Chapter 4 Diagonalizations

4-1 Eigenvalues and Eigenvectors

Eigenvalue and eigenvector: If $Ax=\lambda x$, x is an eigenvector of A corresponding to the eigenvalue λ.

Eigenspace: $E_\lambda=\{x\in V: Ax=\lambda x\}=N(A-\lambda I)$.

Eg. $\begin{bmatrix}-2 & 0\\ 0 & 3\end{bmatrix}\cdot\begin{bmatrix}1\\0\end{bmatrix}=-2\cdot\begin{bmatrix}1\\0\end{bmatrix}$ **and** $\begin{bmatrix}-2 & 0\\ 0 & 3\end{bmatrix}\cdot\begin{bmatrix}0\\1\end{bmatrix}=3\cdot\begin{bmatrix}0\\1\end{bmatrix}$**, then -2 and 3 are the eigenvalues of** $\begin{bmatrix}-2 & 0\\ 0 & 3\end{bmatrix}$**. And** $\begin{bmatrix}1\\0\end{bmatrix}$ **and** $\begin{bmatrix}0\\1\end{bmatrix}$ **are the eigenvectors corresponding to -2 and 3, respectively.**

Theorem One scalar λ is an eigenvalue of $A\Leftrightarrow det(A-\lambda I)=0$.

(Proof) λ is an eigenvalue of $A\Leftrightarrow \exists x\in V$ fulfills $Ax=\lambda x$, $x\neq 0$.

$(A-\lambda I)x=0\Leftrightarrow A-\lambda I$ is not invertible if $x\neq 0\Leftrightarrow det(A-\lambda I)=0$

Eg. Find the eigenvalues and the corresponding eigenvectors of $A=\begin{bmatrix}-2 & 0\\ 0 & 3\end{bmatrix}$**.**

(Sol.) $det(A-\lambda I)=det(\begin{bmatrix}-2-\lambda & 0\\ 0 & 3-\lambda\end{bmatrix})=\lambda^2-\lambda-6=0$, $\lambda=-2, 3$

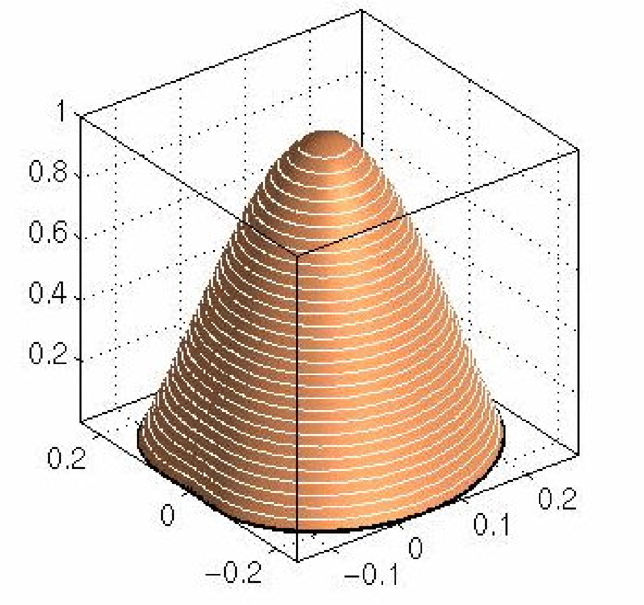
$\lambda_1=-2$, $(A-\lambda_1 I)x_1=\begin{bmatrix}0 & 0\\ 0 & 5\end{bmatrix}\cdot\begin{bmatrix}a\\b\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix}\Rightarrow x_1=\begin{bmatrix}a\\0\end{bmatrix}=a\begin{bmatrix}1\\0\end{bmatrix}$

$\lambda_2=3$, $(A-\lambda_2 I)x_2=\begin{bmatrix}-5 & 0\\ 0 & 0\end{bmatrix}\cdot\begin{bmatrix}c\\d\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix}\Rightarrow x_2=\begin{bmatrix}0\\d\end{bmatrix}=d\begin{bmatrix}0\\1\end{bmatrix}$

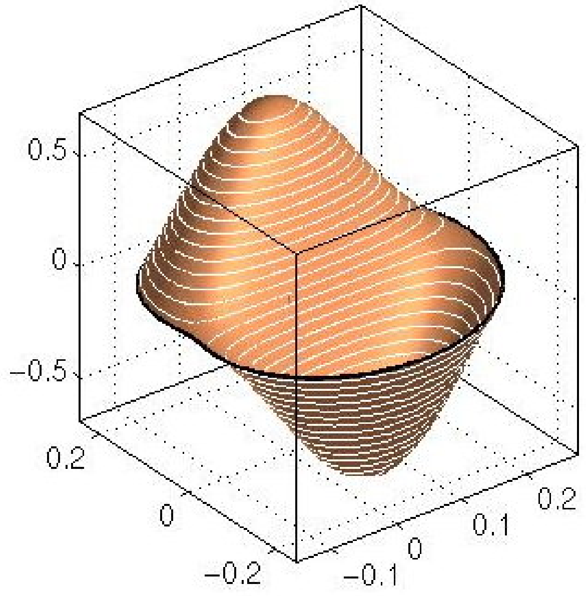
In **Matlab** language, we can use the following instructions to obtain the eigenvalues of a matrix:

```
>>A=[1,4,3;4,9,6;7,1,9]

    A =

         1     4     3

         4     9     6

         7     1     9

>>eig(A)

    ans =

       -1.0205

        5.1344

       14.8861
```

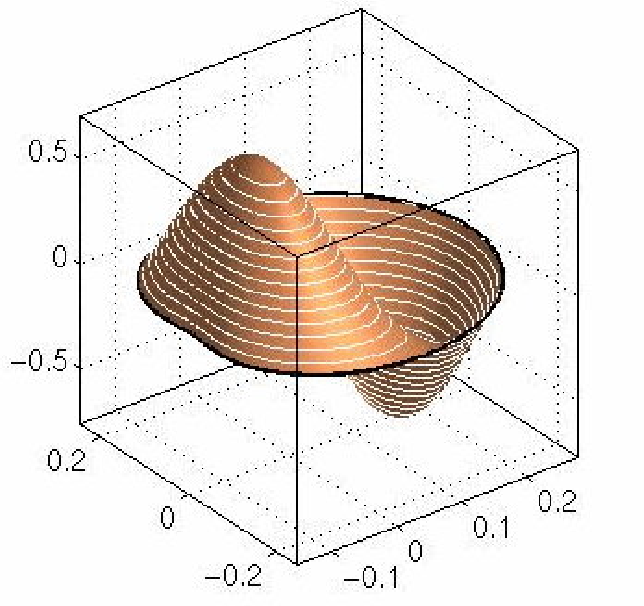
Eg. The eigenfunctions (eigenvectors) and their corresponding eigenvalues of the stationary Helmholtz equation $\frac{\partial^2\psi}{\partial x^2}+\frac{\partial^2\psi}{\partial y^2}=-k^2\psi$ are presented as follows.



