Background

September-12-11 9:34 AM

Fields

Basic theory of vector spaces works over any field.
$\mathbb{Q}, \mathbb{R}, \mathbb{C}, \mathbb{F}_p$

- We will mostly work over $\mathbb{C}$ or $\mathbb{R}$
- Other fields if convenient

Algebraically Closed

$\mathbb{F}$ is called algebraically closed if every polynomial $p(x) \in \mathbb{F}[x]$ factors into linear terms.

$p(x) = c(x - a_1) \dots (x - a_n)$
$x \in \mathbb{F}, n = \deg p$

Fundamental Theorem of Algebra

$\mathbb{C}$ is algebraically closed

Determinants

If $A = [a_{i,j}]_{n\times n}$ then $\det A$ is determined algorithmically.

$det\, I_n = 1$

Determinant is n-linear

Think of $A = [v_1, v_2, \dots, v_n]$ $v_i \in \mathbb{F}^n$

$$\det\left(v_1, v_2, \dots, v_{i-1}, \sum a_j w_j, v_{i+1}, \dots, v_n\right) =$$
$$\sum a_i \det(v_1, \dots, v_{i-1}, w_j, v_{i+1}, \dots, v_n)$$

Determinant is antisymmetric

$\det(v_1, \dots, v_{i-1}, u, v_{i+1}, \dots, v_{j-1}, u, v_{j+1}, \dots, v_n) = 0$
$\Rightarrow$ (except if 1+1=0)
$\det(v_1, \dots, v_{i-1}, v_j, v_{i+1}, \dots, v_{j-1}, v_i, v_{j+1}, \dots, v_n) =$
$-\det(v_1, \dots, v_n)$

Theorem 1

$\det(AB) = \det A \times \det B$

Theorem 2

$\det A = 0 \Leftrightarrow A\ is\ singular$

Linear Transformation and Matrices

V is a vector space (over field $\mathbb{F}$)
$\mathcal{L}(V)$ is the set of all linear transformations from V to V
W another vector space over $\mathbb{F}$
$\mathcal{X}(V, W)$ =linear transformation from V to W

If $\beta(v_1, \dots, v_n)$ is a basis for V
$T \in \mathcal{L}(V)$

$$T\, v_j = \sum_{i=1}^{n} a_{ij}, v_i$$

$[T]_\beta = [a_{ij}]$ is the matrix x of T with respect to β
$x \in V, x = \sum_{i=1}^{n} \square$
$[T_x]_\beta = [a_{ij}](x_1, \dots, x_n) = [T]_\beta [x]_\beta$

Also if $S \in \mathcal{L}(V, W)$
$\beta = \{v_1, \dots, v_n\}$ bases for V
$\beta' = \{w_1, \dots, w_m\}$ bases for W

$$S(v_j) = \sum_{i=1}^{m} a_{ij} w_i \quad i \le j \le n$$

$[S]_\beta^{\beta'} = [a_{ij}]$

Theorem

If $T \in \mathcal{L}(V)$ then $\det[T]_\beta$ is independent of the choice of basis.

So we can define $\det T \ := \det[T]_\beta$

Sketch of Theorem 2

If A is singular (i.e. rank A < n)
Some column $V_{i_o} = \sum_{i \ne i_o} a_i v_i$

$$\det A = \det\left(v_i, v_{i_0 - 1}, \sum_{i \ne i_0} a_i v_i, v_{i_0+1}, v_n\right) = \sum_{i \ne i_0} a_i \det(v_i, \dots, v_{i_0-1}, v_i, v_{i_0+1}, \dots, v_n) = 0$$

If A is invertible,
$1 = \det I = \det(AA^{-1}) = \det A \times \det A^{-1}$
$\therefore \det A \ne 0$

Proof of Theorem

Let $\beta = \{v_1, \dots, v_n\}$ and $\beta' = \{w_1, \dots, w_n\}$ be two bases for V

Write

$$w_j = \sum_{i=1}^{n} a_{ij} v_i$$

$$Q = [a_{ij}] = [I]_{\beta'}^{\beta} = \left[[w_1]_\beta, [w_2]_\beta, \dots, [w_n]_\beta\right]$$

$$If\ x = \sum_{j=1}^{n} x_j w_j = \sum_{j=1}^{n}\sum_{i=1}^{n} a_{ij} x_j\, v_i = \sum_{i=1}^{n}\left(\sum_{j=1}^{n} a_{ij} x_j\right) v_i$$

$[x]_\beta = [a_{ij}][x]_{\beta'} = Q[x]_{\beta'}$

Look at Tx
$[Tx]_\beta = [T]_\beta [x]_\beta = [T]_\beta Q[x]_{\beta'}$
$[Tx]_{\beta'} = Q^{-1}[T]_\beta Q[x]_{\beta'}$

$\therefore \det[T]_{\beta'} = \det Q^{-1}[T]_\beta Q = \det Q^{-1} \times \det[T]_\beta \times \det Q = \det[T]_\beta$
QED

$\det[T]_\beta$ does not depend on which basis is used.

Eigenvalues

9:30 AM

Eigenvalue (a.k.a. characteristic value)

$T \in \mathcal{L}(V) = set\ of\ all\ linear\ transformations\ from\ V\ to\ V$
A scalar $\lambda \in \mathbb{F}$ is an eigenvalue for T if $\exists v \neq 0$ s.t. $Tv = \lambda v$

Eigenvector

Any non-zero vector v s.t. $Tv = \lambda v$ is an eigenvector for (T, λ)

Eigenspace

The space $\ker(T - \lambda I) = \{v: Tv = \lambda v\}$ is the eigenspace for (T, λ)

Theorem

$T \in \mathcal{L}(V)$, The following are equivalent

1. λ is an eigenvalue for T
2. $T - \lambda I$ is singular
3. $\det(T - \lambda I) = 0$

Characteristic Polynomial

The characteristic polynomial of T is
$P_T(x) = \det(xI - T)$

Note

$P_T(x)$ is a monic polynomial of degree $n = \dim V$

Monic: coefficient on highest degree is 1

Spectrum

The spectrum of T is $\sigma(T)$, the set of all eigenvalues.

Corollary

$\sigma(T)$ is the set of zeros of $P_T(x)$

Corollary

$\sigma(T)$ has at most $n = \dim V$ eigenvalues.

Corollary

Similar transformations have the same spectrum

Direct Sums

Say V is the direct sum of V_1 and V_2 if $V_1 \cap V_2 = \{0\}$ and $V_1 + V_2 = V$. Write $V = V_1 \dotplus V_2\ or\ V = V_1 \oplus V_2$

Say V is the direct sum of $V_1, \dots, V_k$ if

1. $V = \sum_{i=1}^{k} V_i$
2. $V_j \cap \left(\sum_{i \neq j} V_i\right) = \{0\}, for\ 1 \leq j \leq k$

Proposition

If $\{0\} \neq V_i$ subspaces of V such that
$$V = \sum_{i=1}^{k} V_i$$
then TFAE (the following are equivalent)

1. Sum is direct: $V = V_1 \dotplus \dots \dotplus V_k$
2. If $0 \neq v_i \in V_i$, then $\{v_1, \dots, v_k\}$ is linearly independent
3. If $w_i \in V_i$ and $\sum_{i=1}^{k} w_i = 0$ then $w_i = 0,\ 1 \leq i \leq k$
4. Every $v \in V$ has a unique expression as
$$v = \sum w_i, w_i \in V_i$$

Corollary

If $V = V_1 \dotplus V_2 \dotplus \dots \dotplus V_k$
Then if you take a basis for V_i, say $v_{i_1}, \dots, v_{i_{d_i}}$
then the union $\{v_{11}, \dots, v_{1d_1}, v_{21}, \dots, v_{k1}, \dots, v_{kd_k}\}$ is a basis for V.

Example

T is diagonal w.r.t. bases $\beta = \{v_1, \dots, v_n\}$ if
$$[T]_\beta = \begin{bmatrix} \lambda_1 & 0 & \dots \\ 0 & \lambda_2 & \dots \\ & \square & \end{bmatrix}$$
So $Tv_i = \lambda_i v_i$
So $\lambda_1, \dots, \lambda_n$ are eigenvalues

If $u \in \{\lambda_1, \dots, \lambda_n\}$ eigenspace for u
$\ker(T - \mu T) = span\{v_i: \lambda_i = \mu\}$

$\mu \neq \{\lambda_1, \dots \lambda_n\}$

Only eigenvalues are $\{\lambda_1, \dots, \lambda_n\}$

$T = diagonal\ (1, 2, 1, 2, 1, 3)$
$\ker T - I = span\ \{v_1, v_3, v_5\}$
$\ker(T - 2I) = span\{v_2, v_4\}$
$\ker(T - 3I) = span\{v_6\}$

Example

$T = \begin{bmatrix} 1 & 3 \\ 0 & 2 \end{bmatrix}$
$T\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$
1 is an eigenvalue, $\ker(T - I) = \mathbb{F}\begin{pmatrix} 1 \\ 0 \end{pmatrix}$
$\mathbb{F} - span\ or\ set\ of\ all\ multiples\ of$

$T\begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 6 \\ 2 \end{pmatrix} = 2\begin{pmatrix} 3 \\ 1 \end{pmatrix}$
2 is an eigenvalue
$\ker(T - 2I) = \mathbb{F}\begin{pmatrix} 3 \\ 1 \end{pmatrix}$

$u \neq \{1, 2\}$
$T - uI = \begin{pmatrix} 1 - \mu & 3 \\ 0 & 2 - \mu \end{pmatrix}$
$$\begin{pmatrix} 1 - \mu & 3 \\ 0 & 2 - \mu \end{pmatrix}\begin{pmatrix} \frac{1}{1 - \mu} & -\frac{3}{(2 - \mu)(1 - \mu)} \\ 0 & \frac{1}{2 - \mu} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$
So it is invertible, so rank is 0, so no more eigenvalues.

Proof of Theorem

1. λ is an eigenvalue for T
$\Leftrightarrow \ker(T - \lambda I) \neq \{0\}$
$\Leftrightarrow$ 2. $T - \lambda I$ is singular
$\Leftrightarrow$ 3. $\det(T - \lambda I) = 0$

Example

$T = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$
Look at
$p(x) = \det(xI - T) = \begin{bmatrix} x & 1 \\ -1 & x \end{bmatrix} = x^2 + 1$
$\mathbb{F} = \mathbb{R}$ no eigenvalues
$\mathbb{F} = \mathbb{C}\ x^2 + 1 = (x + i)(x - i)$
$T - iI = \begin{bmatrix} -i & -1 \\ 1 & -i \end{bmatrix}\begin{pmatrix} 1 \\ -i \end{pmatrix} = 0$
$T + iI = \begin{bmatrix} i & -1 \\ 1 & i \end{bmatrix}\begin{pmatrix} 1 \\ i \end{pmatrix} = 0$
$\pm i\ are\ eigenvalues$

In $\mathbb{R}^2$, T is a rotation

Example

$$T = \begin{bmatrix} 4 & -1 & -1 \\ -2 & 5 & -1 \\ 3 & -3 & 6 \end{bmatrix}$$
$$p(x) = \det(xI - T) = \begin{vmatrix} x - 4 & 1 & 1 \\ 2 & x - 5 & 1 \\ -3 & 3 & x - 6 \end{vmatrix}$$
$= (x - 4)\big((x - 5)(x - 6) - 3\big) - 1(2(x - 6) + 3) + 1\big(6 + 3(x - 5)\big)$
$= (x - 4)(x^2 - 11x + 27) - (2x - 9) + (3x - 9)$
$= x^3 - 15x^2 + 71x - 108$
$= (x - 3)(x - 6)^2$
Eigenvalues are 3, 6
$$T - 3I = \begin{bmatrix} 1 & -1 & -1 \\ -2 & 2 & -1 \\ 3 & -3 & 3 \end{bmatrix} \to \begin{bmatrix} 1 & -1 & -1 \\ 0 & 0 & -3 \\ 0 & 0 & 0 \end{bmatrix}\begin{pmatrix} a \\ a \\ 0 \end{pmatrix} = 0$$
$$T - 6I = \begin{bmatrix} -2 & -1 & -1 \\ -2 & -1 & -1 \\ 3 & 03 & 0 \end{bmatrix} \to \begin{bmatrix} -2 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & -1 & 0 \end{bmatrix}\begin{pmatrix} a \\ a \\ -3a \end{pmatrix} = 0$$

$$T\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = 3\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

$T\begin{pmatrix}1\\1\\-3\end{pmatrix}=6\begin{pmatrix}1\\1\\-3\end{pmatrix}$
Only 2-dimensions of eigenvectors!

Proof of 3rd Corollary

$T\in\mathcal{L}(V), S\ invertible$
$STS^{-1}\ is\ similar\ to\ T$
$P_{STS^{-1}}(x)=\det(xI-STS^{-1})=\det(S(xIS^{-1}S-T)S^{-1})=\det(xI-T)=P_T(x)$

Proof of Proposition

My Proofs

$1\Rightarrow 2$

Suppose $v_i\neq 0\in V_i$ and

$$\sum_{i=0}^{k}a_iv_i=0\ for\ some\ a_i\in\mathbb{F}\ not\ all\ 0$$

Then, for $a_i\neq 0$

$$a_iv_i=-\sum_{j\neq i}a_jv_j$$

$$a_iv_i\in V_i\ and\ -\sum_{j\neq 1}a_jv_j\in\sum_{j\neq i}V_j\ but$$

$$V_i\cap\sum_{j\neq 1}V_j=\{0\},$$

a contradiction since $a_i\neq 0\ and\ v_i\neq 0$.

$2\Rightarrow 3$

$$\sum_{i=1}^{k}w_i=0\Rightarrow w_i$$

w_i are linearly dependent, but by 2 $w_i\neq 0\Rightarrow w_i$ are linearly independent, so $w_i=0\ \forall i$

$3\Rightarrow 4$

By definition of vector sums, for any $v\in V$ there exists at least one set of $v_i\in V_i$ such that $v=\sum_i v_i$
Now suppose there exists $w_i\in V_i$, such that

$$v=\sum_i v_i=\sum_i w_i$$

$$\Rightarrow 0=\sum_i v_i-w_i$$

But $v_i-w_i\in V_i$ therefore by 3, $v_i-w_i=0\Rightarrow v_i=w_i\ \forall 1\leq i\leq k$

$4\Rightarrow 1$

Already have

$$V=\sum_{i=1}^{k}V_i$$

Suppose for some $1\leq j\leq k, \exists e\neq 0\ s.t.$

$e\in V_j\cap\sum_{i\neq j}V_i$, Select $w_i\in V_i\ s.t.\ e=\sum_{i\neq j}w_i$

$Let\ w_j=e\in V_j$

Then

$$e=w_j+\sum_{i\neq j}0=0+\sum_{i\neq j}w_i,$$

This is not unique, a contradiction, so

$$V_j\cap\sum_{i\neq j}V_i=\{0\}$$

since 0 is certainly in both V_j and $\sum_{i\neq j}V_i$
QED

His Proof

$3\Rightarrow 1$

If $v\in V_i\cap\left(\sum_{j\neq 1}V_j\right)$

$v=v_i\in V$

$=\sum_{j\neq i}v_j\quad v_j\in V_j$

$\therefore -v_i+\sum_{j\neq i}v_j=0$

By 3, $v_i=0=v_j$,

$\therefore v_i\cap\sum_{j=i}V_j=\{0\}$

Proof of Corollary

Suppose

$$0=\sum_{i,j}a_{ij}v_{ij}=\sum_i\left(\sum_j a_{ij}\,v_{ij}\right)=\sum_i v_i\ where\ v_i\in V_i$$

by 3, $v_i = 0 \; 1 \le i \le k$
$\{v_{ij}\}$ is a basis for v_i, so all $a_{ij} = 0$
$\{v_{ij}\}_{i=1,j=1}$ *is lin indep.*
Clearly v_i spans V
$\therefore$ *basis*
■

Diagonalization

September-16-11 9:29 AM

Proposition

Let $T \in \mathcal{L}(v)$

$\sigma(T) = \{\lambda_1, \dots, \lambda_k\}$

$W_i = \ker(T - \lambda_i I)$

$$W = \sum_{i=1}^{k} W_i \subseteq V$$

Then $W = W_1 \dot{+} W_2 \dot{+} \cdots \dot{+} W_k$

Diagonalizable

A linear transformation $T \in \mathcal{L}(V)$ is diagonalizable if it has a basis $\beta = \{v_1, \dots, v_n\}$ so that

$$[T]_\beta = \begin{bmatrix} c_i & 0 & 0 & \dots & 0 \\ 0 & c_2 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & c_n \end{bmatrix}$$

is diagonal.

Note

$Tv_i = c_i v_i$

So v_i is an eigenvector

T is diagonalizable $\Leftrightarrow$ V has a basis containing eigenvectors of T

$\sigma(T) = \{c_1, \dots, c_n\} = \{\lambda_1, \dots \lambda_k\}$

$\{c_1, \dots, c_n\}$ −might have repetitions

$\lambda \in \mathbb{F}, \ker(T - \lambda I) = span\ \{v_i : c_i = \lambda\}$

$p \in \mathbb{F}[x]$ polynomial

$$p(T) = \begin{bmatrix} p(c_1) & 0 & 0 & \dots & 0 \\ 0 & p(c_2) & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & p(c_n) \end{bmatrix}$$

Nullity

$nul(T) = \dim \ker T$

Theorem

$T \in \mathcal{L}(V), \sigma(T) = \{\lambda_1, \dots, \lambda_k\}$

TFAE

1. T is diagonalizable
2. $\displaystyle\sum_{i=1}^{k} nul(T - \lambda_i I) = n = \dim V$
3. $\displaystyle p_T(x) = \prod_{i=1}^{k} (x - \lambda_i)^{d_i}$

   where $d_i = nul(T - \lambda_i I)$

Corollary

If T has n distinct eigenvalues, then T is diagonalizable.

Proof of Proposition

Suffices to show that if $w_i \in W_i$ $1 \le i \le k$ and $\sum_{i=1}^{k} w_i = 0$ then $w_i = 0$ for $1 \le i \le k$

(By Proposition in previous lecture)

If $w \in W_i$ then $(T - \lambda_i I)w = 0$ and

$Tw = \lambda_i w, \quad T^2 w = \lambda_i^2 w, \quad \dots$

Therefore for any polynomial $p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_p x^p$

$p(T) = a_o I + a_1 T + a_2 T^2 + \cdots + a_p T^p$

$$p(T)w = \sum_{j=0}^{p} a_j T^j w = \left(\sum_{j=0}^{p} a_j \lambda_i^j\right) w = p(\lambda_i) w$$

Fix i and show $w_i = 0$:

Let $\displaystyle p(x) = \prod_{j \ne i} (x - \lambda_j)$

Let $\displaystyle x = \sum_{j=1}^{k} w_j = 0$

$$0 = p(T)x = p(T)\left(\sum_{j=1}^{k} w_j\right) = \sum_{j=1}^{k} p(\lambda_j) w_j = \left(\prod_{j \ne i} (\lambda_i - \lambda_j)\right) w_i$$

$\displaystyle\prod_{j = i} (\lambda_i - \lambda_j) \ne 0, so\ w_i = 0$

$\therefore w_i = 0\ \forall i, \Rightarrow$ Sum is direct

Example

$$T = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix} \in \mathcal{L}(\mathbb{C}^2)$$

$p(x) = \mathbb{C}[x]$

$$p(T) = \begin{bmatrix} p(0) & 0 & 0 & 0 \\ 0 & p(0) & 0 & 0 \\ 0 & 0 & p(1) & 0 \\ 0 & 0 & 0 & p(2) \end{bmatrix}$$

Let $A(T) = span\{I, T, T^2, T^3, \dots\} \subseteq \mathcal{L}(V)$

$$A(T) = \begin{bmatrix} a & 0 & 0 & 0 \\ 0 & a & 0 & 0 \\ 0 & 0 & b & 0 \\ 0 & 0 & 0 & c \end{bmatrix}, a, b, c \in \mathbb{C}$$

Because given a, b, c $\exists p$ $(of\ degree\ 2)$ s.t. $p(0) = a, p(1) = b, p(2) = c$

$A(T)$ is isomorphic to $C(\{0,1,2\})$, the algebra of functions in $\{0, 1, 2\}$

Question

Which $T \in \mathcal{L}(V)$ are diagonalizable?

Example

$T = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, p_T(x) = x^2 + 1$

No eigenvalues in $\mathbb{R}$ so it is not diagonalizable if $V = \mathbb{R}^2$ but $mV = \mathbb{C}^2, \sigma(T) = \{i, -i\}$

$\therefore \exists v_1, v_2\ Tv_1 = iv_1, Tv_2 = -iv_2$

$\therefore \{v_1, v_2\}$ is a basis $[T]_\beta = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}$

Example

$T = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$

$P_T(x) = \det(xI - T) = \begin{vmatrix} x & -1 \\ 0 & x \end{vmatrix} = x^2$

$\sigma(T) = \{0\}$

$\ker(T) = \mathbb{C}\begin{pmatrix} 1 \\ 0 \end{pmatrix}$

Need two linearly independent eigenvectors to diagonalize T - NOT POSSIBLE.

Proof

$T\ has\ basis\ \beta = \{v_1, \dots, v_n\}$

$$[T]_\beta = \begin{bmatrix} c_i & 0 & 0 & \dots & 0 \\ 0 & c_2 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & c_n \end{bmatrix}$$

1$\Rightarrow$ 2

$\ker(\mathrm{T} - \lambda_i \mathrm{I}) = span\{v_j : c_j = \lambda_i\}$

$nul(T-\lambda_i I) = \left|\{j: c_j = \lambda_i\}\right|$

So $\sum_{i=1}^{k} nul(T-\lambda_i I) = |\{j: 1 \le j \le n\}| = n$

2⇒ 1

$Let\ \ W_i = \ker(T-\lambda_i I)$

$$\sum_{i=1}^{k} W_i = W_i \dotplus \cdots \dotplus W_k$$

$$\dim\left(\sum_{i=1}^{k} W_i\right) = \sum_{i=1}^{k} \dim W_i = \sum_{i=1}^{k} nul\ (T-\lambda_i I) = n\ , by\ (2)$$

$$\therefore \sum W_i = V$$

Take a basis for each W_i

- they are eigenvectors for the eigenvalues λ_i
- put them together, get a basis for V consisting of eigenvectors ⇒ diagonalizable

1 ⇒ 3

$$T = \begin{bmatrix} c_1 & 0 & 0 & \dots & 0 \\ 0 & c_2 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & c_n \end{bmatrix}$$

$nul(T-\lambda_i) = \left|\{j: c_j = \lambda_i\}\right|$

$$p_T(x) = \det(xI - T) = \begin{vmatrix} x-c_i & 0 & 0 & \dots & 0 \\ 0 & x-c_2 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & x-c_n \end{vmatrix} = \prod_{i=1}^{k} (x-\lambda_i)^{d_i}$$

$where\ d_i = \left|\{j: c_j = \lambda_i\}\right| = nul(T-\lambda_i I)$

3 ⇒ 2

$$\sum_{i=1}^{k} nul(T-\lambda_i I) = \sum_{i=1}^{k} d_i = \deg(p_T) = \text{n}$$

Proof of Corollary

$nul(T-\lambda_i I) = 1$ for $1 \le i \le n$ so by 2, T is diagonalizable.

Linear Recursion

September-19-11 10:09 AM

Computational Device

Suppose you are given T as in example * and you need to compute T^n

If D is the diagonal matrix of T

$T = QDQ^{-1}$

$T^n = (QDQ^{-1})^n = QDQ^{-1}QDQ^{-1} \dots QDQ^{-1}$
$= QD^nQ^{-1}$

Linear Recursion

In general, if we have $x_0, x_1, \dots, x_n$ given,

$x_{k+1} = a_0 x_k + a_1 x_{k-1} + \dots + a_n x_{k-n}$

linear recursion

$$\begin{pmatrix} x_{k+1} \\ x_k \\ x_{k-1} \\ \vdots \\ x_{k-n} \end{pmatrix} = \begin{pmatrix} a_0 & a_1 & a_2 & \dots & a_n \\ 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix} \begin{pmatrix} x_k \\ x_{k-1} \\ x_{k-2} \\ \vdots \\ x_{k-n-1} \end{pmatrix}$$

$$p_A(x) = \det(xI - A) = \begin{vmatrix} x - a_0 & -a_1 & -a_2 & \dots & -a_n \\ -1 & x & 0 & \dots & 0 \\ 0 & -1 & x & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & -1 & x \end{vmatrix}$$
$$= x^{n+1} - a_o x^n - a_1 x^{n-1} - a_2 x^{n-2} - \dots - a_n$$

Now try to diagonalize A, and get a formula for x_n

Example *

$$T = \begin{bmatrix} -3 & 3 & -1 & -2 \\ -8 & 2 & 3 & -4 \\ -4 & 2 & 1 & -2 \\ 0 & -4 & 4 & 1 \end{bmatrix}$$

Using Matlab got

$p_T(x) = (x-1)^2 x(x+1)$

So $\sigma(T) = \{1, 0, -1\}$

$$\ker(T - I) = span\left\{\begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \\ -1 \end{pmatrix}\right\}$$

$$\ker(\mathrm{T}) = \mathrm{span}\left\{\begin{pmatrix} 3 \\ 1 \\ 2 \\ -4 \end{pmatrix}\right\}$$

$$\ker(T + I) = span\left\{\begin{pmatrix} -2 \\ 1 \\ -1 \\ 4 \end{pmatrix}\right\}$$

Change of basis matrix:

$$Q = \begin{bmatrix} 0 & 1 & 3 & -2 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 2 & -1 \\ 1 & -1 & -4 & 4 \end{bmatrix}$$

$$Q^{-1}TQ = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} = D$$

Example: Fibonacci Sequence

$x_0 = 0, x_1 = 1$

$x_n = x_{n-1} + x_{n-2}\ for\ n \geq 2$

$$\begin{pmatrix} x_n \\ x_{n+1} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_{n-1} \\ x_n \end{pmatrix}$$

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

$$\begin{pmatrix} x_n \\ x_{n+1} \end{pmatrix} = A^n \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$p_A(x) = \det\begin{pmatrix} x & -1 \\ -1 & x-1 \end{pmatrix} = x(x-1) - 1 = x^2 - x - 1$$

$$\tau = \frac{1 \pm \sqrt{1+4}}{2}$$

$$\tau = \frac{1+\sqrt{5}}{2}, \frac{1-\sqrt{5}}{2} = -\frac{1}{\tau}$$

$$\sigma(A) = \left\{\frac{1+\sqrt{5}}{2}, \frac{1-\sqrt{5}}{2}\right\} = \left\{\tau, -\frac{1}{\tau}\right\}$$

$$A - \tau I = \begin{pmatrix} -\tau & 1 \\ 1 & 1-\tau \end{pmatrix} \begin{pmatrix} 1 \\ \tau \end{pmatrix} = \begin{pmatrix} 0 \\ 1 + \tau - \tau^2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\ker(A - \tau I) = \mathbb{C}\begin{pmatrix} 1 \\ \tau \end{pmatrix}$$

$$A + \frac{1}{\tau} I = \begin{pmatrix} \frac{1}{\tau} & 1 \\ 1 & 1 + \frac{1}{\tau} \end{pmatrix} \begin{pmatrix} \tau \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ \tau - 1 - \frac{1}{\tau} \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{\tau^2 - \tau - 1}{\tau} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\ker\left(A + \frac{1}{\tau} I\right) = \mathbb{C}\begin{pmatrix} \tau \\ -1 \end{pmatrix}$$

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

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

$$Q^{-1} = \frac{1}{-1-\tau^2}\begin{pmatrix} -1 & -\tau \\ -\tau & 1 \end{pmatrix} = \frac{1}{1+\tau^2}\begin{pmatrix} 1 & \tau \\ \tau & -1 \end{pmatrix} = \frac{1}{1+\tau^2} Q$$

$$Q^{-1}AQ = D = \begin{pmatrix} \tau & 0 \\ 0 & -\frac{1}{\tau} \end{pmatrix}$$

$$A^n = QD^nQ^{-1} = \frac{1}{1+\tau^2}\begin{pmatrix}1 & \tau\\ \tau & -1\end{pmatrix}\begin{pmatrix}\tau^n & 0\\ 0 & \frac{(-1)^n}{\tau^n}\end{pmatrix}\begin{pmatrix}1 & \tau\\ \tau & -1\end{pmatrix}$$

$$= \frac{1}{1+\tau^2}\begin{pmatrix}1 & \tau\\ \tau & -1\end{pmatrix}\begin{pmatrix}\tau^n & \tau^{n+1}\\ \frac{(-1)^n}{\tau^{n-1}} & \frac{(-1)^{n+1}}{\tau^n}\end{pmatrix}$$

$$= \frac{1}{1+\tau^2}\begin{pmatrix}\tau^n + \frac{(-1)^n}{\tau^{n-2}} & \tau^{n+1} + \frac{(-1)^{n+1}}{\tau^{n-1}}\\ \tau^n + \frac{(-1)^{n+1}}{\tau_{n+1}} & \tau^{n+2} + \frac{(-1)^{n+2}}{\tau^n}\end{pmatrix}$$

$$\begin{pmatrix}x_n\\ x_{n+1}\end{pmatrix} = A^n\begin{pmatrix}0\\ 1\end{pmatrix} = \begin{pmatrix}\frac{\left(\tau^{n+1} + \frac{(-1)^{n+1}}{\tau^{n+1}}\right)}{1+\tau^2}\\ *\end{pmatrix}$$

$$x_n = \left(\frac{\tau}{1+\tau^2}\right)\left(\tau^n - \left(-\frac{1}{\tau}\right)^n\right)$$

$$\frac{\tau}{1+\tau^2} = \frac{1}{\sqrt{5}}$$

$$x_n = \frac{\tau^n - \left(-\frac{1}{\tau}\right)^n}{\sqrt{5}}$$

$x \geq 2$, x_n is the closest integer to $\frac{\tau^n}{\sqrt{5}}$

Triangular Forms

September-21-11 9:31 AM

Upper Triangular

A matrix T is upper triangular if
$a_{ij} = 0 \; if \; j < i$

Say $T \in \mathcal{L}(V)$ is triangularizable if there is a basis β such that $[T]_\beta$ is upper triangular.

Triangular Determinant

$$\det T = \prod_{i=1}^{n} a_{ii}$$

1. $\sigma(T) = \{a_{11}, a_{22}, \ldots, a_{nn}\}$
2. $p_T(x)$ factors into linear terms.

Invariant Subspace

If $T \in \mathcal{L}(V)$, a subspace $W \subseteq V$ is an invariant subspace for T if $TW \subseteq W$

$W_k = span\,\{\,v_1, v_2, \ldots, v_k\}\; 0 \le k \le n$

Theorem

For $T \in \mathcal{L}(V), TFAE$

1. T is triangularizable
2. $P_T(x)$ factors into linear terms
3. T has a chain of invariant subspaces $\{0\} = W_0 \subset W_1 \subset W_2 \subset \cdots \subset W_n = V$
   With dim $W_k = k$ for $1 \le k \le n$

Corollary

If $\mathbb{F}$ is algebraically closed (such as $\mathbb{C}$) then every $T \in \mathcal{L}(V)$ is triangularizable.

