Lecture 19 (MA4020) LINEAR ALGEBRA

Nov 09, 2021

Section 5. Orthogonal Matrices and Rotations

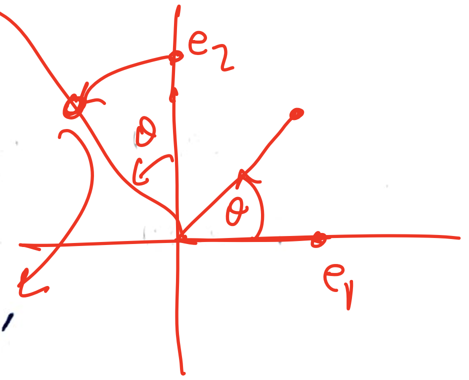
Section 6. Diagonalization

$$\rho_\theta : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$$

$$X \longmapsto \rho_\theta(X) =$$

ρ_θ := Rotation of the plane $\mathbb{R}^2$ through an angle θ

$$\rho_\theta(e_1) = \cos\theta\, e_1 + \sin\theta\, e_2,$$



$$\rho_\theta(e_2) = -\sin\theta\, e_1 + \cos\theta\, e_2,$$

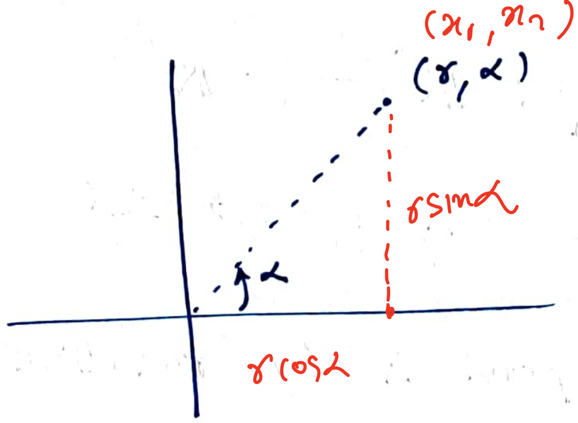
$$\begin{array}{ccc} & \rho_\theta(e_1) & \rho_\theta(e_2) \\ R = & \cos\theta & -\sin\theta \\ & \sin\theta & \cos\theta \end{array}$$

$$\boxed{R = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}} \qquad \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

Matrix of ρ_θ w.r.t. standard basis (e_1, e_2) of $\mathbb{R}^2$.

In general, Let $X \in \mathbb{R}^2$ in polar co-ordinates,

as $X = (r, \alpha)$.



$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} r\cos\alpha \\ r\sin\alpha \end{bmatrix}$$

rectangular co-ordinate

Apply the matrix (left multiplication by $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$

to a vector $X = \begin{bmatrix} r\cos\alpha \\ r\sin\alpha \end{bmatrix}$.

$$\underbrace{\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}} \begin{bmatrix} r\cos\alpha \\ r\sin\alpha \end{bmatrix}$$

$$= \begin{bmatrix} r\cos\theta\cos\alpha - r\sin\theta\sin\alpha \\ r\sin\theta\cos\alpha + r\cos\theta\sin\alpha \end{bmatrix}$$

$$= \begin{bmatrix} r\cos(\alpha+\theta) \\ r\sin(\alpha+\theta) \end{bmatrix}$$

Hence, in polar co-ordinates, we get

$$RX = (r, \alpha+\theta).$$

Thus ~~RX =~~ $\rho_\theta : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$

$$X \longmapsto \underline{RX}$$

is obtained from X by rotation through the angle θ.

$S_\theta : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$

$X \longmapsto (A)X$

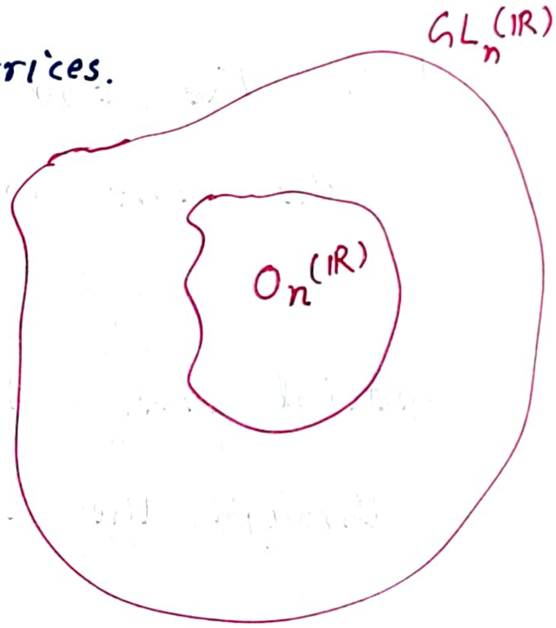
Orthogonal Matrices.