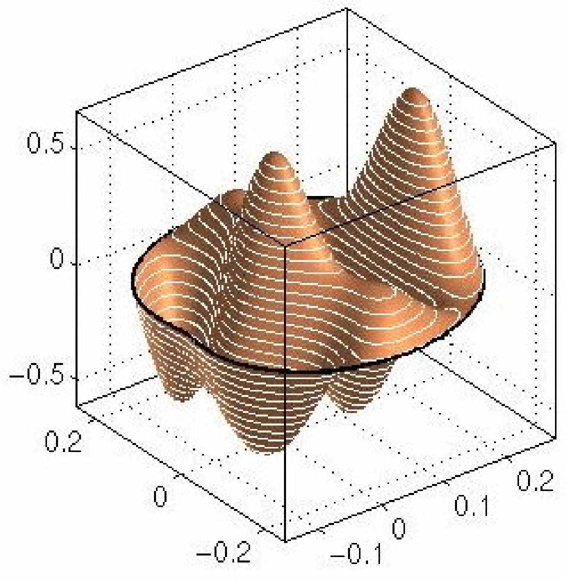
The first eigenfunction, k^2 = 106.6774



The second eigenfunction, k^2 = 254.2339

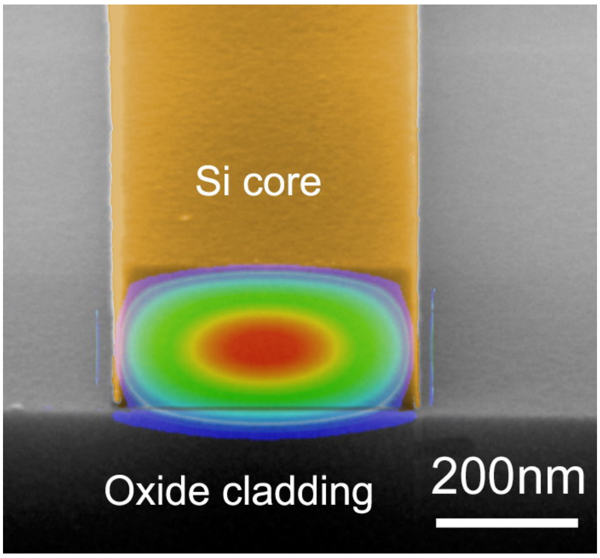


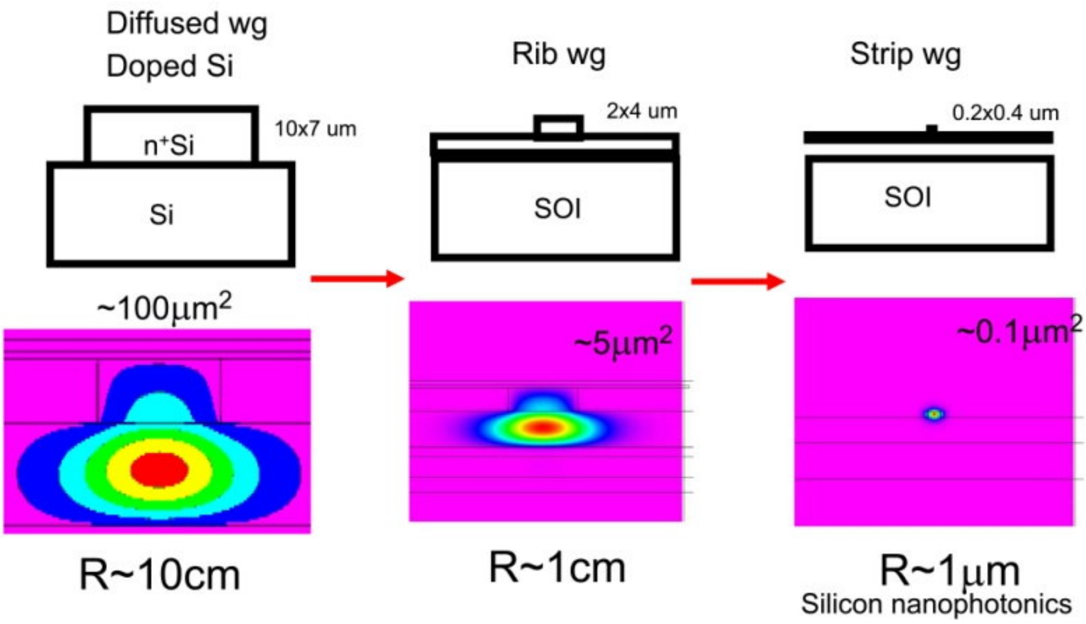
The third eigenfunction, k^2 = 286.0975

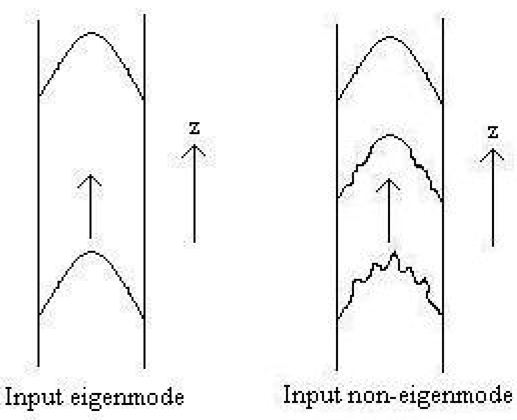


The tenth eigenfunction, k^2 = 960.0726

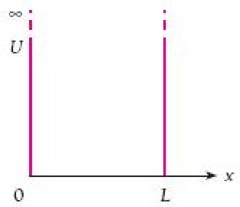
Eg. Given a refractive index distribution $n(x,y)$, the eigenmodal function $\Phi(x,y)$ of an optical waveguide fulfills $\frac{\partial^2\phi}{\partial x^2}+\frac{\partial^2\phi}{\partial y^2}+k^2n(x,y)^2\Phi=\beta^2\Phi$, where β^2 is the eigenvalue and β represents the phase constant of the lossy waveguide or the propagation constant of the lossless waveguide. The eigenmodes of some optical waveguides are presented as follows.



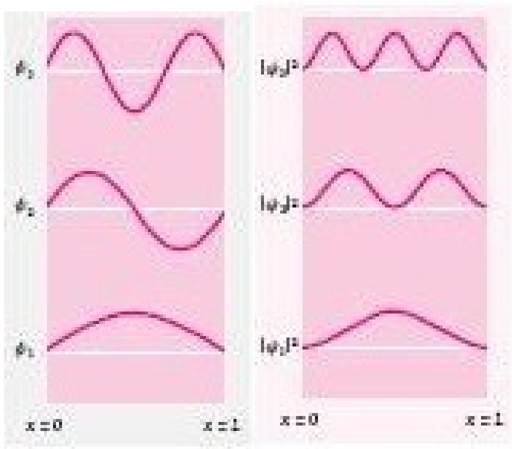


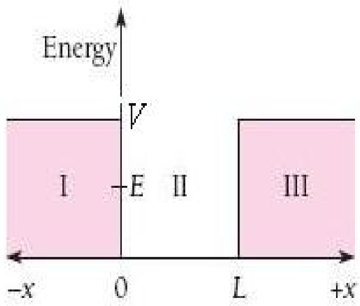


If the eigenmode is injected into an infinitely-long straight waveguide, it can propagate along the waveguide without any deformation. However, in case the input light is not an eigenmode, some optical power loss occurs and then it becomes the eigenmode gradually.



