a_{1j} w_1 + \cdots a_{mj} w_m, \forall j = 1, 2, ..., n$$

Linear Maps (Linear Operators)

Hence,

$$
\begin{aligned}
&\begin{pmatrix} \mathcal{A}_1(v_1) & \mathcal{A}_1(v_2) & \cdots & \mathcal{A}_n(v_n) \end{pmatrix} \\
&= \begin{pmatrix} y_1 & y_2 & \cdots & y_n \end{pmatrix} \\
&= \begin{pmatrix} w_1 & w_2 & \cdots & w_m \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} = [w]A
\end{aligned}
$$

Note that A is a $m \times n$ matrix, and ith column of A is the representation of y_i w.r.t basis of $\mathbb{W}$.

Linear Maps (Linear Operators)

Now, for any $y = Ax \in \mathbb{W}$, there exists a unique $\beta_1, ..., \beta_m \in \mathbb{F}$ w.r.t. basis $\{w_1, ..., w_m\}$ such that

$$\begin{aligned} y &= \beta_1 w_1 + \cdots + \beta_m w_m \\ &= [w]\beta \\ &= \alpha_1 \mathcal{A}(v_1) + \cdots \alpha_n \mathcal{A}(v_n) \\ &= [\mathcal{A}(v)]\alpha \\ &= [w]A\alpha \end{aligned}$$

Hence, $\beta = A\alpha$. Since we have unique α and β, we have the desired result

Linear Maps (Linear Operators)

We will now use A instead of $\mathcal{A}$ for a given basis.

i)The matrix A gives the relation between the representations (coordinates) α and β, not x and y.

ii) This also means that A depends on bases. So, with different bases, we have the different representation of the same operator $\mathcal{A}$. This means that A is not unique under different bases.

The most important case is when $\mathbb{V} = \mathbb{W}$, in which case we can use the same basis for the domain and codomain. For a different basis, we may obtain $\bar{A}$, which is a different representation of $\mathcal{A}$.

Similarity Transformation

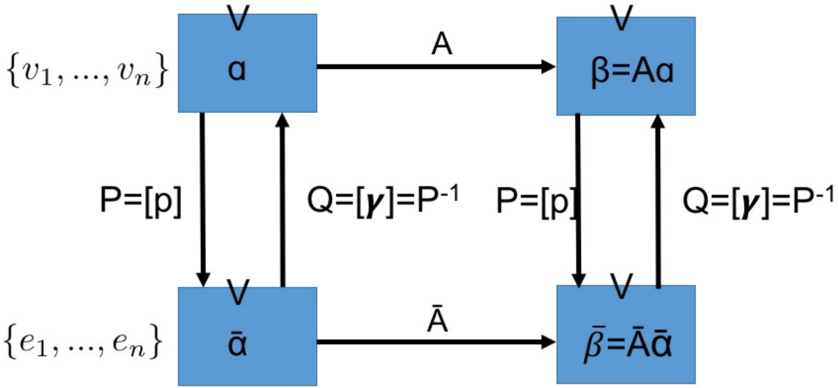
Suppose that $v = \{v_1, ..., v_n\}$ and $e = \{e_1, ..., e_n\}$ are bases for $\mathbb{V}$. Then for the linear operator $\mathcal{A}$, it can have different representations: A w.r.t. v or $\bar{A}$ w.r.t. e.

We now see the relationship between A and $\bar{A}$.

Think about the change of basis, we have shown that given two bases, we can change the coordinate (representation) of a vector by using the transformation matrix.

This means that if the domain and the codomain are the same, then we can use the same transformation matrix

Similarity Transformation



Similarity Transformation

We have

$$\begin{aligned}\bar{\beta} &= \bar{A}\bar{\alpha} \\ &= P\beta = PA\alpha = PAP^{-1}\bar{\alpha}\end{aligned}$$

This implies $\bar{A} = PAP^{-1}$. Similarly, $\bar{A} = Q^{-1}AQ$

Moreover, $A = P^{-1}\bar{A}P = Q\bar{A}Q^{-1}$

Similar Matrix: A and $\bar{A}$ are *similar* if there exists a nonsingular matrix P (or Q) satisfying the above transformation.