A real $n \times n$ matrix A is called orthogonal if $A^t = A^{-1}$, or equivalently, if $A^t A = I_n$.

Notation.

$$O_n = \{ A \in GL_n(\mathbb{R}) \mid A^t A = I_n \}$$

set of all orthogonal matrices.



Question. $\det(A) = ?$

$$\det(A^t A) = \det(I_n) = 1$$

$$\| $$

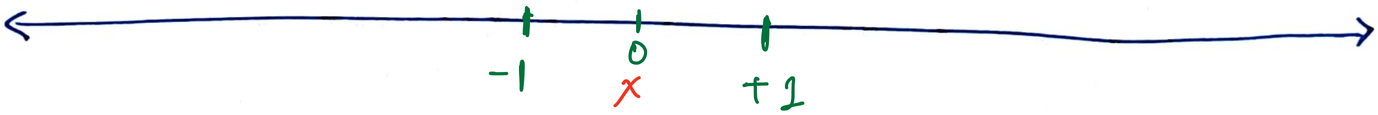
$$\det(A^t) \cdot \det(A)$$

$$\| $$

$$\det(A) \cdot \det(A)$$

$$\Rightarrow \quad \det(A)^2 = 1$$

$$\Rightarrow \quad \det(A) = \pm 1$$

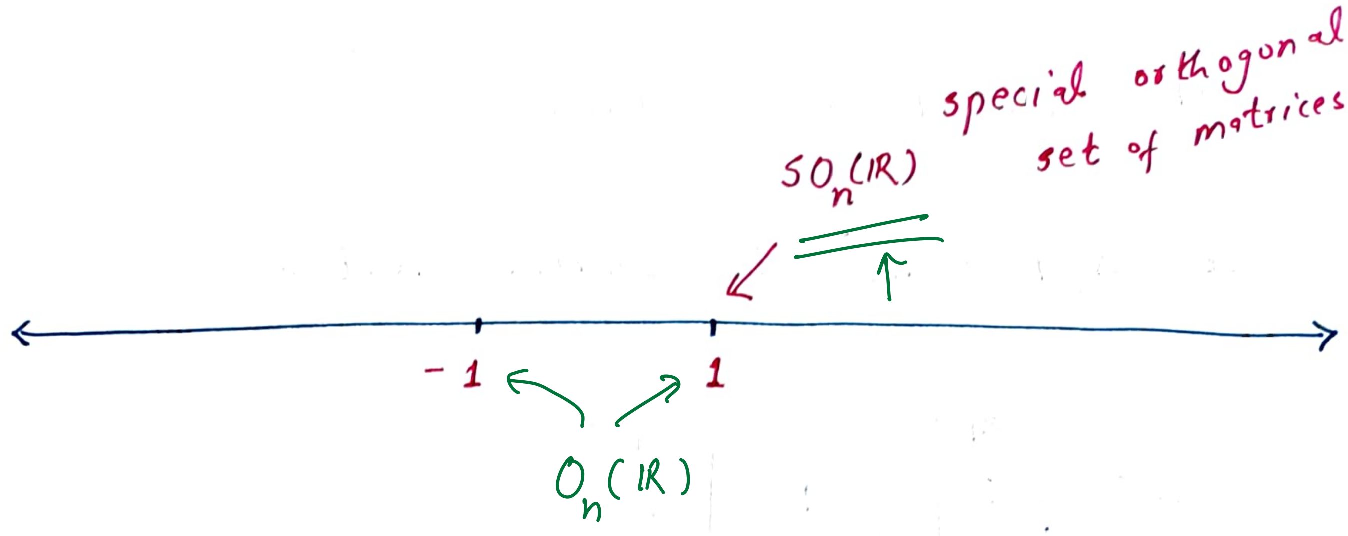


Rotation of $\mathbb{R}^3$ about the origin can be described by a pair (v, θ) consisting of a unit vector v, a vector of length 1, which lies in the axis the rotation, and a non-zero angle θ, the angle of rotation.