Eg. One-dimensional wave function $\Psi(x)$ in a quantum well with infinitely hard walls, $V=\begin{cases} 0, & 0 \le x \le L \\ \infty, & elsewhere \end{cases}$, fulfills $d^2\Psi/dx^2+2mE\Psi/\hbar^2=0$ in $0 \leqq x \leqq L$ with boundary conditions $\Psi(0)=\Psi(L)=0$. It can be transformed into the eigenvalue problem as $-\frac{\hbar^2}{2m}\frac{d^2\Psi}{dx^2}=E\Psi$. It is proved that the eigenvalue E is quantized as $E_n=\frac{n^2\pi^2\hbar^2}{2mL^2}$ and the corresponding eigenfunction is $\Psi_n(x)=\sqrt{2/L}\sin(n\pi x/L)$.



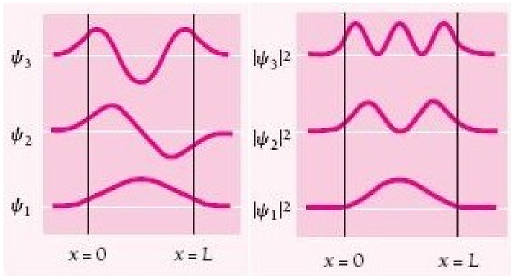


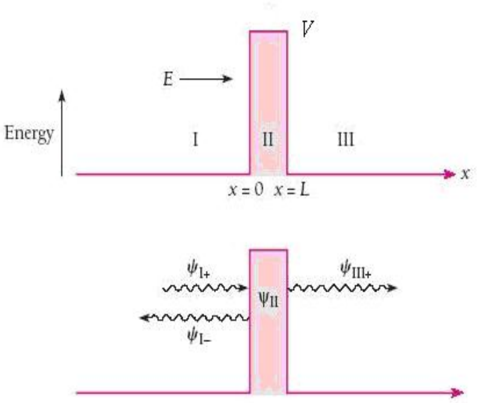
Eg. One-dimensional wave function $\Psi(x)$ in a quantum well with two finite potential walls, $V=\begin{cases} 0, & 0 \le x \le L \\ V, & elsewhere \end{cases}$, fulfills

$$\begin{cases} \frac{d^2\Psi_I}{dx^2}+\frac{2m}{\hbar^2}(E-V)\Psi_I=0 \\ \frac{d^2\Psi_{II}}{dx^2}+\frac{2m}{\hbar^2}\Psi_{II}=0 \\ \frac{d^2\Psi_{III}}{dx^2}+\frac{2m}{\hbar^2}(E-V)\Psi_{III}=0 \end{cases}$$

with boundary conditions: $\Psi_I(0)=\Psi_{II}(0)$, $\Psi_{II}(L)=\Psi_{III}(L)$, $\Psi_I'(0)=\Psi_{II}'(0)$, $\Psi_{II}'(L)=\Psi_{III}'(L)$. And the eigenfunctions have the forms as

$$\begin{cases} \Psi_I(x)=Ce^{\alpha x} \\ \Psi_{II}(x)=A\sin(\frac{\sqrt{2mE}x}{\hbar})+B\cos(\frac{\sqrt{2mE}x}{\hbar}) \\ \Psi_{III}(x)=De^{-\alpha x} \end{cases}.$$





Eg. Tunnel Effect: One-dimensional wave function Ψ(*x*) in a quantum barrier, *V*(*x*)= $\begin{cases} V(>E), & 0 \le x \le L \\ 0, & elsewhere \end{cases}$**, fulfills**

$$\begin{cases} \frac{d^2\Psi_I}{dx^2} + \frac{2m}{\hbar^2}E\Psi_I = 0 \\ \frac{d^2\Psi_{II}}{dx^2} + \frac{2m}{\hbar^2}(E-V)\Psi_{II} = 0 \\ \frac{d^2\Psi_{III}}{dx^2} + \frac{2m}{\hbar^2}E\Psi_{III} = 0 \end{cases}$$

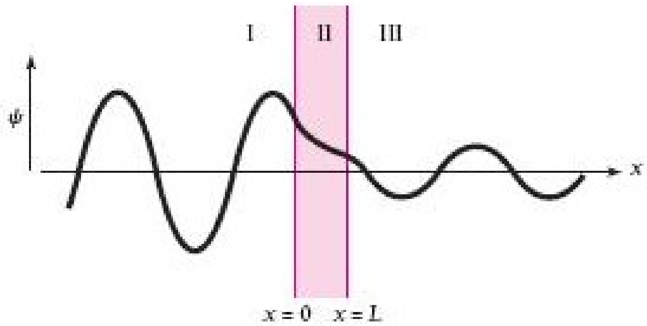
with boundary conditions:

$\Psi_I(0)=\Psi_{II}(0)$, $\Psi_{II}(L)=\Psi_{III}(L)$, $\Psi_I'(0)=\Psi_{II}'(0)$, $\Psi_{II}'(L)=\Psi_{III}'(L)$. And hence the eigenfunctions have the forms as

$$\begin{cases} \Psi_I(x) = Ae^{ik_1x} + Be^{-ik_1x} \\ \Psi_{II}(x) = Ce^{-ik_2x} + De^{-ik_2x} \\ \Psi_{III}(x) = Fe^{ik_3x} \end{cases}$$

, where $k_1=\frac{\sqrt{2mE}}{\hbar}$**,** $k_2=\frac{\sqrt{2m(V-E)}}{\hbar}$**,** $k_3=\frac{\sqrt{2mE}}{\hbar}=k_1$**. The quantum mechanics can prove that the transmission probability is**

$T=|\Psi_{III}^+|^2/|\Psi_I^+|^2=|F|^2/|A|^2 \approx [\frac{16}{4+(K_2/K_1)^2}]\cdot e^{-2k_2L} \approx e^{-2k_2L}$.



Eg. Let λ be an eigenvalue of A and x be an eigenvector belonging to λ, then $f(\lambda)$ is an eigenvalue of $f(A)$ and x be an eigenvector belonging to $f(\lambda)$. [台大電研]

(Proof) $Ax=\lambda x \Rightarrow A^n x=\lambda^n x \Rightarrow f(A)=\sum_{n=-\infty}^{\infty} a_n A^n \Rightarrow f(A)x=\sum_{n=-\infty}^{\infty} a_n A^n x=\sum_{n=-\infty}^{\infty} a_n \lambda^n x=f(\lambda)x$

Eg. Let λ be an eigenvalue of B and x be an eigenvector belonging to λ. Show that e^{λ} is an eigenvalue of e^{B} and x be an eigenvector belonging to e^{λ}. [成大電研]

(Proof) $Bx=\lambda x \Rightarrow B^n x=\lambda^n x \Rightarrow e^B=\sum_{n=0}^{\infty}\frac{1}{n!}B^n \Rightarrow e^B x=\sum_{n=0}^{\infty}\frac{1}{n!}B^n x=\sum_{n=0}^{\infty}\frac{1}{n!}\lambda^n x=e^{\lambda}x$

Eg. For $A=\begin{bmatrix}-2 & 2 & -3\\ 2 & 1 & -6\\ -1 & -2 & 0\end{bmatrix}$, what are the eigenvalues and their corresponding eigenvectors of A^2 and $3A^4+2A^2+5I$? [中正電研]

(Sol.) $\det\begin{bmatrix}-2-\lambda & 2 & -3\\ 2 & 1-\lambda & -6\\ -1 & -2 & -\lambda\end{bmatrix}=-\lambda^3-\lambda^2+21\lambda+45=-(\lambda+3)^2(\lambda-5)=0$

$\lambda_1=5$, $(A-5I)X_1=\begin{bmatrix}-7 & 2 & -3\\ 2 & -4 & -6\\ -1 & -2 & -5\end{bmatrix}\cdot\begin{bmatrix}a_1\\ b_1\\ c_1\end{bmatrix}=0 \Rightarrow x_1=\begin{bmatrix}a_1\\ b_1\\ c_1\end{bmatrix}=k_1\begin{bmatrix}1\\ 2\\ -1\end{bmatrix}$,

$\lambda_2=-3$, $(A+3I)X_2=\begin{bmatrix}1 & 2 & -3\\ 2 & 4 & -6\\ -1 & -2 & 3\end{bmatrix}\cdot\begin{bmatrix}a_2\\ b_2\\ c_2\end{bmatrix}=0 \Rightarrow x_2=k_2\begin{bmatrix}-2\\ 1\\ 0\end{bmatrix}+k_3\begin{bmatrix}3\\ 0\\ 1\end{bmatrix}$.