Determinant of Upper Triangular

$\begin{vmatrix} a & b \\ 0 & d \end{vmatrix} = ad$

For n > 2, take determinant of first column leaves a_{11} *determinant of upper triangular matrix with n -1

So by induction, det $T = a_{11}a_{22} \ldots a_{nn}$

Alternate Proof

$$|a_{ij}| = \sum_{\sigma \in S_n} (-1)^{sign(\sigma)} \prod_{i=1}^{n} a_{i\sigma(i)}$$

If $\sigma \in S_n$ and for some $i, \sigma(1) = j < i$ then $a_{i\sigma(i)} = 0 \Rightarrow \prod_{i=1}^{n} a_{i\sigma(i)} = 0$
Only $\sigma = identity$ satisfies $\sigma(i) \ge i \; \forall i$ because if say $\sigma(i) = 1$ for $1 \le i < i_0$ but $\sigma(i_o) > i_o$ then some j has $\sigma(j) = i_o, but \; j > i_0$
$\therefore \prod a_{i\sigma(i)} = 0$

$$|a_{ij}| = \prod_{i=1}^{n} a_{ii}$$

Types of Invariant Subspaces

If T is upper triangular w.r.t. $\beta = \{v_1, \ldots, v_n\}$
$Tv_1 = a_{11}v_1$ eigenvector
$\therefore W_1 = span\{v_1\}$ is invariant

$W_0 = \{0\}$ is invariant for every T
$W_n = V$ is invariant for every T

$T_{v_2} = a_{22}v_2 + a_{12}v_1 \in span\{v_1, v_2\}$
$T_{v_1} = a_{11}v_1 \in span\{v_1, v_2\}$
$\therefore W_2 = span\,\{\,v_1, v_2\}$ is invariant for T

$W_k = span\,\{\,v_1, v_2, \ldots, v_k\}\; 0 \le k \le n$

$$Tv_j = \sum_{i=1}^{n} a_{ij}v_i = \sum_{i=1}^{j} a_{ij}v_i \in span\{v_1, \ldots, v_j\} = W_j \subseteq W_k \; if \; j \le k$$

$T_{v_j} \in W_k \; 1 \le j \le k$
$\therefore TW_k \subseteq W_k$

Suppose conversely that I have such a chain of invariant subspaces. Pick $0 \ne v_1 \in W_1$
$\dim(W_1) = 1$, so $W_1 = span\,\{v_1\}$
In W_2, pick $v_2 \in W_2$ independent of v_1 so $\{v_1, v_2\}$ is a basis for W_2, since dim $W_2 = 2$
End up with a basis $\beta = \{v_1, \ldots, v_n\}$ such that $W_k = span\,\{v_1, \ldots, v_k\}\; 1 \le k \le n$

Find $[T]_\beta$, $Tv_1 \in W_1$ since $(TW_1 \subseteq W_1)$
$\therefore Tv_1 = a_{11}v_1$
$Tv_2 \in W_2$
$\therefore Tv_2 = a_{22}v_2 + a_{12}v_1$
$T_{v_k} \in W_k$

$$T_{v_k} \in \sum_{i=1}^{k} a_{ik}v_i$$

So $[T]_\beta = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots \\ 0 & a_{22} & a_{23} & \cdots \\ 0 & 0 & a_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$ is triangular

Proof

Already proved 1⇒ 2, 1⇒ 3, and 3⇒ 1
Let's show 2 ⇒ 1 by induction on n.
$n = 1$: $T = [a]_{1\times 1}$ is always upper triangular
$n > 1$: assume theorem for $n - 1$
$P_T(x) = (x - \lambda_1)(x - \lambda_2) \ldots (x - \lambda_n)$
λ_1 is an eigenvalue of T
So we can find an eigenvector $v_1 \ne 0$ so $Tv_1 = \lambda_1 v_1$
Extend v_1 to a basis $\beta_1 = \{v_1, w_2, w_3, \ldots, w_n\}$
Express T in this basis.

$$[T]_{\beta_1} = \begin{bmatrix} \lambda_1 & b_{12} & \cdots & b_{1n} \\ 0 & & & \\ \vdots & & T_1 & \\ 0 & & & \end{bmatrix}$$

$$P_T(x) = \det(xI_n - T) = \det\left(\begin{bmatrix} x - \lambda_1 & -b_{12} & \cdots & -b_{1n} \\ 0 & & & \\ \vdots & & xI_{n-1} - T_1 & \\ 0 & & & \end{bmatrix}\right) = (x - \lambda_1)|xI_{n-1} - T_i|$$
$$= (x - \lambda_1)P_{T_1}(x)$$

$P_T(x) = (x - \lambda_1)P_{T_1}(x) = (x - \lambda_1)(x - \lambda_2) \ldots (x - \lambda_n)$
$\therefore P_{T_1}(x) = (x - \lambda_2) \ldots (x - \lambda_n)$
So $P_{T_1}(x)$ factors into linear terms. By the induction hypothesis, $W = span\,\{\,w_2, \ldots, w_n\}$ has another

basis $\beta' = \{v_2, \dots, v_n\}$ so that $[T_1]_\beta = \begin{bmatrix} a_{22} & \dots & a_{2n} \\ 0 & \dots & a_{3n} \\ \vdots & \ddots & \vdots \\ 0 & \dots & a_{n_n} \end{bmatrix}$ is upper triangular.

So $\beta = \{v_1, v_2, \dots, v_n\}$ is a basis for V and

$$[T]_{\beta_1} = \begin{bmatrix} \lambda_1 & a_{12} & \dots & a_{1n} \\ & a_{22} & \dots & a_{2n} \\ 0 & 0 & \dots & a_{3n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{n_n} \end{bmatrix}$$

So $[T]_\beta$ is upper triangular ■

Cayley-Hamilton Theorem

September-23-11 9:56 AM

Cayley-Hamilton Theorem

$T \in \mathcal{L}(V)$, then $p_T(T) = 0$

Computational Aside

If $T = \begin{bmatrix} A & B \\ 0 & D \end{bmatrix}$, block upper triangular.

$T^2 = \begin{bmatrix} A^2 & AB + BD \\ 0 & D^2 \end{bmatrix}$

$T^3 = \begin{bmatrix} A^3 & A^2B + 2ABD + BD^2 \\ 0 & D^3 \end{bmatrix}$

...

$T^k = \begin{bmatrix} A^k & * \\ 0 & D^k \end{bmatrix}$

$p(x) = a_0 + a_1 x + \cdots + a_d x^d$

$p(T)$

$= \begin{bmatrix} a_0 I_k & 0 \\ 0 & a_0 I_{n-k} \end{bmatrix} + \begin{bmatrix} a_1 A & * \\ 0 * a_1 D & \end{bmatrix} + \cdots$

$+ \begin{bmatrix} a_d A^d & * \\ 0 & a_d D^d \end{bmatrix} = \begin{bmatrix} p(A) & * \\ 0 & p(D) \end{bmatrix}$

Example

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

$p_T(x) = x^2 + 1$ does not factor over $\mathbb{R}$ so it is not triangularizable over $\mathbb{R}$

It does factor over $\mathbb{C}$ so it is triangularizable over $\mathbb{C}$

$T \sim \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}$, $\quad \sim similar$

$p_{T(T)} = T^2 + I = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = 0$

Example

$T = \begin{bmatrix} 2 & 3 & 5 \\ -1 & -3 & -4 \\ 0 & 1 & 1 \end{bmatrix}$

$p_T(x) = \begin{vmatrix} x-2 & -3 & -5 \\ 1 & x+3 & 4 \\ 0 & -1 & x-1 \end{vmatrix} = (x-2)\big((x+3)(x-1)+4\big) - 1\big((-3)(x-1)-5\big) = x^3$

x^3 splits into linear terms so T is triangularizable

$\sigma(T) = \{0\}$ - look for kernel

$\begin{bmatrix} 2 & 3 & 5 \\ -1 & -3 & -4 \\ 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 3 & 4 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$

$\ker T = \mathbb{R}\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$

Take new basis $v_1 = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, v_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, v_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$

$v_2 \mapsto \begin{pmatrix} 3 \\ -3 \\ 1 \end{pmatrix} = 3\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} + (-6)\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + 4\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$

$v_3 \mapsto \begin{pmatrix} 5 \\ -4 \\ 1 \end{pmatrix} = 5\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} - 9\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + 6\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$

$T_{\beta_1} = \begin{bmatrix} 0 & 3 & 5 \\ 0 & -6 & -4 \\ 0 & 4 & 6 \end{bmatrix}$

$T_1 = \begin{bmatrix} -6 & -9 \\ 4 & 6 \end{bmatrix}, p_{T_1}(x) = x^2$

$\ker T_1 = \mathbb{R}\begin{pmatrix} 3 \\ -2 \end{pmatrix}$

New bases

$w_1 = v_1 = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, w_2 = \begin{pmatrix} 0 \\ 3 \\ -2 \end{pmatrix}, w_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$

$Tw_2 = \begin{bmatrix} 2 & 3 & 5 \\ -1 & -3 & -4 \\ 0 & 1 & 1 \end{bmatrix}\begin{pmatrix} 0 \\ 3 \\ -2 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix} = -w_1$

$Tw_3 = \begin{pmatrix} 5 \\ -4 \\ 1 \end{pmatrix} = 5\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} - 3\begin{pmatrix} 0 \\ 3 \\ -2 \end{pmatrix} + 0\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$

$T_\beta = \begin{bmatrix} 0 & -1 & 5 \\ 0 & 0 & -3 \\ 0 & 0 & 0 \end{bmatrix}$

$[T]_\beta$ is upper triangular, diagonal entries all 0 since roots of $p_T(x) = x^3$ are 0, 0, 0

$[T^2]_\beta = \begin{bmatrix} 0 & 0 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

$[T^3]_\beta = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

$T^3 = 0 = p_T(T)$

Proof of Cayley-Hamilton Theorem

First assume $p_T(x)$ splits into linear factors.

Apply triangular theorem, find basis to triangularize T.

So wlog, T is triangular

$p_T(x) = (x - \lambda_1)(x - \lambda_2) \ldots (x - \lambda_n)$

Proceed by induction on n.

n=1

$T = [\lambda_1], p_T(x) = x - \lambda_1, p_T(T) = T - \lambda_1 I = [\lambda_1] - [\lambda_1] = 0$

Assume for $k < n$

Write $T = \begin{bmatrix} \lambda_1 & t_{12} & \cdots & t_{1n} \\ 0 & & & \\ \vdots & & T_1 & \\ 0 & & & \end{bmatrix}$

From the proof of triangularizability

$p_{T_1} = (x - \lambda_2)(x - \lambda_2) \ldots (x - \lambda_n)$

By the induction hypothesis $p_{T_1}(T_1) = 0$

$p_T = (x-\lambda_1)P_{T_1}(x)$

$P_T(T) = (T-\lambda_1 I)P_{T_1}(T) = \begin{bmatrix} 0 & * \\ 0 & T_1-\lambda_1 I \end{bmatrix} P_{T_1}\left(\begin{bmatrix} \lambda_1 & * \\ 0 & T_1 \end{bmatrix}\right) = \begin{bmatrix} 0 & 0 \\ 0 & T_1-\lambda_1 I \end{bmatrix}\begin{bmatrix} p_{T_1}(\lambda_1) & * \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

$= 0$

So by induction $p_T(T) = 0$. For algebraically closed fields.

In general, $p_T(x)$ does not split on $\mathbb{F}[x]$ but there is always a bigger field $\mathbb{G} \supseteq \mathbb{F}$ so that $p_T(x)$ splits on $\mathbb{G}[x]$

$T = [t_{ij}] \in M_n(\mathbb{F})$

Can think of T as an element of $M_n(\mathbb{G})$. $p_T(x)$ splits in $\mathbb{G}[x] \therefore p_T(T) = 0$

But the calculation of $p_T(T)$ happens over $\mathbb{F}$ since all the coefficients $a_k \in \mathbb{F}[x]$

So $p_T(x) = a_0 I + a_1 T + \cdots + a_n T^n$, this is all in $M_n(\mathbb{F})$

$\therefore p_T(T) = 0\ in\ M_n(\mathbb{F})$

Ideals

September-26-11 9:31 AM

Look at $\mathbb{F}[x]$ - the ring of polynomials with coefficients in $\mathbb{F}$

Ideal

An ideal in $\mathbb{F}[x]$ is a non-empty subset $J \subseteq \mathbb{F}[x]$ which is

1) a subspace
2) if $p \in J$ and $q \in \mathbb{F}[x]$ then $pq \in J$

Principal Ideal

A principal ideal is an ideal of the form
$(p_0) = \{p_0 q : q \in \mathbb{F}[x]\}$

Theorem

Every ideal in $\mathbb{F}[x]$ is principal

Lemma

$T \in \mathcal{L}(V)$
$J = \{p \in \mathbb{F}[x] : p(T) = 0\}$ is a non-zero ideal in $\mathbb{F}[x]$

Corollary

$\{p : p(T) = 0\} = (m_T)$

Minimal Polynomial

The unique monic polynomial $m_T(x)$ generating $\{p : p(T) = 0\}$ is the minimal polynomial of T

Theorem

$T \in \mathcal{L}(V)$
Then $m_T(x)$ has the same roots as $p_T(x)$, namely $\sigma(T)$, except for multiplicity. Furthermore, it also has the same irreducible polynomial factors.

Principal Ideal

Check that (p_0) is an ideal

1. $p_0, p_r \in (p_0), \lambda \in F$ then
   $p_o q + p_o r = p_o(q + r) \in p_o$
   $\lambda(p_0 q) = p_o(\lambda q) \in p_o$
   $\therefore (p_0)$ is a vector space

2. If $p_0 q \in (p_0), r \in \mathbb{F}[x]$ then
   $(p_o q)r = p_0(qr) \in p_o$

Proof

Let J be an ideal of $\mathbb{F}[x]$. If $J = \{0\}$, then $J = (0)$.
Otherwise let p_o be a monic polynomial in J of minimal degree.
$p_0 = x^d + a_{d-1}x^{d-1} + \cdots + a_0$

Let q be any non-zero element of J. Use the division algorithm to divide p_0 into q.
$q = p_o q_1 + r$, $\deg(r) < \deg(p_0)$, but p_0 was the element of smallest degree.
$\therefore$ by minimality, $r = 0$, so $q = p_0 q$.
$\therefore J = (p_0)$
*monic generator is unique

Proof of Lemma

$p_T \in J$, so $J \neq \{0\}$ (by Cayley-Hamilton)

If $p, q \in J, \lambda \in \mathbb{F}[x]$
$(p + q)(T) = p(T) + q(T) = 0$
$(\lambda p)(T) = \lambda p(T) = 0$
$\therefore$subspace

$p \in J, q \in \mathbb{F}[x]$ then
$(pq)(T) = p(T)q(T) = 0$ ■

Example

$$T = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$p_T(x) = x^4, m_T(x) = x^2$

$T = diag(1,1,2,2,2,3)$
$p_T(x) = (x-1)^2(x-2)^3(x-3)$
$m_T(x) = (x-1)(x-2)(x-3)$

$$T = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

$p_T(x) = (x-1)^3$
$m_T | p_T$ so $m_T(x) = (x-1)^d, d \in \{1,2,3\}$

$$T - I = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

$$(T - I)^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$(T - I)^3 = 0$
$\therefore m_T = p_T = (x-1)^3$

Proof of Theorem

$m_T | p_T$ so $roots(m_T) \subseteq roots(p_T) = \sigma(T)$
If λ is an eigenvalue of T $\exists v \neq 0$ eigenvector $Tv = \lambda v$
$\therefore T^k v = \lambda^k v, \forall k \geq k$
$\Rightarrow p(T)v = p(\lambda)v$
So $0 = m_T(T)v = m_T(\lambda)v, \therefore m_T(\lambda) = 0$
So $roots(m_T) \supseteq \sigma(T)$
$\therefore roots(m_T) = roots(p_T) = \sigma(T)$

Remark

Over a non-algebraically closed field $\mathbb{F}$ this proof does not show the stronger fact that the same irreducible factors will be in both p_T and m_T

Possible Problem

$T \in \mathcal{L}(\mathbb{R}^4)$
$p_T(x) = (x^2+1)^2$
$m_T | p_T, m_T \neq 1$
$\therefore m_T = x^2 + 1, or\ (x^2+1)^2$

If we can calculate $T \in \mathcal{L}(\mathbb{C}^4)$ then m_T can be $x^2+1, (x^2+1)^2, (x^2+1)(x-i), or\ (x^2+1)(x+i)$
Calculate $m_T(T)$ using a real basis
Take $p(x) = (x^2+1)(x-i)$
$0 = p(T) = (T^2+I)(T_{iI}) = (T^2+I)T - i(T^2+I) = 0 + i0$
$T^2 + I = 0$

Better Proof of Theorem

The minimal polynomial $m_T(x)$ of $T \in \mathcal{L}(V)$ has degree of d if $\{I, T, T^2, \dots, T^{d-1}\}$ is linearly independent, but $\{I, T, T^2, \dots, T^d\}$ is linearly dependent. $m_T(x)$ is given the unique way to express T^d as $\sum_{i=0}^{d-1} a_i T^i$

$$T^d = \sum_{i=0}^{d-1} a_i T_i$$

$$T^{d+k} = \sum_{i=0}^{d-1} a_i T^{i+k} = \sum_{i=0}^{d-1} b_i T_i$$

$\therefore A(T) = span\ \{ I, T, T^2, \dots, \} = span\ \{ I, T, T^2, \dots, T^{d-1}\}$

$\therefore d = \dim(A)$

This unique way to express m_T does not depend on a larger field.

$\therefore m_T(x)$ is unchanged if we enlarge the base field so that $p_T(x)$ splits.

Diag. & Nilpotent

September-28-11 9:45 AM

Theorem

$T \in \mathcal{L}(V)$ and $p_T(x)$ splits then
T is diagonalizable
$\Leftrightarrow$
$m_T(x)$ has only simple roots.
$i.e. \, m_T(x) = (x - \lambda_1)(x - \lambda_2) \dots (x - \lambda_k)$
where $\sigma(T) = \{\lambda_1, \lambda_2, .., \lambda_k\}$

Lemma

$A, B \in \mathcal{L}(V)$
$nul(AB) \leq nul(A) + nul(B)$

Nilpotent Matrices

$T \in \mathcal{L}(V)$ is **nilpotent of order k** if $T^k = 0$ and $T^{k-1} \neq 0$

Proof of Theorem

$" \Rightarrow "$
$T = diag(c_1, c_2, \dots, c_n)$
$\sigma(T) = \{\lambda_1, \lambda_2, \dots, \lambda_k\}$
Rearrange bases so
$T = diag(\lambda_1, \dots, \lambda_1, \lambda_2, \dots, \lambda_2, \dots, \lambda_k, \dots, \lambda_k)$
$m_T(x)$ has $\lambda_1, .., \lambda_k$ as roots
$(T - \lambda_1 I)(T - \lambda_2 I) \dots (T - \lambda_k I)$
$diag(0, \dots, 0, \lambda_2, .., \lambda_2, \dots, \lambda_k, .., \lambda_k) *$
$diag(\lambda_1, \dots, \lambda_1, 0, .., 0, \dots, \lambda_k, .., \lambda_k) *$
$diag(\lambda_1, \dots, \lambda_!, \lambda_2, .., \lambda_2, \dots, 0, .., 0) = diag(0, .., 0) = 0$
$\therefore m_T(x) = (x - \lambda_1) \dots (x - \lambda_k)$

? 2nd Proof of $\Rightarrow$

$\text{nul}(T - \lambda_i) = |\{c_j : c_j = \lambda_i\}|$

$$\sum_{i=1}^{k} nul(T - \lambda_i I) = \sum_{i=1}^{k} |\{c_j : c_j = \lambda_i\}| = |\{c_j\}| = n$$

$$\ker \prod_{i=1}^{k} (T - \lambda_i I) \supseteq \sum \ker(T - \lambda_i I) = V$$

$" \Leftarrow "$

Proof of Lemma

$\ker(AB) \supseteq \ker B$
chose a basis $v_1, \dots, v_b$ for $\ker B$, $b = nul(B)$
Extend to a basis for $\ker(AB)$: $v_1, .., v_b, v_{b+1}, .., v_{b+c}$
$span\{v_{b+1}, \dots, v_{b+c}\} \cap span\{v_1, \dots, v_b\} = \{0\}$
So $B \mid span\{ v_{b+1}, \dots, v_{b+c}\}$ is injective (1-1)
B maps $sp\{v_{b+1}, \dots, v_{b+c}\}$ into $\ker A$
$\therefore nul\ A = \dim \ker A \geq \dim span\{v_{b+1}, \dots, v_{b+c}\}$
$nul\ AB = b + c = nul(B) + c \leq nul(B) + nul(A)$
$\blacksquare$

Back to Theorem

By hypothesis
$0 = m_T(T) = (T - \lambda_1)(T - \lambda_2 I)..(T - \lambda_k I)$

$$n = nul(m_T(T)) \leq \sum_{i=1}^{k} nul(T - \lambda_i I)$$

but know that $\sum_{i=1}^{k} \ker(T - \lambda_i I)$ is a direct sum , so

$$\sum_{i=1}^{k} \text{nul}(T - \lambda_i I) = \dim\left(\sum \ker(T - \lambda_i I)\right) \leq n$$

Example of Nilpotent

$T = \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}, T^2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$
$\{0\} \subset \ker T \subset \ker T^2 = R^2$
$\ker T = \mathbb{R}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$
Chose a new basis
$v_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, v_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$
$Tv_1 = 0, Tv_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix} = v_1$
$\beta = \{v_1, v_2\}$
$[T]_\beta = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$

Example

$$T = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$T^2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$
$$T^3 = 0$$

$$T^d = 0 \Rightarrow T^n = 0, p_T(x) = x^n$$
$$\{0\} \subset \ker T \subset \ker T \subset \ker T^2 = \mathbb{R}^3$$
$$\ker T = \mathbb{R}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$
$$\ker T^2 = span\left\{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}\right\}$$
$$[T]_\beta = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

Jordan Nilpotent

September-30-11 9:41 AM

Jordan Nilpotent

The Jordan nilpotent of order k is

$$J_k = \begin{bmatrix} 0 & 1 & 0 & 0 \\ & 0 & 1 & 0 \\ & & 0 & 1 \\ & & & 0 \end{bmatrix}_{k\times k}$$

$i.e.$ There is a basis $e_1, e_2, \dots, e_k$ and

$J_k e_i = e_{i-1}\ 2 \le i \le k$

$J_k e_1 = 0$

We can get a lot of nilpotent matrices by taking direct sums of Jordan nilpotents (Canonical form) :

$n_1 \le n_2 \le \dots \le n_k$

$J_{n_1} \oplus J_{n_2} \oplus \dots \oplus J_{n_k}$

Complement

If subspace $W_1 \subseteq V$ then a complement of W_1 in V is a subspace $W_2 \subseteq V$ s.t. $W_1 \cap W_2 = \{0\}$ and $W_1 + W_2 = V$.

i.e. $V = W_1 \dotplus W_2$

Extension

Suppose $W_1, W_2 \subseteq Y \subseteq V$

$W_1 \cap W_2 = \{0\}$ but $W_1 + W_2 \subset Y$

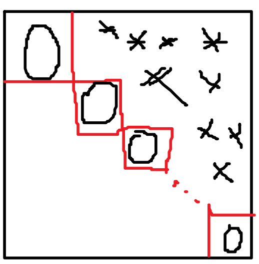
Can find $W_3 \supset W_2$ s.t. $Y = W_1 \dotplus W_3$

Note: Nimpotent

If T is nimpotent of order k, then $m_T(x) = x^k$ and $p_T(x) = x^n, n = \dim V$

Theorem

$T \in \mathcal{L}(V)$ is nilpotent $\Leftrightarrow$ there is a basis in which T is strictly block upper triangular



Better Example

$$A = \begin{bmatrix} 13 & -21 & -5 & 4 \\ 6 & -8 & -4 & 1 \\ 5 & -3 & -7 & -1 \\ 1 & -5 & 3 & 2 \end{bmatrix}$$

$A^3 = 0$

$$\ker A = sp\left\{\begin{bmatrix}2\\1\\1\\0\end{bmatrix}, \begin{bmatrix}1\\1\\0\\2\end{bmatrix}\right\}$$

$$\ker A^2 = sp\left\{\begin{bmatrix}2\\1\\1\\0\end{bmatrix}, \begin{bmatrix}1\\1\\0\\2\end{bmatrix}, \begin{bmatrix}2\\0\\0\\-1\end{bmatrix}\right\}$$

$$\ker A^3 = R^4 = sp\left\{\begin{bmatrix}2\\1\\1\\0\end{bmatrix}, \begin{bmatrix}1\\1\\0\\2\end{bmatrix}, \begin{bmatrix}2\\0\\0\\-1\end{bmatrix}, \begin{bmatrix}0\\0\\0\\1\end{bmatrix}\right\}$$

$$\text{Let } v_4 = \begin{bmatrix}0\\0\\0\\1\end{bmatrix}$$

$$\text{Let } v_3 = Av_4 = \begin{bmatrix}4\\1\\-1\\2\end{bmatrix} \in \ker A^2, v_3 \notin \ker A$$

$$\text{Let } v_2 = Av_3 = \begin{bmatrix}44\\22\\22\\0\end{bmatrix} \in \ker A$$

$$\text{Find a vector } v_1\ s.t.\ \ker A = sp\{v_1, v_2\}, v_1 = \begin{bmatrix}1\\1\\0\\2\end{bmatrix}$$

$$[A]_\beta = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} = [0] \oplus \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} = J_1 \oplus J_3$$

Complement Example

Suppose $V = \mathbb{R}^3$

$$W_1 = span\left\{e_1 = \begin{bmatrix}1\\0\\0\end{bmatrix}, e_2 = \begin{bmatrix}0\\1\\0\end{bmatrix}\right\}$$

then $W_2 = sp\left\{\begin{bmatrix}0\\0\\1\end{bmatrix}\right\}$ is a complement but $W_2' = sp\left\{\begin{bmatrix}1\\2\\3\end{bmatrix}\right\}$ is also a complement

In general $W_2'' = span\left\{\begin{bmatrix}*\\*\\1\end{bmatrix}\right\}$

Find a Complement

To find a complement, choose a basis for W_1, say $\{v_1, \dots, v_k\}$ extend to a basis of V $\{v, \dots, v_k, v_{k+1}, \dots, v_n\}$ let $W_2 = span\ \{v_{k+1}, \dots, v_n\}$

Then W_2 is a complement of W_1

Proof of Extension

Same proof:

Chose basis for W_1, W_2 combine and extend to basis for Y. Remove W_1 basis and have remainder is span of W_3

Proof of Nimpotence note

$T^k\ and\ T^{k-1} \neq 0$

$q(x) = x^k$

$\Rightarrow q(T) = 0 \therefore q \in J = \{p(x) : p(T) = 0\} = (m_T) = \{m_T(x)r(x)\}$

So $m_T |\ x^k \therefore m_T(x) = x^d$ for some $d \le k$

But $T^{k-1} \neq 0$ so $d \ge k \therefore m_T(x) = x^k$

$p_T(x)$ has the same roots $\therefore$ 0 is the only root of p_T

$deq\ (p_T) = n \therefore p_T(x) = x^n$

Proof of Theorem

$\Rightarrow$

Look at

$V_0 = \{0\}, V_1 = \ker T, \dots, V_i = \ker T^i, \dots, V_k = \ker T^k = V$

$\{0\} = V_0 \subset V_1 \subset \dots \subset V_k = V$

If I choose a basis $v_1, \dots, v_{n_1}\ for\ V_1$ and extend to basis $v_1, \dots, v_{n_1}, v_{n_1+1}, \dots, v_{n_2}$

And so on to $v_1, \dots, v_{n_1}, v_{n_1+1}, \dots, v_{n_2}, \dots, v_{(n_{k-1}+1)}, \dots, v_{n_k}$

T is block upper triangular with diagonal blocks = 0.

$\Leftarrow$Strictly block upper triangular

Conversely, if $[T]_\beta$ is strictly block upper triangular then T is nilpotent

Suppose $T = J_{n_1} \oplus J_{n_2} \oplus \cdots \oplus J_{n_k}$
$n_1 \leq n_2 \leq \ \dots \leq n_k$
$\ker J_n = \mathbb{F} e_1$
$\ker J_n^2 = sp\{e_1, e_2\}$
$\ker J_n^i = sp\{e_1, \dots, e_i\}$

$nul\left(J_n^i\right) = \begin{cases} i \ if \ i \leq n \\ n \ if \ i > n \end{cases}$
$nul\left(J_{n_1} \oplus \cdots \oplus J_{n_k}\right) = k$
$nul\left(J_{n_1} \oplus \cdots \oplus J_{n_k}\right)^2$

Example
$T = J_1 \oplus J_1 \oplus J_2 \oplus J_5 \oplus J_7$
$nul\ (T) = 5, nul(T^2) = 8, nul(T^3) = 10, nul(T^4) = 12, nul(T^5) = 14, nul(T^6) = 15, nul(T^7) = 16$
$= \dim V$

$nul\ (T^i) - nul(T^{i-1}) = \left|\{n_j : n_j \geq i\}\right| = \left|\{n_j : n_j = i\}\right| + \left|\{n_j : n_j > i\}\right|$
$= \left|\{n_j : n_j = i\}\right| + \left|\{n_j : n_j \geq i+1\}\right| = \left|\{n_j : n_j = i\}\right| + nul\left(T^{i+1}\right) - nul\left(T^i\right)$
$\therefore \left|\{n_j : n_j = i\}\right| = 2\ nul\left(T^i\right) - nul\left(T^{i+1}\right) - nul\left(T^{i-1}\right)$

Nilpotent Jordan Canonical Form

October-03-11 9:37 AM

Theorem

$T \in \mathcal{L}(V)$ nilpotent of order k, then T is similar to a direct sum of Jordan nilpotents.
$T \sim J_{n_1} \oplus J_{n_2} \oplus \cdots \oplus J_{n_s}$
$k = n_1 \geq n_2 \geq \cdots \geq n_s$
Moreover,
$|\{n_i = j\}| = 2nul(T^j) - nul(T^{j+1}) - nul(T^{j-1})$

Proof of Theorem

(taken from Herstein, Intro to Alg)
Induction on $n = \dim V$
$n = 1 : T = [0] = J_1$

Now assume it holds for $\dim V < n$
$T^k = 0 \neq T^{k-1}$
$\exists u_1 \in V\ s.t. T^{k-1}u_1 \neq 0$

Claim

$\{u_1, Tu_1, T^2u_1, \dots, T^{k-1}u_1\}$ is linearly independent.

$If\ 0 = \sum_{i=0}^{k-1} a_i T^i u_1,\qquad a_i$ not all zero, $then\ \exists i_0\ s.t. a_i = 0\ \forall i < i_0, a_{i_0} \neq 0$

$$0 = T^{k-i_0-1}\left(\sum_{i=0}^{\infty} a_i T^i u_1\right) = a_{i_0}T^{k-1}u_1 + a_{i_0+1}T^k\, u_1 \dots = a_{i_0}T^{k-1}u_1$$

$T^{k-1}u_1 \neq 0 \Rightarrow a_{i_0} = 0$
$\therefore linearly\ independent$

Let $U = sp\{\, u_1, Tu_1, \dots, T^{k-1}u_1)$
$\dim U = k, TU \subseteq U$
$A = T\big|_U$

$$A\begin{cases}(T^i u_1) = T^{i+1}u_1 & 0 \leq i < k-1\\ (T^{k-1}u_1) = 0\end{cases}$$

$\therefore A \sim J_k$

Need to find subspace W s.t.