If P (or Q) exists, the above transformation is called a *similarity transformation*

Note: all the matrix representations (w.r.t the different bases) of the same operator are similar

Similarity Transformation

Example: Second-order differential equation with $x_1 = x$ and $x_2 = \dot{x} = \dot{x}_1$

$$\ddot{x} + 3\dot{x} + 2x = u \Rightarrow \dot{x} = \begin{pmatrix} 0 & 1 \\ -2 & -3 \end{pmatrix} x + \begin{pmatrix} 0 \\ 1 \end{pmatrix} u$$

The eigenvalue and eigenvector of A

$$\det(sI - A) = 0 \Rightarrow s = -1, -2$$

$$Q = \begin{pmatrix} e_1 & e_1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -1 & -2 \end{pmatrix}$$

$$Q^{-1} = P = \begin{pmatrix} 2 & 1 \\ -1 & -1 \end{pmatrix}$$

Let $z = Qx$. Then $\dot{z} = Q\dot{x}$. Hence,

$$\dot{z} = Q^{-1}AQz + Q^{-1}u = \begin{pmatrix} -1 & 0 \\ 0 & -2 \end{pmatrix} z + \begin{pmatrix} 1.41 \\ 2.23 \end{pmatrix} u$$

Range and Null Spaces

Definition: Range Space

Suppose that $(\mathbb{V}, \mathbb{F})$ and $(\mathbb{W}, \mathbb{F})$ are n and m dimensional vector spaces, respectively. The range of a linear operator A from $\mathbb{V}$ to $\mathbb{W}$ is the set $R(A)$ defined by

$$R(A) = \{y \in \mathbb{W} : \ y = Ax \text{ for some } x \in \mathbb{V}\}$$

Theorem: The range of a linear operator A is a subspace of $(\mathbb{W}, \mathbb{F})$.

Proof: 1) We have $R(A) \subset \mathbb{W}$

2) For any $y_1, y_2 \in R(A)$, and $\alpha, \beta \in \mathbb{F}$, we have $\alpha_1 y_1 + \alpha_2 y_2 = \alpha_1 A x_1 + \alpha_2 A x_2 = A(\alpha_1 x_1 + \alpha_2 x_2)$ (why??). Hence, by the definition of $R(A)$, $\alpha_1 y_1 + \alpha_2 y_2 \in R(A)$.

Range and Null Spaces

- $R(A) \subset \mathbb{W}$, and $R(A)$ is a linear space
- Let $x = (x_1, ..., x_n)^T$, and $A = (A_1, ..., A_n)$, where A_i is the ith column of A. Then $Ax = y$ can be written as

$$y = A_1 x_1 + \cdots A_n x_n$$

  $R(A)$ is the set of all the possible linear combinations of the columns of A.
- Note that $R(A)$ is a linear subspace of $(\mathbb{W}, \mathbb{F})$. Hence, $\dim(R(A))$ is the maximum number of linearly independent columns of A.

Range and Null Spaces

Definition: Null Space

Suppose that $(\mathbb{V}, \mathbb{F})$ and $(\mathbb{W}, \mathbb{F})$ are n and m dimensional vector spaces, respectively. The null space of a linear operator A from $\mathbb{V}$ to $\mathbb{W}$ is the set $N(A)$ defined by

$$N(A) = \{x \in \mathbb{V} : \ Ax = 0\}$$

- The null space of a linear operator A is a subspace of $(\mathbb{V}, \mathbb{F})$.
- The null space is a subspace of domain, whereas the range space is a subspace of $(\mathbb{W}, \mathbb{F})$

Range and Null Spaces

Rank and Nullity

- The dimension of $R(A)$ is denoted by $rank(A)$
- The dimension of $N(A)$ is denoted by $nullity(A)$

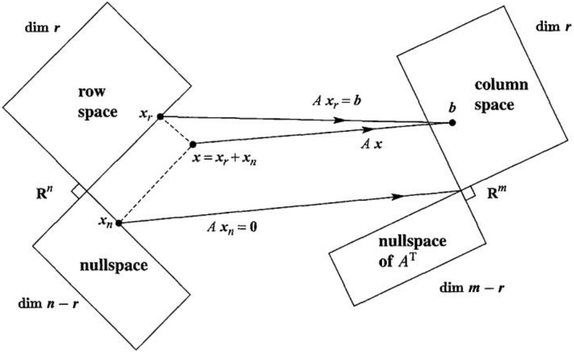
Theorem: Dimension Theorem or Rank-Nullity Theorem
Suppose that $(\mathbb{V}, \mathbb{F})$ and $(\mathbb{W}, \mathbb{F})$ are n and m dimensional vector spaces, respectively. Then for a linear operator A from $\mathbb{V}$ to $\mathbb{W}$, we have