Note. The two pairs (v, θ) and $(-v, -\theta)$ represent the same rotation.

Special case. The matrix representing a rotation through the angle θ about the vector e_1 is

$$\left[\begin{array}{c|cc} 1 & 0 & 0 \\ \hline 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{array}\right]$$



Theorem. The rotations of $\mathbb{R}^2$ or $\mathbb{R}^3$ about the origin are the linear operators whose matrices w.r.t. standard basis are orthogonal and have determinant 1.

In other words, a matrix A represents a rotation of $\mathbb{R}^2$ (or $\mathbb{R}^3$) if and only if

$$A \in SO_2(\mathbb{R}) \quad (\text{or } SO_3(\mathbb{R})).$$

Proof:-

(Not included in this lecture)

DOT PRODUCT OF VECTORS.

Let $X, Y \in \mathbb{R}^n$ be column vectors.

$$X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad Y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

$$(X \cdot Y) = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$$

Special case: $X, Y \in \mathbb{R}^2$

$$(X \cdot X) = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1^2 + x_2^2$$

Similarly, $X \in \mathbb{R}^3$

$$(X \cdot X) = x_1^2 + x_2^2 + x_3^2$$

square of the length of the vector w.r.t. origin.

The dot product of column vectors X and Y in $\mathbb{R}^n$ is defined as

$$(X \cdot Y) = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n .$$

$\updownarrow$ same as

$$X^t Y .$$

$$(X \cdot X) = x_1^2 + \cdots + x_n^2 = |X|^2 .$$

The distance between two vectors X, Y is defined as the the length $|X - Y|$ of vector $X - Y$.

$$(X \cdot Y) = |X|\,|Y| \cos\theta \qquad (\text{in } \mathbb{R}^2)$$

where θ is the angle between the vectors.

[Proof. See Textbook , page 126]

X is orthogonal to Y if $(X \cdot Y) = 0$

$\Big\downarrow$ extends to $\mathbb{R}^n$ as definition.

Definition:

X is orthogonal to Y if $(X \cdot Y) = 0$

Proposition: The following conditions on a real $n \times n$ matrix A are equivalent:

(a) A is orthogonal. $(A^t A = I_n)$

(b) Multiplication by A preserves dot product, i.e. $(AX \cdot AY) = (X \cdot Y)$ for all column vectors $X, Y \in \mathbb{R}^n$.

(c) The columns of A are mutually orthogonal unit vectors.

Proof.

(a) $\Longrightarrow$ (b)

$$(X \cdot Y) := X^t Y = X^t A^t A Y = (AX)^t AX$$
$$= (AX \cdot AY)$$

(b) $\Rightarrow$ (a)

Since $(X \cdot Y) = (AX \cdot AY)$,

$$\Downarrow$$

$$X^t Y = X^t A^t A Y \quad \text{for all } X, Y.$$

Rewrite it as

$$X^t \underbrace{(I - A^t A)}_{\text{Call it } B = (b_{ij})} Y = 0 \quad \forall\ X, Y$$

Note that $e_i^t B e_j = b_{ij}$.

If $X^t B Y = 0$ for all X, Y, then

$$e_i^t B e_j = 0 \quad \text{for all } i, j$$

$$\Rightarrow \quad b_{ij} = 0 \quad \text{for all } i, j$$

$$\Rightarrow \quad B = 0$$

$$\Rightarrow \quad I = A^t A$$

$\Rightarrow$ A is orthogonal matrix.

Now to prove (a) $\Longleftrightarrow$ (c).

A_j : j^{th} column of A

$$A = \begin{bmatrix} & & \\ & & \\ & & \end{bmatrix}$$

The (i,j) th entry of the product matrix $A^t A$ is $(A_i \cdot A_j)$.

Thus, $A^t A = I$

$$\Longleftrightarrow (A_i \cdot A_j) = \begin{cases} 1 & \text{if } i=j \\ 0 & \text{if } i \neq j \end{cases}$$

columns of A have length 1 and are orthogonal.

Definition. A rigid motion or isometry of $\mathbb{R}^n$ is a map

$$m : \mathbb{R}^n \longrightarrow \mathbb{R}^n$$

which is distance preserving.