1) $U \cap W = \{0\}$
2) U+W=V
3) $TW \subseteq W$

$\Rightarrow V = U \dot{+} W$
$\Rightarrow T \sim T\big|_U \oplus T\big|_W$
$0 = T^k = \left(T\big|_U\right)^k \oplus \left(T\big|_W\right)^k$
$B = \left(T\big|_W\right)$ is nilpotent of order $\leq k$
By induction, $B \sim J_{n_2} \oplus J_{n_3} \oplus \cdots \oplus J_{n_s}$
$\therefore T \sim J_k \oplus J_{n_2} \oplus \cdots \oplus J_{n_s}$

Take a maximal subspace W satisfying

1) $U \cap W = \{0\}$
2) $TW \subseteq W$

So $U \dot{+} W$ is direct
Claim: If $Tv \in U + W, so\ Tv = u + w\ u \in U, w \in W$ then $u = \sum_{i=1}^{k-1} a_i T^i u_1$

Let $u = \sum_{i=0}^{k-1} a_i T^i u_1$
$Tv = u + w$
$\therefore 0 = T^{k-1}(Tv) = T^{k-1}u + T^{(k-1)}w$
$\qquad\qquad \in U \qquad \in W$ because $TU \subseteq U, TW \subseteq W$
$U \cap W = \{0\} \therefore T^{k-1}u = 0, T^{k-1}w = 0$
$0 = T^{k-1}a_0u_1 \Rightarrow a_o = 0$

Claim $U + W = V$
Suppose otherwise. Pick $v \notin U + W$
Look at $v \notin U + W, Tv, T^2v, \dots, T^{k-1}v, T^kv = 0 \in U + W$
$\therefore \exists v_1 = T^iv \notin U + W, but\ Tv_1 \in U + W$
$Tv_1 = u_2 + w_2\ \ , u_2 \in U, w_2 \in W$

$$u_2 = \sum_{i=1}^{k-1} a_i T^i u_1 = T\left(\sum_{i=0}^{k-2} a_{i+1}T^i u_1\right) = Tu_3$$

$Let\ v_2 = v_1 - u_3 \notin U + W$
$Tv_2 = Tv_1 - Tu_3 = (u_2 + w_2) - u_2 = w_2 \in W$
Let $W' = span\,\{W, v_2\} \supset W$
$TW' = span\,\{TW, Tv_2\} \subseteq W \subseteq W'$
$W' \cap U = \{0\}$
(otherwise $\alpha v_2 + w \in W = u \in U \Rightarrow \alpha = u - w \in U + W \Rightarrow \alpha = 0 \Rightarrow W = 0, U = 0$

So W is not maximal w.r.t 1), 3) a contradiction. So $U + W = V \therefore V = U \dot{+} W$
This completes the proof. ■

2nd Proof

More constructive
Let $N_i = \ker T^i\ 0 \leq i \leq k$
$\{0\} = N_0 \subset N_1 \subset N_2 \subset \cdots \subset N_k = V$
Choose a complement W_k to N_{k-1}: $N_{k-1} \dot{+} W_k = V$
Choose a basis $w_1, \dots, w_{r_1}\ for\ W_k$
$w_j, T_{w_j}, \dots, T^{k-1}w_j$ all non-zero
As first proof, they are linearly independent

$T\Big|_{span\{w_j,\ldots,T^{k-1}w_j\}} \sim J_k$

Claim

$Tw_1, Tw_2, \ldots, T_{w_r}$ are linearly independent, and $sp\{Tw_1, \ldots, T_{w_r}\} \cap N_{k-2} = \{0\}$

Proof

Suppose $\sum_{i=1}^{r} a_i T w_i = v \in N_{k-2}$

$\therefore T^{k-2} \sum_{i=1}^{r} a_i T w_i = T^{k-2} v = 0 = T^{k-1}\left(\sum_{i=1}^{r} a_i w_i\right)$

$\therefore \sum_{i=1}^{r} a_i w_i \in N_{i-1} \cap W_k = \{0\}$

$\{w_i\}$ lin. indep. $\Rightarrow a_i = 0$

$\therefore \{Tw_i\}\ lin.independent, sp\{Tw_1, .., Tw_{r_1}\} \cap N_{k-2} = \{0\}$

$N_{k-2} \dotplus sp\{Tw_1, \ldots, Tw_{r_1}\} \subseteq N_{k-1}$

Find W_{k-1} s.t. $N_{k-2} \dotplus span\{Tw_1, \ldots, Tw_{r_1}\} \dotplus W_{k-1} = N_{k-1}$

Choose a basis for $W_{k-1}\{w_{r_1+1}, \ldots, w_{r_2}\}$

Claim

Suppose $N_j = N_{j-1} \dotplus U_j, j \geq 2$. U_j has basis $u_1, \ldots, u_m$

then $\{Tu_1, \ldots, Tu_m\}$ is linearly independent and $sp\{Tu_1, .., Tu_m\} \cap N_{j-2} = \{0\}$

Proof

If $\sum_{i=1}^{m} a_i T u_i = v \in N_{j-2} \Rightarrow T^{j-2}\left(\sum a_i T u_i\right) = T^{j-2} v = 0 \Rightarrow T^{k-1}\left(\sum a_i u_i\right)$

$\Rightarrow \sum a_i u_i \in N_{j-1} \cap U_j = \{0\}$

$\therefore a_i = 0, v = 0$

Then I can extend $\{Tu_1, \ldots, Tu_m\}$ to a complement of N_{j-2} inside N_{j-1} by adding new basis vectors $v_{r_{k-j}+1}, \ldots, v_{r_{k+1-j}}$

This process builds the Jordan form. Get $\dim V - \dim(N_{k-1})$ blocks of length k

Our formula was

$2nul(T^k) - nul(T^{k+1}) - nul(T^{k-1}) = 2n - n - \dim(N_{k-1}) = \dim V - \dim(N_{k-1})$

$N_j = N_{j-1} \dotplus U_j$

$\dim U_j = \dim N_j - \dim N_{j-1} = \#\ of\ Jordan\ blocks\ of\ size \geq j$

$nul(T^j) - nul(T^{j-1}) = |\{n_j \geq j\}|$

$nul(T^{j+1}) - nul(T^j) = |\{n_i > j\}|$

$2\,nul(T^j) - nul(T^{j+1}) - nul(T^{j-1}) = |\{n_i = j\}|$

The Algebra of Nilpotent Transformation

October-05-11 10:05 AM

Homomorphism

A homomorphism between two algebras A and B over a ring K is a map $F: A \to B$ with the following properties:
$\forall k \in K, x, y \in A$

1) $F(xk) = kF(x)$
2) $F(x+y) = F(x) + F(y)$
3) $F(xy) = F(x)F(y)$

Modulo Polynomials

If $m \in \mathbb{F}[x]$, (m) ideal of all multiples of m.
Say $p \equiv q \; mod(m)$ if $p - q \in (m) \equiv m|(p-q)$
Make $\mathbb{F}[x]/(m)$ into a ring. Elements are equivalence classes.
$[p] = \{q \equiv p \; mod\;(m)\}$
$[p] \pm [q] = [p \pm q]$
$[p][q] = [pq]$

Check that this is well-defined.
If $p_1 \equiv p_2 \; mod(m), q_1 \equiv q_2 \; mod(m)$
$(p_1 \pm q_1) - (p_2 \pm q_2) = (p_1 - p_2) + (q_1 - q_2) \in (m)$
$p_1 \pm q_1 \equiv p_2 \pm q_2$
$p_2q_2 - p_1q_1 = (p_2 - p_1)q_2 + p_1(q_2 - q_1) \in (m)$
$p_2q_2 \equiv p_1q_1$

Algebra

An algebra is a set A which is

1) A vector space over a field $\mathbb{F}$
2) Has an associative multiplication
3) Distributive law
   $a(x \pm y) = ax \pm ay, \quad a, x, y \in A$
   $\lambda(x+y) = \lambda x + \lambda y, \quad \lambda \in \mathbb{F}$

Algebra of Nilpotent Transformation

$T \in \mathcal{L}(V)$
$A(T) = sp\{I, T, T^2, T^3, \dots\} = \{p(T) : p \in \mathbb{F}[x]\}$
There is a map from
$\mathbb{F}[x] \to A(T), \qquad \Phi: p \mapsto p(T)$

This is a homomorphism. i.e.
$\alpha, \beta \in \mathbb{F}, p, q \in \mathbb{F}[x]$
$(\alpha p + \beta q) \mapsto (\alpha p + \beta q)(T) = \alpha p(T) + \beta q(T)$
$(pq) \mapsto (pq)(T) = p(T)q(T)$

Lemma

If $T^d = 0 \neq T^{d-1}, p \in \mathbb{F}[x]$ then

1) $p(T)$ is invertible $\Leftrightarrow p(0) \neq 0$
2) $p(T) = 0 \Leftrightarrow x^d | p$

Equivalence Class

$$T = J_k = \begin{bmatrix} 0 & 1 & & & & \\ & 0 & 1 & & & \\ & & \ddots & \ddots & & \\ & & & 0 & 1 & \\ & & & & 0 & 1 \end{bmatrix}_{k \times k}$$

$p(x) = a_0 + a_1x + \cdots + a_mx^m$
$p(T) = a_0I + a_1T + a_2T^2 + \cdots + a_mT^m$

$$= \begin{bmatrix} a_0 & & \\ & \ddots & \\ & & a_0 \end{bmatrix} + \begin{bmatrix} 0 & a_1 & & \\ & \ddots & \ddots & \\ & & 0 & a_1 \\ & & & 0 \end{bmatrix} + \cdots + \begin{bmatrix} 0 & \cdots & a_{k-1} \\ & \ddots & \vdots \\ & & 0 \end{bmatrix}$$

$$= \begin{bmatrix} a_0 & a_1 & \cdots & a_{k-1} \\ & \ddots & \ddots & \vdots \\ & & a_0 & a_1 \\ & & & a_0 \end{bmatrix}$$

If q is some polynomial $q(x) = b_0 + b_1x + \cdots + b_mx^m$
$p(T) = q(T)$
$\Leftrightarrow a_i = b_i \; for \; 0 \le i \le k-1$
$\Leftrightarrow x^k | (p(x) - q(x))$
$\Leftrightarrow p \equiv q \; mod(x^k)$

Algebra of Nilpotent Transformation Explanation

$T^d = 0 \neq T^{d-1}$
map is linear, preserves product
Show $p(T) = \Phi(\mathrm{p}) = \Phi(\mathrm{q}) = \mathrm{q}(\mathrm{T}) \Leftrightarrow \mathrm{p} - \mathrm{q} \in (\mathrm{x}^\mathrm{d}) \Leftrightarrow \mathrm{x}^\mathrm{d}|\mathrm{p} - \mathrm{q}$

$m \in \mathbb{F}[x]$
$\mathbb{F}[x]/(m)$ is a "quotient ring" of polynomials modulo m.
$p \equiv q \Leftrightarrow m | p - q$
$\Psi: \mathbb{F}[x] \to \mathbb{F}[x]/(x^d)$ is a homomorphism
Showed if $p_1 \equiv p_2, q_1 \equiv q_2 \; (mod\; x^d)$ then $\alpha p_1 + \beta q_1 \equiv \alpha p_2 + \beta q_2$ and $p_1q_1 \equiv p_2q_2 \; (mod\;(x^d))$
$\therefore$ maps are well defined

$\ker\Phi = (x^d) = \ker\Psi$
$\mathbb{F}[x] \to^{\Phi} A(T)$
$\mathbb{F}[x] \to^{\Psi} \mathbb{F}[x]/(x^d)$
$\mathbb{F}[x] \to^{\Phi^\sim} A(T)$
Can defined $\Phi^\sim$ by $\Phi^\sim([p]) = p(T)$
Well defined $p_1 \equiv p_2 \; (mod\; x^d)$ then $x^d | p_1 - p_2$
$(p_1 - p_2)(x) = x^d r(x)$
$p_1(T) - p_2(T) = T^d r(T) = 0$
$\therefore p_1(T) = p_2(T)$
$\therefore \Phi^\sim$ is well defined

Claim: $\Phi^\sim$ is 1-1 and onto
$\Phi^\sim([p]) = 0 \Leftrightarrow p(T) = 0$

Proof

2) $p_T(x) = x^d$
   $p(T) = 0 \Leftrightarrow x^d p$
1) Write $p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_kx^k,$
   $p(0) = a_0$

   If $p(0) = a_0 = 0$ then $p(x) = xq(x)$
   $\therefore p(T) = Tq(T)$
   T is not invertible $\therefore p(T)$ is not invertible

   If $p(0) = a_0 \neq 0$
   $p(x) = a_0(1 + xq(x))$
   $p(T) = a_0(I + Tq(T))$

Proof 1:

T upper triangular, 0 on diagonal

$$p(T) = \begin{bmatrix} a_0 & \cdots \\ & \ddots \\ & & a_0 \end{bmatrix}$$

$\therefore \sigma(p(T)) = \{a_0\} \neq 0 \therefore invertible$

Proof 2:

Let $\beta = a_0^{-1}\left(I - Tq(T) + (Tq(T))^2 - \cdots + (-1)^d T^d q(T)^d\right)$

$p(T)\beta = a_0\left(I + T(q(T))\right)\dfrac{1}{a_0}\left(I - Tq(T) + (Tq(T))^2 - \cdots + (-1)^d T^d q(T)^d\right)$

$= I - Tq(T) + (Tq(T))^2 - \cdots + (-1)^d T^d q(T)^d + Tq(T) - (Tq(T))^2 - \cdots$
$+ (-1)^d T^{d+1} q(T)^{d+1} = I + (-d)^d T^{d+1} q(T)^{d+1} = I$

$\Phi^{\sim}$ is 1-1
$\Phi^{\sim}$ is onto, $\Phi^{\sim}([p]) = p(T) \in A(T)$

If $\Phi^{\sim}([p]) = \Phi^{\sim}([q]) \Leftrightarrow \Phi^{\sim}([p-q]) = 0 \Leftrightarrow x^d | p-q \Leftrightarrow [p-q] = 0 \Leftrightarrow [p] = [q]$
$\Phi^{\sim}$ is an isomorphism
(It is a bijection, homomorphism, and $\Phi^{\sim}$ is a homomorphism)

Did this for $T = J_d = \begin{bmatrix} 0 & 1 & & & & \\ & 0 & 1 & & & \\ & & \ddots & \ddots & & \\ & & & 0 & 1 & \\ & & & & 0 & 1 \end{bmatrix}_d$

General case
$T = J_{n_1} \oplus J_{n_2} \oplus \cdots \oplus J_{n_s}$
$n = n_1 \geq n_2 \geq \cdots \geq n_s$

$T = \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{bmatrix}_{d\times d} \oplus \begin{bmatrix} 0 & 1 & \\ & 0 & 1 \\ & & 0 \end{bmatrix} \oplus \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \oplus [0] \oplus [0]$

$p(x) = a_0 + a_1 x + a_2 x_2 + \cdots + a_k x^k$

$p(T) = \begin{bmatrix} a_0 & a_1 & a_2 & \dots & a_{d_T} \\ & a_0 & a_1 & & \\ & & \ddots & \ddots & \ddots \\ & & & a_0 & a_1 \\ & & & & a_0 \end{bmatrix}_{d\times d} \oplus \begin{bmatrix} a_0 & a_1 & a_2 \\ & a_0 & a_1 \\ & & a_0 \end{bmatrix} \oplus \begin{bmatrix} a_0 & a_1 \\ 0 & a_0 \end{bmatrix} \oplus [a_0] \oplus [a_0]$

$p(T) \mapsto p(J_d)$, $\qquad p(T) \in A(T), p(J_d) A(J_d)$
$A(J_d) \mapsto A(T)$

Jordan Forms

October-07-11 10:09 AM

Jordan Block

A Jordan block is a matrix $J(\lambda, k) = \lambda I_k + J_k = \begin{bmatrix} \lambda & 1 & \dots & \\ & \ddots & \ddots & \vdots \\ & & \lambda & 1 \\ & & & \lambda \end{bmatrix}$

Jordan Form

A Jordan form is a direct sum of Jordan blocks

From the nilpotent case, we get

Corollary

If $T \in \mathcal{L}(V)$ and $p_T(x) = (x-\lambda)^n$ then $m_T(x) = (x-\lambda)^d$ where $\ker(T-\lambda I)^{d-1} \subset \ker(T-\lambda I)^d = \ker(T-\lambda I)^{d+1}$ and T is similar to
$T \sim J(\lambda, n_1) \oplus J(\lambda, n_2) \oplus \dots \oplus J(\lambda, n_s), d = n_1 \le n_2 \le \dots \le n_s$
Moreover,
$\left|\{u_j = i\}\right| = 2nul(T-\lambda I)^i - nul(T-\lambda I)^i - nul\left(T - \lambda^{i-1}\right)$

Lemma

If $T \in \mathcal{L}(V), \lambda \in \mathbb{F}$ then $N_j = \ker(T-\lambda I)^j$ and $R_j = range(T-\lambda I)^i$ are invariant subspaces for T (and for any A s.t. $AT = TA$)

Proof of Corollary

$p_T(x) = (x-\lambda)^n \Leftrightarrow p_{T-\lambda I}(x) = x^n \Leftrightarrow T - \lambda I$ is nilpotent

Goal

The goal is to prove that if $p_T(x)$ splits into linear terms $p_T(x) = \prod_{i=1}^k (x-\lambda_i)^{e_i}$ then V splits as a direct sum
$V = V_1 \dotplus V_2 \dotplus \dots \dotplus V_k$ where $V_i = \ker(T-\lambda_i I)^{e_i}$

Then T is similar to
$T \sim \left(T\big|_{V_1}\right) \oplus \left(T\big|_{V_2}\right) \oplus \dots \oplus \left(T\big|_{V_k}\right) = T_1 \oplus T_2 \oplus \dots \oplus T_k$
$\left(T_j - \lambda_j I\right)^{e_j} V_j = \{0\}$
So $\left(T_j - \lambda_j I\right)^{e_j} = 0$
$\left(T_j - \lambda_j I\right) \sim J\left(\lambda_j, n_{j,1}\right) \oplus \dots \oplus J\left(\lambda_j, n_{j,s_i}\right)$

Proof of Lemma

$x \in N_j$, then $(T-\lambda I)^j x = 0$
$AT = TA$ then $(T-\lambda I)^j Ax = A(T-\lambda I)^j x = 0$
$\therefore Ax \in \ker(T-\lambda I)^j$

If $y \in Ran(T-\lambda I)^j, y = (T-\lambda I)^j x$
$Ay = A(T-\lambda I)^j x = (T-\lambda I)^j (Ax) \in ran(T-\lambda I)^j$

$J_d, ker J_d = sp\{e_1, \dots, e_i\}$
ran $J_d = sp\{e_{n-i}, e_{n-i+1}, \dots, e_i\}$

Jordan Form Theorem

October-12-11 9:32 AM

Lemma

$T \in \mathcal{L}(V)$ s.t. $(T - \lambda I)^d = 0$ then if $p \in \mathbb{F}[x]$,
$p(T)$ is invertible
$\Leftrightarrow$
$p(\lambda) \neq 0$

Lemma

$T \in \mathcal{L}(V), \lambda \in \sigma(T)$
Let $N_i = \ker(T - \lambda I)^i$
$R_i = ran(T - \lambda I)^i, i \geq 0$
Suppose $\{0\} = N_0 \subsetneq N_1 \subsetneq \cdots \subsetneq N_d = N_{d+1}$

Then $N_{d+j} = N_d \; \forall j \geq 1$
and $V = R_0 \supset R_1 \supset \cdots \supset R_d = R_{d+j} \; \forall j \geq 1$
and $V = N_d \dotplus R_d$

Lemma

$T \in \mathcal{L}(V)$
$\ker(T - \lambda I)^{d-1} \subsetneq \ker(T - \lambda I)^d = \ker(T - \lambda I)^{d+1}$
Then $m_T(x) = (x - \lambda)^d n(x)$ where $n(\lambda) \neq 0$

Theorem

$T \in \mathcal{L}(V)$
Assume $p_T(x)$ splits into linear factors

$$p_T(x) = \prod_{i=1}^{s} (x - \lambda_i)^{e_i}$$

Let $m_T(x) = \prod_{i=1}^{s} (x - \lambda_i)^{d_i}$
$V_i = \ker(T - \lambda_i I)^{d_i}$
Then $V = V_1 \dotplus V_2 \dotplus \cdots \dotplus V_s$

Corollary

If $p_T(x)$ splits
$V = V_1 \dotplus \cdots \dotplus V_s$
$T_i = T\big|_{V_i} \in \mathcal{L}(V_i)$
then $(T_i - \lambda_i I)^{d_i} = 0$
$T \sim T_1 \oplus T_2 \oplus \cdots \oplus T_s$

Proof of Lemma

$T - \lambda I$ is nilpotent
$T - \lambda I \sim J_{n_1} \oplus \cdots \oplus J_{n_s}$
$T \sim J(\lambda, n_1) \oplus \cdots \oplus J(\lambda, n_s)$

Expand p around $x = \lambda$
$p(x) = a_0 (= p(\lambda)) + a_1(x - \lambda) + a_2(x - \lambda)^2 + \cdots + a_n(x - \lambda)^n$
$p(T) = p(\lambda)I + a_1(T - \lambda I) + \ldots + a_n(T - \lambda I)^n = p(\lambda)I + (T - \lambda I)q(T)$
$(T - \lambda I)q(T)$ is strictly upper triangular
Invertible $\Leftrightarrow p(\lambda) \neq 0$

Example

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

$N_1 = sp\{e_1\}, R_1 = sp\{e_1, e_2, e_4, e_5, e_6\}$
$N_2 = sp\{e_1, e_2\}, R_2 = sp\{e_1, e_4, e_5, e_6\}$
$N_3 = sp\{e_1, e_2, e_3\}, R_3 = sp\{e_3, e_5, e_6\}$
$N_4 = sp\{e_1, e_2, e_3\}, R_3 = sp\{e_3, e_5, e_6\}$
$\vdots$

Proof of Lemma

$N_{d+1} = N_d$, Proceed by induction
Assume $N_{d+j} = N_{d+j-1}$
take $v \in N_{d+j+1}$
$\therefore (T - \lambda I)v \in N_{d+j} = N_{d+j-1}$
$\therefore (T - \lambda I)^{d+j-1}(T - \lambda I)v = 0 = (T - \lambda I)^{d+j} v \Rightarrow v \in N_{d+j}$
$\dim(N_i) + \dim(R_i) = n$
$\therefore N_i \subsetneq N_{i+1} \Leftrightarrow R_i \supset R_{i+1}$
So $R_{d+j} = R_d \; \forall j \geq 1$

Claim

$N_d \cap R_d = \{0\}$
Take $v \in R_d \therefore \exists x \in V \; s.t. \; v = (T - \lambda I)^d x$
$v \in N_d \therefore 0 = (T - \lambda I)^d v = (T - \lambda I)^{2d} x$
$\therefore x \in N_{2d} = N_d$
So $v = (T - \lambda I)^d \, x = 0$

$N_d \cap R_d = \{0\}$
So $\dim N_d + R_d = \dim N_d + \dim R_d = n$
$\therefore N_d \dotplus R_d = V$

Proof of Lemma

Factor $m_T(x) = (x - \lambda)^e n(x)$ where $n(\lambda) \neq 0$
Let $N_d = \ker(T - \lambda I)^d$
From Lemma, $n(T)|_{N_d}$ is invertible on $\mathcal{L}(V)$

Claim: $e \geq d$
Take $v \in N_d \setminus N_{d-1} \therefore (T - \lambda I)^{d-1} v \neq 0$
$\therefore n(T)(T - \lambda I)^{d-1} v \neq 0$
$\therefore n(T)(T - \lambda I)^{d-1} \neq 0$
$\therefore e \geq d$ because $0 = m_T(T) = n(T)(T - \lambda I)^e$

Claim $e = d$
Since $0 = m_T(T)v = (T - \lambda I)^e n(T) v$
$\Rightarrow n_T(T)v \in N_e = N_d \; (since \; e \geq d)$
$\Rightarrow (T - \lambda I)^d n(T) v = 0$
$\Rightarrow (T - \lambda I)^d n(T) = 0$
$m_T | \, (x - \lambda)^d n(x)$ or $e = d$

Proof of Theorem

Let $R_1 = ran(T - \lambda_1 I)^{d_1}$, Know $V = V_1 \dotplus R_1$
Claim: $V_i \subseteq R_1 \; for \; i \geq 2$
$[(x - \lambda_i)^{d_i}](\lambda_1) = (\lambda_1 - \lambda_i)^{d_i} \neq 0$

V_1 and R_1 are invariant for T and hence invariant for $(T-\lambda_i I)^{d_i}$
$(T-\lambda_i I)^{d_i}\Big|_{V_1}$ is invertible

Take $v \in V_i, i \geq 2$. Write $v = n + r, n \in N_1, r \in R_1$
$0 = (T-\lambda_i I)^{d_i} v = (T-\lambda_i I)^{d_i} n + (T-\lambda_i I)^{d_i} r = 0 + 0$
(Because of direct sum, both terms are 0)
Since $(T-\lambda_i I)|_{V_1}$ is invertible, $n = 0 \therefore v = r \in R_1$

Now we can prove the theorem by induction on $n = \dim V$
$n = 1: T = [\lambda]$
$\lambda_1 = \lambda, V_1 = V$ Done

Assume result for $m < n$
$V = V_1 \dotplus R_1, T = T\Big|_{V_1} \dotplus T\Big|_{R_1} = T_1 \oplus S$
$(T_1 - \lambda_1 I)^{d_1} = 0$
S acts in R_1, $\dim R_1 < n$
$T \sim \begin{bmatrix} T_1 & 0 \\ 0 & S \end{bmatrix}$ on $V = N_1 \dotplus R_1$
$p_T(x) = p_{T_1}(x) p_S(x)$
$p_{T_1}(x) = (x-\lambda_1)^{e_1}, e_1 = \dim V_1$
$p_S(x) = (x-\lambda_2)^{e_2}(x-\lambda_3)^{e_3} \ldots (x-\lambda_s)^{e_s}$
By induction Hypothesis
$R_1 = V_2 \dotplus V_3 \dotplus \cdots \dotplus V_s$
$\therefore\ \ker(S-\lambda_i I)^{d_i} = \ker(T-\lambda_i I)^{d_i} \subseteq R_i$

Applications of Jordan Forms

October-14-11 9:43 AM

Jordan Form Theorem

$\mathbb{F}$ algebraically closed (or $p_T(x)$ splits into linear terms)

$T \in \mathcal{L}(V), p_T(x) = \prod_{i=1}^{s} (x - \lambda_i)^{e_i}$

Then T is similar to

$$\sum_{i=1}^{s \oplus} \sum_{j=1}^{k_i \oplus} J(\lambda_i, n_{i,j})$$

where $n_{i1} \geq n_{ik_i}, \sum_{j=1}^{k_i} n_{ij} = e_i$

Moreover, for each i, $|\{n_{i,j} = r\}| = 2nul(T - \lambda_i I)^r - nul(T - \lambda_i I)^{r+1} - nul(T - \lambda_i I)^{r-1}$

Note

Jordan blocks can be used to answer similarity-invariant questions.

Proof of Jordan Form Theorem

Already been done

$V = V_1 \dotplus V_2 \dotplus \cdots \dotplus V_s$ where $V_i = \ker(T - \lambda_i I)^{e_i}$

Each V_i is invariant for T, and $T_i = T|_{V_i}$, then $(T_i - \lambda_i I) = 0$

$\therefore T_i \sim \sum_{j=1}^{k_i \oplus} J(\lambda_i, n_{i,j}), \qquad \sum n_{i,j} = \dim V_i = e_i$

Cardinality of # $\{n_{ij}, = r\}$ was done

Example

Which $A \in \mathcal{M}_3(\mathbb{C})$ satisfy $A^3 = I$?

If $A^3 = I$ and $A \sim B$ $B = SAS^{-1}$ then $B^3 = SA^3S^{-1} = SS^{-1} = I$

Look for similarity classes of solutions

Say $A \sim \sum^{\oplus} J(\lambda_i, k_i)$

$A^3 \sim \sum^{\oplus} J(\lambda_i, k_i)^3$

Look at $J(\lambda, k)^3 = (\lambda_i I + J_k)^3 = \lambda^3 I + 3\lambda^2 J_k + 3\lambda J_k^2 + J_k^3$

Need $\lambda^3 = 1$ and $3\lambda^2 = 0$ or $k = 1$

$\therefore \lambda \in \left\{1, e^{i\pi\frac{1}{3}}, e^{-i\pi\frac{1}{3}}\right\}$ and $k = 1$

So A is diagonalizable $A \sim diag(\lambda_1, \lambda_2, \lambda_3)$, $\lambda_i^3 = 1$

Count similar classes:

All λ_i same 3

2 same 1 other 3×2

3 different 1

$= 10$

Example

Find all A with $p_A(x) = (x-4)^4(x+1)^3$ and $m_A(x) = (x-4)^3(x+1)^2$

$\Rightarrow \dim V = 7 = \deg p_A$

$nul\ (A - 4I)^4 = nul(A - 4)^3$

$nul(A + I)^3 = nul(A + I)^2$

Size of largest Jordan block is 3 (from $m_a(x)$)

$\Rightarrow A \sim J(4,3) \oplus J(4,1) \oplus J(-1,2) \oplus J(-1,1)$

Example

Find all A with $p_A(x) = (x+2)^4(x-1)^3$

and $m_{A(x)} = (x+2)^2(x-1)$

$\dim V = 4 + 3 = 7 = \deg p_A$

$\sigma(A) = \{-2, 1\}$

$nul\ \left((A + 2I)^7\right) = nul\left((A + 2I)^2\right) = 4$

$nul(A - I)^7 = nul\left((A - I)^1\right) = 3$

$A \sim J(-2,2) \oplus J(-2,k_2) \oplus J(-2,k_3)$

$2 + k_2 + k_3 = 4$

$\oplus J(1,1) \oplus J(1,1) \oplus J(1,1)$

Two choices $k_2 = 2$ or $k_2 = k_3 = 1$

Gives

$\begin{bmatrix} -2 & 1 \\ 0 & -2 \end{bmatrix} \oplus \begin{bmatrix} -2 & 1 \\ 0 & -2 \end{bmatrix} \oplus I_3$

or

$\begin{bmatrix} -2 & 1 \\ 0 & -2 \end{bmatrix} \oplus [-2] \oplus [-2] \oplus I_3$

The similarity classes of these are the solutions

Example

Which matrices have square roots?

Suppose $A \sim \sum_{i=1}^{n \oplus} J(\lambda_i, k_i)$

Then $A^2 \sim \sum^{\oplus} J(\lambda_i, k_i)^2$

$B = J(\lambda, k)^2 = \begin{bmatrix} \lambda & 1 & \cdots & \\ & \ddots & \ddots & \\ & \cdots & & \lambda \end{bmatrix}^2 = \begin{bmatrix} \lambda^2 & 2\lambda & 1 & \cdots & 0 \\ \vdots & & & & \end{bmatrix}$

$\sigma(B) = \{\lambda^2\}$. If $\lambda \neq 0$ then $(B - \lambda^2 I) = \begin{bmatrix} 0 & 2\lambda & 1 & \cdots & 0 \\ & & \vdots & & \end{bmatrix}$

$(B - \lambda^2 I)^{k-1} = \begin{bmatrix} 0 & 0 & \cdots & 0 & (2\lambda)^{(k-1)} \\ & & & \vdots & \end{bmatrix}$

Jordan form for B is $J(\lambda^2, k)$

Conversely, if $\lambda \neq 0$ $J(\lambda^2, k)$ has a square root.

$S \begin{bmatrix} \lambda^2 & 1 & 0 & \cdots & 0 \\ \vdots & & & & \end{bmatrix} S^{-1} = \begin{bmatrix} \lambda^2 & 2\lambda & 1 & \cdots & 0 \\ \vdots & & & & \end{bmatrix} = \begin{bmatrix} \lambda & 1 & \cdots & \\ & \ddots & \ddots & \\ & \cdots & & \lambda \end{bmatrix}^2$

$$S^{-1}\begin{bmatrix} \lambda & 1 & \dots & \\ & \ddots & \ddots & \\ & \dots & & \lambda \end{bmatrix} S$$

$\lambda = 0$

$J_k^2 = \begin{bmatrix} 0 & 0 & 1 & \dots & 0 \\ \vdots & & & & \end{bmatrix}$

$$\begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \oplus \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

If $k \geq 2$

$J_k^2 \sim J_{\lceil \frac{k}{2} \rceil} \oplus J_{\lfloor \frac{k}{2} \rfloor}$

So if A is a square, the nilpotent part of A must come in pairs of size differing by 0 or 1

Plus we can have as many $J_1 s$ as we want

So e.g. $A \sim J(1,7) \oplus J(2,9) \oplus J(0,5) \oplus J(0,4) \oplus J(0,3) \oplus J(0,3) \oplus J(0,2) \oplus J(0,1) \oplus J(0,1)$

Is a square

The Algebra A(T)

October-17-11 9:30 AM

Generalized Eigenspace

$V_i = \ker(T-\lambda_i)^{e_i}$

Idempotent

A map E is idempotent iff $E^2 = E$
Projections are idempotent

Proposition

$T \in \mathcal{L}(V), p_T(x)$ splits, $p_T(x) = \prod_{i=1}^{s}(x-\lambda_i)^{e_i}$

Let $V_i = \ker(T-\lambda_i)^{e_i}$
Then the idempotents E_i in $\mathcal{L}(V)$ given by
$V = V_1 \oplus V_2 \oplus \cdots \oplus V_s$
$E_i(v) = E_i\left(\sum_{j=1}^{s} v_j\right) = v_i, 1 \le i \le s$ belong to $A(T)$

Chinese Remainder Theorem

$m_1, m_2, \dots, m_s \in \mathbb{N}$ relatively prime
$(\gcd(m_i, m_j) = 1\ for\ i \neq j)$

Then $x \equiv a_i\ (mod\ m_i)$ has a unique solution
$x \equiv a\left(mod \prod_{i=1}^{s} m_i\right)$ for every choice of a_i
$\mathbb{Z} \to \mathbb{Z}/m\mathbb{Z} \to \mathbb{Z}/m_i\mathbb{Z} \oplus \mathbb{Z}/m_2\mathbb{Z} \oplus \cdots \oplus \mathbb{Z}/m_s\mathbb{Z}$
$n \mapsto n\ (mod\ m)$
$\mapsto (n\ mod(m_1), n\ mod\ (m_2), \dots, n\ mod\ (m_s))$

CRT says
$\mathbb{Z}/m\mathbb{Z} \to \mathbb{Z}/m_i\mathbb{Z} \oplus \mathbb{Z}/m_2\mathbb{Z} \oplus \cdots \oplus \mathbb{Z}/m_s\mathbb{Z}$
is a bijection.

Chinese Remainder Theorem for Polynomials

If $m_i(x) \in \mathbb{F}[x], 1 \le i \le s, \gcd(m_i, m_j) = 1\ i \neq j$
then if $p_i \in \mathbb{F}[x]$, the equation $p \equiv p_2\ mod(m_i)$ has a unique solution modulo $m = m_1 m_2 \dots m_s$

Theorem

$T \in \mathcal{L}(V), p_T$ splits $m_T = \prod_{i=1}^{s}(x-\lambda_i)^{a_i}$

Then $A(T) \cong A\left(T\big|_{V_1}\right) \oplus A\left(T\big|_{V_2}\right) \oplus \cdots A\left(T\big|_{V_s}\right)$
$A(T) \leftrightarrow \mathbb{F}[x]/(m_t)$
$A\left(T\big|_{V_1}\right) \oplus A\left(T\big|_{V_2}\right) \oplus \cdots A\left(T\big|_{V_s}\right)$
$\leftrightarrow \mathbb{F}[x]/(m_1) \oplus \cdots \oplus \mathbb{F}[x]/(m_s)$

The Algebra A(T) Description

$T \in \mathcal{L}(V)$

$A(T) = span\{I, T, T^2, \dots, T^{n-1}, \dots\}$
$p_T(x) = x^n + \cdots$
Cayley-Hamilton Theorem: $p_T(T) = 0$
$T^n = -\sum_{i=0}^{n-1} a_i T^i \in span\{T, T, \dots, T^{n-1}\}$
$T^{n+k} = -\sum_{i=0}^{n-1} a_i T^{i+k} \in sp\{I, \dots, T^{n+k-1}\} = sp\{I, \dots, T^{n-1}\}$
by induction.
In fact $m_T(T) = 0, m_T | p_T \deg m_T = d \le n$
$T^d == \sum_{i=0}^{d-1} b_i T^i$
Same argument shows $A(T) = sp\{\ I, T, \dots, T^{d-1}\}$ $\dim A(T) = d = \deg m_T$

$p, q \in \mathbb{F}[x]\ \ p(T) = q(T) \Leftrightarrow (p-q)(T) = 0 \Leftrightarrow m_T | (p-q) \Leftrightarrow p \equiv q\ mod(m_T)$

$\mathbb{F}[x] \to A(T): p \mapsto p(T)$ is a homomorphism; It is linear and multiplicative.
$\mathbb{F}[x] \to \mathbb{F}[x]/(m_T): p \mapsto [p]$ is a homomorphism
$\mathbb{F}[x]/(m_T) \to A(T): [p] \to p(T)$ is an isomorphism.