For A^2, $\lambda_1=5^2=25$ and $x_1=\begin{bmatrix}a_1\\ b_1\\ c_1\end{bmatrix}=k_1\begin{bmatrix}1\\ 2\\ -1\end{bmatrix}$, $\lambda_2=(-3)^2=9$ and $x_2=k_2\begin{bmatrix}-2\\ 1\\ 0\end{bmatrix}+k_3\begin{bmatrix}3\\ 0\\ 1\end{bmatrix}$.

For $3A^4+2A^2+5I$, $\lambda_1=3\cdot 5^4+2\cdot 5^2+5=1930$ and $x_1=\begin{bmatrix}a_1\\ b_1\\ c_1\end{bmatrix}=k_1\begin{bmatrix}1\\ 2\\ -1\end{bmatrix}$, $\lambda_2=3\cdot(-3)^4+2\cdot(-3)^2+5=266$ and

$x_2=k_2\begin{bmatrix}-2\\ 1\\ 0\end{bmatrix}+k_3\begin{bmatrix}3\\ 0\\ 1\end{bmatrix}$.

Eg. (a) Find out the eigenvalues and the corresponding eigenvectors of $A^{999}+2A^{998}+A^{997}$, where $A=\begin{bmatrix}1&1&0\\0&0&1\\0&0&1\end{bmatrix}$. (b) Calculate $A^{999}-3A^{998}+3A^{997}-A^{996}$.

(Sol.) $\det\begin{bmatrix}1-\lambda&1&0\\0&-\lambda&1\\0&0&1-\lambda\end{bmatrix}=-(1-\lambda)^2(-\lambda)=0$, $\lambda_1=1$ and $x_1=\begin{bmatrix}1\\0\\0\end{bmatrix}$, $\lambda_2=0$ and $x_2=\begin{bmatrix}1\\-1\\0\end{bmatrix}$

(a) For $A^{999}+2A^{998}+A^{997}$, $\lambda_1=1^{999}+2\times1^{998}+1^{997}=4$ and $x_1=\begin{bmatrix}1\\0\\0\end{bmatrix}$, $\lambda_2=0$ and $x_2=\begin{bmatrix}1\\-1\\0\end{bmatrix}$

(b) $A^3-2A^2+A=0$, $A^{999}-3A^{998}+3A^{997}-A^{996}=A^{995}(A-I)(A^3-2A^2+A)=0$

Eg. $V=M_{n\times n}(F)$, T: $V\rightarrow V$ is a linear transformation given by $T(A)=A^t$. (a) Show that T has only eigenvalues 1 and -1. (b) Find the eigenvectors corresponding to 1 and -1, respectively.

(Sol.)

(a) $T(A)=A^t$, $T^2(A)=T(A^t)=A$, $(T^2-I)(A)=0$, then $\lambda^2-1=0\Rightarrow\lambda=1$ and -1.

(b) Let E_1 be the eigenvector corresponding to 1, we have $T(E_1)=E_1=E_1^t$, then E_1 is a symmetric matrix. And the bases of the symmetric-matrix set are

$$\begin{bmatrix}1&&&\\&0&&0\\&&0&\\&0&&\ddots&\\&&&&0\end{bmatrix},\begin{bmatrix}0&0&\cdots&&0\\&1&&&\\&&0&&\\&0&&\ddots&\\&&&&0\end{bmatrix},\ldots,\begin{bmatrix}0&&&\\&0&&0\\&&\ddots&\\&0&&0\\&&&&1\end{bmatrix},$$

$$\begin{bmatrix}0&1&&&\\1&0&&0&\\&&0&&\\&0&&\ddots&\\&&&&0\end{bmatrix},\begin{bmatrix}0&0&1&0&\cdots\\0&0&0&&\\1&0&0&&\\&0&&\ddots&\\&&&&0\end{bmatrix},\cdots,\begin{bmatrix}0&0&\cdots&0&1\\\vdots&0&&0&\\&&\ddots&&\\0&0&&0&\\1&&&&0\end{bmatrix},$$

$$\begin{bmatrix}0&0&0&0&\cdots\\0&0&1&0&\\0&1&0&&\\0&0&&\ddots&\\&&&&0\end{bmatrix},\begin{bmatrix}0&0&0&0&0\\0&0&0&1&\\0&0&0&&\\0&1&&\ddots&\\&&&&0\end{bmatrix},\cdots,\begin{bmatrix}0&&&\\&0&&0&1\\&&\ddots&&\\&0&&0&\\&1&&&0\end{bmatrix},$$

$\vdots$

On the other hand, let E_2 be the eigenvector corresponding to -1, we have $T(E_2)=-E_2=E_2^t$, then E_2 is a skew-symmetric matrix. And the bases of the skew-symmetric-matrix set are

$$\begin{bmatrix} 0 & 1 & & & \\ -1 & 0 & & 0 & \\ & & 0 & & \\ & 0 & & \ddots & \\ & & & & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 & 0 & \cdots \\ 0 & 0 & 0 & & \\ -1 & 0 & 0 & & \\ & 0 & & \ddots & \\ & & & & 0 \end{bmatrix}, \cdots, \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \\ \vdots & 0 & & 0 & \\ & & \ddots & & \\ 0 & 0 & & 0 & \\ -1 & & & & 0 \end{bmatrix},$$

$$\begin{bmatrix} 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & \\ 0 & -1 & 0 & & \\ 0 & 0 & & \ddots & \\ & & & & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & \\ 0 & 0 & 0 & & \\ 0 & -1 & & \ddots & \\ & & & & 0 \end{bmatrix}, \cdots, \begin{bmatrix} 0 & & & & \\ & 0 & & 0 & 1 \\ & & \ddots & & \\ & 0 & & 0 & \\ & -1 & & & 0 \end{bmatrix},$$

$$\vdots$$

Matrices commute: $AB = BA$.

Theorem If A commutes with B, then A and B have the same eigenvectors. [2006 台大電研]

(Proof) $\forall$ x and y fulfill $Ax=\lambda_A x$ and $By=\lambda_B y$. Let $C=AB-BA=0$, and we have

$Cx=(AB-BA)x=ABx-BAx=ABx-B\lambda_A x=ABx-\lambda_A Bx= (A-\lambda_A I)Bx=0$, $A(Bx)=\lambda_A(Bx)$,

$\therefore$ Bx is an eigenvector of $A \Rightarrow Bx=\lambda_B' x \Rightarrow x$ is also an eigenvector of B.