$$rank(A) + nullity(A) = \dim(\mathbb{V}) = n$$

Proof: HW

Range and Null Spaces

- $R(A^T) = \{x \in \mathbb{V} : \ x = A^T y, \text{ for some } y \in \mathbb{W}\}$ (row space)
- $N(A^T) = \{y \in \mathbb{W} : \ A^T y = 0\}$
- $rank(A^T) + nullity(A^T) = m$
- $rank(A)$ = number of linearly independent columns of A
- $rank(A^T)$ = number of linearly independent raws of A
- $rank(A) = rank(A^T)$
- A square matrix matrix A in nonsingular if and only if all the rows and columns of A are linearly independent



Range and Null Spaces

Example:

$$A_1 = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 8 & 12 \end{pmatrix}, \; A_2 = \begin{pmatrix} 1 & 3 \\ 2 & 6 \\ 3 & 9 \end{pmatrix}$$

We can check

$$x_1 \begin{pmatrix} 1 \\ 4 \end{pmatrix} + x_2 \begin{pmatrix} 2 \\ 8 \end{pmatrix} + x_3 \begin{pmatrix} 3 \\ 12 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \; x = (1, 1, -1)^T, \; (2, 0, -1)^T$$

$$v_1 \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + v_2 \begin{pmatrix} 3 \\ 6 \\ 9 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \; v = (3, -1)^T$$

Hence, $nullity(A_1) = 2, \; rank(A_1) = 1, \; nullity(A_2) = 1, \; rank(A_2) = 1$

Invariant Subspace

Invariant subspace

Let A be a linear operator of $(\mathbb{V}, \mathbb{F})$ into itself. A subspace $\mathbb{Y}$ of $\mathbb{V}$ is said to be an invariant subspace of $\mathbb{V}$ under A, or an A-invariant subspace of $\mathbb{V}$ if

$$A(\mathbb{Y}) \subset \mathbb{Y} \Rightarrow Ay \in \mathbb{Y}, \ \forall y \in \mathbb{Y}$$

Trivial Example:
$A : \mathbb{R}^n \to \mathbb{R}^n$: $\mathbb{R}^n$ is an invariant subspace

Eigenvalues and Eigenvectors

Definition: Eigenvalues and Eigenvectors

Let A be a linear operator that maps $(\mathbb{C}^n, \mathbb{C})$ into itself. Then a scalar $\lambda \in \mathbb{C}$ is called an eigenvalue of A if there exists a nonzero $x \in \mathbb{C}^n$ such that

$$Ax = \lambda x$$

Any nonzero x satisfying $Ax = \lambda x$ is called an eigenvector of A associated with the eigenvalue λ

- $(A - \lambda I)x = 0$ or $(\lambda I - A)x = 0$
- For a nonzero $x \in \mathbb{C}^n$, λ is an eigenvalue of A if and only if $\det(\lambda I - A) = 0$ or $\det(A - \lambda I) = 0$
- $\det(\lambda I - A)$ is called the characteristic polynomial of A
- The degree of the characteristic polynomial of A is n. Hence, the $n \times n$ matrix A has n eigenvalues

Eigenvalues and Eigenvectors

Question: For a $n \times n$ square matrix A, are all the eigenvalues of A distinct?

Answer: Not necessarily. There might be a repetition of some eigenvalues. Namely, it can be $\lambda_i = \lambda_j$ for some $i \neq j$ $i, j = 1, 2, .., n$

Case I: We consider the case when all the eigenvalues of A are distinct

Theorem:

Let $\lambda_1, ..., \lambda_n$ be distinct eigenvalues of A, and $v_1, ..., v_n$ are associated eigenvectors. Then the set $\{v_1, ..., v_n\}$ is linearly independent, and for a $n \times n$ matrix $V = [v] = (v_1, ..., v_n)$, $rank(V) = n$.