Proof 1 of Proposition

Let $m_T(x) = \prod_{i=1}^{s}(x-\lambda_i)^{d_i}$
$V_j = \ker(T-\lambda_j I)^{d_j}$
So for a polynomial $p(T)$ to satisfy $p(T)v = 0\ \forall v_j \in V_j$ need $(x-\lambda_j)^{d_j} | p$
Let $q_i(x) = \prod_{j \neq i}(x-\lambda_j)^{d_j}$
Then $q_i(T)v_j = 0\ \forall v_j \in V_j, j \neq i$
Look at $q_i(T)|_{V_i}$. $T|_{V_i} = \lambda_i I + N_i$, N_i nilpotent
$q_i(T)\Big|_{V_i} = q_i\left(T\Big|_{V_i}\right) \Rightarrow q_i(\lambda_i) = \prod_{j \neq i}(\lambda_i - \lambda_j)^{d_j} \neq 0$
By Lemma, $q_i(T)_{V_i}$ is invertible. Moreover, the inverse is a polynomial of T
$\Big($recall, $N = T - \lambda_i I$ $nilpotent\ q_i(N) = a_0(I + Nr(N)) \Rightarrow q_i(N)^{-1}$
$= \frac{1}{a_o}(I - Nr(N) + N^2 r(N)^2 - \cdots)$ terminates $N^d = 0\Big)$
So there is a polynomial $r_i \in \mathbb{F}[x]$ s.t. $e_i(T) = q_i(T)r_i(T)|_{V_i} = I|_{V_i}$

Let $e_i(x) = q_i(x)r_i(x)$
Let $E_i = e_i(T) \in A(T)$
$v_j \in V_j, j \neq i, \qquad E_i v_j = r_i(T)q_i(T)v_j = 0$
$E_i v_i = v_i$
$\therefore E_i\left(\sum_{j=1}^{s} v_j\right) = v_i$
$E_i^2 v = E_i v = v_i \Rightarrow E_i^2 = E_i$

Proof 2 of Proposition

Consider $q_1, \dots, q_s$, q_i defined as before
$\gcd(q_1, q_2, \dots, q_s) = 1 \Rightarrow \sum E_i = I$

By the Euclidian Algorithm $\exists r_i \in \mathbb{F}[x]$ s.t. $\sum_{i=1}^{s} q_i r_i = 1$
Let $e_i = q_i r_i$, and $E_i = e_i(T)$
$E_i v = E_i(v_1 + \cdots + v_s) = r_i(T)q_i(T)(v_1 + \cdots + v_s) = E_i v_i \in V_i$ $(q_i(T)v_j = 0, j \neq i)$
$v = Iv = \left(\sum_{i=1}^{n} E_i\right)v = \sum_{i=1}^{n} E_i v_i$
Direct sum $V = \sum_{i=1}^{s}{}^{+} V_i$ $\therefore$ unique decomposition
$v_i = E_i v_i, \qquad i = 1, 2, \dots, s$
$\therefore E_i^2 = E_i$ has range V_i and kernel $\sum_{j \neq i} V_j$

Example of CRT

$m = 6, m_1 = 2, m_3 = 3$

$\mathbb{Z} \to \mathbb{Z}/6\mathbb{Z}$

$\mathbb{Z}$	$\mathbb{Z}/6\mathbb{Z}$	$(\mathbb{Z}/2\mathbb{Z}, \mathbb{Z}/3\mathbb{Z})$
0	[0]	(0,0)
1	[1]	(1, 1)
2	[2]	(0, 2)
3	[3]	(1, 0)
4	[4]	(0, 1)
5	[5]	(1, 2)
$\vdots$	$\vdots$	$\vdots$

Proof 3 of Proposition

By Proof 2 we get $e_i = q_i r_i \in \mathbb{F}[x]\ s.t. \sum_{i=1}^{s} e_i(x) = 1$

Let $m_i(x) = (x - \lambda_i)^{d_i}, \gcd(m_i, m_j) = 1\ \forall i \neq j$

Let $m = m_1(x)m_2(x) \dots m_s(x) = m_T(x)$

Now $e_i \equiv 0\ mod(m_j), j \neq i$

$$1 = \sum_{j=1}^{s} e_j = e_i\ \ (mod\ m_i)$$

$$\therefore e_i \equiv \begin{cases} 0\ (mod\ m)\ j \neq i \\ 1\ (mod\ m_i) i = j \end{cases} r$$

To solve $\{p \equiv p_i\ (mod_{m_i})\ i \leq i \leq s\}$

Let $p = \sum_{i=1}^{s} p_i e_i(x)$, $\quad p \equiv p_i(x) \cdot 1 + \sum_{j \neq i} p_j(x) \cdot 0 \equiv p_i\ (mod\ m_i)$

$p \equiv q\ (mod\ m_i) i \leq i \leq s$

$\Leftrightarrow m_i | (p - q)\ 1 \leq i \leq s \Leftrightarrow m_i | (p - q) \Leftrightarrow p \equiv q\ (mod\ m)$

Jordan Form Application

October-19-11 9:25 AM

Proposition

$T \in \mathcal{L}(V), p_T$ *splits*
Then T can be expressed uniquely as $T = D + N$ where D is diagonalizable and N is nilpotent and $DN = ND$.

Cyclic Vectors

$T \in \mathcal{L}(V)$ has a **cyclic vector** x if $sp\{x, Tx, T^2x, \dots, \} = V$
T is **cyclic** if it has a cyclic vector.

T has a cyclic vector iff $m_T = p_T$

Theorem

$T \in \mathcal{L}(V)$ TFAE

1) T is cyclic
2) $m_T = p_T$
3) T has a single Jordan block for each eigenvalue

Remark

1 $\Leftrightarrow$ 2 is always true, does not require $p_T(x)$ to split.

Example use of Jordan Form

$T \in \mathcal{L}(V), m_T = \prod (x-\lambda_i)^{d_i}$

$A(T) \cong \mathbb{F}[x]/(m_T) \cong \sum^{\oplus} \mathbb{F}[x]/\left((x-\lambda_i)^{d_i}\right)$

$V_i = \ker(T-\lambda_i I)^{d_i}$
$V = V_1 \dotplus V_2 \dotplus \cdots \dotplus V_s$
$T_i = T\big|_{V_i}, m_{T_i} = (x-\lambda_i)^{d_i}$
$T \sim T_1 \oplus T_2 \oplus \cdots \oplus T_s$
$p(T) \sim p(T_1) \oplus p(T_2) \oplus \cdots \oplus p(T_s)$
but $p(T_i) = q(T_i)\ iff\ p \equiv q\ \left(mod\ (x-\lambda_i)^{d_i}\right)$

Express $p(x)$ as a Taylor around λ_i
$p(x) = a_0 + a_1(x-\lambda_i) + a_2(x-\lambda_i)^2 + \cdots$

$T_i \sim \sum_{i=1}^{k_i} \lambda_i I + J_{n_{ij}}$

$$J = \begin{bmatrix} 3 & 1 & 0 & 0 \\ 0 & 3 & 1 & 0 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

$p(x) = 1 + 2x^2 + x^3$
$p(3) = 1 + 29 + 27 = 46$
$p'(x) = 4x + 3x^2$
$p'(3) = 12 + 27 = 39$
$p''(x) = 4 + 6x, p''(3) = 22$
$p^{(3)}(x) = 6$

$$p(x) = p(3) + p'(3)(x-3) + \frac{p''(3)}{2!}(x-3)^2 + \frac{p^{(3)}}{3!}(x-3)^3$$
$$= 49 + 39(x-3) + 11(x-3)^2 + (x-3)^3$$

$$p(J) = \begin{bmatrix} 46 & 39 & 11 & 1 \\ 0 & 46 & 39 & 11 \\ 0 & 0 & 46 & 39 \\ 0 & 0 & 0 & 49 \end{bmatrix}$$

Proof of Proposition

$$T \sim \sum_{i=1}^{s\ \oplus} T_i \sim \sum_{i=1}^{s} \sum_{j=1}^{k_i} \lambda_i I + J_{n_{ij}}$$

$$D \sim \sum_{i=1}^{s} \sum_{j=1}^{k_i\ \oplus} \lambda_i I$$

$$D = \sum_{i=1}^{s} \lambda_i E_i, E_i \text{ idempotent } ran(E_i) = V_i, \ker(E_i) = \sum_{j \neq i} V_j$$

D is a polynomial in T, $D = \sum \lambda_i E_i = (\sum \lambda_i e_i)(T)$
$\therefore TD = DT$
D is diagonalizable

$N = T - D \sim \sum_{i=1}^{s} \sum_{j=1}^{k_i} J_{n_{ij}}$ is nilpotent

N is also in $A(T)$

Uniqueness

Suppose $T = D_1 + N_1$, D_1 diag, N_1 nilpotent $D_1N_1 = N_1D_1$
D_1 commutes with $D_1 + N_1 = T \therefore D_1$ commutes with $A(T)$
$\therefore D_1$ commutes with D, N
Similarly, N_1 commutes with D,N
D_1 commutes with E_i. If $v_i \in V_i, v_i = E_i v_i$
$D_1 v_i = D_1 E_i v_i = E_i D_1 v_i \in ran\ E_i = v_i$
So V_i is invariant for D_1 (and N_1)
$D_1 = D_1\big|_{V_1} \oplus D_1\big|_{V_2} \oplus \cdots \oplus D_1\big|_{V_s}$
$D = \lambda_1 I\big|_{V_1} \oplus \lambda_2 I\big|_{V_2} \oplus \cdots \oplus \lambda_s I\big|_{V_s}$
Each $D_1|_{V_1}$ is diagonalizable so $(D_1 - \lambda_i I)|_{V_i}$ is diagonalizable
$\therefore D_1 - D$ is diagonalizable $\sim diag(\mu_1, \mu_2, \dots, \mu_s)$
$D_1 + N_1 = T = D + N$

$\therefore D_1 - D = N - N_1$

$(N - N_1)^{2n} = \sum_{j=0}^{2n} (-1)^j \binom{2n}{j} N^j N_1^{2n-j} = 0$

(Because N, N_1 commute, first =)

$(Second\ =)\ j \geq n\ N_j = 0, j \leq n \Rightarrow\ 2n - j \geq n \therefore N_1^{2n-j} = 0$

$0 = (D_1 - D)^{2n} \sim diag(\mu_1^{2n} \& \mu_2^{2n}, \dots, \mu_n^{2n}) \therefore \mu_i^{2n} = 0 \Rightarrow \mu_i = 0 \Rightarrow D_1 = D$

$\therefore N_1 = T - D_1 = N$

■

Cyclic Vectors

If $m_T(x) = x^d + a_{(d-1)}x^{d-1} + \dots + a_0$

$0 = T^d + a_{d-1}T^{d-1} + \dots + a_1 T + a_0 I$

$T^d = -a_{d-1}T^{d-1} - \dots - a_1 T - a_0 I$

$\therefore T^d x \in sp\{x, Tx, \dots, T^{d-1}x\}$

So $sp\{x, Tx, \dots\} = sp\{x, Tx, \dots, T^{d-1}x\}$, where $d = \deg m_T(x)$

$\dim sp\{x, Tx, \dots, T^{d-1}x\} \leq d$

A necessary condition for T to be cyclic is $\deg m_T = n, i.e. m_T = p_T$

Note that $m_T = p_T \Leftrightarrow$ there is a single Jordan block for each eigenvalue.

$m_T(x)$

$= \prod_{i=0}^{n} (x - \lambda_i)^{d_i}$, where d_i is the size of the largest Jordan block for λ_i

$T \sim \sum_{i=1}^{s}{}^{\oplus} (\lambda_i I + J_{d_i})$

A Jordan block with basis $\{e_1, \dots, e_k\}$ has a cyclic vector e_k

Let $v_i \in V_i$ be a cyclic vector for $T|_{V_i}$

Let $v = v_1 + v_2 + \dots + v_s$

Claim: v is cyclic for T

$E_i \in A(T)$ So $v_i = E_i v \in A(T)v = sp\{v, Tv, \dots\}$

$\therefore T^k v_i \in A(T)v \Rightarrow V_i \subseteq A(T)v \Rightarrow V = \sum V_i = A(T)v$

Linear Recursion Revisited

October-21-11 9:31 AM

Linear Recursion Formulae

Given $x_0, x_1, \dots, x_{k-1}$ and the linear recursion $x_{k+n} + a_{n-1}x_{k+n-1} + a_{n-2}x_{k+n-2} + \ \dots + a_0 x_n = 0$
Find a formula for x_k

Let $A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -a_0 & -a_1 & \dots & -a_{n-1} \end{bmatrix}$

$$\begin{bmatrix} x_k \\ x_{k+1} \\ \vdots \\ x_{k+n-1} \end{bmatrix} = A^k \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_{n-1} \end{bmatrix}$$

$$P_A(x) = \begin{vmatrix} x & -1 & 0 & 0 \\ 0 & x & -1 & 0 \\ 0 & 0 & x & -1 \\ a_0 & a_1 & \dots & x + a_{n-1} \end{vmatrix} = \begin{vmatrix} x & 0 & 0 & 0 \\ 0 & x & 0 & 0 \\ 0 & 0 & x & 0 \\ a_0 & a_1 + \dfrac{a_0}{x} & \dots & x + a_{n-1} + \dfrac{a_{n-2}}{x} + \dots + \dfrac{a_0}{x^{n-1}} \end{vmatrix} = x^n + a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + \dots + a_0$$

Factor $p_A(x) = \prod_{i=1}^{s} (x - \lambda_i)^{d_i}$

Case 1: n distinct roots ∴ A is diagonalizable
$A \sim diag(\lambda_1, \lambda_2, \dots, \lambda_n)$

Let $v_i = \begin{pmatrix} 1 \\ \lambda_i \\ \lambda_i^2 \\ \vdots \\ \lambda_i^{n-1} \end{pmatrix} \Rightarrow Av_i = \begin{pmatrix} \lambda_i \\ \lambda_i^2 \\ \dots \\ \lambda_i^{n-1} \\ -a_0 - a_1\lambda_1 - \dots - a_{n-1}\lambda_i^{n-1} \end{pmatrix}$

$-a_0 - a_1\lambda_1 - \dots - a_{n-1}\lambda_i^{n-1} = \lambda_i^n - p_A(\lambda_i) = \lambda_i^n$

$$Av_i = \begin{pmatrix} \lambda_i \\ \lambda_i^2 \\ \vdots \\ \lambda_i^{n-1} \\ \lambda_i^n \end{pmatrix} = \lambda_i v_i$$

So $v_1, \dots, v_n$ is the basis that diagonalizes A.

Express $\begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_{n-1} \end{pmatrix} = b_1 v_1 + \dots + b_n v_n$

$$\begin{pmatrix} xk \\ x_{k-1} \\ \vdots \\ x_{k+n-1} \end{pmatrix} = A^k \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_{n-1} \end{pmatrix} = A^k(b_1v_1 + \dots + b_nv_n) = b_1\lambda_1^k v_1 + b_2\lambda_2^k v_2 + \dots + b_n\lambda_n^k v_n = \begin{pmatrix} b_1\lambda_1^k + b_2\lambda_2^k + \dots + b_n\lambda_n^k \\ \vdots \\ \vdots \end{pmatrix}$$

So $\boxed{x_k = b_1\lambda_1^k + \dots + b_n\lambda_n^k}$
The set of possible sequences we get is the linear span of $(1, \lambda_i, \lambda_i^2, \lambda_i^3, \dots)$

Note

If $p \in \mathbb{C}[x]$ has repeated roots, say $p(x) = (x - \lambda)^2 q(x)$
Then $p'(x) = 2(x - \lambda)q(x) + (x - \lambda)^2 q'(x) = (x - \lambda)r(x)$
If $p(x) = (x - \lambda)q(x), q(\lambda) \neq 0$
$p'(x) = q(x) + (x - \lambda)q'(x)$
$p'(\lambda) = q(x) \neq 0$

So p, p' have a common factor $(x - \lambda)$ iff λ is a root of p of multiplicity ≥ 2
∴ p has simple roots $\Leftrightarrow \gcd(p, p') = 1$

Case 2
Repeated roots:
$$p_A(x) = \prod_{i=1}^{s} (x - \lambda_i)^{d_i}$$
A has a cyclic vector e_n

$$A^2 = A\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -a_0 & -a_1 & \dots & -a_{n-1}\end{bmatrix}\begin{bmatrix}0\\0\\ \vdots \\ 1\end{bmatrix} = A\begin{pmatrix}0\\0\\ \vdots\\0\\1\\ *\end{pmatrix} = \begin{pmatrix}0\\ \vdots\\0\\1\\ *\\ *\end{pmatrix}$$

$\therefore$ only one Jordan block for each eigenvalue

$$A \sim \sum_{i=1}^{s}{}^{\oplus} J(\lambda_i, d_i)$$

Pick $v_{i,0} \in \ker(A-\lambda_i I)^{d_i}$ but not in $\ker(A-\lambda_i I)^{d_i-1}$
Let $v_{i,j} = (A-\lambda_i I)^j v_{i,0},\ 1 \le j \le d_i - 1$
$\{v_{i,0}, \dots, v_{i,d_i-1}\}$ is a basis for Jordan block $\lambda_i I + J_{d_i}$
So $\{v_{i,j} : 1 \le i \le s, 0 \le j \le d_i - 1\}$ is a basis for V

Write $\begin{bmatrix} x_0 \\ \vdots \\ x_{n-1}\end{bmatrix} = \sum b_{ij} v_{ij}$
What is $A^k v_{ij}$?

$$\lambda I + J_d = \begin{bmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{bmatrix}, v_{i,0} = \begin{bmatrix}0\\0\\ \vdots\\0\\1\end{bmatrix}, v_{i,j} = \begin{bmatrix}\vdots\\0\\1\\0\\ \vdots\\0\end{bmatrix}$$

$$\begin{bmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{bmatrix}^k = (\lambda_i + J_d)^k = \sum_{i=0}^{k}\binom{k}{i}\lambda^{k-1}J_d^i = \sum_{i=0}^{d-1}\binom{k}{i}\lambda^{k-1}J_d^i$$

$$= \begin{bmatrix} \lambda^k & k\lambda^{k-1} & \dots & \binom{k}{d-1}\lambda^{k+1-d_i} \\ 0 & \lambda_k & \dots & \binom{k}{d-2}\lambda^{k+1-d_i} \\ & & \vdots & \\ 0 & & & \lambda^k \end{bmatrix}$$

$$A^k v_{i,0} = \lambda^k v_{i,0} + k\lambda_{k-1}v_{i,1} + \dots + \binom{k}{d-1}\lambda^{k+1-d_i}$$

$$A^k v_{i,j} = \lambda^k v_{i,j} + k\lambda^{k-1}v_{i,j+1} + \dots + \binom{k}{?}\lambda^{k-?}v_{i,d_i-1}$$

$$\begin{pmatrix} x_0 \\ \vdots \\ x_{n-1}\end{pmatrix} = \sum b_{ij}v_{ij}$$

$$\begin{pmatrix} x^k \\ \vdots \\ x_{k+n-1}\end{pmatrix} = A^k\begin{pmatrix} x_0 \\ \vdots \\ x_{n-1}\end{pmatrix} = \sum b_i A^k v_{i,j} = \sum b_i\left(\lambda^k v_{i,j} + k\lambda^{k-1}v_{i,j+1} + \cdots\right)$$

$$x_k = \sum_{i,j} b_{i,j}\left(\lambda_i v_{i,j}^{(1)} + k\lambda^{k-1}v_{i,j+1}^{(1)} + \cdots\right) = \sum_{i,j}\lambda_i^k(c_{i,0} + c_{i,1}k + c_{i,2}k^2 + \dots + c_{i,d_i-1}k^{d_i-1} = \sum_i \lambda_i^k q_i(k), \deg q_i < d_i$$

General Solution

$$x_k = \sum_i \lambda_i^k q_i(k)$$

has n unknowns $q_i(x) = c_{i,0} + c_{i,1}\lambda + \dots + c_{i,(d_i-1)}\lambda^{d_i-1}$
Know $x_0, \dots, x_{n-1}$ solve for c_i

Solution space is spanned by
$(1, \lambda_i, \lambda_i^2, \lambda_i^3, \dots)$
$(0, \lambda_i, 2\lambda_i^2, 3\lambda_i^3, \dots)$
$(0, \lambda_i, 2^{d_i-1}\lambda_i^2, 3^{d_i-1}\lambda_i^3, \dots)$

Markov Chains

October-24-11 11:25 AM

Discrete State Space

A discrete state space Σ is a finite set of possible states.

A **discrete** process provides probabilities for transition between states at discrete time intervals.

A process is **stationary** if the transition probabilities are time independent.

A discrete stationary process is called a **Markov process**.

Regular Markov Process

A Markov process is regular if there is an N so $(A^N)_{ij} > 0 \ \forall i,j$
i.e. It is possible over time to move from any state to any other.

Lemma

$A = (a_{ij}) \in \mathcal{L}(V)$

Let $\rho(A) = \max_{1\le i\le n} \sum_{j=1}^{n} |a_{ij}|$ (max of row sum)

Then $\sigma(A) \le \{\lambda: |\lambda| \le \rho(A)\}$

Theorem

$A = (a_{ij})$ is a transition matrix.

Then $1 \in \sigma(A) \subseteq \bar{\mathbb{D}} = \{\lambda: |\lambda| \le 1\}$

Moreover, if A is regular then $\sigma(A) \subseteq \{1\} \cup \mathbb{D} = \{1\} \cup \{\lambda: |\lambda| < 1\}$ and $nul(A-I) = nul(A-I)^2 = 1$

Euclidean Norm

$$\|A\|_2 = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}^2}$$

Usual Euclidean norm on $\mathbb{R}^{n^2}$

Claim

$\|AB\|_2 \le \|A\|_2\|B\|_2$

Proof

$\|AB\|_2^2$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n}\left(\sum_{k=1}^{n} a_{ik}b_{kj}\right)^2 \le_{CS} \sum_{i=1}^{n}\sum_{j=1}^{n}\left(\sum_{k=1}^{n} a_{ik}^2\right)\left(\sum_{l=1}^{n} b_{lj}^2\right)$$

By Cauchy-Schwarz inequality

$$= \left(\sum_{i=1}^{n}\sum_{k=1}^{n} a_{ik}^2\right)\left(\sum_{j=1}^{n}\sum_{l=1}^{n} b_{lj}^2\right) = \|A\|_2^2\|B\|_2^2$$

Corollary

If A is a regular transition matrix, then A^m converges to $L = vu^t$ where $Av = v$, v has entries $\sum_i v_i = 1$ and $u^t = (1,1,\dots,1)$

This is the idempotent in $\mathcal{A}(A)$ with range $\ker(A-I)$.

Moreover, if w is any probability vector then $\lim_{n\to\infty} A^n w = v$

Label the states $\Sigma = \{1,2,\dots,n\}$. The probability of moving from state j to state I is $p_{ij} \ge 0$. So $\sum_{i=1}^{n} p_{ij} = 1 \ \forall j$

Let $A = [p_{ij}]_{n\times n} = \begin{bmatrix}p_{11}\\p_{21}\\ \vdots\\p_{n1}\end{bmatrix} \dots \begin{bmatrix}p_{1n}\\p_{2n}\\ \vdots\\p_{nn}\end{bmatrix}$ Column sums are 1

What is the limiting behaviour as time $\to \infty$?

Initial state $p_0 = \begin{pmatrix}\alpha_1\\ \dots\\ \alpha_n\end{pmatrix}$ At time 1 $p_1 = Ap_0$, $p_{n+1} = Ap_n \ \forall n \ge 1$

Interested in $\lim_{n\to\infty} A^n p_0$

Example

A microorganism has 3 possible reproductive states: Male, Female, and Neuter.
Male one day → M 2/3 time, N 1/3 time next day
Female one day → F 1/2 time, N 1/2 time next day
Neuter one day → M 1/6, F 1/2, N 1/3

$$A = \begin{bmatrix}\frac{2}{3} & 0 & \frac{1}{6}\\ 0 & \frac{1}{2} & \frac{1}{2}\\ \frac{1}{3} & \frac{1}{2} & \frac{1}{3}\end{bmatrix}. \text{ Initially } p_0 = \begin{bmatrix}m_0\\f_0\\n_0\end{bmatrix}, p_n = A^n p_0$$

$$A^T\begin{bmatrix}1\\1\\1\end{bmatrix} = \begin{bmatrix}\frac{2}{3} & 0 & \frac{1}{3}\\ 0 & \frac{1}{2} & \frac{1}{2}\\ \frac{1}{6} & \frac{1}{2} & \frac{1}{3}\end{bmatrix}\begin{bmatrix}1\\1\\1\end{bmatrix} = \begin{bmatrix}1\\1\\1\end{bmatrix} \text{ In general } A^t\begin{bmatrix}1\\1\\1\end{bmatrix} = \begin{bmatrix}1\\1\\1\end{bmatrix}$$

so 1 is always an eigenvalue since $\sigma(A^T) = \sigma(A)$

$$p_A(x) = (x-1)\left(x^2 - \frac{1}{2}x - \frac{1}{12}\right), \qquad \sigma(A) = \left\{1, \frac{1 \pm \sqrt{\frac{7}{3}}}{4}\right\} \therefore \text{Diagonalizable}$$

$$A = S^{-1}\begin{bmatrix}1 & 0 & 0\\ 0 & \frac{1+\sqrt{\frac{7}{3}}}{4} & 0\\ 0 & 0 & \frac{1-\sqrt{\frac{7}{3}}}{4}\end{bmatrix} S \text{ As } n \to \infty, \qquad A^n = S^{-1}\begin{bmatrix}1&0&0\\0&0&0\\0&0&0\end{bmatrix} S = L$$

$L = L^2$ is the idempotent in $\mathcal{A}(A)$ with range $span(v$ where $Av = v$ and v is a probability vector.

$$\ker(A-I)\begin{bmatrix}-\frac{1}{3} & 0 & \frac{1}{6}\\ 0 & -\frac{1}{2} & \frac{1}{2}\\ \frac{1}{3} & \frac{1}{2} & -\frac{2}{3}\end{bmatrix} \mapsto \begin{bmatrix}2 & 0 & -1\\ 0 & 1 & -1\\ 0 & 0 & 0\end{bmatrix}\begin{bmatrix}1\\2\\2\end{bmatrix} = 0$$

Normalize $\begin{bmatrix}1\\2\\2\end{bmatrix}$ to get the probability vector $v = \begin{bmatrix}0.2\\0.4\\0.4\end{bmatrix}$

Have vectors v, v_2, v_3 a basis s.t.

$$Av = v, \qquad Av_2 = \frac{1+\sqrt{\frac{7}{3}}}{4}v_2, \qquad Av_3 = \frac{1-\sqrt{\frac{7}{3}}}{4}v_3$$

If $p_0 = a_1 v + a_2 v_2 + a_3 v_3$

$$p_n = A^n p_0 = a_1 v + \left(\frac{1+\sqrt{\frac{7}{3}}}{4}\right)^n v_2 + \left(\frac{1-\sqrt{\frac{7}{3}}}{4}\right)^n v_3 \to a_1 v$$

$u = \begin{bmatrix}1\\1\\1\end{bmatrix}$, $A^T u = u$ and $u^T A = u^T$

$u^T p_0 = m_0 + f_0 + n_0 = 1$

$u^T p_n = u^T(A^n p_0) = (u^T A^n)p_0 = u^T p_0 = 1$, and $p_n = \begin{bmatrix}m_n \ge 0\\ f_n \ge 0\\ n_n \ge 0\end{bmatrix}$ because $a_{ij} \ge 0$

So p_n is a probability vector.

$a_1 v = p_n = \lim_{n\to\infty} A^n p_0, \qquad 1 = u^T(a_1 v) = a_1 \Rightarrow a_1 = 1$

Therefore in the limit as $n \to \infty$ is 20% M, 40% F, 40%N

Proof of Lemma

Suppose $\lambda \in \sigma(A)$, $Av = \lambda v$, $v \ne 0$

$v = \begin{bmatrix}v_1\\ \vdots\\ v_n\end{bmatrix}$. Pick i_0 such that $|v_{i_0}| \ge |v_i| \forall 1 \le i \le n$

$$|\lambda v_{i_0}| = \left|\sum_{j=1}^{n} a_{i_0 j}v_j\right| \le \sum_{j=1}^{n}|a_{i_0 j}||v_j| \le \left(\sum_{j=1}^{n}|a_{i_0 j}|\right)|v_{i_0}| \le \rho(A)|v_{i_0}|$$

$\therefore |\lambda| \le \rho(A)$

Proof of Theorem

$u^T = (1, 1, \dots, 1)$ then $u^T A = u^T$ because column sums are all 1. So $A^T u = u$, or $1 \in \sigma(A^T) = \sigma(A)$ - Since $A \sim A^T$ so $\det(A) = \det(A^T)$
$\rho(A^T) = \max\{1, 1, .., 1\} = 1 \therefore \sigma(A) = \sigma(A^T) \subseteq \overline{\mathbb{D}}$ by Lemma

Proved first part, now prove that $(1, \dots, 1)^T$ is the only eigenvector for 1 or -1

A is regular so $\exists N$ such that $A^N = (c_{ij}), c_{ij} > 0$
Observe that A^{N+1} has strictly positive entries.