$Cy=(AB-BA)y=ABy-BAy=A\lambda_B y-BAy=\lambda_B Ay-BAy=(\lambda_B I-B)Ay=0$, $B(Ay)=\lambda_B(Ay)$,

$\therefore$ Ay is an eigenvector of $B \Rightarrow Ay=\lambda_A' y \Rightarrow y$ is also an eigenvector of A.

4-2 Diagonalizations of Matrices

Diagonalizable matrix: $A \in M_{m\times n}(F)$ if $D=Q^{-1}AQ$ is diagonal.

Characteristic polynomial of A: $f(\lambda)=det(A-\lambda I)$.

Multiplicity: Let λ' be an eigenvalue of a linear matrix with characteristic polynomial $f(\lambda)$. The multiplicity of λ' is the largest positive number k for which $(\lambda-\lambda')^k$ is the factor of $f(\lambda)$.

Split in F: $f(x)$ in $P(F)$ if $\exists a_0, a_1, \ldots, a_n$ such that $f(x)=a_0(x-a_1)\ldots(x-a_n)$.

Theorem A is diagonalizable $\Leftrightarrow$

$$\begin{cases} 1.\ \text{The characteristic polynomial of } A \text{ splits} \\ 2.\ \text{The multiplicity of } \lambda \text{ is equal to } n-Rank(A-\lambda I) \text{ for each eigenvalue } \lambda \text{ of } A \end{cases}$$

Diagonalization of matrix A: Find a matrix Q such that $D=Q^{-1}AQ$ is diagonal. In fact, the columns of Q are the eigenvectors and the diagonal elements of A are the eigenvalues. On the other hand, $A=QDQ^{-1}$.

Theorem (a) Let $\lambda_1, \ldots, \lambda_n$ be distinct eigenvalues of A, corresponding to eigenvectors $x_1, \ldots,$ and x_n, respectively. Then $\{x_1, \ldots, x_n\}$ is linearly independent. [中央資電所]

(b) For $dim(V)=dim(A)=n$, and $\lambda_1, \ldots, \lambda_n$ are distinct, then A is diagonalizable.

(Proof) (a) By induction method on k:

$$(A-\lambda_k I)(a_1x_1+\cdots+a_kx_k)=a_1(\lambda_1-\lambda_k)x_1+\cdots+a_{k-1}(\lambda_{k-1}-\lambda_k)x_{k-1}=0$$

$\because$ $\lambda_1-\lambda_k \neq 0$, $\lambda_2-\lambda_k \neq 0$, …, $\lambda_{k-1}-\lambda_k \neq 0$, $\therefore a_1=a_2=\ldots=a_{k-1}=0$

$\Rightarrow \{x_1,\ldots,x_{k-1}\}$ is linearly independent.

Eg. Find all the eigenvectors of $A=\begin{bmatrix}1&1\\4&1\end{bmatrix}$ and diagonalize it.

(Sol.) $\det\left(\begin{bmatrix}1-\lambda&1\\4&1-\lambda\end{bmatrix}\right)=(\lambda-3)(\lambda+1)$, $\lambda=3,-1$: distinct, then A is diagonalizable.

$$\lambda_1=3\Rightarrow[A-\lambda_1 I]\cdot\begin{bmatrix}\varepsilon_1\\\varepsilon_2\end{bmatrix}=\begin{bmatrix}-2&1\\4&-2\end{bmatrix}\cdot\begin{bmatrix}\varepsilon_1\\\varepsilon_2\end{bmatrix}=0\Rightarrow x_1=\begin{bmatrix}\varepsilon_1\\\varepsilon_2\end{bmatrix}=s\begin{bmatrix}1\\2\end{bmatrix},\quad s\neq 0$$

$$\lambda_2=-1\Rightarrow[A-\lambda_2 I]\cdot\begin{bmatrix}\varepsilon_1'\\\varepsilon_2'\end{bmatrix}=\begin{bmatrix}2&1\\4&2\end{bmatrix}\cdot\begin{bmatrix}\varepsilon_1'\\\varepsilon_2'\end{bmatrix}=0\Rightarrow x_2=\begin{bmatrix}\varepsilon_1'\\\varepsilon_2'\end{bmatrix}=t\begin{bmatrix}1\\-2\end{bmatrix},\quad t\neq 0$$

$\therefore$ Eigenvectors are $\begin{bmatrix}1\\2\end{bmatrix}$, $\begin{bmatrix}1\\-2\end{bmatrix}$, and $\underbrace{\begin{bmatrix}1&1\\2&-2\end{bmatrix}^{-1}}_{Q^{-1}}\underbrace{\begin{bmatrix}1&1\\4&1\end{bmatrix}}_{A}\underbrace{\begin{bmatrix}1&1\\2&-2\end{bmatrix}}_{Q}=\underbrace{\begin{bmatrix}3&0\\0&-1\end{bmatrix}}_{D}$

Eg. Show that $A=\begin{bmatrix}3&1&0\\0&3&0\\0&0&4\end{bmatrix}$ is not diagonalizable.

(Sol.) $f(\lambda)=-(\lambda-4)(\lambda-3)^2$, $\therefore$ The multiplicity of λ_2 is 2.

$n-Rank(A-\lambda I)=3-Rank\left(\begin{bmatrix}0&1&0\\0&0&0\\0&0&1\end{bmatrix}\right)$=3-2=1≠2, $\therefore$ It is not diagonalizable.

Eg. Is $A=\begin{bmatrix}0&-2&-3\\1&3&3\\0&0&1\end{bmatrix}$ diagonalizable? If it is diagonalizable, please diagonalize it.

(Sol.) $f(\lambda)=-(\lambda-1)^2(\lambda-2)$, λ=1, 1, 2.

For λ_1=1, multiplicity = 2, n-$Rank(A-\lambda_1 I)$=3–$Rank\left(\begin{bmatrix}-1&-2&-3\\1&2&3\\0&0&0\end{bmatrix}\right)$=3–1=2, $\therefore$ A is diagonalizable.

$$(A-\lambda_1 I)x_1=0\Rightarrow x_1=\begin{bmatrix}\varepsilon_1\\\varepsilon_2\\\varepsilon_3\end{bmatrix}=s\begin{bmatrix}-2\\1\\0\end{bmatrix}+t\begin{bmatrix}-3\\0\\1\end{bmatrix}$$

For λ_2=2, multiplicity = 1, n–$Rank(A-\lambda_2 I)$=3 –$Rank\left(\begin{bmatrix}-2&-2&-3\\1&1&3\\0&0&-1\end{bmatrix}\right)$ = 3 – 2 =1

$$(A-\lambda_2 I)x_2=0\Rightarrow x_2=\begin{bmatrix}\varepsilon_1'\\\varepsilon_2'\\\varepsilon_3'\end{bmatrix}=u\begin{bmatrix}-1\\1\\0\end{bmatrix},\ \therefore\ Q=\begin{bmatrix}-2&-3&-1\\1&0&1\\0&1&0\end{bmatrix},\ Q^{-1}AQ=\begin{bmatrix}1&0&0\\0&1&0\\0&0&2\end{bmatrix}$$

Eg. For $A=\begin{bmatrix}-3 & -2\\ 4 & 3\end{bmatrix}$ and $B=\begin{bmatrix}-1 & -1\\ 4 & 3\end{bmatrix}$, which is diagonalizable? [文化電機轉學考]

(Ans.) A is diagonalizable.