In other words; isometry is a map satisfying the following condition:

If X, Y are points of $\mathbb{R}^n$, then the distance from X to Y is equal to the distance from $m(X)$ to $m(Y)$:

$$|m(X) - m(Y)| = |X - Y|.$$

Such a rigid motion carries a triangle to a congruent triangle, and therefore it preserves angles and shapes in general.

Proposition. Let $m : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ be a map. The following conditions on m are equivalent:

(a) m is a rigid motion which fixes the origin.

(b) m preserves dot product; i.e.,

$$\forall\ X, Y \in \mathbb{R}^n, \quad \big(m(X) \cdot m(Y)\big) = (X \cdot Y).$$

(c) m is left multiplication by an orthogonal matrix.

Remark. A rigid motion $m : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ which fixes the origin is a linear operator.

[follows from equivalence of (a) and (c)]

Proof of proposition.

(a) $\Longrightarrow$ (b).

$$\Big[\ |m(X) - m(Y)| = |X - Y| \quad \text{and} \quad m(0) = 0\ \Big]$$

$$\Downarrow$$

$$\big(m(X) - m(Y) \cdot m(X) - m(Y)\big) = (X - Y \cdot X - Y)$$

for all $X, Y \in \mathbb{R}^n$.

$$X \cdot e_j = (m(X) \cdot m(e_j)) \quad \text{for any } X \in \mathbb{R}^n$$

If $m(e_j) = e_j$, then

$$x_j = (X \cdot e_j) = (m(X) \cdot m(e_j)) = (m(X) \cdot e_j) = m(x_j) \quad \text{for all } j$$

Hence $X = m(X)$, and hence m is the identity map.

(b) $\Rightarrow$ (c) $\leftarrow$ (m is left multiplication by an orthogonal matrix)

$\uparrow$

(m preserves dot product)

<u>Definition.</u> A basis consisting of mutually orthogonal unit vectors is called an orthonormal basis.

An orthonormal matrix is one whose columns form an orthonormal basis.

Proof of (b) $\Rightarrow$ (c).

Note that $m(e_1), \ldots, m(e_n)$ are orthonormal basis vectors.

$$\begin{cases} (m(e_i) \cdot m(e_i)) = 1 \\ (m(e_i) \cdot m(e_j)) = 0 \quad \text{if } i \neq j. \end{cases}$$

Let $\mathcal{B}' = (m(e_1), \ldots, m(e_n))$.

Let $A = [\mathcal{B}']$, then A is an orthogonal matrix.

Orthogonal matrices forms a group,
Hence A^{-1} is also orthogonal.
$\Rightarrow$ multiplication by A^{-1} preserves dot product

Thus $A^{-1}m$ ~~pro~~ preserves dot product,

and it fixes each of the basis vectors e_i.

Thus $A^{-1}m$ is the identity map.

Thus m is the left multiplication by

A.

(c) $\Rightarrow$ (a).

If $m : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a linear operator whose matrix A is orthogonal, then

$$m(X) - m(Y) = m(X-Y).$$

Hence

$$|m(X) - m(Y)| = |m(X-Y)|$$
$$= |X-Y|$$

So, m is a rigid motion.

Since a linear operator also fixes 0,

$\Rightarrow$ m is a rigid motion fixing the origin.

Proposition. Every rigid motion m is the composition of an orthogonal linear operator and a translation. In other words, it has the form $m(X) = AX + b$ for some orthogonal matrix A and some vector b.

Discussion is incomplete : See Artin page no. 128, 129, 130

Section 6. Diagonalization.

Proposition. $T : V \to V$ over $\mathbb{C}$

(a) Vector space form: Let T be a linear operator on a finite-dimensional complex vector space V. There is a basis $\mathcal{B}$ of V such that the matrix A of T is upper triangular.

(b) Matrix form. Every complex $n \times n$ matrix A is similar to an upper triangular matrix.

In other words, there is a matrix $P \in GL_n(\mathbb{C})$ such that $P \in GL_n(\mathbb{C})$ such that PAP^{-1} is upper triangular.

Proof. We have seen before equivalence of two assertions (a) and (b).

$p(t) = \det(t \cdot I - A)$

T has at least one eigenvalue, and hence it has an eigenvector, call it v_1'.

Extend v_1' to a basis $\mathcal{B}' = (v_1', \ldots, v_n')$ for V.