Eigenvalues and Eigenvectors

We prove the first part. Assume that the set $\{v_1, ..., v_n\}$ is linearly dependent. Then there exist nonzero $\alpha_1, ..., \alpha_n \in \mathbb{C}$ such that

$$\alpha_1 v_1 + \cdots + \alpha_n v_n = 0$$

Without loss of generality, assume that $\alpha_1 \neq 0$. Note that

$$(A - \lambda_2 I)(A - \lambda_3 I) \cdots (A - \lambda_n I) \sum_{i=1}^{n} \alpha_i v_i = 0$$

We have

$$\begin{aligned}
&(A - \lambda_j I) v_i = (\lambda_i - \lambda_j) v_i, \; j \neq i \\
&(A - \lambda_i I) v_i = 0 \\
&\alpha_1 (\lambda_1 - \lambda_2)(\lambda_1 - \lambda_3) \cdots (\lambda_1 - \lambda_n) v_1 = 0
\end{aligned}$$

Since all the eigenvalues are distinct, we have

$$\alpha_1 \prod_{i=2}^{n} (\lambda_1 - \lambda_i) v_1 = 0$$

which implies $\alpha_1 = 0$. This is a contradiction. Hence, $\{v_1, ..., v_n\}$ is linearly independent

Eigenvalues and Eigenvectors

Theorem: A is diagonalizable if and only if it has n linearly independent eigenvectors.

A consequence of diagonalization

$$\begin{aligned} A = Q^{-1}\Lambda Q \Rightarrow \det(\lambda I - A) &= \det(\lambda I - Q^{-1}\Lambda Q) \\ &= \det(Q^{-1}(\lambda I - \Lambda)Q) = \det(\lambda I - \Lambda) \end{aligned}$$

Example: Second-order differential equation with $x_1 = x$ and $x_2 = \dot{x} = \dot{x}_1$

$$\ddot{x} + 3\dot{x} + 2x = u \Rightarrow \dot{x} = \begin{pmatrix} 0 & 1 \\ -2 & -3 \end{pmatrix} x + \begin{pmatrix} 0 \\ 1 \end{pmatrix} u$$

The eigenvalue and eigenvector of A

$$\det(sI - A) = 0 \Rightarrow s = -1, -2$$

$$Q = \begin{pmatrix} e_1 & e_1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -1 & -2 \end{pmatrix}, \; Q^{-1} = P = \begin{pmatrix} 2 & 1 \\ -1 & -1 \end{pmatrix}$$

Let $z = Qx$. Then $\dot{z} = Q\dot{x}$. Hence,

$$\dot{z} = Q^{-1}AQz + Q^{-1}u = \begin{pmatrix} -1 & 0 \\ 0 & -2 \end{pmatrix} z + \begin{pmatrix} 1.41 \\ 2.23 \end{pmatrix} u$$

Eigenvalues and Eigenvectors

Case II: We consider the case when eigenvalues of A are not distinct

Example I: The case when A has repeated eigenvalues, but has linearly independent eigenvectors $\Rightarrow$ A is diagonalizable (theorem)

$$A = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \;\; \det(\lambda I - A) = 0, \;\; \lambda = 1, 1, 2$$

$$v_1 = (1, 1, 0)^T, \;\; v_2 = (0, 1, 0)^T, \;\; v_3 = (-1, 0, 1)^T$$

$[v_1 v_2 v_3]$ forms a basis

Example II: The case when A has repeated eigenvalues, and has linear dependent eigenvectors

$$A = \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 2 \end{pmatrix}, \;\; \det(\lambda I - A) = 0, \;\; \lambda = 1, 1, 2$$

$$v_1 = (1, 0, 0)^T, \;\; v_2 = (1, 0, 0)^T, \;\; v_3 = (5, 3, 1)^T$$

$[v_1 v_2 v_3]$ is not sufficient to form a basis

Jordan Form

As we have seen, there is repeated eigenvalues of A, A can be diagonalizable or not depending on the corresponding eigenvectors.

Assume that A has an eigenvalue of λ_i with multiplicity of m_i. Then the number of linearly independent eigenvectors associated with an eigenvalue λ_i is equal to the dimension of $N(\lambda_i I - A)$. Hence, we have

$$q_i = n - rank(\lambda_i I - A) = nullity(\lambda_i I - A)$$

Previous examples: $\lambda_i = 1$ with $m_i = 2$

$$(I - A) \Rightarrow \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{pmatrix}, \quad \begin{pmatrix} 0 & -1 & -2 \\ 0 & 0 & -3 \\ 0 & 0 & -2 \end{pmatrix}$$

Jordan Form

We want to diagonalize a matrix A when A has repeated eigenvalues $\Rightarrow$ We want to fine some other eigenvectors of A

Generalized eigenvectors:
A vector v is said to be a generalized eigenvector of grade $k \geq 1$ if and only if

$$(A - \lambda I)^k v = 0,\ (A - \lambda I)^{k-1} v \neq 0$$

$k = 1$: v is an eigenvector of A
Example II

$$A = \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 2 \end{pmatrix},\ \det(\lambda I - A) = 0,\ \lambda = 1, 1, 2$$

$$v_1 = (1, 0, 0)^T = v_2,\ v_3 = (5, 3, 1)^T$$

We compute the generalized eigenvector of A associated with $\lambda = 1$.