Eg. If A has eigenvalues 0 and 1, corresponding to the orthonormal eigenvectors $\frac{1}{\sqrt{5}}\begin{bmatrix}1\\2\end{bmatrix}$ and $\frac{1}{\sqrt{5}}\begin{bmatrix}2\\-1\end{bmatrix}$, (a) how can you tell in advance that A is symmetric? (b)What is A? [台大電機期中考、大同電研]

(Sol.) (a) $D=\begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix}=Q^{-1}AQ \Rightarrow A=QDQ^{-1}$, where $Q=\begin{bmatrix}\frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}}\\ \frac{2}{\sqrt{5}} & \frac{-1}{\sqrt{5}}\end{bmatrix}$, $Q^{-1}=\begin{bmatrix}\frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}}\\ \frac{2}{\sqrt{5}} & \frac{-1}{\sqrt{5}}\end{bmatrix}=Q=Q^t$

$\because A^t=(QDQ^{-1})^t=(QDQ^t)^t=Q^{tt}D^tQ^t=QDQ^{-1}=A$, $\therefore$ A is symmetric.

(b) $A=\begin{bmatrix}\frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}}\\ \frac{2}{\sqrt{5}} & \frac{-1}{\sqrt{5}}\end{bmatrix}\cdot\begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix}\cdot\begin{bmatrix}\frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}}\\ \frac{2}{\sqrt{5}} & \frac{-1}{\sqrt{5}}\end{bmatrix}=\begin{bmatrix}0 & \frac{2}{\sqrt{5}}\\ 0 & \frac{-1}{\sqrt{5}}\end{bmatrix}\cdot\begin{bmatrix}\frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}}\\ \frac{2}{\sqrt{5}} & \frac{-1}{\sqrt{5}}\end{bmatrix}=\begin{bmatrix}\frac{4}{5} & \frac{-2}{5}\\ \frac{-2}{5} & \frac{1}{5}\end{bmatrix}$

Eg. For $A=\begin{bmatrix}1 & 4\\ 2 & 3\end{bmatrix}$, find A^n. [1990 交大電子所]

(Sol.) $D=\begin{bmatrix}5 & 0\\ 0 & -1\end{bmatrix}=\begin{bmatrix}\frac{1}{3} & \frac{2}{3}\\ -\frac{1}{3} & \frac{1}{3}\end{bmatrix}\cdot\begin{bmatrix}1 & 4\\ 2 & 3\end{bmatrix}\cdot\begin{bmatrix}1 & -2\\ 1 & 1\end{bmatrix}$, $D^n=\begin{bmatrix}\frac{1}{3} & \frac{2}{3}\\ -\frac{1}{3} & \frac{1}{3}\end{bmatrix}\cdot\begin{bmatrix}1 & 4\\ 2 & 3\end{bmatrix}^n\cdot\begin{bmatrix}1 & -2\\ 1 & 1\end{bmatrix}=\begin{bmatrix}5^n & 0\\ 0 & (-1)^n\end{bmatrix}$

$\therefore \begin{bmatrix}1 & 4\\ 2 & 3\end{bmatrix}^n=\begin{bmatrix}1 & -2\\ 1 & 1\end{bmatrix}\cdot\begin{bmatrix}5^n & 0\\ 0 & (-1)^n\end{bmatrix}\cdot\begin{bmatrix}\frac{1}{3} & \frac{2}{3}\\ -\frac{1}{3} & \frac{1}{3}\end{bmatrix}=\begin{bmatrix}\frac{5^n+2(-1)^n}{3} & \frac{2\cdot 5^n-2(-1)^n}{3}\\ \frac{5^n-(-1)^n}{3} & \frac{2\cdot 5^n+(-1)^n}{3}\end{bmatrix}$

Eg. Find A^{99} if $A\in M_{2\times 2}(R)$ and A has the eigenvalues 1 and –1. [台大電研]

(Sol.) $A=QDQ^{-1}$, $A^{99}=QD^{99}Q^{-1}$, $D^{99}=\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}^{99}=\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}=D$, $\therefore A^{99}=QD^{99}Q^{-1}=QDQ^{-1}=A$

Eg. Let $A=\begin{bmatrix}0 & -2\\ 1 & 3\end{bmatrix}$. Compute A^n for any positive integer n. [台大電研]

4-3 Matrix Functions

Caley-Hamilton Theorem Let $f(\lambda)$ be the characteristic polynomial of $A \in M_{n\times n}(F)$, then $f(A)=0$.

(Proof) $det(A-\lambda I)\cdot I=f(\lambda)\cdot I=adj(A-\lambda I)\cdot(A-\lambda I)$ (**Note:** $adj(B)\cdot B=det(B)\cdot I$)

Suppose $adj(A-\lambda I)=C_{n-1}\lambda^{n-1}+C_{n-2}\lambda^{n-2}+...+C_0$, where $C_i \in M_{n\times n}(F), \ \forall\, 0\le i\le n-1$

$$f(\lambda)\cdot I=(\lambda^n+a_{n-1}\lambda^{n-1}+...+a_o)\cdot I=(C_{n-1}\lambda^{n-1}+C_{n-2}\lambda^{n-2}+...+C_o)\cdot(A-\lambda I)$$

$$\begin{cases} C_{n-1}=-I \\ C_{n-2}-AC_{n-1}=-a_{n-1}I \\ C_{n-3}-AC_{n-2}=-a_{n-2}I \\ \vdots \\ C_0-AC_1=-a_1 I \\ -AC_0=-a_0 I \end{cases} \Rightarrow \begin{cases} A^nC_{n-1}=-A^n \\ A^{n-1}C_{n-2}-A^nC_{n-1}=-a_{n-1}A^{n-1} \\ \vdots \\ \vdots \\ AC_0-A^2C_1=-a_1A \\ -AC_0=-a_0I \end{cases} \Rightarrow A^n+a_{n-1}A^{n-1}+...+a_0I=0,\ \therefore f(A)=0$$

Application of Caley-Hamilton Theorem: Computing matrix functions

Eg. For $A=\begin{bmatrix}-3 & 2\\ 4 & 4\end{bmatrix}$, find $A^4+A^3-22A^2-38A+I$. [2015 台師大電機所]

(Sol.) $f(\lambda)=\lambda^2-\lambda-20=0 \Rightarrow f(A)=A^2-A-20I=0$

$g(A)=A^4+A^3-22A^2-38A+I=h(A)(A^2-A-20I)+aA+bI=aA+bI$

$$\begin{cases} g(5I)=(5I)^4+(5I)^3-22(5I)^2-38(5I)+I=11I=(5a+b)I \\ g(-4I)=(-4I)^4+(-4I)^3-22(-4I)^2-38(-4I)+I=-7I=(-4a+b)I \end{cases} \Rightarrow \begin{cases} a=2 \\ b=1 \end{cases}$$

$$\Rightarrow A^4+A^3-22A^2-38A+I=2\begin{bmatrix}-3 & 2\\ 4 & 4\end{bmatrix}+\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}=\begin{bmatrix}-5 & 4\\ 8 & 9\end{bmatrix}$$

Eg. For $B=\begin{bmatrix}-1 & -1\\ 4 & 3\end{bmatrix}$, find B^4-3B^2+2I.