Then

$$A' = \begin{array}{c} \begin{matrix} T(v_1') & T(v_2') & \cdots & T(v_n') \end{matrix} \\ \left[\begin{array}{c|ccc} \lambda_1 & * & \cdots & * \\ \hline 0 & & & \\ 0 & & B & \\ \vdots & & & \\ 0 & & & \end{array}\right] \end{array} \quad B: n-1 \times n-1$$

$T(v_1') = \lambda_1 v_1'$

In Matrix form: Given any $n \times n$ matrix A, there is a $P \in GL_n(\mathbb{C})$ such that $A' = PAP^{-1}$ has the form (as above).

Now: prove by induction on n:

If $n = 1$, Nothing to prove.

Assume that the assertion is true for $n-1$, i.e., existence of some $Q \in GL_{n-1}(\mathbb{C})$ such that QBQ^{-1} is upper triangular.

Let Q_1 be the $n \times n$ matrix

$$Q_1 = \left[\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ 0 & & Q & \\ \vdots & & & \\ 0 & & & \end{array}\right]$$

Then $(Q_1 P) A (Q_1 P)^{-1}$

$$= Q_1 \underbrace{(P A P^{-1})}_{\downarrow} Q_1^{-1}$$

$$= \underline{\underline{Q_1 A' Q_1^{-1}}}$$

$$\left[\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & Q & \\ 0 & & & \end{array}\right] \left(\begin{array}{c|ccc} \lambda_1 & * & \cdots & * \\ \hline 0 & & & \\ \vdots & & B & \\ 0 & & & \end{array}\right) \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & Q & \\ 0 & & & \end{array}\right)^{-1}$$

$$= \left(\begin{array}{c|ccc} \lambda_1 & * & \cdots & * \\ \hline 0 & & & \\ \vdots & & QBQ^{-1} & \\ 0 & & & \end{array}\right),$$

upper triangular (by induction hypothesis)

Corollary.

$p(t)$ $\underset{}{\overset{\text{with}}{\text{———}}}$ triangular matrix

characteristic polynomial

$T: V \to V$

Corollary. Let F be a field.

(a) *Vector space form:* Let T be a linear operator on a finite-dimensional vector space V over F, and suppose that the characteristic polynomial of T factors into linear factor in the field F. Then there is a basis $\mathcal{B}$ of V such that the matrix A of T is triangular.

$$p(t) = (t - a_1) \cdots (t - a_n)$$

(b) *Matrix form:* Let A be an $n \times n$ matrix whose characteristic polynomial factors into linear factors in the field F. There is a matrix $P \in GL_n(F)$ such that PAP^{-1} ~~has~~ is triangular.

Proof. The proof is the same as in the previous Proposition, except that to make the induction step, one has to verify that

$$\boxed{p_B(t) = \frac{p_A(t)}{t-\lambda_1} \qquad \text{— (i)}}$$

$$\begin{cases} p_A(t) = \det(t\cdot I - A) \\ p_B(t) = \det(t\cdot I - B) \end{cases}$$

(i) holds since,

$$\det(tI - A') = \det(t\cdot I - A) = p_A(t)$$

$$\| $$

$$(t-\lambda_1)\cdot \det(t\cdot I - B)$$

$$\| $$

$$(t-\lambda_1)\cdot p_B(t) \qquad A' = \left[\begin{array}{c|c} t-\lambda_1 & * \\ \hline 0 & (tI - B) \end{array}\right]$$

So our hypothesis that the charecteristic polynomial factors into linear factors carries over from A to B.

Proposition. Let $v_1, \dots, v_r \in V$ be eigenvectors for a linear operator T, with distinct eigenvalues $\lambda_1, \dots, \lambda_r$. Then the set $(v_1, \dots, v_r)$ is linearly independent.

Proof.

~~Su~~ Induction on r:

$$\begin{pmatrix} | & & | \\ T(v_1) & \cdots & T(v_r) \\ | & & | \end{pmatrix}$$

Base case: $r = 2$.

Claim: (v_1, v_2) is L.I.