Jordan Form

$$B := (A - I)^2 = \begin{pmatrix} 0 & 0 & 5 \\ 0 & 0 & 3 \\ 0 & 0 & 1 \end{pmatrix}, \ Bv_2 = 0, \ v_2 = (0, 1, 0)^T$$

Note that $(A - I)v \neq 0$. Then

$$Q = (v_1, v_2, v_3) = \begin{pmatrix} 1 & 0 & 5 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{pmatrix}, \ J = Q^{-1}AQ = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

J: Jordan matrix

$$N := J - \mathrm{diag}\{1, 1, 2\} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

N: Nilpotent

Jordan Form

Assume that A is 5×5 matrix with λ_1 with multiplicity of 4, and with λ_2 with multiplicity of 1 (simple). Then we may have the following Jordan forms

$$A_1 = \begin{pmatrix} \lambda_1 & 1 & 0 & 0 & 0 \\ 0 & \lambda_1 & 1 & 0 & \\ 0 & 0 & \lambda_1 & 1 & 0 \\ 0 & 0 & 0 & \lambda_1 & 0 \\ 0 & 0 & 0 & 0 & \lambda_2 \end{pmatrix}, \quad A_2 = \begin{pmatrix} \lambda_1 & 1 & 0 & 0 & 0 \\ 0 & \lambda_1 & 1 & 0 & \\ 0 & 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & 0 & \lambda_1 & 0 \\ 0 & 0 & 0 & 0 & \lambda_2 \end{pmatrix}$$

$$A_3 = \begin{pmatrix} \lambda_1 & 1 & 0 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 & \\ 0 & 0 & \lambda_1 & 1 & 0 \\ 0 & 0 & 0 & \lambda_1 & 0 \\ 0 & 0 & 0 & 0 & \lambda_2 \end{pmatrix}, \quad A_4 = \begin{pmatrix} \lambda_1 & 1 & 0 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 & \\ 0 & 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & 0 & \lambda_1 & 0 \\ 0 & 0 & 0 & 0 & \lambda_2 \end{pmatrix}$$

$$A_5 = \mathrm{diag}\{\lambda_1, \lambda_1, \lambda_1 \lambda_1, \lambda_2\}$$

J_1: order 4, J_2: order 3
J_3: two order 2, J_4: two order 1, order 2

Jordan Form

Theorem:
Suppose that $A \in \mathbb{C}^{n \times n}$. Then there exists a nonsingular matrix $T \in \mathbb{C}^{n \times n}$ and an integer $1 \leq p \leq n$ such that

$$T^{-1}AT = J = \begin{pmatrix} J_1 & & & 0 \\ & J_2 & & \\ & & \ddots & \\ 0 & & & J_p \end{pmatrix},$$

where J_k are Jordan matrix (or Jordan block) with order n_k

Caley-Hamilton Theorem

Caley-Hamilon Theorem

Let us define the characteristic polynomial of A

$$\Delta(\lambda) = \det(\lambda I - A) = \lambda^n + \alpha_1 \lambda^{n-1} + \alpha_2 \lambda^{n-2} + \cdots + \alpha_{n-1}\lambda + \alpha_n$$

Then

$$\Delta(A) = A^n + \alpha_1 A^{n-1} + \cdots + \alpha_{n-1} A + \alpha_n I = 0$$

Simple proof: when A is diagonalizable (general proof: HW)
Similarity transformation implies that

$$\Lambda = Q^{-1} A Q \Rightarrow A = Q \Lambda Q^{-1}$$
$$\det(A) = \det(Q)\det(\Lambda)\det(Q)^{-1} = \det(\Lambda)$$
$$A^2 = Q\Lambda^2 Q^{-1}, ..., A^k = Q\Lambda^k Q^{-1}$$

Consider

$$\Delta(A) = Q\Big[\Lambda^n + \alpha_1 \Lambda^{n-1} + \cdots + \alpha_{n-1}\Lambda + \alpha_n I\Big] Q^{-1} = 0$$

Caley-Hamilton Theorem

Note that

$$A^n = -\alpha_1 A^{n-1} - \cdots - \alpha_n I$$

It is a linear combination of $\{A^{n-1}, ..., I\}$!!!