(Sol.) $f(\lambda)=\lambda^2-2\lambda+1=0 \Rightarrow f(B)=B^2-2B+I=0$

$g(B)=B^4-3B^2+2I=cB+dI \Rightarrow g(I)=0=cI+dI$

$$g'(B)=4B^3-6B=cI \Rightarrow g'(I)=4I^3-6I=-2I=cI,\ \ c=-2,\ d=2 \Rightarrow g(B)=-2B+2I=\begin{bmatrix}4 & 2\\ -8 & 4\end{bmatrix}$$

Eg. $A=\begin{bmatrix}1 & 2 & 1\\ 0 & -1 & 1\\ 0 & 1 & 0\end{bmatrix}$, compute $A^8-2A^6+3A^4+A^2-6I$. [2006 台科大電子所]

(Sol.) $det(A-\lambda I)=-\lambda^3+2\lambda-1=0$, $\therefore -A^3+2A-I=0$

$$A^8-2A^6+3A^4+A^2-6I=(-A^3+2A-I)(-A^5+A^2-3A+2I)+8A^2-7A-4I=8A^2-7A-4I=\begin{bmatrix}-3 & -6 & 17\\ 0 & 19 & -15\\ 0 & -15 & 4\end{bmatrix}$$

Eg. For $A=\begin{bmatrix}1 & 1\\ 0 & 0\end{bmatrix}$, find e^A and A^{100}.

(Sol.) $f(\lambda)=\lambda(\lambda-1)=0\Rightarrow A(A-I)=0$. $g(A)=e^A=h(A)(A^2-A)+aA+bI=aA+bI$

$$\Rightarrow\begin{cases}g(I)=e^I=\begin{bmatrix}e & 0\\ 0 & e\end{bmatrix}=eI=(a+b)I\\ g(0)=e^0=I=bI\Rightarrow b=1\end{cases}\Rightarrow\begin{cases}a=e-1\\ b=1\end{cases}\Rightarrow e^A=(e-1)\begin{bmatrix}1 & 1\\ 0 & 0\end{bmatrix}+\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}=\begin{bmatrix}e & e-1\\ 0 & 1\end{bmatrix}$$

$h(A)=A^{100}=aA+bI$

$$\Rightarrow\begin{cases}h(I)=I^{100}=I=(a+b)I\\ h(0)=0^{100}=0=bI\Rightarrow b=0\end{cases}\Rightarrow\begin{cases}a=1\\ b=0\end{cases}\Rightarrow A^{100}=1\cdot\begin{bmatrix}1 & 1\\ 0 & 0\end{bmatrix}+0\cdot\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}=\begin{bmatrix}1 & 1\\ 0 & 0\end{bmatrix}$$

Eg. For $A=\begin{bmatrix}3 & -1\\ 1 & 1\end{bmatrix}$, find A^{-1}, A^3, $A^{1/2}$, and $\sin(A)$.

(Sol.) $f(\lambda)=(\lambda-2)^2=\lambda^2-4\lambda+4=0\Rightarrow A^2-4A+4I=0$

(a) $h(A)=A^{-1}=g(A)(A^2-4A+4I)+aA+bI=aA+bI$

$h'(A)=-A^{-2}=aI$

$$\begin{cases}h(2I)=\frac{1}{2}I=(2a+b)I\\ h'(2I)=-(2I)^{-2}=-\frac{1}{4}I=aI\end{cases}\Rightarrow\begin{cases}a=-\frac{1}{4}\\ b=1\end{cases}\Rightarrow A^{-1}=\frac{-1}{4}A+I=\begin{bmatrix}\frac{1}{4} & \frac{1}{4}\\ -\frac{1}{4} & \frac{3}{4}\end{bmatrix}$$

(b) $A^3=aA+bI=h(A)$, $\begin{cases}h(2I)=8I=(2a+b)I\\ h'(2I)=12I=aI\end{cases}\Rightarrow\begin{cases}a=12\\ b=-16\end{cases}\Rightarrow A^3=12A-16I=\begin{bmatrix}20 & -12\\ 12 & -4\end{bmatrix}$

(c) $\sqrt{A}=aA+bI=h(A)$, $\begin{cases}h(2I)=\sqrt{2}I=(2a+b)I\\ h'(2I)=\frac{I}{2\sqrt{2}}=aI\end{cases}\Rightarrow\begin{cases}a=\frac{\sqrt{2}}{4}\\ b=\frac{\sqrt{2}}{2}\end{cases}\Rightarrow\sqrt{A}=\frac{\sqrt{2}}{4}A+\frac{\sqrt{2}}{2}I=\frac{\sqrt{2}}{4}\begin{bmatrix}5 & -1\\ 1 & 3\end{bmatrix}$

(d) $\sin(A)=aA+bI=h(A)$, $\begin{cases}h(2I)=\sin(2)\cdot I=(2a+b)I\\ h'(2I)=\cos(2)\cdot I=aI\end{cases}\Rightarrow\begin{cases}a=\cos(2)\\ b=\sin(2)-2\cos(2)\end{cases}$

$\therefore \sin(A)=\cos(2)\cdot\begin{bmatrix}3 & -1\\ 1 & 1\end{bmatrix}+\{\sin(2)-2\cos(2)\}\cdot\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}=\begin{bmatrix}\sin(2)+\cos(2) & -\cos(2)\\ \cos(2) & \sin(2)-\cos(2)\end{bmatrix}$

Eg. Compute e^c for $C=\begin{bmatrix}3 & 1\\ -2 & 0\end{bmatrix}$. [文化電機轉學考]

(Sol.) $f(\lambda)=\lambda(\lambda-3)+2=(\lambda-1)(\lambda-2)=0\Rightarrow(C-2I)(C-I)=0$

$g(C)=e^C=aC+bI$,

$$\begin{cases}g(2I)=e^{2I}=\begin{bmatrix}e^2 & 0\\ 0 & e^2\end{bmatrix}=e^2I=(2a+b)I\\ g(I)=e^I=eI=(a+b)I\Rightarrow a+b=e\end{cases}\Rightarrow\begin{cases}a=e^2-e\\ b=2e-e^2\end{cases}\Rightarrow e^c=(e^2-e)\begin{bmatrix}3 & 1\\ -2 & 0\end{bmatrix}+(2e-e^2)\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$$

$$\Rightarrow e^c=\begin{bmatrix}-e+2e^2 & -e+e^2\\ 2e-2e^2 & 2e-e^2\end{bmatrix}$$

Eg. For $A=\begin{bmatrix}2&2&1\\1&3&1\\1&2&2\end{bmatrix}$**, find (a) the eigenvalues and the corresponding eigenvectors, (b)** e^A**=?**

[1990中山機研]

Eg. Find A^{10} **if** $A=\begin{bmatrix}1&1&1\\0&1&0\\1&1&1\end{bmatrix}$**. [台科大電研]**

Matrix limit: $\lim_{m\to\infty}(A_m)_{ij}=L_{ij}$, write $\lim_{m\to\infty}A=L$.