$T(v_1) = \lambda_1 v_1$

$T(v_2) = \lambda_2 v_2$

Suppose (v_1, v_2) is L.D

$$\Rightarrow \quad a_1 v_1 + a_2 v_2 = 0 \qquad \text{— (1)}$$

$$\Downarrow$$

$$T(a_1 v_1 + a_2 v_2) = T(0)$$

$$\|$$

$$a_1 \lambda_1 v_1 + a_2 \lambda_2 v_2 = 0 \qquad \text{— (2)}$$

$$\begin{array}{r} a_1 \lambda_2 v_1 + a_2 \lambda_2 v_2 = 0 \\ - \quad a_1 \lambda_1 v_1 \underline{+} a_2 \lambda_2 v_2 = 0 \\ \hline a_1 (\lambda_2 - \lambda_1) v_1 \quad = 0 \end{array} \qquad \Rightarrow \quad a_1 = 0$$

$$(\lambda_2 - \lambda_1) \neq 0, \quad v_1 \neq 0$$

$a_1 = 0$

$$\Downarrow$$

$$a_2 = 0$$

Hence (v_1, v_2) is L.I.

Assume that $(v_1, \dots, v_{r-1})$ is L.I.

We want to show $(v_1, \dots, v_{r-1}, v_r)$ is L.I.

Let $$a_1 v_1 + \dots + a_r v_r = 0 \qquad \text{--- (iii)}$$

Applying T, we get

$$a_1 \lambda_1 v_1 + \dots + a_r \lambda_r v_r = 0 \qquad \text{--- (iv)}$$

(iii) $\times \lambda_r$ − (iv), we get

$(v_1, \dots, v_{r-1})$

$$a_1 \underbrace{(\lambda_r - \lambda_1)}_{\neq 0} \cancel{v_1} v_1 + \dots + a_{r-1} \underbrace{(\lambda_r - \lambda_{r-1})}_{\neq 0} v_{r-1} = 0$$

$\lambda_r - \lambda_j \neq 0$

$$\Rightarrow \quad a_1 = \dots = a_{r-1} = 0$$

$$\Downarrow$$

$$a_r = 0 \quad (\because v_r \neq 0)$$

Thus eigenvectors $v_1, \dots, v_r$ w.r.t. distinct eigenvalues $\lambda_1, \dots, \lambda_r$ are linearly independent.

Theorem. Let T be a linear operator on a vector space V of dimension n over a field F. $p(t) = \det(tI - A) = \prod_{}^{n} (x - \lambda_i)$

Assume that its characteristic polynomial has n distinct roots in F. Then there is a basis for V w.r.t. which the matrix of T is diagonal.

Proof. The proof follows from previous discussion.

$$\begin{bmatrix} \lambda_* & & 0 \\ & \ddots & \\ 0 & & \lambda_* \end{bmatrix}$$

R

$$p(t) = \det(t \cdot I - A)$$

$$= (x - a_1)^{r_1} (x - a_2)^{r_2} \cdots (x - a_k)^{r_k}$$

The study of this case leads to "Jordon canonical form" for a matrix.

Recall:

$A = \begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix}$,

$A^{100} = \begin{bmatrix} ? & ? \\ ? & ? \end{bmatrix}$

Compute A^{100}.

$\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}^{100} = \begin{bmatrix} \lambda_1^{100} & 0 \\ 0 & \lambda_2^{100} \end{bmatrix}$

If A is similar to a diagonal matrix A', then

$$\boxed{A' = P A P^{-1}} \quad \text{for some } P \in GL_2(\mathbb{R})$$

$$\Downarrow$$

$$A = P^{-1} A' P$$

$$\Downarrow$$

$$A^{100} = (P^{-1} A' P)^{100}$$

$$= P^{-1} A' P \cdot P^{-1} A' P \cdots \cdot P^{-1} A' P \quad (100 \text{ times})$$

$$= \boxed{P^{-1} \cdot (A')^{100} P}$$

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

$= -\frac{1}{3} \begin{bmatrix} -1 & 2 \\ -1 & 1 \end{bmatrix}$

$$= -\frac{1}{3} \begin{bmatrix} -1 & -2 \\ -1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 5 & 0 \\ 0 & 2 \end{bmatrix}^{100} \begin{bmatrix} -1 & -2 \\ -1 & 1 \end{bmatrix}$$

(Put $k=100$ here) $= \frac{1}{3} \begin{bmatrix} 5^k + 2 \cdot 2^k & 2(5^k - 2^k) \\ 5^k - 2^k & 2 \cdot 5^k + 2^k \end{bmatrix}$

Corollary 6.7

$A \longrightarrow$ eigenvectors

basis consisting of eigenvectors

if $P = [\quad]^{-1}$ invertible matrix

$A' = PAP^{-1}$ diagonal