Suppose $|\lambda| = 1, A^T u = \lambda u,\ u = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \neq 0$

Repeat argument in Lemma for $(A^N)^T$ and $(A^{N+1})^T$
$(A^N)^T = (c_{ij})^T$ has row sums $= 1$
Pick i_0 s.t. $|u_{i_0}| \ge |u_i|\ \forall i$

$$|u_{i_0}| = |\lambda^N||u_{i_0}| =_1 \left|\sum_{i=1}^n c_{ii_0} u_i\right| \le_2 \sum_{i=1}^n c_{ii_0}|u_i| \le_3 \left(\sum_{i=1}^n c_{ii_0}\right)|u_{i_0}| =_4 |u_{i_0}|$$

1: Since $\lambda^N u = (A^N)^T u$
2: Since $c_{ii_0} > 0$ do not need absolute values about them.
3: An equality iff $u_{i_0} = u_i \forall i$
4: $(A^N)^T$ has row sums 1
This is an equality therefore if $u_{i_0} > 0$ then $u_i \ge 0\ \forall i$.
3 must be made equal so $u_i = u_{i_0} \forall i$ so 2 is also an equality.

$$\therefore u_i = u_{i_0} \Rightarrow u \in sp\left\{\begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}\right\}$$

$$A^T \mu \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} = \mu \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$$

$\therefore \lambda u = A^T u = u \Rightarrow \lambda = 1$
So $\sigma(A) \subseteq \{1\} \cup \mathbb{D}$
$nul(A - I) = nul(A^T - I) = 1$
$\therefore$ Single Jordan block for 1

$$A \sim (I_k + J_k) \oplus \sum_{i=1}^s J(\lambda_i, k_i), |\lambda_i| < 1$$

$I_k + J_k = S^{-1} A S$

$$(S^{-1}AS)^m = (I + J_k)^m \oplus \sum_{i=1}^s J(\lambda_i, k_i)^m,$$

For $|\lambda| < 1$

$$J(\lambda, k)^m = (\lambda I_k + J_k)^m = \lambda^m I_k + \binom{m}{1}\lambda^{m-1} J_k + \binom{m}{2}\lambda^{m-2} J_k^2 + \dots + \binom{m}{k-1}\lambda^{m+1-k} J_k^{k-1}$$

$$= \begin{pmatrix} \lambda^m & m\lambda^{m-1} & \dots & \binom{m}{k-1}\lambda^{m+1-k} \\ & \ddots & & \vdots \end{pmatrix} \to 0\ as\ m \to \infty$$

$$(I_k + J_k)^m = \begin{pmatrix} 1 & m & \dots & \binom{m}{k-1} \\ & \ddots & & \vdots \end{pmatrix}$$

$m = 1$: $(1) \to (1)$
$m \ge 2$: $\binom{m}{1}\ \|(I + J_k)^2\| \ge m \to \infty$

On the other hand
$\|(S^{-1}AS)^m\|_2 = \|S^{-1}A^m S\|_2 \le \|S^{-1}\|_2 \|A^m\|_2 \|S\|$
A^m is a transition matrix so

$$\sum_{i=1}^n b_{ij} = 1 \ge b_{ij} \ge 0$$

$b_{ij}^2 \le b_i$

$$\text{So } \|A^m\|_2^2 = \sum_{j=1}^n \sum_{i=1}^n b_{ij}^2 \le \sum_{j=1}^n \sum_{i=1}^n b_{ij} = n$$

$$\left\|(I + J_k)^m + \sum^{\oplus} J(\lambda_i, k_i)^m\right\| = \|S^{-1}AS\| \le \sqrt{n}\|S\|_2 \|S^{-1}\|_2$$

$$\left\|(I + J_k)^m + \sum^{\oplus} J(\lambda_i, k_i)^m\right\| \ge m \text{ If } nul(A - I)^2 \ge 2$$

$\therefore nul(A - I)^2 = 1$

Proof of Corollary

The last argument shows that

$$(S^{-1}AS)^m = (1) \oplus \sum_{i=1}^s J(\lambda_i, k_i)^m \to (1) \oplus 0$$

This is the idempotent in A(T) with range $\ker(T - I)$
$A^m = ST^m S^{-1} \to S((1) \oplus 0)S^{-1} = L$
L is the idempotent in A(A) with range $\ker(A - I)$

So $\ker L = span\{\ker(A-\lambda_i)^{d_i}, 1 \le i \le s\}$
Let $v \in \ker(A-I)$

Know $u = \begin{pmatrix}1\\1\\\vdots\\1\end{pmatrix}$ is an eigenvector for A^T, eigenvalue 1

So $u^T A = u^T$. Look at $A^m\left(\frac{1}{n}u\right)$

$u^T\left(A^m \frac{1}{n}u\right) = (u^T A^m)\frac{1}{n}u = u^T \frac{1}{n}u = \frac{n}{n} = 1$

$\frac{1}{n}u$ is a probability vector (w prob. vector $\Leftrightarrow w_i \ge 0, u^t w = \sum w_i = 1$)

$u^T\left(A^m \frac{1}{n}u\right) = 1$

$(A^m)_{ij} \ge 0 \Rightarrow \left(A^m \frac{1}{n}u\right)_i \ge 0\ \forall i$

Eventually $(A^m)_{ij} > 0 \Rightarrow \left(A^m \frac{1}{n}u\right) > 0$

$L\frac{1}{n}u = \lim_{m\to\infty} A^m\ \frac{1}{n}u = cv, \qquad probability\ vector$

$ran\ L = \ker(A-I) = Iv$

Normalize v so that $u^T v = 1 \Rightarrow \therefore c = 1$

$A^m\left(\frac{1}{n}u\right) \to v$

$v = Av = A^m v = (b_{ij})\begin{pmatrix}v_1\\\vdots\\v_n\end{pmatrix} = \begin{pmatrix}v_1\\\vdots\\v_n\end{pmatrix}$

For m large $m_{ij} > 0, v_i \ge 0$

$\therefore v_i = \sum_{j=1}^{n} b_{ij}v_i > 0$

$L \in A(A)$

$LA = \lim_{m\to\infty} A^m A = \lim_{m\to\infty} A^{m+1} = L$

$AL = \lim_{m\to\infty} A^{m+1}\, L$

Write

$L = [\alpha_1 \& \alpha_2 \& \dots \& \alpha_n]\ , \alpha_i \in \mathbb{R}^n$

$L = AL = [A\alpha_1 \& A\alpha_2 \& \dots \& A\alpha_n]$

$\therefore A\alpha_i = \alpha_i, so\ \alpha_i = c_i v$

Similarly,

$L = \begin{bmatrix}\beta_1^T\\\beta_2^T\\\vdots\\\beta_n^T\end{bmatrix}, \beta \in \mathbb{R}^n$

$L = LA = \begin{bmatrix}\beta_1^T A\\\beta_2^T A\\\vdots\\\beta_n^T A\end{bmatrix}$

$\therefore \beta_i^T A = \beta_i^T$ or $A^T\beta_i = \beta_i$

$\therefore \beta_i = d_i u, u = \begin{pmatrix}1\\1\\\vdots\\1\end{pmatrix}$

So each row of L has all entries the same.

If $v = \begin{bmatrix}v_1\\v_2\\\vdots\\v_n\end{bmatrix} \Rightarrow L = c\begin{bmatrix}v_1 & v_1 & \dots & v_1\\v_2 & v_2 & \dots & v_2\\\vdots & \vdots & \ddots & \vdots\\v_n & v_n & \dots & v_n\end{bmatrix}$

L is a transition matrix $\therefore c = 1$

$L = \begin{bmatrix}v_1 & v_1 & \dots & v_1\\v_2 & v_2 & \dots & v_2\\\vdots & \vdots & \ddots & \vdots\\v_n & v_n & \dots & v_n\end{bmatrix} = \begin{bmatrix}v_1\\v_2\\\vdots\\v_n\end{bmatrix}[1 \quad 1 \quad \dots \quad 1]$

$w = \begin{bmatrix}w_1\\w_2\\\vdots\\w_n\end{bmatrix}$ probability vector

$\lim_{m\to\infty} A^m w = Lw = (vu^T)w = v(u^T w) = v$

Markov Chain Example

October-28-11 9:30 AM

Example: Hardy-Weinberg Law

A certain gene has a dominant form G and a recessive form g. Each individual has either GG, Gg, or gg. At time 0, the probability distribution of these types is (p_0, q_0, r_0).

Assume:

1) The distribution is the same for both sexes
2) This gene does not affect reproductive capability

p_0 of time, father is GG. Probabilities for offspring in terms of mother's type:

$$\begin{array}{ccc} GG & Gg & gg \end{array}$$
$$\begin{bmatrix} 1 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

q_0 of time, father is Gg. Probability of offspring is

$$\begin{array}{ccc} GG & Gg & gg \end{array}$$
$$\begin{bmatrix} \frac{1}{2} & \frac{1}{4} & 0 \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{4} & \frac{1}{2} \end{bmatrix}$$

r_0 of time, father is gg. Probability of offspring is

$$\begin{bmatrix} 0 & 0 & 0 \\ 1 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 1 \end{bmatrix}$$

Total probability:

$$p_0\begin{bmatrix} 1 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 1 \\ 0 & 0 & 0 \end{bmatrix} + q_0\begin{bmatrix} \frac{1}{2} & \frac{1}{4} & 0 \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{4} & \frac{1}{2} \end{bmatrix} + r_0\begin{bmatrix} 0 & 0 & 0 \\ 1 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 1 \end{bmatrix} = \begin{bmatrix} p_0 + \frac{1}{2}q_0 & \frac{1}{2}p_0 + \frac{1}{4}q_0 & 0 \\ \frac{1}{2}q_0 + r_0 & \frac{1}{2}p_0 + \frac{1}{2}q_0 + \frac{1}{2}r_0 & p_0 + \frac{1}{2}q_0 \\ 0 & \frac{1}{4}q_0 + \frac{1}{2}r_0 & \frac{1}{2}q_0 + r_0 \end{bmatrix} = M$$

Let $\alpha_0 = p_0 + \frac{1}{2}q_0$, $\beta_0 = \frac{1}{2}q_0 + r_0$

$$M = \begin{bmatrix} \alpha_0 & \frac{1}{2}\alpha_0 & 0 \\ \beta_0 & \frac{1}{2} & \alpha_0 \\ 0 & \frac{1}{2}\beta_0 & \beta_0 \end{bmatrix}$$

To find the new probability distribution for the next generation, apply this to the probability distribution of females.

$$\begin{bmatrix} \alpha_0 & \frac{1}{2}\alpha_0 & 0 \\ \beta_0 & \frac{1}{2} & \alpha_0 \\ 0 & \frac{1}{2}\beta_0 & \beta_0 \end{bmatrix}\begin{bmatrix} p_0 \\ q_0 \\ r_0 \end{bmatrix} = \begin{bmatrix} \alpha_0\left(p_0 + \frac{1}{2}q_0\right) \\ \beta_0 p_0 + \frac{1}{2}q_0 + \alpha_0 r_0 \\ \beta_0\left(\frac{1}{2}q_0 + r_0\right) \end{bmatrix} = \begin{bmatrix} \alpha_0^2 \\ 2\alpha_0\beta_0 \\ \beta_0^2 \end{bmatrix}$$

Get a new transition matrix for a new generation (by applying the above with $\begin{bmatrix} \alpha_0^2 \\ 2\alpha_0\beta_0 \\ \beta_0^2 \end{bmatrix}$, substituted for $\begin{bmatrix} p_0 \\ q_0 \\ r_0 \end{bmatrix}$.

$$\alpha_1 = p_1 + \frac{1}{2}q_1 = \alpha_0^2 + \frac{1}{2}2\alpha_0\beta_0 = \alpha_0(\alpha_0 + \beta_0) = \alpha_0$$
$$\beta_1 = r_1 + \frac{1}{2}q_1 = \beta_0^2 + \alpha\beta = \beta_0$$

So the new transition matrix

$$\begin{bmatrix} \alpha_1 & \frac{1}{2}\alpha_1 & 0 \\ \beta_1 & \frac{1}{2} & \alpha_1 \\ 0 & \frac{1}{2}\beta_1 & \beta_1 \end{bmatrix} = \begin{bmatrix} \alpha_0 & \frac{1}{2}\alpha_0 & 0 \\ \beta_0 & \frac{1}{2} & \alpha_0 \\ 0 & \frac{1}{2}\beta_0 & \beta_0 \end{bmatrix}$$

$\therefore$ system is Markov.

In 2nd generation, new probabilities:

$$\begin{bmatrix} p_2 \\ q_2 \\ r_2 \end{bmatrix} = \begin{bmatrix} \alpha_0 & \frac{1}{2}\alpha_0 & 0 \\ \beta_0 & \frac{1}{2} & \alpha_0 \\ 0 & \frac{1}{2}\beta_0 & \beta_0 \end{bmatrix}\begin{bmatrix} \alpha_0 \\ 2\alpha_0\beta_0 \\ \beta_0^2 \end{bmatrix} = \begin{bmatrix} \alpha_0^3 + \alpha_0^2\beta_0 \\ \alpha_0^2\beta_0 + \alpha_0\beta_0 \\ \alpha_0\beta_0^2 + \beta_0^3 \end{bmatrix} = \begin{bmatrix} \alpha_0^2(\alpha_0 + \beta_0) \\ \alpha_0\beta_0(\alpha_0 + 1 + \beta_0) \\ \beta_0^2 \end{bmatrix} = \begin{bmatrix} \alpha_0^2 \\ 2\alpha_0\beta_0 \\ \beta_0^2 \end{bmatrix} = \begin{bmatrix} p_1 \\ q_1 \\ r_1 \end{bmatrix}$$

Stabilizes after 1 generation.

Inner Product Space

October-28-11 9:55 AM

Inner Product

An inner product on a vector space V over $\mathbb{F} = \mathbb{C}$ or $\mathbb{R}$ is a function $\langle *,* \rangle: V \times V \to \mathbb{F}$ s.t.

1. $\langle \alpha v + \beta w, u \rangle = \alpha \langle v, u \rangle + \beta \langle w, u \rangle$
   Linear in first variable
2. $\langle v, w \rangle = \overline{\langle w, v \rangle}$
3. $\langle v, v \rangle > 0 \; if \; v \neq 0$
   Positive Definite

$2 \Rightarrow$
$\langle u, \alpha v + \beta w \rangle = \bar{\alpha} \langle u, v \rangle + \bar{\beta} \langle u, w \rangle$

Norm

The norm on $(V, \langle , \rangle)$ is $\|v\| = \sqrt{\langle v, v \rangle}$

Theorem

$v, u \in V, \alpha \in \mathbb{F}$

1) $\|\alpha v\| = |\alpha| \|v\|$
2) $\|v\| \geq 0, \|v\| = 0 \Leftrightarrow v = 0$
3) Cauchy-Schwarz inequality
   $|\langle u, v \rangle| \leq \|u\| \cdot \|v\|$
   Equality $\Leftrightarrow u, v$ collinear
4) Triangle inequality
   $\|u + v\| \leq \|u\| + \|v\|$
   Equality $\Rightarrow u, v$ collinear

Conjugate in 2nd Variable

$2 \Rightarrow$
$< u, \alpha v + \beta w > = \overline{< \alpha v + \beta w, u >} = \overline{\alpha < v, u > + \beta < w, u >} = \bar{\alpha}\, \overline{< v, u >} + \bar{\beta}\, \overline{< w, u >}$
$= \bar{\alpha} < u, v > + \bar{\beta} < u, w >$
Conjugate linear in second variable.
Sesquilinear form (1/2 linear)

Examples

1) $V = \mathbb{C}^n, < (x_i), (y_i) > = \sum_{i=1}^{n} x_i \overline{y_i}$
2) $V = \mathbb{R}^n, < (x_i), (y_i) > = \sum_{i=1}^{n} x_i y_i \quad (dot\ product)$
3) $V = \mathbb{C}^2, \; < \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} > = x_1 \overline{y_1} - x_1 \overline{y_2} - x_2 \overline{y_1} + 3 x_2 \overline{y_2}$
   Check properties:
   1. Linear in 1st variable
   2. Symmetric
   3. $< \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \geq |x_1|^2 - x_1 \overline{x_2} - x_2 \overline{x_1} + 3|x_2|^2 = |x_1 - x_2| \overline{|x_1 - x_2|} + 2|x_2|^2$
      $= |x_1 - x_2|^2 + 2|x_2|^2 \geq 0$
      And equals 0 iff $x_1, x_2 = 0$, So positive definite.
4) $V = C[0,1]$ (Continuous functions from [0,1] to [0,1])
   $< f, g > = \int_0^1 f(x) \overline{g(x)} dx$

   1. Linear in 1st variable
   2. Symmetric
   3. $< f, f > = \int_0^1 |f(x)|^2 dx$
      $If \; f \neq 0, f(x_0) \neq 0$ by continuity $|f(x)| \geq \delta > 0$ on $(x_0 - r, x_0 + r)$
      $\therefore \int |f(x)|^2 dx \; \geq \int_{x_0 - r}^{x_0 + r} \delta^2 dx > 0$

Proof of Theorem

1,2 easy

3. wlog $v \neq 0$.
   $0 \leq \|u + \alpha v\|^2 = \langle u + \alpha v, u + \alpha v \rangle = \langle u, u \rangle + \alpha \langle v, u \rangle + \bar{\alpha} \langle u, v \rangle + |\alpha|^2 \langle v, v \rangle$
   Take $\alpha = t \langle u, v \rangle, \; t \in \mathbb{R}$
   $= \langle u, u \rangle + t |\langle u, v \rangle|^2 + t |\langle u, v \rangle|^2 + t^2 |\langle u, v \rangle|^2 \|v\|$
   Quadratic; minimized if $t = \frac{1}{m^2}$, setting $t = \frac{1}{\|v\|^2}$
   $0 \leq \|u\|^2 - \frac{2}{\|v\|^2} |\langle u, v \rangle|^2 + \frac{|\langle u, v \rangle|^2 \|v\|^2}{\|v\|^4} = \|u\|^2 - \frac{|\langle u, v \rangle|^2}{\|v\|^2}$
   $\therefore |\langle u, v \rangle|^2 \leq \|u\|^2 \|v\|^2$
   Equality $\Rightarrow 0 = \left\| u - \frac{<u,v>}{\|v\|^2} v \right\|^2 \; \Rightarrow$ u is a multiple of v
4. $\|x + y\|^2 = < x + y, x + y > = < x, x > + < x, y > + < y, x > + < y, y >$
   $= \|x\|^2 + 2Re(< x, y >) + \|y\|^2 \leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2$
   $equality \Leftrightarrow x, y \; collinear \; and < x, y > \geq 0$

Example

$$\left| \sum_{i=1}^{n} x_i \overline{y_i} \right| \leq \left(\sum |x_i|^2 \right) \left(\sum |y_i|^2 \right)$$

Example

$$\left| \int_0^1 f(x) g(x) dx \right| \leq \left(\int_0^1 |f(x)|^2 dx \right) \left(\int_0^1 |g(x)|^2 dx \right)$$

Orthogonality

October-31-11 9:35 AM

Orthogonal

Say u is orthogonal to v $(u \perp v)$ if $\langle u, v\rangle = 0$

Orthonormal

A set $\{e_i\}_{i\in I}$ is orthonormal if

$$\langle e_i, e_j\rangle = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$$

If $M \subseteq V$, let $M^{\perp} = \{v \in V : \langle v, m\rangle = 0 \ \forall m \in M\}$

Remarks

1. If $u \perp v$, then $\|u+v\|^2 = <u+v, u+v> = \|u\|^2 + 2Re <u,v> + \|v\|^2 = \|u\|^2 + \|v\|^2$
   Pythagorean Law
2. $M^{\perp}$ is a subspace
   If $u, v \in M^{\perp}, \alpha, \beta \in \mathbb{C}, m \in M$
   $\langle \alpha u + \beta v, m\rangle = \alpha\langle u, m\rangle + \beta\langle v, m\rangle = 0$

Lemma

Let $\{e_1, \dots, e_n\}$ be an orthonormal (o.n.) set, and $x \in span\,\{e_1, \dots, e_n\}$ then

$$x = \sum_{i=1}^{n} \langle x, e_i\rangle e_i = \sum_{i=1}^{n} \alpha_i e_i$$

If $y \in \sum_{i=1}^{n} \beta_i e_i$, then $\langle x, y\rangle = \sum_{i=1}^{n} \alpha_i \bar{\beta_i}$

and $\|x\| = \sqrt{\sum_{i=1}^{n} |a_i|^2}$

Note

If $\{e_1, \dots, e_n\}$ are orthonormal, and $v \in V$, then

$$v - \sum_{i=1}^{n} \langle v, e_i\rangle e_i \perp sp\{e_1, \dots, e_n\}$$

Gram-Schmidt Process

Start with a set of vectors $\{v_1, v_2, \dots, v_m\}$
Build an o.n. set with the same span.

1. Throw out v_j if $v_j \in sp\{v_1, \dots, v_{j-1}\}$
   So wlog $\{v_1, \dots, v_m\}$ is independent
2. Let $e_1 = \frac{v_1}{\|v_1\|}$
   Let $e_2 = \dfrac{v_2 - \langle v_2, e_1\rangle e_1}{\|v_2 - \langle v_2, e_2\rangle e_1\|}$
   ...
3. If $e_1, \dots, e_{k-1}$ are defined and $o.n.$ Let
   $e_k = \dfrac{v_k - \sum_{i=1}^{k-1}\langle v, e_i\rangle e_i}{\left\|v_k - \sum_{i=1}^{k-1}\langle v, e_i\rangle e_i\right\|}$

$span\{e_1, \dots, e_k\} = span\{v_1, \dots, v_k\}$

Lemma

If $\{e_i\}$ are orthonormal, then they are linearly independent.

Lemma

Every finite dimensional subspace $M \subseteq V$ has an orthonormal basis.

Proof of Lemma

Write $x = \sum_{i=1}^{n} \alpha_i e_i$

$$\langle x, e_j\rangle = \left\langle \sum a_i e_i, e_j \right\rangle = \sum_{i=1}^{n} a_i \langle e_i, e_j\rangle = \alpha_j$$

$$\langle x, y\rangle = \left\langle \sum_{i=1}^{n} \alpha_i e_i, \sum_{j=1}^{n} \beta_j e_j \right\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i \bar{\beta_j} \langle e_i, e_j\rangle = \sum_{i=1}^{n} \alpha_i \bar{\beta_i}$$

Example

$H = C[0,1]$ with

$$\langle f, g\rangle = \int_0^1 f(x)\overline{g(x)}dx$$

Let $e_n(x) = e^{2\pi i n x}, n \in dZ$

$$\langle e_n, e_m\rangle = \int_0^1 e^{2\pi inx}\overline{e^{2\pi imx}}dx = \int_0^1 e^{2\pi i(n-m)x}dx$$

$$= \begin{cases} 1, & n = m \\ \dfrac{1}{2\pi i(n-m)} e^{2\pi i(n-m)x}\Big|_0^1 = 1 - 1 = 0, & n \neq m \end{cases}$$

So $\{e_n, n \in \mathbb{Z}\}$ is orthonormal
If $c \in C[0,1]$ get a series

$$\sum_{n=-\infty}^{\infty} \langle f, e_n\rangle e_n = \sum_{-\infty}^{\infty} \hat{f}(n) e^{e\pi in}$$

Fourier Series

Proof of Lemma

If $0 = \sum_{i=1}^{n} a_i e_i$

then $0 = \|0\| = \left\|\sum_{i=1}^{n} a_i e_i\right\| = \sum_{i=1}^{n} |a_i|^2$

$\therefore a_i = 0 \forall i$

Proof of Lemma

Take a basis $\{v_1, \dots, v_k\}$ for M and apply the Gram-Schmidt Process to get an orthonormal basis.

Proof of Theorem(Projection)

1.
$ran\ P = sp\{e_1, \dots, e_n\} = M$
$\ker P = \{v : \langle v, e_i\rangle = 0\ for\ 1 \le i \le n\} = \{e_1, \dots, e_n\}^{\perp}$
$= (sp\{e_1, \dots, e_n\})^{\perp} = M^{\perp}$

If $w \in M, w = \sum_{i=1}^{n} a_i e_i$

$$Pw = \sum \langle w, e_i\rangle e_i = \sum_{i=1}^{n} a_i e_i = w$$

$P^2 v = P(Pv) = Pv$
$\therefore$Projection onto M

2.
$v \in V, Pv \in M$
$\langle v - Pv, e_i\rangle = 0\ for\ 1 \le i \le n$
$\therefore v - Pv \in M^{\perp}$
$v = Pv + (v - Pv)$
$\|v\|^2 = \|Pv\|^2 + \|v - Pv\|^2$ (Pythagorean)
Suppose $m \in M$
$v - m = (Pv - m) + (v - Pv)$
$\therefore \|v - m\|^2 = \|Pv - m\|^2 + \|v - Pv\|^2 \ge \|v - Pv\|^2$
equality $\Leftrightarrow m = Pv$
$\therefore Pv$ is the unique closest point
$\therefore Pv$ is the only projection onto M because Pv =the closest point on M ■

Lemma
Every finite dimensional subspace $M \subseteq V$ has an orthonormal basis.

Projection
V inner product space. $P \in \mathcal{L}(V)$ is a projection if $P = P^2$ (idempotent) s.t. $\ker P \perp ran\ P$

Theorem (Projection)
Let M be a finite dimensional subspace of V with orthonormal basis $\{e_1, \dots, e_n\}$. Define $P \in \mathcal{L}(V)$ by

$$Pv = \sum_{i=1}^{n} \langle v, e_i \rangle e_i$$

Then:

1) P is the projection of V onto M
   (i.e. $ran\ P = M, \ker P = M^{\perp}, P = P^2$)
2) $v \in V, \|v\|^2 = \|Pv\|^2 + \|v - Pv\|^2$
3) Pv is the unique closest point in M closest to v

Corollary - Bessel's Inequality
If V is an inner product space and $\{e_n : n \in S\}$ is orthonormal then

$$\sum_{n \in S} |\langle v, e_n \rangle|^2 \leq \|v\|^2 \quad \forall v \in V$$

Corollary
$f \in C[0,1], \quad \{e^{2\pi i n x} : n \in \mathbb{Z}\}\ orthonormal$

So if $a_n = \int_0^1 f(x) \overline{e^{2\pi i n x}} dx$

then $\sum_{h=-\infty}^{\infty} |a_n|^2 \leq \int_0^1 |f(x)|^2 dx$

$\therefore Pv$ is the unique closest point
$\therefore Pv$ is the only projection onto M because $Pv =$the closest point on M■
$I - P$ is written $P^{\perp}$ and $P^{\perp}$ is the projection onto $M^{\perp}$

Proof of Corollary
If S is finite, not problem
Let $M = sp\{e_n : n \in S\}$

$$Pv = \sum_{n \in S} \langle v, e_n \rangle e_n$$

and $\|v\|^2 \geq \|Pv\|^2 = \sum_{n \in S} |\langle v, e_n \rangle|^2$

If S is infinite for each finite $F \subseteq S$ let $M_F = sp\{e_n, n \in F\}$
$P_F\ , projction\ onto\ M_F$
Then $\|v\|^2 \geq \|P_F v\|^2 = \sum_{n \in F} |\langle v, e_n \rangle|^2$

$\therefore \|v\|^2 \geq \sup_{F \subseteq S, finite} \sum_{n \in F} |\langle v, e_n \rangle|^2 = \sum_{n \in S} |\langle v, e_n \rangle|^2$
At most $\|v\|^2$ coefficients $\langle v, e_n \rangle$ have $|\langle v, e_n \rangle| \geq 1$
Otherwise $\exists$finite $N > \|v\|^2$ and $|F| = N$ s.t. $|\langle v, e_n \rangle| \geq 1,\ n \in F$

$$\Rightarrow \sum_{n \in F} |\langle v, e_n \rangle|^2 = N > \|v\|^2$$

At most $4^k \|v\|^2$ coefficients with $|\langle v, e_n \rangle| \geq \frac{1}{2^k}$

$$F_k = \left\{ n : |\langle v, e_n \rangle| \geq \frac{1}{2^k} \right\}$$

$$\|v\|^2 \geq \sum_{F_k} |\langle v, e_n \rangle|^2 \geq \frac{|F_k|}{4^k}$$

$\therefore |F_k| \leq 4^k \|v\|^2$

So $\{n : \langle v, e_n \rangle \neq 0\} = \bigcup_{k \geq 0} \{k : |\langle v, e_n \rangle| \geq 2^{-k}\}$
Is countable
List them $n_1, n_2, n_3, \dots$

$$\sum_{i=1}^{\infty} \left|\langle v, e_{n_i} \rangle\right|^2 = \lim_{k \to \infty} \sum_{i=1}^{k} \left|\langle v, e_{n_i} \rangle\right|^2$$

$$\therefore \sum_{n \in S} |\langle v, e_n \rangle|^2 \leq \|v\|^2$$

Canonical Forms in Inner Product Spaces

November-02-11 9:33 AM

Theorem

If V is a complex inner product space, $\dim V < \infty, T \in \mathcal{L}(V)$.
Then there is an orthonormal basis $\beta = \{e_1, \dots, e_n\}$ such that $[T]_\beta$ is upper triangular.

Adjoint

V inner product space, $T \in \mathcal{L}(V)$
The adjoint of T is the linear map T^* such that
$\langle T^*v, w\rangle = \langle v, Tw\rangle\ \forall v, w \in V$

Fix an orthonormal basis $\xi = \{e_1, \dots, e_n\}$
$[T]_\xi = [t_{ij}]_{n\times n}$
$t_{ij} = \langle Te_j, e_i\rangle$
Then $[T^*]_\xi = [\overline{t_{ji}}]_{n\times n}$

Proposition

If $S, T \in \mathcal{L}(V)$ then

1) $(S^*)^* = S$
2) $(\alpha S + \beta T)^* = \bar{\alpha}S^* + \bar{\beta}T^*$
3) $I^* = I$
4) $(ST)^* = T^*S^*$

Hermitian (Self-Adjoint)

$T \in \mathcal{L}(V)$ is Hermitian or self-adjoint if $T = T^*$

If $T = [t_{ij}] = T^* = [\overline{t_{ji}}]$
Then $\overline{t_{ji}} = t_{ij}$ and $t_{ii} = \overline{t_{ii}} \in \mathbb{R}$

If we check that $[T]_\beta = [T^*]_\beta$ then it has $[T]_\xi = [T^*]_\xi$ on every basis.

Reason:

$T = T^* \Leftrightarrow \langle Tu, v\rangle = \langle u, Tv\rangle\ \forall u, v \in V$
This is basis independent.

Theorem

If $T \in \mathcal{L}(V)$, V finite and a $\mathbb{C}$ inner product space, and $T = T^*$, then there is an orthonormal basis ξ such that
$$[T]_\xi = \begin{bmatrix} d_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & d_n \end{bmatrix} \text{ is diagonal with } d_i \in \mathbb{R}$$
So $\sigma(T) \subseteq \mathbb{R}$ and $\ker(T - \lambda_i I) \perp \ker(T - \lambda_j I)$ if $\lambda_i \neq \lambda_j \in \sigma(T)$

Corollary

If V is a finite $\mathbb{R}$ —inner product space. $T \in \mathcal{L}(V)$ s.t. $T = T^*$ then there is an orthonormal basis ξ such that
$$[T]_\xi = \begin{bmatrix} d_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & d_n \end{bmatrix} \text{ is diagonal}$$

Proof of Theorem

Since $\mathbb{C}$ is algebraically closed, $p_T(x)$ splits into linear terms. Hence there is a basis $\{v_1, \dots, v_n\}$ such that T is upper triangular with respect to $\{v_i\}$

Apply Gram-Schmidt process to $\{v_1, \dots, v_n\}$ to an orthonormal basis $\{e_1, \dots, e_n\}$
$Tv_1 = t_{11}v_1$
Since $e_1 = \dfrac{v_1}{\|v_1\|}, Te_1 = t_{11}e_1$
$T_2v_2 = t_{22}e_2 + t_{12}e_1$
$e_2 = \dfrac{v_2 - \langle v_2, e_1\rangle e_1}{\|v_2 - \langle v_2, e_1\rangle e_1\|} = a_1v_1 + a_2v_2$
$Te_2 = a_1Tv_1 + a_2Tv_2 \in sp\{v_1, v_2\}$

T upper Δ with respect to $\{v_1, \dots, v_n\}$ means $M_k = sp\{v_1, v_2, \dots, v_k\}$ is invariant for T
But $span\{e_1, \dots, e_k\} = span\{v_1, \dots, v_k\}$
$\therefore Te_k \in M_k \left(i.e. Te_k = \sum_{i=1}^{k} b_{ik}e_i\right)$
So $[T]_\beta$ is upper triangular.

What is T*?

Fix an orthonormal basis $\xi = \{e_1, \dots, e_n\}$
$[T]_\xi = [t_{ij}]_{n\times n}$
$Te_j = \sum_{i=1}^{n} t_{ij}, e_i \Rightarrow \langle Te_j, e_i\rangle = t_{ij}$
$\langle T^*e_j, e_i\rangle = \langle e_j, Te_i\rangle = \overline{\langle Te_i, e_j\rangle} = \overline{t_{ji}}$
So $[T^*]_\xi = [\overline{t_{ji}}]$
Conjugate transpose of T

So we can define a linear transformation
$T^* \in \mathcal{L}(V)$ with $[T^*]_\xi = [\overline{t_{ji}}]$
Need to check that the identity holds for all vectors $v, w \in T$
Take $v = \sum_{i=1}^{n} \alpha_i e_i, \quad w = \sum_{j=1}^{n} \beta_j e_j$
Calculate
$$\langle T^*v, w\rangle = \left\langle T^*\sum_{i=1}^{n}\alpha_i e_i, \sum_{j=1}^{n}\beta_j e_j\right\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\bar{\beta}_j\langle T^*e_i, e_j\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\bar{\beta}_j\overline{\langle Te_j, e_i\rangle}$$
$$= \overline{\sum_{i=1}^{n}\sum_{j=1}^{n}\bar{\alpha}_i b_j\langle Te_j, e_i\rangle} = \overline{\left\langle T\sum_{j=1}^{n}\beta_j e_j, \sum_{i=1}^{n}\alpha_i e_i\right\rangle} = \overline{\langle Tw, v\rangle} = \langle v, Tw\rangle$$
So T^* is a well defined linear map.

Proof of Proposition

1.
Fix an orthonormal basis ξ
$[S]_\xi = [s_{ij}]$
$[S^*]_\xi = [\overline{s_{ji}}]$
$[S^{**}]_\xi = [\overline{\overline{s_{ij}}}] = [S]_\xi$

2.
$[\alpha S]_\xi = [\alpha s_{ij}]$
$[(\alpha S)^*]_\xi = [\overline{\alpha s_{ji}}] = \bar{\alpha}[\overline{s_{ji}}] = \alpha[S^*]_\xi$

$[T]_\xi = [t_{ij}]$
$[\alpha S + \beta T]_\xi = [\alpha s_{ij} + \beta t_{ij}]_\xi$
$[(\alpha S + \beta T)^*]_\xi = [\overline{\alpha S_{ji} + \beta t_{ji}}] = \bar{\alpha}[\overline{s_{ji}}] + \bar{\beta}[\overline{t_{ji}}] = \bar{\alpha}[S^*]_\xi + \bar{\beta}[T^*]_\xi$

3.
$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & 1 \end{bmatrix} = I^*$$

4.
$S = [s_{ij}]_{n\times n}, \quad T = [t_{ij}]_{n\times n}$
$S^* = [\overline{s_{ji}}], \quad T = [\overline{t_{ji}}]$
$ST = \left[\sum_{k=1}^{n} s_{ik}t_{kj}\right]_{n\times n}$
$\therefore (ST)^* = \left[\sum_{k=1}^{n} \overline{s_{jk}t_{ki}}\right]$

$$T^*S^* = \left[\sum_{k=1}^{n} \overline{t_{ki}}\,\overline{s_{jk}}\right] = (ST)^* \blacksquare$$

Proof of Theorem

Since V is a $\mathbb{C}$-vector space there is an orthonormal basis ξ such that $[T]_\xi$ is upper triangular.

$$[T]_\xi = \begin{bmatrix} t_{11} & \cdots & t_{1n} \\ \vdots & \ddots & \vdots \\ 0 & \cdots & t_{nn} \end{bmatrix} = [T^*]_\xi = \begin{bmatrix} \overline{t_{11}} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ \overline{t_{n1}} & \cdots & \overline{t_{nn}} \end{bmatrix}$$

If $i < j, t_{ij} = 0$ If $i = j, t_{ii} = \overline{t_{ii}} \in \mathbb{R}$

$$\therefore [T]_\xi = \begin{bmatrix} t_{11} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & t_{nn} \end{bmatrix}, t_{ii} \in \mathbb{R}$$

$\sigma(T) = \{t_{ii} : 1 \le i \le n\} \subseteq \mathbb{R}$

$\ker(T - \lambda_i I) = sp\{e_j : t_{jj} = \lambda_i\}$ are pairwise orthogonal. ■

Proof of Corollary

Fix an orthonormal basis β, $T = [t_{ij}]_\beta = [t_{ji}]_\beta$

Think of T as acting on $\mathbb{C}^n$

$T = T^*$ so by Theorem $p_T(x) = \prod_{i=1}^{n}(x - \lambda_i)$ and $\lambda_i \in \mathbb{R}$

So p_T splits in $\mathbb{R}[x]$

∴T is triangularizable over $\mathbb{R}$ $\exists \zeta$ $s.t.$ $[T]_\zeta$ is upper triangular

Apply Gram-Schmidt to basis to get an orthonormal basis ξ and $[T]_\xi$ is upper Triangular and self adjoint, so the same argument shows $[T]_\xi$ is diagonal. ■

Unitary Maps

November-04-11

Unitary and Orthogonal Maps

V, W $\mathbb{C}$- inner product spaces.
$U \in \mathcal{L}(V, W)$ is called **unitary** iff it is invertible and preserves inner product: $\langle Uv_1, Uv_2\rangle_W = \langle v_1, v_2\rangle_V$

If V, W are $\mathbb{R}$-inner product spaces, call such a map **orthogonal**.