Theorem $A\in M_{n\times n}(C)$ **and** A **is diagonalizable. If** λ **is an eigenvalue of** A**, and** $|\lambda|<1$ **or** $\lambda=1$**, then** $\lim_{m\to\infty}A^m$ **exists.**

(Proof) $A=Q\begin{bmatrix}\lambda_1&&\\&\ddots&\\&&\lambda_n\end{bmatrix}Q^{-1}$, $A^m=Q\begin{bmatrix}\lambda_1^m&&\\&\ddots&\\&&\lambda_n^m\end{bmatrix}Q^{-1}$, $\lim_{m\to\infty}A^m=Q\begin{bmatrix}\lim_{m\to\infty}\lambda_1^m&&\\&\ddots&\\&&\lim_{m\to\infty}\lambda_n^m\end{bmatrix}Q^{-1}$

For $|\lambda_i|<1$ or $\lambda_i=1$, for i = 1, … , n, then $\lim_{m\to\infty}\lambda_i^m<\infty$, $\therefore$ $\lim_{m\to\infty}A^m$ exits.

Theorem If $D=P^{-1}AP$**, then** $e^A=Pe^DP^{-1}$**. Note:** $e^{A+B}\neq e^A\cdot e^B$ for some $A, B\in M_{n\times n}(F)$.

Eg. $A=\begin{bmatrix}0&1\\0&0\end{bmatrix}$, $B=\begin{bmatrix}0&0\\1&0\end{bmatrix}$, $e^A=\begin{bmatrix}1&1\\0&1\end{bmatrix}$, $e^B=\begin{bmatrix}1&0\\1&1\end{bmatrix}$

$$e^A\cdot e^B=\begin{bmatrix}2&1\\1&1\end{bmatrix}\neq e^{A+B}=e^{\begin{bmatrix}0&1\\1&0\end{bmatrix}}=\begin{bmatrix}\frac{1}{2}(e+e^{-1})&\frac{1}{2}(e-e^{-1})\\\frac{1}{2}(e-e^{-1})&\frac{1}{2}(e+e^{-1})\end{bmatrix}$$

Transition (Stochastic) matrix: (a) All the column entries are nonnegative. (b) The sum of the column entries is 1.

Regular transition matrix: All entries are positive in some power of a transition matrix.

Eg. $\begin{bmatrix}0.90&0.02\\0.10&0.98\end{bmatrix}$ **is a regular transition matrix.** $\begin{bmatrix}1&0.4&0.1&0\\0&0.3&0.5&0\\0&0&0.2&0\\0&0.3&0.2&1\end{bmatrix}$ **is a transition matrix but not a regular transition matrix.**

Probability vector: A column vector containing nonnegative entries whose sum is 1.

Theorem (a) The product of two transition matrices is a transition matrix.

(b) The product of a transition matrix and a probability vector is a probability vector.

(c) If λ is an eigenvalue of a transition matrix, then $|\lambda| \le 1$

(d) Every transition matrix has 1 as an eigenvalue. [台大電研]

(Proof)

(a) For $C=AB$, $\sum_i C_{ik} = \sum_i \sum_j A_{ij} B_{jk} = \sum_j \sum_i A_{ij} B_{jk} = \sum_j B_{jk} \sum_i A_{ij} = \sum_j B_{jk} = 1$

(b) For $y = Ax$, $\sum_i y_i = \sum_i \sum_j A_{ij} x_j = \sum_j x_j \sum_i A_{ij} = \sum_j x_j = 1$

(d) $u = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$, $\because$ $M^t u = Mu = 1u$, $\therefore$ 1 is an eigenvalue of M^t or M.

Invariant subspace: T is a linear operator on V. A subspace W of V is called a T-invariant subspace of V if $T(W) \subseteq W$, that is, if $T(x) \in W$ for all $x \in W$.

Theorem T is a linear operator on V, and $V=W_1 \oplus W_2 \oplus W_3 \cdots \oplus W_k$, where W_i is a T-invariant subspace of V, $\forall 1 \le i \le k$. Let $\begin{cases} f(\lambda) \text{ denote the characteristic polynomial of } V \\ f_i(\lambda) \text{ denote the characteristic polynomial of } W_i \end{cases}$, then $f(\lambda)=f_1(\lambda)f_2(\lambda)\cdots f_k(\lambda)$.

(Proof) By induction in k:

1. $k=2$, Let $\begin{cases} \beta_1 = \{x_1, ..., x_l\} \text{ denote the basis for } W_1 \\ \beta_2 = \{x_{l+1}, ..., x_n\} \text{ denote the basis for } W_2 \end{cases}$

$T(W_1) \subseteq W_1$, $T(W_2) \subseteq W_2$, and $V = W_1 \oplus W_2$, $\beta = \beta_1 \cup \beta_2$

$$\begin{cases} T(x_1) \in \text{span}(\{x_1, ..., x_l\}) \\ T(x_2) \in \ldots\ldots \\ \quad \vdots \\ T(x_l) \in \ldots\ldots \end{cases} \text{ and } \begin{cases} T(x_{l+1}) \in \text{span}(\{x_{l+1}, ..., x_n\}) \\ T(x_{l+2}) \in \ldots\ldots \\ \quad \vdots \\ T(x_n) \in \ldots\ldots \end{cases}$$

i. e., $$\begin{cases} T(x_1) = a_{11}x_1 + a_{21}x_2 + \ldots + a_{l1}x_l \\ T(x_2) = a_{12}x_1 + a_{22}x_2 + \ldots + a_{l2}x_l \\ \quad \vdots \\ T(x_l) = a_{1l}x_1 + a_{2l}x_2 + \ldots + a_{ll}x_l \end{cases} \text{ and } \begin{cases} T(x_{l+1}) = a_{l+1,1}x_{l+1} + a_{l+1,2}x_{l+2} + \ldots + a_{n1}x_n \\ T(x_{l+2}) = a_{l+2,1}x_{l+1} + a_{l+2,2}x_{l+2} + \ldots + a_{n2}x_n \\ \quad \vdots \\ T(x_n) = a_{n,1}x_{l+1} + a_{n,2}x_{l+2} + \ldots + a_{nn}x_n \end{cases}$$

$$\Rightarrow [T]_\beta = A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1l} & 0 & \cdots & 0 \\ a_{21} & & & \vdots & \vdots & & \vdots \\ \vdots & & & \vdots & \vdots & & \vdots \\ a_{l1} & \cdots & \cdots & a_{ll} & 0 & \cdots & 0 \\ 0 & \cdots & \cdots & 0 & a_{l+1,1} & \cdots & a_{l+1,n} \\ \vdots & & & & & & \\ 0 & \cdots & \cdots & 0 & a_{n1} & \cdots & a_{nn} \end{bmatrix} = \begin{bmatrix} B_1 & O \\ O & B_2 \end{bmatrix}$$

$$\therefore \quad f(\lambda) = \det(A - \lambda I) = \det(B_1 - \lambda I) \cdot \det(B_2 - \lambda I) = f_1(\lambda) \cdot f_2(\lambda)$$

2. Assume that the theorem is true for k-1 summation.
3. Let $W = W_1 + \ldots + W_{k-1}$, then $V = W \oplus W_k$ by the proof in (2) $\Rightarrow f(\lambda) = g(\lambda) \cdot f_k(\lambda)$ in which $g(\lambda) = f_1(\lambda) \cdot \ldots \cdot f_{k-1}(\lambda)$

Direct sum of matrices: $[A] \oplus [B] \oplus [C] = \begin{bmatrix} A & 0 & 0 \\ 0 & B & 0 \\ 0 & 0 & C \end{bmatrix}$

Eg. $\begin{bmatrix} 1 & 2 \\ 3 & -4 \end{bmatrix} \oplus \begin{bmatrix} 1 & 1 \\ -2 & -2 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 0 & 0 \\ 3 & -4 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & -2 & -2 \end{bmatrix}$