Similarly

$$A^{n+1} = -\alpha_1 A^n - \cdots - \alpha_n A$$

It is a linear combination of $\{A^{n-1}, ..., I\}$!!!, which is also a linear combination of $\{A^{n-1}, ..., I\}$

Given any polynomial function $f(\lambda)$, $f(A)$ can be expressed as follows

$$f(A) = \beta_0 I + \beta_1 A + \cdots + \beta_{n-1} A^{n-1}$$

We can extend $f(\lambda)$ to any function, that is, $f(\lambda)$ is not necessarily a polynomial.

Caley-Hamilton Theorem

For A with $n \times n$, let $\lambda_1, ..., \lambda_m$ be distinct eigenvalues with multiplicity $n_1, ..., n_m$, respectively. (note that $\sum_{i=1}^{m} n_i = n$)

Theorem (Theorem 3.5 of the textbook)

Given $f(\lambda)$ and $n \times n$ matrix A with

$$\Delta(A) = \Pi_{i=1}^{m}(\lambda - \lambda_i)^{n_i}$$

Define

$$h(\lambda) = \beta_0 + \beta_1\lambda + \cdots + \beta_{n-1}\lambda^{n-1}$$

The coefficients $\{\beta_0, ..., \beta_{n-1}\}$ are to be solved from the following set of n equations:

$$\frac{df^l(\lambda)}{d\lambda}|_{\lambda=\lambda_i} = \frac{dh^l(\lambda)}{d\lambda}|_{\lambda=\lambda_i}, \ l = 0, 1, ..., n_i - 1, \ i = 1, 2, ..., m$$

Then

$$f(A) = h(A)$$

Caley-Hamilton Theorem

Example:

$$A = \begin{pmatrix} 0 & 1 \\ -1 & -2 \end{pmatrix}$$

We want to compute A^{100}. $\det(\lambda I - A) = \lambda^2 + 2\lambda + 1 = (\lambda + 1)^2$. We let $h(\lambda) = \beta_0 + \beta_1\lambda$, and $f(\lambda) = \lambda^{100}$.

We have

$$\begin{aligned}
&f(-1) = h(-1) \Rightarrow (-1)^{100} = \beta_0 - \beta_1 \\
&f'(-1) = h'(-1) \Rightarrow 100(-1)^{99} = \beta_1 \\
&\beta_1 = -100,\ \beta_0 = 1 + \beta_1 = -99,\ h(\lambda) = -99 - 100\lambda \\
&f(A) = A^{100} = h(A) = -99I - 100A = \begin{pmatrix} -99 & -100 \\ 100 & 101 \end{pmatrix}
\end{aligned}$$

Caley-Hamilton Theorem

Example:

$$A = \begin{pmatrix} 0 & 0 & -2 \\ 0 & 1 & 0 \\ 1 & 0 & 3 \end{pmatrix}$$

We want to compute e^{At}. $\det(\lambda I - A) = (\lambda - 1)^2(\lambda - 2)$. Let $h(\lambda) = \beta_0 + \beta_1\lambda + \beta_2\lambda^2$, and $f(\lambda) = e^{\lambda t}$. Then

$$\begin{aligned}
&f(1) = h(1) \Rightarrow e^t = \beta_0 + \beta_1 + \beta_2 \\
&f'(1) = h'(1) \Rightarrow te^t = \beta_1 + 2\beta_2 \\
&f(2) = h(2) \Rightarrow e^{2t} = \beta_0 + 2\beta_1 + 4\beta_2 \\
&\beta_0 = -2te^t + e^{2t},\ \beta_1 = 3te^t + 2e^t - 2e^t \\
&\beta_2 = e^{2t} - e^t + te^t
\end{aligned}$$

We have

$$f(A) = e^{At} = h(A) = \beta_0 I + \beta_1 A + \beta_2 A^2$$

Positive (semi)-Definite Matrix

A matrix A is symmetric if $A = A^T$

For any $n \times n$ matrix A, and $x \in \mathbb{R}^n$, $x^T A x$ is called a quadratic form