Theorem

If $\dim V = \dim W < \infty,\ U \in \mathcal{L}(V, W)$, TFAE

1) U is unitary
2)
   a. U preserves inner product
   b. U is isometric (preserves norm)
3)
   a. U sends every orthonormal basis of V to an orthonormal basis for W
   b. U sends some orthonormal basis of V to an orthonormal basis of W

Remark

If $V = \mathbb{C} = sp\{e_1\}, W = \mathbb{C}^2 = sp\{f_1, f_2\}$
$T(\alpha e_1) = \alpha f_1$ preserves inner product but not onto so not invertible.

Proposition

$U \in \mathcal{L}(V, W)$ is unitary $\Leftrightarrow$
$U^*U = I_V$ and $UU^* = I_W \Leftrightarrow$
$U^{-1} = U^*$

Unitarily Equivalent

Say two transformations $S \in \mathcal{L}(V)$ and $T \in \mathcal{L}(W)$ are **unitarily equivalent** iff $\exists$ unitary $U \in \mathcal{L}(V, W)$ s.t. $T = USU^{-1} = USU^*$

Corollary

If T is self-adjoint ($T = T^*$) then $T \cong D$ (T unitarily equivalent to D) where D is diagonalizable with real entries.

Just a restatement of theorem that T is diagonalizable with respect to an orthonormal basis $\{f_1, \ldots, f_n\}$ say $Tf_i = d_i f_i,\ d_i \in \mathbb{R}$

Say $T = [t_{ij}]$ in $\{e_1, \ldots, e_n\}$ orthonormal basis. Let $Ue_i = f_i$ $1 \le i \le n$
Then U is unitary (takes one orthonormal basis to another) and
$(U^*TU)e_i = U^*Tf_i = U^*d_i f_i = d_i e_i$
($U^* = U^{-1}$, so $U^* f_i = e_i$)
$\therefore D = U^*TU = diag(d_1, d_2, \ldots, d_n)$

Proof of Theorem

$1 \Rightarrow 2a$ By definition

$2a \Rightarrow 2b$
$\|Uv\|^2 = \langle Uv, Uv\rangle = \langle v, v\rangle = \|v\|^2$

$2b \Rightarrow 2a$
Assignment 5 #5a
$\langle Uv_1, v_2\rangle = \frac{1}{4}(\|v_1 + v_2\|^2 - \|v_1 - v_2\|^2 + i\|v_1 + iv_2\|^2 - i\|v_1 - iv_2\|^2)$

$2a \Rightarrow 3a$
If $\{e_1, \ldots, e_n\}$ is an orthonormal basis for V, Let $f_i = Ue_i$
$\langle f_i, f_j\rangle = \langle Ue_i, U_j\rangle = \langle e_i, e_j\rangle = \delta_{ij} \therefore \{f_i\}$ is orthonormal
Since $\dim W = \dim V$, $\{f_i\}$ is an orthonormal basis.

$3a \Rightarrow 3b$ Obvious

$3b \Rightarrow a$
Let $\{e_1, \ldots, e_n\}$ be an orthonormal basis such that $f_i = Ue_i$ is an orthonormal basis for W.
U takes a basis for V to a basis for W $\therefore$U is invertible
Let $v_1 = \sum \alpha_i e_i, v_2 = \sum \beta_j e_j$

$\langle v_1, v_2\rangle = \sum_{i=1}^{n} \alpha_i \overline{\beta_i}$

$Uv_1 = \sum \alpha_i f_i, \quad Uv_2 = \sum \beta_j f_j$

$\therefore \langle Uv_1, Uv_2\rangle = \left\langle \sum \alpha_i f_i, \sum \beta_j f_j \right\rangle = \sum \alpha_i \overline{\beta_i} = \langle v_1, v_2\rangle$

So it preserves inner product. $\therefore$ U is unitary ■

Proof of Proposition

3nd and 2rd statements are clearly equivalent.
$\Rightarrow$
Let $v_1, v_2 \in V, w_i = Uv_i$
$\langle v_1, U^*w_2\rangle = \langle Uv_1, w_2\rangle = \langle Uv_1, Uv_2\rangle = \langle v_1, v_2\rangle = \langle v, U^{-1}w_2\rangle$
$\langle v_1, U^*w_2 - U^{-1}w_2\rangle = 0\ \forall v_1 \in V$
$\therefore U^*w_2 = U^{-1}w_2, \forall w_2 \in UV = V\ i.e.\ U^* = U^{-1}$
$\Leftarrow$
U is invertible and
$\langle Uv_1, Uv_2\rangle = \langle U^*Uv_1, v_2\rangle = \langle v_1, v_2\rangle$ preserves $\langle , \rangle$ ■

Normal Maps

November-07-11 9:40 AM

Definition

$N \in \mathcal{L}(V)$ is normal if $N^*N = NN^*$

Theorem

$T \in \mathcal{L}(V)$ is normal $\Leftrightarrow$
There is an orthonormal basis which diagonalizes T.

Corollary

If T is normal and

$\sigma(T) = \{\lambda_1, \dots, \lambda_s\}$ then $m_T(x) = \prod_{i=1}^{s}(x - \lambda_i)$

and $V_i = \ker(T - \lambda_i I)$ are pairwise orthogonal

Corollary

If U is unitary, then
$\sigma(Y) \subseteq \mathbb{T} = \{\lambda : |\lambda| = 1\}$
and U is diagonalizable w.r.t. some o.n. basis.

Corollary

If N is normal $\sigma(N) = \{\lambda_1, \dots, \lambda_s\}$ and $V_i = \ker(N - \lambda_i I)$
The idempotent $E_i \in \mathcal{A}(N)$ onto V_i is the orthonormal projection of V onto V_i. Moreover $N = \sum_{i=1}^{s} \lambda_i E_i$

Corollary

If p is a polynomial, N normal write $N = \sum_{i=1}^{s} \lambda_i E_i$, E_i as above

Then $p(N) = \sum_{i=1}^{s} p(\lambda_i) E_i$

Rank 1 Matrices

Suppose $T \in \mathcal{L}(V, W)$ and $rank(T) = 1$
Let $K = \ker T \subseteq V$
$\mathrm{n} = \dim V = nul\ (T) + rank\ (T) = \dim K + 1$
$\therefore \dim K = n - 1$

Pick a unit vector $e \in V, e \perp K$. Let $w = Te\ (\neq 0\ since\ e \neq K)$
$V = K \oplus K^\perp = K \oplus \mathbb{F}e$
If $v \in V, v = k + \lambda e,\ k \in K, \lambda \in \mathbb{F}$

$Tv = T(k + \lambda e) = \lambda Te = \lambda w$

Think of $e = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}$ as a $n \times 1$ matrix

So $e \in \mathcal{L}(\mathbb{F}, V)$ by $e(\lambda) = \lambda e$
$e^* = [\overline{\alpha_1}, \overline{\alpha_2}, \dots, \overline{\alpha_n}] \in \mathcal{L}(V, \mathbb{F})$ is a $1 \times n$ matrix

If $v \in V, \quad v = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$

$$e^*v = [\overline{\alpha_1}, \overline{\alpha_2}, \dots, \overline{\alpha_n}] \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \sum_{i=1}^{n} \bar{\alpha}_i v_i = \langle v, e \rangle$$

$e^*(k + \lambda e) = 0 + \lambda \|e\|^2 = \lambda$

$$we^* = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} [\overline{\alpha_1}, \overline{\alpha_2}, \dots, \overline{\alpha_n}] = \begin{bmatrix} w_1\overline{\alpha_1} & w_1\overline{\alpha_2} & \dots & w_1\overline{\alpha_n} \\ w_2\overline{\alpha_1} & w_2\overline{\alpha_2} & \dots & w_2\overline{\alpha_n} \\ \vdots & \vdots & \ddots & \vdots \\ w_n\overline{\alpha_1} & w_n\overline{\alpha_2} & \dots & w_n\overline{\alpha_n} \end{bmatrix}$$

$we^* \in \mathcal{L}(\mathbb{F}, W) \cdot \mathcal{L}(V, \mathbb{F}) = \mathcal{L}(V, W)$
$(we^*)(k + \lambda e) = \lambda w = T(k + \lambda e)$
$T = we^* = Tee^*$

Example of Normal Maps

1. $T = T^*$ are normal $(TT = TT)$
2. Unitaries are normal $(U^*U = I = UU^*)$
3. If D is diagonal w.r.t an orthonormal basis
   $D = diag(d_1, d_2, \dots,), D^* = (\overline{d_1}, \overline{d_2}, \dots, \overline{d_n})$
   $D^*D = DD^* = diag(|d_1|^2, |d_2|^2, \dots, |d_n|^2)$

Proof of Theorem

$\Leftarrow$ Example 3
$\Rightarrow$ If T is normal then $\|Tx\| = \|T^*x\|\ \forall x \in V$ because:
$\|Tx\|^2 = \langle Tx, Tx \rangle = \langle x, T^*Tx \rangle = \langle x, TT^*x \rangle = \langle T^*x, T^*x \rangle = \|T^*x\|^2$

Choose an orthonormal basis $\{e_1, \dots, e_n\}$ so that $[T]_\beta$ is upper Δ

$$T = \begin{bmatrix} t_{11} & t_{12} & \dots & t_{1n} \\ 0 & t_{22} & \dots & t_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & t_{nn} \end{bmatrix}, T^* = \begin{bmatrix} \overline{t_{11}} & 0 & \dots & 0 \\ \overline{t_{12}} & \overline{t_{22}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \overline{t_{1n}} & \overline{t_{1n}} & \dots & \overline{t_{nn}} \end{bmatrix}$$

$\|Te_1\|^2 = \|t_{11}e_i\|^2 = |t_{11}|^2$

$$\|Te_1\|^2 = \|T^*e_1\|^2 = \|\overline{t_{11}}e_1 + \overline{t_{12}}e_2 + \dots + \overline{t_{1n}}e_n\|^2 = \sum_{j=1}^{n} |t_{1j}|^2 = |t_{11}|^2 + \sum_{j=2}^{n} |t_{ij}|^2$$

$\therefore t_{1j} = 0\ for\ 2 \leq j \leq n$

Repeat $\|Te_2\| = |t_{22}| = \|T^*e_2\| = \sqrt{\sum_{j=2}^{n} |t_{2j}|^2}$
$\therefore t_{2j} = 0\ 3 \leq j \leq n$
$\therefore T$ is diagonal ■

Proof of Corollary

Since T is diagonalizable wrt some basis, $m_T(x) = \prod(x - \lambda_i)$ has only simple roots.
Say $\{e_i\}_{i=1}^n$ orthonormal, $Te_i = d_i e_i$
$V_j = \ker(T - \lambda_j I) = sp\{e_i : d_i = \lambda_j\}$
$\therefore V_j$ are pairwise $\perp$

Proof of Corollary

U normal $\therefore$ diagonalizable
Say $Ue_i = d_i e_i,\ \{e_i\}$ orthonormal
$\|Ue_i\| = \|e_i\| = 1$
$\|Ue_i\| = |d_i|\|e_i\| = |d_i|$
$\therefore |d_i| = 1$

Proof of Corollary

E_i is the projection onto V_i

The range of E_i is V_i and

$ker(E_i) = \sum_{j \neq i} V_j = V_i^\perp$

$V_i = sp\{e_k : d_k = \lambda_i\}$

$\sum_{j \neq i} V_j = sp\{e_k : d_k \neq \lambda_i\} = V_i^\perp$

$NE_i = E_i N = \lambda_i E_i$

So $N = N\left(\sum_{i=1}^{s} E_i\right) = \sum_{i=1}^{s} \lambda_i E_i$

Example

Orthogonal projection on to $\mathbb{F}e$
$Te = e$ so

$T = ee^* = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} [\overline{\alpha_1} \quad \dots \quad \overline{\alpha_n}] = [\alpha_i \overline{\alpha_j}]$

Polar Decomposition

November-09-11 9:30 AM

Complex
$z \in \mathbb{C}, z = re^{i\theta}, \quad r = |z|, |e^{i\theta}| = 1$

Positive
$T \in \mathcal{L}(V), V$ $\mathbb{C}$ −vector space is **positive** if $T = T^*$ and $\sigma(T) \subseteq [0, \infty)$
Write $T \geq 0$

Proposition
If $T \in \mathcal{L}(V)$ then $T^*T \geq 0$

Square Root
T^*T can be diagonalized with orthonormal basis $\xi = \{e_1, e_2, \dots, e_n\}$
$[T^*T]_\xi = diag(d_1, d_2, \dots, d_n), \quad d_i \geq 0$
$\sqrt{d_i}$ the square root of d_i
$[A]_\xi = diag(\sqrt{d_1}, \sqrt{d_2}, \dots, \sqrt{d_n})$ and $A^2 = T^*T$
i.e. A is the square root of T^*T call this $|T|$ (absolute value of T)

Fact (Homework)
The square root of T^*T is unique

Want to write $T = U|T|$

Partial Isometry
A partial isometry is a map $U \in \mathcal{L}(V, W)$ such that $U|_{\ker U^\perp}$ is isometric (preserves norm)

Examples
$U: \mathbb{C}^2 \to \mathbb{C}^3$ by $U(x, y) = (x, y, 0)$
$U^*: \mathbb{C}^3 \to \mathbb{C}^2$ by $U(x, y, z) = (x, y)$ −not unitary
U unitary is a partial isometry

Proposition
$U \in \mathcal{L}(V, W)$ TFAE
1. U is a partial isometry
2. U^*U is a projection (onto $(\ker U)^\perp$)
3. UU^* is a projection (onto ran U)
4. $U = UU^*U$

Theorem (Polar Decomposition)
If $T \in \mathcal{L}(V, W)$ then there is a unique partial isometry U with $\ker U = \ker T$ such that $T = U|T|$ $\left(|T| = \sqrt{T^*T}\right)$

S-Numbers
The s-numbers of $T \in \mathcal{L}(V, W)$ are the eigenvalues of $|T|$ (including multiplicity) in decreasing order.

Geometry of how T acts
$|T| = diag(s_1, s_2, \dots, s_n)$ wrt $\{e_1, e_n\}$
If considering the action on a unit sphere, T stretches it onto an ellipsoid (axis length defined by s-numbers). U is a partial rotation in space.

Proof of Proposition
$(T^*T)^* = T^*T^{**} = T^*T$

If $T^*Tx = \lambda x, \|x\| = 1$
$\lambda = \langle \lambda x, x\rangle = \langle T^*Tx, x\rangle = \langle Tx, Tx\rangle = \|Tx\|^2 \geq 0$
$\therefore T^*T \geq 0$

Proof of Proposition
1⇒2
$\ker U \supseteq \ker U^*U$
$x \in \ker U^*U \Rightarrow 0 = \langle U^*Ux, x\rangle = \langle Ux, Ux\rangle = \|Ux\|^2$
$\therefore x \in \ker U, \ker U \subseteq \ker U^*U$
$\therefore \ker U = \ker U^*U$

If $x \perp \ker U$ then $\|Ux\| = \|x\|$
$\langle x, x\rangle = \|x\|^2 = \|Ux\|^2 = \langle Ux, Ux\rangle = \langle U^*Ux, x\rangle$

ran $(U^*U) \perp \ker U$ since $y \in \ker U$:
$\langle U^*Ux, y\rangle = \langle Ux, Uy\rangle = \langle Ux, 0\rangle = 0$

$x, y \in (\ker U)^\perp$
$\langle U^*Ux, y\rangle = \langle Ux, Uy\rangle = \langle x, y\rangle$ (because of isomorphic)
$U^*Ux \in (\ker U)^\perp$

Take orthonormal basis $\{e_1, \dots, e_k\}$ for $(\ker U)^\perp$
$\langle U^*Ue_i, e_j\rangle = \langle e_i, e_j\rangle = \delta_{ij}$
$\therefore U^*Ue_i = \sum \langle U^*Ue_i, e_j\rangle e_i = e_i$

$\therefore U^*Ux = x$ for $x \in (\ker U)^\perp$
$\therefore U^*U$ is the projection onto $(\ker U)^\perp$

$2 \Rightarrow 1$, if $x \in (\ker U)^\perp$
$\|Ux\|^2 = \langle Ux, Ux\rangle = \langle U^*Ux, x\rangle = \langle x, x\rangle = \|x\|^2$

$1 \Rightarrow 3$
Claim:
If U is a partial isometry so is U^*

Claim
$\ker U^* = (ran\ U)^\perp$

Proof of Claim
If $y \perp ran\ U$, then $0 = \langle y, Ux\rangle\ \forall x \in V$
$0 = \langle U^*y, x\rangle$, Take $x = U^*y$
$0 = \langle U^*y, U^*y\rangle = \|U^*y\|^2$

If $y \in ker\ U^*, x \in V$
$\langle y, Ux\rangle = \langle U^*y, x\rangle = 0$
$\therefore y \perp ran\ U$
$\therefore \ker U^* \perp ran\ U$ ■

On the ran U
$U^*(Ux) = P^\perp_{\ker U} x$
$y \in ran\ U$ replace x by U^*Ux becomes $x - U^*Ux \in \ker U$
$0 = Ux - UU^*Ux \Rightarrow Ux = UU^*Ux$ (2⇒4)
$y = Ux, x = U^*Ux, U^*y = U^*Ux = x$

$y = Ux, \quad x = U^*Ux$
$U^*y = x$
$\|U^*y\| = \|x\| = \|Ux\| = \|y\|$
U^* is a partial isometry
⇔
$UU^* = U^{**}U^*$ is a projection

4⇒ 2
$U = UU^*U$
$\therefore U^*U = U^*UU^*U = (U^*U)^2$
Self adjoint, idempotent ∴ projection

Proof of Polar Decomposition Theorem
Diagonalize $|T| = diag(s_1, s_2, \dots, s_n), \quad s_1 \geq s_2 \geq \dots \geq s_n \geq 0$

Claim
$\|Tx\| = \||T|x\|\ \forall x \in V$
Proof
$\||T|x\|^2 = \langle |T|x, |T|x\rangle = \langle |T|^2x, x\rangle = \langle T^*Tx, x\rangle = \langle Tx, Tx\rangle = \|Tx\|^2$
■

$\ker |T| = \ker T = sp\{e_i : s_i = 0\}$
$ran\ |T| = sp\{e_i : s_i > 0\} = (\ker T)^\perp$
Define U on ran $|T|$ by $U(|T|x) = Tx$

U is isometric on $ran\ |T|$ by Claim (above)
Since $ran\ |T| = (\ker T)^{\perp}$ have that $U_{(\ker T)^{\perp}}$ is isometric. Hence U is a partial isometry.

Define $U|_{\ker T} = 0$
$U\left(\sum a_i e_i\right) = U\left(\sum_{i=1}^{k} a_i e_i\right), \sum_{i=1}^{k} a_i e_i \in ran\ T$
U is a partial isometry $T = U|T|$

Remark
$\{e_1, \ldots, e_k\}$ orthonormal basis for $(\ker T)^{\perp}$. Let $f_i = Ue_i, 1 \leq i \leq k$
f_i are orthonormal in W
$|T| = \sum_{i=1}^{k} s_i e_i e_i^*, e_i e_i^*\ is\ projection\ to\ \mathbb{C}e_i$
$T = U|T| = \sum_{i=1}^{k} s_i(f_i e_i^*)$, rank 1 projection sends $e_i \mapsto f_i$
$U = \sum_{i=1}^{k} f_i e_i^*$

Least Square Approximation

November-11-11 9:30 AM

An experiment is run to test whether the output, y is a linear function of the input variables: $x_1, \dots, x_n$
Run the experiment m times ($m \gg n$) to get a bunch of data.

x_1	x_2	...	x_n	y_n
x_{11}	x_{12}		x_{1n}	y_1
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_{m1}	...	...	...	y_m

Looking for $a_1, \dots, a_n \in \mathbb{R}\ or\ \mathbb{C}$ so that
$\sum_{j=1}^{n} a_j x_{ij} \approx y_i$ for $1 \le i \le m$

$$\text{minimize}_{a_1,\dots,a_n}\left(\sqrt{\sum_{i=1}^{m}\left|y_i - \sum_{y=1}^{n} a_j x_{ij}\right|^2}\right)$$

Let $X_1 = \begin{bmatrix} x_{11} \\ x_{12} \\ \vdots \\ x_{1m} \end{bmatrix}, \dots, X_j = \begin{bmatrix} x_{j1} \\ x_{j2} \\ \vdots \\ x_{jm} \end{bmatrix}, \quad 1 \le j \le n, \quad Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}$

Problem becomes

$$\text{minimize}_{a_1,\dots,a_n}\left\|Y - \sum_{j=1}^{n} a_j X_j\right\|_2 = dist(Y, span\{X_1, \dots, X_n\}) = \left\|Y - P_{sp\{X_j\}}Y\right\|_2$$

We must choose $a_1, \dots, a_n$ so that $\sum_{j=1}^{n} a_j X_j = P_{sp\{X_j\}}Y$

These are the scalars such that $\left\langle Y - \sum_{j=1}^{n} a_j X_j, X_i \right\rangle = 0, \quad 1 \le i \le n$

$$\left\langle Y - \sum_{j=1}^{n} a_j X_j, X_i \right\rangle = \langle Y, X_i\rangle - \sum_{j=1}^{n} a_j \langle X_j, X_i\rangle = X_i^* Y - \sum_{j=1}^{n} a_i X_j^* X_i$$

Let $X = [X_1, \dots, X_n]$, then $X^*Y = \begin{bmatrix} X_1^* \\ X_2^* \\ \vdots \\ X_n^* \end{bmatrix} Y = \begin{bmatrix} X_1^* Y \\ X_2^* Y \\ \vdots \\ X_n^* Y \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^{n} a_j X_1^* X_j \\ \sum_{j=1}^{n} a_j X_2^* X_j \\ \vdots \\ \sum_{j=1}^{n} a_j X_n^* X_j \end{bmatrix}$

$$X^*X = \begin{bmatrix} X_1^* \\ X_2^* \\ \vdots \\ X_n^* \end{bmatrix} [X_1 \quad X_2 \quad \dots \quad X_n] = \begin{bmatrix} X_1^* X_1 & \dots & X_1^* X_n \\ \vdots & \ddots & \vdots \\ X_n^* X_1 & \dots & X_n^* X_n \end{bmatrix}$$

$$\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^{n} a_j X_1^* X_j \\ \sum_{j=1}^{n} a_j X_2^* X_j \\ \vdots \\ \sum_{j=1}^{n} a_j X_n^* X_j \end{bmatrix} = X^*Xa = X^*Y$$

If $X_1, \dots, X_n$ are linearly independent then X has rank n.

Claim
$rank(X^*X) = rank\ X$
Proof
$rank(X) = \dim(domain) - nul(X) = n - nul(X)$

Example

x_1	x_2	y	ax
7	3	1.6	1.86
9	2	2.1	1.94
5	5	2.0	2.02
4	6	2.2	2.10
3	1	0.8	0.73
3	2	1.1	0.98

$X^*X = \begin{bmatrix} 189 & 97 \\ 97 & 79 \end{bmatrix}, X^*Y = \begin{bmatrix} 54.6 \\ 35.2 \end{bmatrix}$
$(X^*X)^{-1} = \begin{pmatrix} 0.0143 & -0.0176 \\ -0.0176 & 0.0324 \end{pmatrix}$
$a = \begin{bmatrix} 0.161 \\ 0.243 \end{bmatrix}$

$rank(X^*X) = n - nul\, X^*X$
If $x \in \ker X$ then $X^*Xx = X^*0 = 0, so\ x \in \ker X^*X$
If $x \in \ker X^*X,\ 0 = \langle X^*Xx, x\rangle = \langle Xx, Xx\rangle = \|Xx\|^2,\ so\ x \in \ker X$
■

$\therefore$ If $X_1, \dots, X_n$ is linearly independent then X^*X is invertible. d
$X^*Xa = X^*Y$
$\therefore a = (X^*X)^{-1}X^*Y$

Sesquilinear Forms

November-11-11 10:09 AM

Sesquilinear Form

V $\mathbb{C}$ vector space.
A function $F: V \times V \to \mathbb{C}$ is a sesquilinear form if it is linear in the first variable and conjugate linear in the second variable.
$F(a_1 v_1 + a_2 v_2, w) = a_1 F(v_1 w) + a_2 F(v_2, w)$
$F(v, a_1 w_1 + a_2 w_2) = \overline{a_1} F(v, w_1) + \overline{a_2} F(v, w_2)$

Definitions

Say F is **Hermitian** if $F(w, v) = \overline{F(v, w)}$
F is **non-negative** if F is Hermitian and $F(v, v) \geq 0$
F is **positive** if $F \geq 0$ and $F(v, v) > 0$ for $v \neq 0$

Theorem

If $F: V \times V \to \mathbb{C}$ is sesquilinear form, then there is a unique $T_F \in \mathcal{L}(V)$ such that $F(v, w) = \langle T_F v, w \rangle$ for $v, w \in V$

Moreover, the map $F \mapsto T_F$ is a linear isomorphism from the vector space of sesquilinear forms onto $\mathcal{L}(V)$

Principal Axis Theorem

If $F(x, y)$ is a Hermitian sesquilinear form then $\exists$ an orthonormal basis $\{e_1, \ldots, e_n\}$ and $d_i \in \mathbb{R}\ s.t.$

$$F\left(\sum \alpha_i e_i, \sum \beta_i e_i\right) = \sum_{i=1}^{n} d_i \alpha_i \overline{\beta_i}$$

e_i are principal axes.

Symmetric Quadratic Form

A symmetric quadratic form on $\mathbb{R}^n$ is

$$q(x_1, \ldots, x_n) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j, \qquad \text{where } a_{ij} = a_{ji} \in \mathbb{R}$$

Any quadratic form in $\mathbb{R}^n$

$$q(x) = \sum \sum b_{ij} x_i x_j$$

Replace b_{ij} by $a_{ij} = \dfrac{b_{ij} + b_{ji}}{2}$ now it is symmetric.

Diagonalization

Again, this quadratic form can be diagonalized
$A = [a_{ij}] = A^*$
$\exists$ o.n. basis $\{e_1, \ldots, e_n\}$ of $\mathbb{R}^n$ consisting of eigenvalues
$Ae_i = d_i e_i, \quad 1 \leq i \leq n, \quad d_i \in \mathbb{R}$

$$e_i = \begin{bmatrix} c_{1i} \\ c_{2i} \\ \vdots \\ c_{ni} \end{bmatrix}, \qquad U = [e_1 \quad e_2 \quad \cdots \quad e_n] = [c_{ij}]_{n \times n}, \qquad U \text{ orthogonal}$$

$U^* A U = diag(d_1, \ldots, d_n) = D$

$$q(x_1, \ldots, x_n) = \left\langle A \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \right\rangle = \left\langle UDU^* \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \right\rangle$$

$$= \left\langle DU^* \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, U^* \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \right\rangle$$

$$U^* \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} e_1^* \\ \vdots \\ e_n^* \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \langle x, e_1 \rangle \\ \vdots \\ \langle x, e_n \rangle \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} c_{i1} x_i \\ \vdots \\ \sum_{i=1}^{n} c_{in} x_i \end{bmatrix}, c_{ij} \in \mathbb{R}$$

$$q(x_1, \ldots, x_n) = \sum_{j=1}^{n} d_j \left(\sum_{i=1}^{n} c_{ij} x_i \right)^2$$

Proof

Fix an orthonormal basis $\xi = \{e_1, \ldots, e_n\}$ for V. F sesquilinear form.
Need $\langle Te_j, e_i \rangle = F(e_j, e_i),\ 1 \leq i, j \leq n$
Let $[T]_\xi = [t_{ij}]_{n \times n}$ where $t_{ij} = \langle Te_j, e_i \rangle$
T is the unique map on $\mathcal{L}(V)$ such that $\langle Te_j, e_i \rangle = F(e_j, e_i),\ 1 \leq i, j \leq n$

Let $v = \sum_{i=1}^{n} a_i e_i, \qquad w = \sum_{i=1}^{n} b_i e_i$

$$\langle Tv, w \rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} a_j \overline{b_i} \langle Te_j, e_i \rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} a_j \overline{b_i} F(e_j, e_i) = \sum_{i=1}^{n} \overline{b_i} F\left(\sum_{j=1}^{n} a_j e_j, e_i \right)$$

$$= F\left(\sum a_j e_j, \sum b_i e_i \right) = F(v, w)$$

Show T_F is uniquely determined by F, $F \mapsto T_F$ is linear.
$T_F = 0 \Leftrightarrow F = 0 \therefore 1\ to\ 1$
Onto $if\ T \in \mathcal{L}(V)$, define $F(v, w) = \langle Tv, w \rangle$ is sesquilinear
So $F \mapsto T, onto$

Proof of Principal Axis Theorem

$F(x, y) = \langle Ax, y \rangle = \langle x, A^* y \rangle$
$F(x, y) = \overline{F(y, x)} = \overline{\langle Ay, x \rangle} = \langle x, Ay \rangle$
$\therefore A = A^*$ is Hermitian
A is diagonalizable w.r.t orthonormal basis
$\xi = \{e_1, \ldots, e_n\}$
$[A]_\xi = diag\ (d_1, \ldots, d_n), \qquad d_i \in \mathbb{R}$

$$F\left(\sum \alpha_i e_i, \sum \beta_i e_i \right) = \left\langle A \sum \alpha_i e_i, \sum \beta_i e_i \right\rangle = \left\langle \sum d_i \alpha_i e_i, \sum \beta_i e_i \right\rangle = \sum d_i \alpha_i \overline{b_i}$$

Conics

November-14-11 10:07 AM

Ellipse

Take two points F_1, F_2, with separation $2c$. Pick $a > c$
Ellipse is $\{P = (x, y) : |P - F_1| + |P - F_2| = 2a\}$

$$\boxed{\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1}$$
$b^2 = a^2 - c^2, \qquad c^2 = a^2 + b^2$

Hyperbola

Take two points F_1, F_2 with separation $2c$
Hyperbola is $\{(x, y) : |PF_1| - |PF_2| = 2a\}$

$F_1 = (-c, 0), \qquad F_2 = (c, 0)$
$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$$
$c^2 = a^2 + b^2$

Parabola

Focus and line. The set of points equidistant to focus and line.

Formula of an Ellipse

Translate so $F_1 = (-c, 0),\ F_2 = (c, 0)$
$$\{(x, y): |(x + c, y)| + |(x - c, y)| = 2a\} = \left\{(x, y): \sqrt{(x + c)^2 + y^2} + \sqrt{(x - c)^2 + y^2} = 2a\right\}$$

$$\sqrt{(x + c)^2 + y^2} = 2a - \sqrt{(x - c)^2 + y^2}$$
$$(x + c)^2 + y^2 = 4a^2 - 4a\sqrt{(x - c)^2 + y^2} + (x - c)^2 + y^2$$
$$4a\sqrt{(x - c)^2 + y^2} = 4a^2 - 4cx$$
$$x^2 - 2cx + c^2(x - c)^2 + y^2 = a^2 - 2cx + \frac{c^2x^2}{a^2}$$
$$\frac{a^2 - c^2}{a^2}x^2 + y^2 = a^2 - c^2$$
$$\frac{x^2}{a^2} + \frac{y^2}{a^2 - c^2} = 1$$

General Conic

$ax^2 + bxy + cy^2 + dx + ey + f = 0$

$ax^2 + bxy + cy^2$ is the quadratic form
$$A = \begin{bmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c \end{bmatrix}$$
$$\left\langle A\begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix}\right\rangle = ax^2 + bxy + cy^2$$

Diagonalize w.r.t. orthonormal basis:
Eigenvectors $v_1 = \begin{pmatrix} \alpha_1 \\ \beta_1 \end{pmatrix}, \qquad v_2 = \begin{pmatrix} \alpha_2 \\ \beta_2 \end{pmatrix}$
$Av_1 = \lambda_1 v_1$
$Av_2 = \lambda_2 v_2$

$U = \begin{pmatrix} \alpha_1 & \alpha_2 \\ \beta_1 & \beta_2 \end{pmatrix}$ orthogonal matrix
$U^* = \begin{pmatrix} \alpha_1 & \beta_1 \\ \alpha_1 & \beta_2 \end{pmatrix}$

$U^*AU = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} = \mathrm{D}, \qquad \mathrm{A} = \mathrm{UDU}^*$
So
$$ax^2 + bxy + cy^2 = \left\langle A\begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix}\right\rangle = \left\langle UDU^*\begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix}\right\rangle = \left\langle DU^*\begin{pmatrix} x \\ y \end{pmatrix}, U^*\begin{pmatrix} x \\ y \end{pmatrix}\right\rangle$$
$$= \left\langle D\begin{pmatrix} \alpha_1 x + \beta_1 y \\ \alpha_2 x + \beta_2 y \end{pmatrix}, \begin{pmatrix} \alpha_1 x + \beta_1 y \\ \alpha_2 x + \beta_1 y \end{pmatrix}\right\rangle = \lambda_1(\alpha_1 x + \beta_1 y)^2 + \lambda_2(\alpha_2 x + \beta_2 y)^2$$

$\lambda_1\lambda_2 = \det D = \det A$
$\lambda_1\lambda_2 > 0$ ellipse
$\lambda_1\lambda_2 = 0$ parabola
$\lambda_1\lambda_2 < 0$ hyperbola

Write $\begin{pmatrix} d \\ e \end{pmatrix} = d'\begin{pmatrix} \alpha_1 \\ \beta_1 \end{pmatrix} + e'\begin{pmatrix} \alpha_2 \\ \beta_2 \end{pmatrix}$
$dx + ey = d'(\alpha_1 x + \beta_1 y) + e'(\alpha_2 x + \beta_2 y)$

The equation
$ax^2 + bxy + cy^2 + dx + dy + f = 0$
becomes
$$\lambda_1(\alpha_1 x + \beta_1 y)^2 + \lambda_2(\alpha_2 x + \beta_2 y)^2 + d'(\alpha_1 x + \beta_1 y) + e'(\alpha_2 x + \beta_2 y)i + f = 0$$
$$\lambda_1\left(\alpha_1 x + \beta_1 y + \frac{d'}{2\lambda_1}\right)^2 + \lambda_2\left(\alpha_2 x + \beta_2 y + \frac{e'}{e\lambda_2}\right)^2 = \left(\frac{d'^2}{2\lambda_1} + \frac{e'^2}{2\lambda_2} - f\right) = f'$$

$$\lambda_1(\alpha_1 x + \beta_1 y)^2 + \lambda_1\frac{2d'}{2\lambda_1}(\alpha_1 x + \beta_1 y) + \frac{d'^2}{4\lambda_1} + \lambda_2(\alpha_2 x + \beta_2 y)^2 + \alpha_2\frac{2e'}{2\lambda_2}(\alpha_2 x + \beta_2 y) + \frac{e'^2}{4\lambda_2}$$

Translate to eliminate constants
$$\frac{d'}{2\lambda_1}, \frac{e'}{2\lambda_2}$$
Rotate by U to get
$\lambda_1 x^2 + \lambda_2 y^2 = f'$
$$\frac{x^2}{\alpha^2} + \frac{y^2}{\beta_2^2} = 1$$

Duality

November-16-11 10:00 AM

Dual Space

If V is a vector space over $\mathbb{F}$ then the dual space of V is
$V^* = \mathcal{L}(V, \mathbb{F})$. Elements of V^* are called **linear functionals.**

Fix a basis $\beta = \{v_1, \dots, v_i, \dots, v_n\}$ for V

Define $\delta_j \in V^*$ by $\delta_j\left(\sum_{i=1}^{n} \alpha_i v_i\right) = \alpha_j$

$$\delta_j(v_i) = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases} = \delta_{ij}$$

Kronecker Delta

Proposition

$\dim V^* = \dim V$ and $\{\delta_1, \dots, \delta_n\}$ is a basis for V^*
(Called the **dual basis** of $\{v_1, \dots, v_n\}$)

Note

$V^{**} = \mathcal{L}(V^*, \mathbb{F})$
If $v \in V$ define $\hat{v} \in V^{**}$ by $\hat{v}(\varphi) \coloneqq \varphi(v),\ \varphi \in V^*$

$$\hat{v}(a\varphi + b\psi) = (a\varphi + b\psi)(v) = a\varphi(v) + b\psi(v) = a\hat{v}(\varphi) + b\hat{v}(\psi)$$

Thus there is a natural linear map
$i: V \to V^{**}$ by $i(v) = \hat{v}$
This is linear.

Theorem

The natural map $i: V \to V^{**}$ is an isomorphism.

Remark

This fails dramatically for infinite dimensional vectors spaces.

Example

Let $c_{00} = \{\text{sequences } (x_1, x_2, x_3, \dots)\ x_i = 0 \text{ except for finitely often}\}$
$e_i = (0, \dots, 0, 1, 0, \dots)$ is a basis for c_{00}

$$\varphi \in C_{00}^*, \qquad \varphi(e_i) = \alpha_i, \qquad \varphi = \sum \alpha_i \delta_i$$

$c_{00}^* = s = \{all\ sequences\ (\alpha_1, \alpha_2, \dots)\}$
$\dim S = 2^{\aleph_0}$
S^* is humongous.

Isomorphism

Since we have an isomorphism $i: V \to V^{**}$ we say
$V^{**} = V$ and identify $i(v)$ with v.

V is **reflexive**

Dual Space Basis

Suppose $\varphi \in V^*$
Let $\varphi(v_i) = \beta_i,\ 1 \le i \le n$

$$\psi = \sum_{j=1}^{n} \beta_j \delta_j \in V^*$$

$$\psi(v_i) \sum_{j=1}^{n} \beta_j \delta_i(v_i) = \beta_i$$

A linear map is determined by what it does to a basis, so $\varphi = \psi$

Proof of Proposition

I expressed every $\varphi \in V^*$ as a linear combination of $\delta_1, \dots, \delta_n$ which are linearly independent.

$$0 = \sum_{i=1}^{n} a_i \delta_i$$

$$0 = \left(\sum_{i=1}^{n} a_i f_i\right)(v_j) = a_j$$

$\Rightarrow a_1 = a_2 = \dots = a_n = 0$
So $\delta_1, \dots, \delta_n$ are linearly independent $span\ V^*$ ∴ is a basis.
$\dim V^* = n = \dim V$ ∎

Proof of Theorem

Fix a basis $v_1, \dots, v_n$ for V
Construct the dual basis $\delta_1, \dots, \delta_n$ for V^*
Construct the dual dual basis $\varepsilon_1, \dots, \varepsilon_n$ for V^{**}

$\hat{v}_i(\delta_j) = \delta_j(v_i) = \delta_{ij}$
$\varepsilon_i(\delta_j) = \delta_{ij}$
So $\hat{v}_i$ and e_i agree on a basis ∴ $\hat{v}_i = e_i$

So $i\left(\sum_{j=1}^{n} a_j v_j\right) = \sum_{j=1}^{n} a_j \varepsilon_j$ is 1-1 and onto ∎

Duality on Inner Product Spaces

November-18-11 9:31 AM

Theorem

Let V be an inner product space. Then for each $\varphi \in V^*$ there is a unique $w \in V$ s.t. $\varphi(v) = \langle v, w\rangle\ \forall v \in V$

The map which sends $\varphi \mapsto w$ is a conjugate linear map of V^* onto V.

Corollary

V inner product space, we convert V^* to an inner product space by

$$\left\langle \sum \alpha_i\delta_i, \sum \beta_i\delta_i \right\rangle = \sum_{i=1}^{n} \alpha_i\overline{\beta_i}$$

If $\varphi \in V^*$ then $\|\varphi\|_{V^*} = \sup_{\substack{\|v\|\le 1\\ v\in V}} |\varphi(v)|$

Notation

$$\left(\sum \alpha_i e_i, \sum \beta_j\delta_j\right) = \sum_{j=1}^{n} \beta_i\delta_j\left(\sum_{i=1}^{n} \alpha_i e_i\right)$$

Definition

Let V be a finite dimensional vector space.
If $S \subseteq V$ let $S^\perp = \{\varphi \in V^* : \varphi(s) = 0\ \forall s \in S\}$
This is the **annihilator** of S

Proposition

$S \subseteq V$ then

1. $S^\perp$ is a subspace of V^*
2. $S^{\perp\perp} = span(S)$
3. $\dim S^\perp + \dim S^{\perp\perp} = \dim V$

Relationship between perps.

H inner product space
H^* conjugate linear isometric v... to H
$\varphi \in H^*,\ \exists!\, y \in H\ s.t.\ \varphi(x) = \langle x, y\rangle,\ \varphi \to y$ conjugate linear
$M \subset H, M^\perp = H(-)M = \{y : \langle x, y\rangle = 0\ \forall x \in M\}$
$M^\perp = M^0 = \{\varphi : \varphi(x) = 0\ \forall x \in M\}$

Proof

Let $\xi = \{e_1, \dots, e_n\}$ be an orthonormal basis for V. Let $\delta_1, \dots, \delta_n$ be the dual basis for V^*
If $\varphi \in V^*$, let $\varphi(e_i) = \beta_i, 1 \le i \le n$

So $\varphi = \sum_{i=1}^{n} \beta_i\delta_i$ because $\left(\sum_{j=1}^{n} \beta_j\delta_j\right)(e_i) = \beta_i$

Want $w \in V$ s.t. $\langle e_i, w\rangle = \beta_i,\ 1 \le i \le n$

$$\left\langle e_i, \sum_{j=1}^{n} \overline{\beta_j} e_j \right\rangle = \beta_i$$

So define $T: V^* \to V$ by

$$T\left(\sum_{i=1}^{n} \beta_i\delta_i\right) = \sum_{i=1}^{n} \overline{\beta_i} e_i$$

$$T\varphi = w = \sum_{i=1}^{n} \overline{\beta_i} e_i$$

$$\langle v, w\rangle = \left\langle \sum \alpha_i e_i, \sum \overline{\beta_i} e_i \right\rangle = \sum \alpha_i\beta_i = \varphi(v)$$

T is not linear-it is conjugate linear. T is 1-1 and onto ■

Proof of Corollary

Clearly this makes V^* an inner product space

Let $\varphi = \sum_{j=1}^{n} \beta_j\delta_j \in V^*$

$$\|\varphi\|_{V^*} = \sqrt{\sum_{j=1}^{n} |\beta_j|^2}$$

If $v \in V$, $\quad v = \sum_{i=1}^{n} \alpha_i e_i$

$$|\varphi(v)| = \left|\left(\sum \alpha_i e_i, \sum \beta_j\delta_j\right)\right| = \left|\sum_{i=1}^{n} \alpha_i\beta_i\right| \le \sqrt{\sum |\alpha_i|^2}\sqrt{\sum |\beta_i|^2} = \|v\|_V\|\varphi\|_{V^*}$$

So get:
$\sup_{\substack{v\in V\\ \|v\|\le 1}} |\varphi(v)| \le \sup_{\|v\|\le 1} \|v\|\|\varphi\|_{V^*} = \|\varphi\|_{V^*}$
To get equality, take

$$v = \frac{\sum_{i=1}^{n} \overline{\beta_i} e_i}{\sqrt{\sum |\beta_i|^2}}, \qquad \varphi(v) = \frac{\sum_{i=1}^{n} \overline{\beta_i}\beta_i}{\sqrt{\sum |\beta_i|^2}} = \sqrt{\sum |b_i|^2} = \|\varphi\|_{V^*}$$

Proof of Proposition

1.

$0 \in S^\perp$
If $\varphi, \psi \in S^\perp$, $\quad s \in S, \quad \alpha, \beta \in F$
$(\alpha\varphi + \beta\psi)(s) = \alpha\varphi(s) + \beta\psi(s) = 0$

2.

$S^{\perp\perp}$ is a subspace of $V^{**} = V$ which contains S because $s \in S,\ \varphi \in S^\perp$
$i(s) \sim s(\varphi) = \varphi(s) = 0$
So $S^{\perp\perp} \supseteq span(S)$

Suppose $v \notin span(S)$
Take a basis for S, say $v_1, \dots, v_k$ ($\dim S = k$) and extend too a basis $v_1, \dots, v_k, v, v_{k+2}, \dots, v_n$
Note, used v in the basis.

Let $\delta_1, \dots, \delta_n$ be the dual basis of V^*
$\delta_{k+1}(v_i) = 0, \quad 1 \le i \le k \Rightarrow \delta_{k+1} \in S^\perp$
$\delta_{k+1}(v) = 1 \ne 0, \quad \therefore v \notin S^{\perp\perp}$
So $S^{\perp\perp} \subseteq span\, S$ ∴ equal

3.

Claim:
$S^\perp = span\{\delta_{k+1}, \dots, \delta_n\}, \quad j \ge k+1:\ \delta_j(v_i) = 0\ for\ 1 \le i \le k \Rightarrow \delta_j \in S^\perp$
So $span\{\delta_{k+1}, \dots, \delta_n\} \subseteq S^\perp$

Let $\varphi = \sum_{j=1}^{n} \beta_i\delta_i \in S^\perp$

$0 = \varphi(v_i) = \beta_i \Rightarrow \varphi \in sp\{\delta_{k+1}, \dots, \delta_n\}, \quad i \le i \le k$
$\dim S = k, \dim S^\perp = n - k, n + n - k = n$

Transpose

November-21-11 9:39 AM

Transpose Map

If $T \in \mathcal{L}(V,W)$ define the **transpose** of T to be the map $T^t \in \mathcal{L}(W^*,V^*)$ by
$(T^t\varphi)(v) = \varphi(Tv)$
$T^t\varphi = \varphi \circ T \in \mathcal{L}(V,\mathbb{F})$

Claim

T^t is a linear map

Claim

"transpose" is a linear map
$(\alpha S + \beta T)^t = \alpha S^t + \beta T^t$

Theorem

$T \in \mathcal{L}(V,W), T^t \in \mathcal{L}(W^*,V^*)$

1. If $\beta = \{v_1,\dots,v_m\}$ basis for V, $\beta' = \{\delta_1,\dots,\delta_m\}$ for V^*
   $\mathcal{C} = \{w_1,\dots,w_n\}$ basis for W, $\mathcal{C}' = \{\varepsilon_1,\dots,\varepsilon_n\}$ for W^*

   If $[T]_\beta^{\mathcal{C}} = [t_{ij}]_{m\times n}$, then $[T]_{\mathcal{C}'}^{\beta'} = [t_{ji}]_{n\times m}$
2. $T \mapsto T^t$ is a linear isomorphism of $\mathcal{L}(V,W)$ onto $\mathcal{L}(W^*,V^*)$
3. $ran\, T^t = (\ker T)^\perp$ and $\ker T^t = (ran\, T)^\perp$
4. $rank\, T^t = rank\, T$

Proof of Claim

$T^t(\alpha\varphi + \beta\psi)(v) = (\alpha\varphi + \beta\psi)(Tv) = \alpha\varphi(Tv) + \beta\psi(Tv) = (\alpha T^t\varphi + \beta T^t\psi)(v)$

Proof of Claim

$\varphi \in W^*, v \in V$
$(\alpha S + \beta T)^t(\varphi)(v) = \psi\big((\alpha S + \beta T)(v)\big) = \psi(\alpha Sv + \beta Tv) = \alpha\varphi(Sv) + \beta\psi(Tv)$
$= \alpha(S^t\varphi)(v) + \beta(T^t\psi)(v) = (\alpha S^t + \beta T^t)(\varphi)(\psi)$

Proof of Theorem

1

$\left([T^t]_{\mathcal{C}}^{\beta'}\right)_{ij} = a_{ij}$ where $(T^t\varepsilon_j)(v_i) = \left(\sum_{k=1}^m a_{kj}\delta_k\right)(v_i) = a_{ij}$

$(T^t\varepsilon_j)(v_i) = \varepsilon_j(Tv_i) = \varepsilon_j\left(\sum_{k=1}^n t_{ki}w_k\right) = t_{ji}$

$\therefore [T^t]_{\mathcal{C}'}^{\beta'} = [t_{ji}] = \left([T]_\beta^{\mathcal{C}}\right)^t$
The matrix of the transpose is the transpose of the matrix.

2

$E_{ij} = [b_{kl}]$ *where* $b = 1$ *if* $k = i, j = l$ *and* $b = 0$ *otherwise*
E_{ij} is a basis for $\mathcal{L}(V,W)$. $E_i = w_i\delta_j$
$E_{ij}^t = E_{ji}$ sends a basis for $\mathcal{L}(W^*,V^*)$ to a basis for $\mathcal{L}(V,W)$. $\therefore$ 1-1 and onto.

3

$\varphi \in \ker T^t \in W^* \Leftrightarrow 0 = T^t\varphi \in V^*$
$\Leftrightarrow 0 = T^t\varphi(v)\forall v \in V = \varphi(Tv) \Leftrightarrow \varphi \in (ran\, T)^\perp$

$\therefore v \in \ker T = \ker T^{tt} = (ran\, T^t)^\perp$
$\therefore (\ker T)^\perp = (ran\, T^t)^{\perp\perp} = ran\, T^t$

4

$rank\, T^t = \dim ran\, T^t = \dim(\ker T)^\perp = \dim V - \dim\ker T = ran\, T$

Since
$M \subseteq V$, basis for M, extend for V. Dual space $\delta_1,\dots,\delta_n$
$M^\perp = sp(\delta_{k+1},\dots,\delta_n\} \Rightarrow \dim M^\perp = n - \dim M$

Quotient Spaces

November-21-11 10:02 AM

Quotient Space

V vector space, M subspace of V
Say $v_1 \equiv v_2$ iff $v_1 - v_2 \in M$
$\frac{V}{M}$ is the set of equivalence classes $\dot{v} = v + M$
Make $\frac{V}{M}$ into a vector space by
$t\dot{v} = t(v + M) = tv + M$
$\dot{v} + \dot{w} = (v \dot{+} w)$

$\frac{V}{M}$ is the called the **quotient space** of V by M.

The map $\Pi: V \to \frac{V}{M}$ by $\Pi(v) = \dot{v}$ is called the **quotient map.**

Proposition

$\Pi \in \mathcal{L}\left(V, \frac{V}{M}\right)$ is surjective and $\ker \Pi = M$.

Theorem

If M is a subspace of V then $M^* \cong \frac{V^*}{M^\perp}$ (isomorphic to) and $\left(\frac{V}{M}\right)^* \cong M^\perp$

Relations

$V^* \to_R M^*$
$V^* \to_q \left(\frac{V^*}{M^\perp}\right) \to_{\tilde{R}} M^*$
$\tilde{R}(\varphi + M^\perp) = R\varphi$ well defined because of
$\varphi_1, \varphi_2 \in \dot{\varphi}, \qquad \varphi_1 - \varphi_2 = \psi \in M^\perp$
$\varphi_2\Big|_M = \varphi_1\Big|_M + \psi\Big|_M = \varphi_1\Big|_M$
$\therefore \tilde{R}\ 1-1$

Proof of Well Definition

If $v_1 \equiv v_2$ then $v_1 - v_2 = m \in M$
$\therefore tv_1 - tv_2 = tm \in M$
$\therefore tv_1 \equiv tv_2$
So $t\dot{v}$ is independent of choice of representative.
$If\ v_1 \equiv v_2, w_1 \equiv w_2$ say $w_1 - w_2 = n \in M$
$v_2 + w_2 = v_1 + m + w_1 + n = (v_1 + w_1) + (m + n), \qquad (m + n) \in M$
$\therefore v_2 + w_2 \equiv v_1 \dot{+} w_1$
So $\dot{v} + \dot{w} = (v \dot{+} w)$ is well defined.

Proof of Proposition

Π is linear, surjective by definition.
$\ker \Pi = \{v : \dot{v} \neq 0\} = \{v : v \in M\} = M$

Proof of Theorem

Let $\Pi: V \to \frac{V}{M}$ be the quotient map, then $\Pi^t: \left(\frac{V}{M}\right)^* \to V^*$
$\ker \Pi^t = (ran\ \Pi)^\perp = \{0\}$
$\therefore \Pi^t$ is injective

$ran\ \Pi^t = (\ker \Pi)^\perp = M^\perp$
So Π^t maps $\left(\frac{V}{M}\right)^*$ 1-1 and onto $M^\perp$. $\therefore$ Linear isomorphism
The connection is given by :
Take $\varphi \in \left(\frac{V}{M}\right)^*$, $\Pi^t\varphi = \varphi \circ \Pi \in V^*$
$(\varphi \circ \Pi)(m) = \varphi(\dot{0}) = 0\ \forall m \in M$

So $\left(\frac{V^*}{M^\perp}\right)^* \cong M^{\perp\perp} = M$
$\therefore \frac{V^*}{M^\perp} = \left(\frac{V^*}{M^\perp}\right)^{**} \cong M^*$

If $\varphi \in V^*$ the restriction map $R\varphi = \varphi|_M$ is a linear map of V^* onto M^*
$\ker R = \{\varphi : \varphi|_M = 0\} = M^\perp$

Convex Sets

November-23-11 9:33 AM

Convexity

A subset C of $\mathbb{R}$ or $\mathbb{C}$ is convex if $\forall c_1, c_2 \in C\ \forall 0 \le t \le 1, (1-t)c_1 + tc_2 \in C$

Hyperplane

H is a hyperplane if $\exists \varphi \in V^*, \varphi \neq 0$ such that $H = \{v : Re\ \varphi(v) = a\}$
A **half space** is a set of form $H^+ = \{v : Re\ \varphi(v) \ge a\}$
Note: H and H^+ are convex.

Proposition

1. The intersection of convex sets is convex.
2. If $S \subseteq V$, $conv(S)$ is the smallest convex set containing S
$$conv(S) = \left\{\sum_{i=1}^{r} t_i s_i : r \in \mathbb{N}, s_i \in S, t_i \ge 0, \sum_{i=1}^{r} t_i = 1\right\}$$

Theorem (Carathéodory)

If V is a real vector space of dimension n, $S \subseteq V$ then every point in $conv(S)$ is a convex combination of $n+1$ points in S

Remark

1. If V is a complex vector space of dimensions n, then it is a real vector space of dimension 2n. So $2n+1$ points are needed.
2. In $\mathbb{R}^n$ take $S = \{0, e_1, e_2, \dots, e_n\}$ the point
$$\frac{1}{n+1}0 + \sum_{i=1}^{n} \frac{1}{n+1} e_i \in S \text{ requires } n+1 \text{ points.}$$

Corollary

If $S \subseteq V$ is compact, $\dim V = n < \infty$ then $conv(S)$ is compact.

Remark: From Calculus

A set $C \subseteq \mathbb{R}^n$ is sequentially compact if every sequence $\{c_n : n \ge 1\}$ of points in C has a convergent subsequence $\lim_{k\to\infty} c_{n_k} = c, c \in C$

Heine-Bore Theorem

$C \subseteq \mathbb{R}^n$ is compact $\Leftrightarrow$ C is closed and bounded

Extreme Value theorem

If C compact, $f: C \to \mathbb{R}$ is continuous then f attains its maximum and minimum values.

Proof of Proposition

1.
$C_i, i \in I$ are convex sets in V
$$C = \bigcap_{i\in I} C_i, \qquad c_1, c_2 \in C, \qquad 0 \le t \le 1$$
$c_1, c_2 \in C_i \Rightarrow (1-t)c_1 + tc_2 \in C_i\ \ \forall i$
$\therefore c_1, c_2, \in C$

2.
$conv(S)$ exists - it is the intersection of all convex sets containing S
Claim
$$\sum_{i=1}^{r} t_i s_i \in conv(S), \qquad s_1 \in conv(S)$$
Suppose
$$v_k = \sum_{i=1}^{k} \left(\frac{t_i}{\sum_{j=1}^{k} t_j}\right) s_i \in conv(S)$$
True for $k = 1$
If true for k then
$$v_{k+1} = \left(\frac{\sum_{i=1}^{k} t_i}{\sum_{i=1}^{k+1} t_i}\right) v_k + \left(\frac{t_{k+1}}{\sum_{i=1}^{k+1} t_i}\right) s_{k+1} \in conv(S)$$
Convex combinations of 2 points of $conv(S)$

By induction $v_r = \sum_{i=1}^{r} t_i s_i \in conv(S)$

If $\sum_{i=1}^{r} t_i s_i, \sum_{j=1}^{r'} t'_j s'_j, \qquad t_i t'_j \ge 0, \qquad \sum_{i=1}^{r} t_i = 1 = \sum_{j=1}^{r'} t'_j$

For $0 \le u \le 1, \qquad (1-u)\sum_{i=1}^{r} t_i s_i + u \sum_{i=1}^{r'} t'_i s'_i = 1$

So the convex combination of two convex combination of two convex combinations of points in S is a convex combinations of points in S
$\therefore \left\{\sum_{i=1}^{r} t_i s_i : r \ge 1, t_i \ge 0, \sum t_i = 1, s_i \in S\right\}$ is the smallest convext set $\supseteq S$

Proof of Theorem

Take a point $v \in conv(S)$. Can write $v = \sum_{i=1}^{r} t_i s_i, \qquad s_i \in S, t_i \ge 0, \sum t_i = 1$

Claim

If $r \ge n+2$, we can find another convex combination equal to v using fewer of the $\{s_i\}$'s.
wlog, $t_i > 0$ (if $t_{i_0} = 0$ throw s_{i_0} out of the set)
The set $\{s_1 - s_r, s_2 - s_r, \dots, s_{r-1} - s_r\}$ has $r - 1 \ge n+1$ elements $\Rightarrow$ linearly dependent.
$\therefore \exists a_i \in \mathbb{R}$, not all zero such that
$$0 = \sum_{i=1}^{r-1} a_i(s_i - s_r) = \sum_{i=1}^{r-1} a_i s_i + a_r s_r \text{ where } a_r = -\sum_{i=1}^{r-1} a_i$$
So $\sum_{i=1}^{r} a_i = 0$ and $\vec{0} = \sum_{i=1}^{r} a_i s_i$

Let $J = \{i : a_i < 0\}$, Let $\delta = \min_{i\in J}\left\{\frac{t_i}{|a_i|}\right\} = \frac{t_{i_0}}{|a_{i_0}|}, for\ some\ i_0 \in J$
$$v = \sum_{i=1}^{r} t_i s_i + \delta \sum_{i=1}^{r} a_i s_i = \sum_{i=1}^{r} (t_i + \delta_i a_i) s_i$$
$i \in J$: $t_i + \delta a_i \ge t_i + \frac{t_i}{|a_i|} a_i = t_i - t_i = 0$
i_i: $t_{i_0} + \delta a_{i_0} = t_{i_0} - t_{i_0} = 0$
$i \notin J$: $t_i + \delta a_i \ge t_i \ge 0$
$$\sum_{i=1}^{r} (t_i + \delta a_i) = \sum_{i=1}^{r} t_i + \delta \sum_{i=1}^{r} a_i = 1 + \delta 0 = 1$$

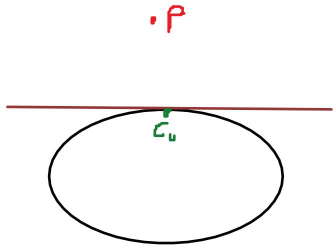
This new combination does not need s_{i_0} because the coefficient is 0. So we have reduced r to $r-1$.

Proof of Corollary

Every $v \in conv(S)$ is the convex combination of $n+1$ points in S
$$S^{n+1} = \{(s_1, s_2, \dots, s_{n+1}) : s_i \in S\}, \qquad \Delta_{n+1} = \left\{(t_1, \dots, t_{n+1}) : t_i \ge 0, \sum_{i=1}^{n+1} t_i = 1\right\}$$
$S^{n+1} \times \Delta_{n+1} \subseteq V^{n+1} \times \mathbb{R}^{n+1}, S^{n+1}$ compact
$$f: S^{n+1} \times \Delta_{n+1} \to V_{n+1}, \qquad f\big((s_1, s_2, \dots, s_{n+1}, t_1, t_2, \dots, t_{n+1})\big) = \sum_{i=1}^{n+1} t_i s_i$$
f is continuous
The continuous image of a compact set is compact (by EVT)
$conv(S) = f(S^{n+1} \times \Delta_{n+1})$ is compact

Convexity

November-25-11 9:32 AM



$H = \{x: Re(\varphi(x)) = Re(\varphi(c_0))\}$

Theorem

Let V be a finite dimensional inner produce space ($\mathbb{F} = \mathbb{R}\ or\ \mathbb{C}$).
$C \subseteq V$ closed convex set, $p \in V, p \notin C$
Then there is a unique point $c_0 \in C$ closet to p.
Let $\varphi(x) = \langle x, p - c_0\rangle$
Then $Re\ \varphi(p) > Re\ \varphi(c_0) \geq Re\ \varphi(c)\ \forall c \in C$
$i.e. C \subseteq \{x: Re\ \varphi(x) \leq Re\ \varphi(c_0)\}$, this is called a **half space**

Separation Theorem

V finite dimensional vector space over $\mathbb{R}$ or $\mathbb{C}$
$C \subseteq V$ closed convex set, $p \in V, p \notin C$

Then $\exists \varphi \in V^*$ such that
$Re\ \varphi(p) > \sup_{c \in C} Re\ \varphi(c)$

Corollary

If C is a closed subset of V then C is the intersection of all closed half spaces which contain it.

Proof

Define $f: C \to \mathbb{R}$ by $f(c) = \|p - c\|^2$
f is continuous, $f(c) > 0$
Pick $c_1 \in C$ the closest point lies in $C \cap \overline{B_{\|p-c_1\|}(p)}$, which is closed and bounded.
So f achieves its minimum value by the extreme value theorem.
So there is at least one closest point c_0

Uniqueness

Suppose $c_0, c_1 \in C$ are both closest
$\|p - c_0\| = \|p - c_1\| = \delta \leq \|p - c\| \forall c \in C$
But then $\dfrac{c_0 + c_1}{2} \in C$ and if $c_0 \neq c_1$ then $\left\|p - \dfrac{c_0 + c_1}{2}\right\| < \delta$, by geometry
Alternatively
$$\left\|p - \frac{c_0 + c_1}{2}\right\|^2 = \left\langle \frac{p - c_0}{2} + \frac{p - c_1}{2}, \frac{p - c_0}{2} + \frac{p - c_1}{2} \right\rangle$$
$$= \left\|\frac{p - c_0}{2}\right\|^2 + 2\ Re \left\langle \frac{p - c_1}{2}, \frac{p - c_0}{2}\right\rangle + \left\|\frac{p - c_1}{2}\right\|^2 \leq \frac{1}{4}\delta^2 + 2\left\|\frac{p - c_1}{2}\right\| \left\|\frac{p - c_0}{2}\right\| + \frac{1}{4}\delta^2 = \delta^2$$
Inequality is Cauchy-Schwartz and must hold with equality
$\therefore \dfrac{p - c_1}{2} = t\dfrac{p - c_0}{2}, t > 0, but\ t = 1 \therefore c_1 = c_0$
So the closest point is unique.,

$\varphi(x) = \langle x, p - c_0\rangle$
$\varphi(p - c_0) = \|p - c_0\|^2 > 0$
$\varphi(p - c_0) = \varphi(p) - \varphi(c_0)$
$\therefore Re\ \varphi(p) = Re\ \varphi(c_0) + \|p - c_0\|^2 > Re\ \varphi(c_0)$

Claim

$Re\ \varphi(c) \leq Re\ \varphi(c_0)\ \forall c \in C$
If not, $\exists c_2 \in C$ s.t.
$Re\ \varphi(c_2) = Re\ \varphi(c_0) + \varepsilon, \quad \varepsilon > 0$
$Re\ \varphi(p - c_2) = Re\ \varphi(p) - Re\ \varphi(c_2) = Re\ \varphi(p) - Re\ \varphi(c_0) + \varepsilon = Re\ \varphi(p - c_0) - \varepsilon$
$= \|p - c_0\|^2 - \varepsilon$

Look at $f(t) = \|p - ((1-t)c_0 + tc_2)\|^2$
$= \langle (1-t)(p - c_0 + t(p - c_2), (1-t)((p - c_0) + t(p - c_2)\rangle$
$= (1-t)^2\|p - c_0\|^2 + 2\ Re(t(1-t)\langle p - c_2, p - c_0\rangle) + t^2\|p - c_2\|^2$
$= (1 - 2t - t^2)\|p - c_0\|^2 + 2(t - t^2)Re\ \varphi(p - c_2) + t^2\|p - c_2\|^2$
$= (1-t)^2\|p - c_0\|^2 - 2(t - t^2)\varepsilon + t^2\|p - c_2\|^2$

$f'(t) = -2t\|p - c_0\|^2 - (2 - 4t)\varepsilon + 2t\|p - c_2\|^2$
$f'(0) = -2\varepsilon$, decreasing
So for $t > 0$, small, $f(t) < f(0)$ so c_0 is not the smallest point. ■

Proof of Separation Theorem

Pick a basis $\{v_1, \ldots, v_n\}$ for V. Impose an inner product:
$$\left\langle \sum \alpha_i v_i, \sum \beta_i v_i \right\rangle = \sum_{i=1}^{n} \alpha_i \overline{\beta_i}$$
Use previous Theorem to get $\varphi \in V^*$ such that
$Re\ \varphi(p) > Re\ \varphi(c_0) = \sup_{c \in C} Re\ \varphi(c)$

Proof of Corollary

Let $\{A_\alpha\}$ be the set of all closed half spaces such that $H \supseteq C$
Clearly $C \subseteq \cap H_\alpha$
But if $p \notin C, \exists \varphi \in V^*\ s.t.$
$Re\ \varphi(p) > \sup_{c \in C} Re\ \varphi(c) = C$
$H = \{x: Re\ \varphi \leq L\}$ half space
$C \subseteq H, p \notin H, \therefore p \notin \bigcap H_\alpha$ ■

Normed Vector Spaces

November-28-11 9:30 AM

$F = \mathbb{R}, \mathbb{C}$

Norm

A norm on a vector space V over $\mathbb{F}$ is a function $\|\cdot\|: V \to [0, \infty)$ such that

1) $\|v\| \geq 0, \quad \|v\| = 0 \Leftrightarrow v = 0$ (positive definite)
2) $\|tv\| = |t|\|v\| \; \forall t \in \mathbb{F}$, (homogeneous)
3) $\|v + w\| \leq \|v\| + \|w\|$, (triangle inequality)

Unit Ball

B_v or $\overline{B_1(0)} = \{v: \|v\| \leq 1\}$

Proposition

$(V, \|\cdot\|)$ normed vector space then B_v is convex, $0 \in B_V$, **balanced** $(if\ v \in B_v,\ tv \in B_v\ \forall |t| = 1)$. Hence $|t| \leq 1$ by convexity.