Definition: A symmetric matrix A is said to be

- positive definite if $x^T A x > 0$ for all $x \neq 0$
- positive semi-definite if $x^T A x \geq 0$ for all $x \neq 0$

Theorem:

- A symmetric matrix A is positive definite (positive semi-definite) if and only if all eigenvalues of A are positive (nonnegative)
- A symmetric positive definite matrix is invertible
- For a symmetric positive definite matrix A, $\det(A) > 0$

Norm and Inner Product

Normed vector space (Normed linear space): Length of the vector

Let $(\mathbb{V}, \mathbb{F})$ be a n-dimensional vector space. A function $\|x\| : \mathbb{V} \to \mathbb{R}$, where $x \in \mathbb{V}$, is said to be a norm if the following properties hold

- $\|x\| \geq 0$ and $\|x\| = 0$ if and only if $x = 0$ (separate points)
- $\|\alpha x\| = |\alpha| \|x\|$ (absolute homogeneity)
- $\|x + y\| \leq \|x\| + \|y\|$ (triangular inequality)

Example: Let $(\mathbb{R}^n, \mathbb{R})$. Then the norm can be chosen as

$$\|x\|_1 := \sum_{i=1}^{n} |x_i|, \ \|x\|_2 := \left(\sum_{i=1}^{n} |x_i|^2\right)^{1/2}, \ \|x\|_\infty := \max_i |x_i|$$

Example: signal norm for the real-valued continuous function $f(t)$

$$\|f\|_p = \left(\int_0^t |f(t)|^p dt\right)^{1/p}$$

where $1 \leq p < \infty$

Norm and Inner Product

Inner Product: measure angle of two vectors

An inner product between two vectors, $\langle x, y\rangle$, on the vector space $(\mathbb{V}, \mathbb{C})$ is a function that maps from $\mathbb{V} \times \mathbb{V}$ to $\mathbb{C}$ such that the following properties hold

- $\overline{\langle x, y\rangle} = \langle y, x\rangle$ complex conjugate
- $\langle x, \alpha_1 y_1 + \alpha_2 y_2\rangle = \alpha_1\langle x, y_1\rangle + \alpha_2\langle x, y_2\rangle, \forall x, y_1, y_2 \in \mathbb{V},\ \alpha_1, \alpha_2 \in \mathbb{C}$
- $\langle x, x\rangle \geq 0$ and $\langle x, x\rangle = 0$ if and only if $x = 0$

Example: Let $(\mathbb{R}^n, \mathbb{R})$. Then the inner product is

$$\|x\|_2^2 = \langle x, x\rangle = \sum_{i=1}^{n} |x_i|^2, \ \ \langle x, y\rangle = x^T y = \sum_{i=1}^{n} x_i y_i$$

Example: signal inner product for the real-valued continuous function $f(t)$

$$\|f\|_2^2 = \int_0^t |f(t)|^2 dt, \ \langle f, g\rangle = \int_0^t f(t)g(t)dt$$

where $1 \leq p < \infty$

Orthogonal and Orthonormal Vectors

Two vectors x and y are said to be orthogonal if and only if

$$\langle x, y \rangle = 0.$$

If $\langle x, y \rangle = x^T y$, then x and y are orthogonal if and only if $x^T y = 0$.
Note that if $x^T y = 0$, then the angle between two vectors is 90°

Fact: Let A be a symmetric matrix, that is, $A = A^T$. If A has distinct eigenvalues, then the corresponding eigenvectors are orthogonal.

Example:

$$A_1 = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, \; A_2 = \begin{pmatrix} 2 & 1 \\ 0.5 & 2 \end{pmatrix}$$

$$\det(\lambda I - A_1) = 0, \quad \Rightarrow \lambda = 1, 3, \; (v_1, v_2) = \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}$$

$$v_1^T v_2 = 0, \; \text{orthogonal}$$

Orthogonal and Orthonormal Vectors

A set of vectors, $x_1, x_2, ..., x_m$ is said to be orthonormal if

$$x_i^T x_j = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}$$

In the textbook, the Schmidt orthonormalization procedure for a set of linearly independent vectors $x_1, x_2, ..., x_m$ is explained (page 60). Please see the textbook.

Conclusions

Linera algebra