Example

If V, W are normed vector spaces, then $\mathcal{L}(V, W)$ can be normed by

$\|T\| = \sup_{\|v\|_V \leq 1} \|Tv\|_W$

1)

$\|Tv\| \geq 0 \Rightarrow \|T\| \geq 0$

$\|T\| = 0 \Rightarrow \|Tv\| = 0\ \forall v \Rightarrow Tv = 0\ \forall v \Rightarrow T = 0$

2)

$\|tT\| = \sup_{\|v\|_V \leq 1} \|tTv\|_W = \sup_{\|v\|_V \leq 1} |t|\|Tv\|_W = |t|\|T\|$

3)

$S, T \in \mathcal{L}(V, W)$

$\|S + T\| = \sup_{\|v\|_V \leq 1} \|(S + T)v\|_W \leq \sup_{\|v\|_V \leq 1} \|Sv\|_W + \|Tv\|_W$

$\leq \sup_{\|v\| \leq 1} \|Sv\| + \sup_{\|v\| \leq 1} \|Tv\| = \|S\| + \|T\|$

Special Cases

1) $W = \mathbb{F},\ \mathcal{L}(V, \mathbb{F}) = V^*$ **dual norm** on V^*
   $\|\varphi\| = \sup_{\|v\| \leq 1} |\varphi(v)|$
2) $W = V,\ \mathcal{L}(V)$ **algebra**
   $\|ST\| \leq \sup_{\|v\| \leq 1} \|S(Tv)\| \leq \sup_{\|w\| \leq \|T\|} \|Sw\| = \|T\| \sup_{\|w\| \leq 1} \|Sw\|$
   $= \|T\| \cdot \|S\|$
3) $T \in \mathcal{L}(V, W), v \in V$
   $\|Tv\| = \left\| T\, \|v\| \left(\frac{v}{\|v\|}\right) \right\| = \|v\| \left\| T\left(\frac{v}{\|v\|}\right) \right\| \leq \|T\| \cdot \|v\|$

Lemma

V finite dimensional normal space.

Let $T: (F^n, \|\cdot\|_2) \to V$ be a linear isomorphism

Then T is uniformly continuous

Theorem

V finite dimensional normal vector space

$T: \mathbb{F}^n \to V$ linear isomorphism

Then $\exists$ constants $0 < c < C < \infty$ such that

$c\|v\| \leq \|Tv\| \leq C\|v\|\ \forall v \in V$

Equivalent

Say two norms $\|\cdot\|_a$ and $\|\cdot\|_b$ are **equivalent** if $\exists 0 < c_1, c_2$ such that

$c_1\|v\|_a \leq \|v\|_b \leq c_2\|v\|_a$

$\frac{1}{c_2}\|v\|_b \leq \|v\|_a \leq \frac{1}{c_1}\|v\|_b$

$\forall v \in V$

Corollary

If V is a finite dimensional normed vector space then any two norms on V are equivalent.

Convergence

Say a sequence $v_n \in V$ converges to v_0 if $\lim_{n \to \infty}\|v_n - v_0\| = 0$

Corollary says that convergence in a finite dimensional normal space is independent of choice of the norm.

So $(V, \|\cdot\|_a)$ and $(V, \|\cdot\|_b)$ have the same closed sets, hence the same open sets.

B_v is a closed balanced convex set containing 0 on the interior.

If $\|v_n\| \leq 1, v_n \to v_0 \Rightarrow \|v_0\| \leq 1$

$(\varepsilon > 0, \exists n\ \|v_n - v_0\| < t \therefore \|v_0\| \leq \|v_n\| + \|v_0 - v_n\| \leq 1 + \varepsilon$ Let $\varepsilon \to 0$

$\|\cdot\|$ is continuous in the norm

Examples

1

$\|v\| = \sqrt{\langle v, v \rangle}$

$V = \mathbb{C}^n$ usual inner product

B_V is unit ball in Euclidean norm

2

$V = \mathbb{C}^n, v = (a_1, \dots, a_n)$

$\|v\|_\infty = \max\{|a_i|, i \leq i \leq n\}$

Satisfies 1, 2

$\|v + w\| = \|a_1 + b_1, \dots, a_n + b_n\|_\infty = \max(|a_i + b_i|\} \leq \max(|a_i| + |b_i|)$

$\leq \max(|a_i|) + \max(|b_i|) = \|v\|_\infty + \|w\|_\infty$

$B_{l_n^\infty(\mathbb{R})} = [-1,1]^n = \{(a_i): |a_i| \leq 1\}$

$B_{l_n^\infty(\mathbb{C})} = \mathbb{D}^n = \{(a_i): |a_i| \leq 1\}$

3

$l_n^1, \quad V = \mathbb{C}^n\ or\ \mathbb{R}^n$

$\|v\|_1 = \sum_{i=1}^n |a_i|$

Satisfies 1, 2

$\|v + w\|_1 = \sum_{i=1}^n |a_i + b_i| \leq \sum_{i=1}^n |a_i| + |b_i| = \|v\|_1 + \|w\|_1$

4

$l_n^p, \quad 1 < p < \infty, V = \mathbb{F}^n$

$\|v\|_p = \left(\sum_{i=1}^n |a_i|^p\right)^{\frac{1}{p}}$

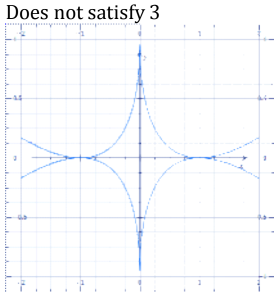
Satisfies 1, 2

Satisfies 3 but hard to prove

Ex: $p = \frac{1}{2}$

$\|v\|_{\frac{1}{2}} = \left(\sqrt{|a_1|} + \sqrt{|a_2|}\right)^2$

Does not satisfy 3



Proof of Proposition

Balanced follows from 2

Convex follows from 3, 2

$\|v\| \leq 1, \|w\| \leq 1, 0 \leq t \leq 1$

$\|tv + (1 - t)w\| \leq \|tv\| + \|(q - t)w\| \leq |t| \times 1 + |1 - t| \times 1 = 1$

Proof of Lemma

Let $e_1, \dots, e_n$ be the standard basis of $\mathbb{F}^n$.

Let $v_1 = Te_i$ this is a basis for V

$w = (a_1, \dots, a_n) = \sum_{i=1}^n e_i$

$\|Tw\| = \left\| \sum_{i=1}^n a_i v_i \right\| \leq \sum_{i=1}^n \|a_i v_i\| = \sum_{i=1}^n |a_i| \|v_i\| \leq \sqrt{\sum_{i=1}^h |a_i|^2} \sqrt{\sum_{i=1}^n \|v_i\|^2}$

$\|T\| = \sup_{\|w\| \leq 1} \|Tw\| \leq \sqrt{\sum_{i=1}^n \|v_i\|^2} = L$

$\therefore \|Tw_1 - Tw_2\| = \|T(w_1 - w_2)\| \leq \|T\|\|w_1 - w_2\| \leq L\, \|w_1 - w_2\|$

This is a Lipschitz function.

If $\varepsilon > 0, let\ \delta = \frac{\varepsilon}{L},\ \|w_1 - w_2\| < \delta \Rightarrow \|Tw_1 - Tw_2\| < L\delta = \varepsilon$

$\therefore T$ is uniformly continuous ■

Proof of Theorem

Lemma shows $C = \|T\| < \infty$

Let $S = \{w \in \mathbb{F}^n: \|w\|_2 = 1\}$, unit space

T is 1-1, s $Ts \neq 0\ \forall x \in S$

T is continuous, S is compact so by Extreme Value Theorem the minimum values is attained:

$\inf_{w \in S} \|T \cdot w\| = \|Tw_0\| = c \neq 0$

B_v is a closed balanced convex set containing 0 on the interior.
If $\|v_n\| \leq 1, v_n \to v_0 \Rightarrow \|v_0\| \leq 1$
$(\varepsilon > 0, \exists n\ \|v_n - v_0\| < t \therefore \|v_0\| \leq \|v_n\| + \|v_0 - v_n\| \leq 1 + \varepsilon$ Let $\varepsilon \to 0$
$\|\cdot\|$ is continuous in the norm

$\{v : \|v\| \geq 1\}$ is closed so $\{v : \|v\| < 1\} = B_1(0)$ is open.

T is 1-1, s $Ts \neq 0\ \forall x \in S$
T is continuous, S is compact so by Extreme Value Theorem the minimum values is attained:
$\inf_{w \in S} \|T \cdot w\| = \|Tw_0\| = c \neq 0$
By Homogeniety $c\|w\|_2 \leq \|Tw\| \leq C\|w\|_2$

Proof of Corollary

Let $T: \mathbb{F}^n \to V$ isometric
Use Theorem, get $0 < c_1, C_1, c_2, C_2$
$c_1\|w\|_2 \leq \|Tw\|_a \leq C_1\|w\|_2$
$c_2\|w\|_2 \leq \|Tw\|_b \leq C_2\|w\|_2$
$\frac{c_2}{C_1}\|Tw\|_a \leq c_2\|w\|_2 \leq \|Tw\|_b \leq C_2\|w\|_2 \leq \frac{C_2}{c_1}\|Tv\|_a$ ■

Norms

November-30-11 9:30 AM

V normed vector space

$V^* = \mathcal{L}(V, \mathbb{F})$ has the dual norm

$\|\varphi\| = \sup_{\|v\| \le 1} |\varphi(v)|$

V^{**} has a norm, $\quad i: V \to V^*, \quad \hat{v}(\varphi) = i(v)(\varphi) = \varphi(v)$

$\|\hat{v}\|_{V^{**}} = \sup_{\substack{\varphi \in V^* \\ \|\varphi\| \le 1}} |\hat{v}(\varphi)| = \sup_{\|\varphi\| \le 1} |\varphi(v)| \le \sup_{\|\varphi\| \le 1} \|\varphi\|_{V^*} \|v\|_V = \|v\|_V$

Theorem

The natural injection $i: V \to V^{**}$ is isometric.

i.e. $\|i(v)\|_{V^{**}} = \|v\|_V$

Corollary

If $v \in V$, then $\exists \varphi \in V^*$ with $\|\varphi\| \le 1$ and $\varphi(v) = \|v\|$

Quotient Norm

If M is a subspace of a finite dimensional subspace V, put the **quotient norm** on $\frac{V}{M}$ by

$\dot{v} = [v]_M = v + M = \{w : w \equiv v \bmod M\} = \{w : w - v \in M\}$

$\|\dot{v}\|_{\frac{V}{M}} = \inf_{m \in M} \|v + m\| = \inf\{\|w\| : w \in [v]\} = dist(v, M)$

Proposition

The quotient norm is a norm.

Question

If $M \subseteq V$ showed $M^* \cong \frac{V^*}{M^\perp}$, $M^\perp = \{\varphi \in V^* : \varphi|_M = 0\}$, $\left(\frac{V}{M}\right)^* \cong M^\perp$

These are linear isomorphisms.

Are they isometric when V is normed?

Lemma

If $T \in \mathcal{L}(V, W)$ is an isometric isomorphism, then $T^t \in \mathcal{L}(W^*, V^*)$ is also in isometric isomorphism.

Theorem

V finite dimensional normed space, $M \subseteq V$ subspace. Then the linear isomorphisms

$M^* \cong \left(\frac{V^*}{M^\perp}\right)$ and $\left(\frac{V}{M}\right)^* \cong M^\perp$ are isometric.

Corollary

If $M \subseteq V, f \in M^*$ then $\exists \varphi \in V^*$ s.t. $\varphi|_M = f$ and $\|\varphi\| = \|f\|$

Proof of Theorem

Have $\|\hat{v}\|_{V^{**}} \le \|v\|_V \Rightarrow \sup B_V \subseteq B_{V^{**}}$

Suppose $v \in V, \|v\| > 1$. By the separation theorem $\exists \varphi \in V^*$ such that

$Re\ \varphi(v) > \sup_{x \in B_v} Re\ \varphi(x) = \sup_{\substack{x \in B_v \\ |\lambda| = 1}} Re\ \varphi(\lambda x) = \sup_{x \in B_V} \sup_{|\lambda|=1} Re\ \lambda\varphi(x) = \sup_{x \in B_V} |\varphi(x)| = \|\varphi\|$

Let $\psi = \frac{\varphi}{\|\varphi\|}, \quad \|\psi\| = 1, \quad |\psi(v)| \ge Re\ \psi(v) > \frac{\|\varphi\|}{\|\varphi\|} = 1$

So $\|\hat{v}\| = \sup_{\|\varphi\|_{V^*} \le 1} |\hat{v}(\varphi)| \ge |\hat{v}(\psi)| > 1$

Thus $\|v\| > 1 \Rightarrow \|\hat{v}\| > 1$

$\therefore B_V \supseteq B_{V^{**}} \Rightarrow B_V = B_{V^{**}} \Rightarrow \|\hat{v}\|_{V^{**}} = \|v\|_V$

because $\|v\| = \inf\{t \ge 0 : v \in tB_V\} = \inf\{t \ge 0 : \hat{v} \in tB_{V^{**}}\}$

Proof of Corollary

$\|v\| = \|\hat{v}\| = \sup_{|\varphi| \le 1} |\hat{v}(\varphi)| = \sup_{|\varphi| \le 1} |\varphi(v)| = |\varphi_0(v)|, \quad$ attained by EVT

Choose $|\lambda| = 1$ such that $\lambda\varphi_0(v) = |\varphi_0(v)| = \|v\|$

Take $\varphi = \lambda\varphi_0$

Proof of Quotient Norm

1) $\|\dot{v}\| \ge 0, \|\dot{v}\| = 0 \Leftrightarrow dist(v, M) = 0 \Leftrightarrow v \in M \Leftrightarrow \dot{v} = \dot{0}$
2) $\|\dot{(tv)}\| = \|t\dot{v}\| = dist(tv, M) = |t| dist(v, M) = |t| \|\dot{v}\|$
3) $\|\dot{(v + w)}\| = \inf_{m \in M} \|v + w + m\| = \inf_{m_1 m_2 \in M} \|(v + m_1) + (w + m_2)\|$

   $\le \inf_{\substack{m_1 \in M \\ m_2 \in M}} \|v + m_1\| + \|w + m_2\| = \|\dot{v}\| + \|\dot{w}\|$

So $\frac{V}{M}$ has a norm ∎

Proof of Lemma

$T: V \to W$ is 1-1, onto and $\|Tv\| = \|v\|\ \forall v \in V$

$\therefore T(B_V) = B_W$. Now let $\varphi \in W^*$

$\|T^t \varphi\|_{V^*} = \sup_{v \in B_V} |(T^t\varphi)(v)| = \sup_{v \in B_V} |\varphi(Tv)| = \sup_{w \in B_W} |\varphi(w)| = \|\varphi\|_{W^*}$

So T^t is isometric

$\ker T^t = (ran\ T)^\perp = W^\perp = \{0\} \therefore 1 - 1$

$ran\ T^t = (\ker T)^\perp = \{0_V\}^\perp = V^* \therefore$ onto ∎

Proof of Theorem

Recall the quotient map $\Pi: V \to \frac{V}{M}, \ \pi(v) = \dot{v}, \ \Pi$ is onto, $\ker \Pi = M$

$\Pi^t: \left(\frac{V}{M}\right)^* \to V^*, \quad \ker \Pi^t = (ran\ \Pi)^\perp = \left(\frac{V}{M}\right)^\perp = \{0\}, \quad ran\ \Pi^t = (\ker \Pi)^\perp = M^\perp$

So Π^t maps $\left(\frac{V}{M}\right)^*$ $1 - 1$ *and onto* $M^\perp \therefore$ linear isomorphism

Take $f \in \left(\frac{V}{M}\right)^*, \quad \Pi^t f = \varphi = f \circ \Pi \in M^\perp$

$\|f\|_{\left(\frac{V}{M}\right)^*} = \sup_{\substack{\dot{v} \in \frac{V}{M} \\ \|\dot{v}\| \le 1}} |f(\dot{v})| = \sup_{\substack{v \in V \\ dist(v,M) \le 1}} |f(\Pi(v))| = \sup_{\substack{v \in V \\ dist(v,M) \le 1}} |\varphi(v)| = \sup_{\substack{v \in V \\ m \in M \\ dist(v,M) \le 1}} |\varphi(v + M)|$

If $dist(v, M) \le 1$ then $\exists m \in M$ so $\|v + m\| \le 1$ so $v \in B_V + M$

Conversely, if $v \in B_V + M$ then $dist\ (v, M) \le 1$

$\|f\|_{\left(\frac{V}{M}\right)^*} = \sup_{\substack{v \in V \\ m \in M \\ dist(v,M) \le 1}} |\varphi(v + M)| = \sup_{\substack{\|v\| \le 1 \\ m \in M}} |\varphi(v + m)| = \sup_{\|v\| \le 1} |\varphi(v)| = \|\varphi\|$

So Π^t is an isometric isomorphism of $\left(\frac{V}{M}\right)^*$ onto $M^\perp$

Apply that to $M^\perp \subseteq V^*$

$\left(\frac{V^*}{M^\perp}\right) \cong (M^\perp)^\perp \subseteq V^{**}$ which is isomorphic to $M \subseteq V$

So we have an isometric isomorphism

$J: \left(\frac{V^*}{M^\perp}\right)^* \to M$ by new lemma $J^t: M^* \to \left(\frac{V^*}{M^\perp}\right)^{**} = \frac{V^*}{M^\perp}$ ∎

Proof of Corollary

$f \in M^* \cong \frac{V^*}{M^\perp}$ is isometric isomorphism

$\exists \varphi \in V^*\ s.t. f \leftrightarrow \dot{\varphi} = \varphi + M^\perp$

So $\varphi\Big|_M = f, \|f\| = \|\dot{\varphi}\| = \inf_{\psi \in M^\perp} \|\varphi + \psi\|$

Since $\dim V \le \infty$, this inf is attainable from EVT

$\|f\| = \|\varphi + \psi_0\|, \ \varphi + \psi_0$ is the desired extensions ∎

$(\varphi + \psi_0)\Big|_M = \varphi\Big|_M + \varphi_0\Big|_M = f + 0 = f$

Norms in Matrices

December-02-11 9:53 AM

Matrix Norm

V normed finite dimensional.
A norm on $\mathcal{L}(V)$ usually should have an additional property
4) $\|ST\| \le \|S\|\|T\|$

Trace Norm

$T \in \mathcal{L}(V)$. V finite dimensional inner product space.
Polar decomposition
$T = UD$
$D = \sqrt{T^*T} \cong diag(s_1, s_2, \dots, s_n), \qquad s_1 \ge s_2 \ge \dots \ge s_n \ge 0$
S-numbers of T, $s_i = s_i(T)$

$$\|T\|_1 = \sum_{i=1}^{n} s_i(T)$$

1) $\|T\|_1 \ge 0, if\ \|T\| = 0 \Rightarrow s_i = 0 \forall i \Rightarrow D = 0 \Rightarrow T = 0$
2) $s_i(tT) = ts_i(T)$ since $tT = U(tD)$

Lemma 1

If $\{e_i\}_1^n, \{f_i\}_1^n$ are orthonormal bases for V, then

$$\sum_{i=1}^{n} |\langle Te_i, f_i\rangle| \le \|T\|_1$$

Corollary

$\|S + T\|_1 \le \|S\|_1 + \|T\|_1$
Hence $\|\cdot\|_1$ is a norm

Lemma 2

$T \in \mathcal{L}(V), \qquad 1 \le j \le n$
$s_j(T) = \inf_{rank(F) \le j-1} \|T - F\|_\infty = dist(T, \mathcal{F}_{j-1})$ matrix of rank $\le j - 1$

Corollary

If $A, T \in \mathcal{L}(V)$, then
$s_j(AT) \le \|A\|_\infty s_j(T)$,
$s_j(TA) \le \|A\|_\infty s_j(T)$

Corollary²

$A, T \in \mathcal{L}(V)$ then
$\|AT\|_1 \le \|A\|_\infty \|T\|_1 \le \|A\|_1 \|T\|_1$
$\|TA\|_1 \le \|T\|_1 \|A\|_\infty$

Therefore $\|\cdot\|_1$ is a matrix norm

Remark

Same argument shows that
$\|AT\|_2 \le \|A_\infty\| \|T\|_2, \qquad \|TA\|_2 \le \|T_2\| \|A\|_\infty$

Theorem

The dual of $(\mathcal{L}(V), \|\cdot\|_\infty)$ is $(\mathcal{L}(V), \ \|\cdot\|_1)$ via a paring
$\varphi_T(A) = Tr(AT)$

Remark 1

$\|\cdot\|_1$ is unitarily invariant
If $T \in \mathcal{L}(V), \ U, V$ unitary then $\|UTV\|_1 = \|T\|_1$

Remark 2

Ky Fan Norms

$$\|T\|_{KF_K} = \sum_{i=1}^{k} s_i(T)$$

is a unitarily invariant matrix norm

Theorem (Ky Fan)

Every unitarily invariant matrix norm on $\mathcal{M}_n$ is a convex combination of the Ky Fan norms.

Examples

1) $\|T\| = \sup_{\|v\| \le 1} \|Tv\| < \infty$ by EVT
Restrict to an inner product space $(V, \langle , \rangle)$
2) $\|T\| = \|T\|_\infty = \sup_{\|v\|=1} \|Tv\|$
Polar decomposition T, $\sqrt{T^*T} = D$ unique positive square root
D is diagonalizable. $\exists$ orthonormal basis $\{u_1, \dots, u_n\}$
$Du_i = s_i u_i \ 1 \le i \le n, \qquad s_1 \ge s_2 \ge \dots \ge s_n \ge 0$
U partial isometry, $U: ran\ D \to ran\ T$ isometrically, $T = UD$
Let $v_i = Uu_i$ $\{v_i | s_i > 0\}$ is orthonormal

$$T = \sum_{i=1}^{n} s_i v_i u_i^*$$

$$\|T\|_\infty = \sup_{\|v\|=1} \|Tv\| = \sup_{\|v\|=1} \|UDv\| = \sup_{\|v\|=1} \|Dv\| = \sup_{\substack{v=\sum a_i u_i \\ \sum|a_i|^2=1}} \left\| \sum s_i a_i u_i \right\|$$

$$= \sup_{\sum|a_i|^2=1} \sqrt{\sum_{i=1}^{n} s_i^2 |a_i|^2} = s_1 \sup_{\sum|a_i|^2=1} \sqrt{\sum |a_i|^2} = s_1$$

3) $\|T\|_2$ fix an orthonormal basis $\{e_1, \dots, e_n\} = \xi$
$T = [T]_\xi = [t_{ij}]$

Define $\|T\|_2 = \sqrt{\sum_{i,j=1}^{n} |t_{ij}|^2}$

Makes $\mathcal{M}_n$ into an inner product space

$[S] = [s_{ij}], \qquad \langle [S], [T] \rangle = \sum_{i,j=1}^{n} s_{ij} \bar{t}_{ij}$

$[T^*]_\xi = [\overline{t_{ji}}], \qquad [ST^*]_\xi = \left[\sum_{k=1}^{n} s_{ik} \overline{t_{jk}} \right]$ has $\sum_{k=1}^{n} s_{ik} \overline{t_{ik}}$ on diagonal (i, i)

$\therefore \langle [S], [T] \rangle = tr(ST^*)$

$$\|ST\|_2^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} \left| \sum_{k=1}^{n} s_{ik} t_{kj} \right|^2 \le \sum_{i=1}^{n} \sum_{j=1}^{n} \left(\sum_{k=1}^{n} |s_{ik}|^2 \right) \left(\sum_{l=1}^{n} |t_{lj}|^2 \right)$$

$$= \left(\sum_{i=1}^{n} \sum_{k=1}^{n} |s_{ik}|^2 \right) \left(\sum_{j=1}^{n} \sum_{l=1}^{n} |t_{lj}|^2 \right) = \|S\|_2^2 \|T\|_2^2$$

If U, V are unitary
$\|UTV\|_2^2 = \langle UTV, UTV \rangle = tr\big((UTV)(UTV)^*\big) = tr(UTVV^*T^*U^*) = tr(UTT^*U^*)$
$= tr(U^*UTT^*) = tr(TT^*) = \|T\|_2^2$

So $\|UTV\|_2 = \|T\|$ (unitarily invariant norm) (so is $\|T\|_\infty$)
In particular, this definition does not depend on choice of o.n. basis.

If $f_1, \dots, f_n$ o.n. basis ζ. Let $Ue_i = f_i$,
$[a_{ij}] = [T]_\zeta = U[T]_\xi U^* = U[t_{ij}]U^*$

$$\sqrt{\sum_{i,j=1}^{n} |a_{ij}|^2} = \|UTU^*\| = \|T\|^2 = \sqrt{\sum_{i,j=1}^{n} |t_{ij}|^2}$$

$T = UD$ polar decomposition, $Uu_i = v_i, \ \ 1 \le i \le k, s_k > 0, s_{k+1} = 0$
extend $v_1, \dots, v_k$ to orthonormal basis. Define $Vu_i = v_i, 1 \le i \le n$ Unitary
$T = UD = VD$

$$\|T\|_2 = \|UD\|_2 = \|D\|_2 = \sqrt{\sum_{i=1}^{n} s_i^2}, \qquad \text{where } [D]_U = diag(s_1, s_2, \dots, s_n)$$

Proof of Lemma 1

$T = UD$
Choose an orthonormal basis $\{u_i\}_1^n$ which diagonalizes D. $\qquad Du_i = s_i u_i, \ 1 \le i \le n$

Let $v_i = Uu_i, \qquad 1 \le i \le \begin{matrix} k\ if\ s_{k+1} = 0 \\ n\ if\ s_n > 0 \end{matrix}$

$$T = \sum_{j=1}^{k} s_j (v_j u_j^*)$$

$$\sum_{i=1}^{n} |\langle Te_i, f_i \rangle| = \sum_{i=1}^{n} \left| \sum_{j=1}^{k} s_j \langle e_i, u_j \rangle \langle v_j, f_i \rangle \right|$$

$$\le \sum_{j=1}^{k} s_j \sum_{i=1}^{n} |\langle e_i, u_j \rangle| |\langle v, f_i \rangle| \le_{C.S.} \sum_{j=1}^{n} s_j \sqrt{\sum_{i=1}^{n} |\langle u_j, e_i \rangle|^2} \sqrt{\sum_{i=1}^{n} |\langle v_j, f_i \rangle|^2} = \sum_{j=1}^{n} s_j \|u_j\| \|v_j\| = \sum_{j=1}^{k} s_j$$

$= \|T\|_1$

Proof of Corollary

$S + T = UE, \qquad E = |S + T| = \sqrt{(S + T)^*(S + T)}$

$$S+T=\sum_{i=1}^{n} s_i(S+T)v_i u_i^*,\qquad \{u_i\}_1^n,\{v_i\}_1^n \text{ orthonormal}$$

$$\|S+T\|_1=\sum_{i=1}^{n} s_i=\sum_{i=1}^{n}\langle (S+T)u_i,v_i\rangle \le \left|\sum_{i=1}^{n}\langle Su_i,v_i\rangle\right|+\left|\sum_{i=1}^{n}\langle Tu_i,v_i\rangle\right| \le_{Lemma\ 1} \|S\|_1+\|T\|_1$$

So $\Delta \le$ holds hence $\|\cdot\|_1$ is a norm ■

Proof of Lemma 2

Write $T=\sum_{i=1}^{n} s_i(v_i u_i^*)$, $\qquad Let\ F_j=\sum_{(i=1)}^{j-1} s_i(v_i u_i^*)\in \mathcal{F}_{j-1}$

Let $T-F_j=\sum_{i=j}^{n} s_i(v_i u_i^*)=U\,diag\{0,0,\dots,0,s_j,\dots,s_n\}$

$\|T-F_j\|=\|T-F_j\|_\infty=\max s_i(T-F_j)=s_j, \therefore dist(T,\mathcal{F}_{j-1})\le s_j$

Suppose $rank(F)\le j-1, nul(F)\ge n-(j-1)=n+1-j$
$\dim(sp\{u_1,\dots,e_n\})+nul(F)\ge j+n-(j-1)=n+1$
$\therefore \dim(sp\{u_1,\dots,u_j\}\cap \ker F)\ge 1$

Pick $x\in sp\{u_1,\dots,u_j\}\cap \ker F$, $\|x\|=1$, $\qquad x=\sum_{i=1}^{j} au_i\in \ker F$

$$\therefore \|T-F\|\ge \|(T-F)x\|=\|Tx\|=\left\|\sum_{i=1}^{j}(s_j a_i)v_i\right\|=\sqrt{\sum_{i=1}^{n} s_i^2|a_i|^2}\ge s_j\sqrt{\sum |a_i|^2}=s_j\|x\|$$

$=s_j$

Proof of Corollary

$s_j(AT)=dist(AT,\mathcal{F}_{j-1})\le \|At-Af_j\|_\infty=\|A(T-F_j)\|\le \|A\|_\infty\|T-F_j\|_\infty=\|A\|_\infty s_j(T)$
Other side is similar.

Proof of Corollary²

$$\|AT\|_1=\sum_{i-1}^{n} s_i(AT)\le \sum_{i=1}^{n}\|A\|_\infty s_i(T)=\|A\|_\infty\|T\|_1$$
Other side is similar

Proof of Theorem

Choose orthonormal basis $\xi=\{e_1,\dots,e_n\}$ matrix units E_{ij} basis for $\mathcal{L}(V)$, $1\le i,j\le n$
$\varphi\in\mathcal{L}(V)^*$, $\qquad$ Let $t_{ij}=\varphi(E_{ij})$, $\qquad$ Let $T=[t_{ji}]_\xi$
So if $[A]_\xi=[a_{ij}]$, $A\in\mathcal{L}(V)$

$$tr(AT)=\sum_{i=1}^{n}[AT]_{ii}=\sum_{i=1}^{n}\sum_{j=1}^{n}[A]_{ij}[T]_{ji}=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}t_{ij}$$

$$A=\sum a_{ij}E_{ij},\qquad \varphi(A)=\sum a_{ij}\varphi(E_{ij})=\sum a_{ij}t_{ij}$$

So $\varphi(A)=Tr(AT)=\varphi_T(A)$

$$\|\varphi\|=\sup_{\|A\|_\infty\le 1}|\varphi(A)|=\sup_{\|A\|_\infty\le 1}|Tr(AT)|$$

$$=\sup_{\|A\|_\infty\le 1}\left|\sum_{i=1}^{n}\langle ATe_i,e_i\rangle\right|\le_{\text{Lemma 1}}\sup_{\|A\|_\infty\le 1}\|AT\|_1\le_{\text{Corollary}^2}\sup_{\|A\|_\infty\le 1}\|A\|_\infty\|T\|_1=\|T\|_1$$

$T=UD$, $\qquad$ Let $A=U^*$, $\qquad \|A\|_\infty=1$
$\varphi_T(T)=Tr(U^*UD)=Tr(D)=Tr(diag(s_1,s_2,\dots,s_n))=\|T\|_1$
$\therefore \|\varphi_T\|\ge\|T\|_1\ \therefore \|\varphi_T\|=\|T\|_1$

Proof of Remark 1

$\|UTV\|_1\le \|U\|_\infty\|T\|_1\|V\|_\infty=\|T\|_1$
$\|T\|_1=\|U^*(UTV)V^*\|_1\le \|U^*\|_\infty\|UTV\|_1=\|UTV\|_1$