ECE 707: Control Systems Design (2)

T. Kirubarajan

Department of Electrical and Computer Engineering
McMaster University
Hamilton, Ontario, Canada

These viewgraphs are based on the text
"Linear System: Theory and Design" by Chi-Tsong Chen
Oxford University Press, 1999.

Norm of $\underline{x}$, $||\underline{x}||$

$||\underline{x}||$ is a real scalar that must satisfy:

(i) $||\underline{x}|| > 0 \;\Leftrightarrow\; \underline{x} \neq \underline{0}$; $||\underline{x}|| = 0 \Leftrightarrow \; \underline{x} = \underline{0}$.

(ii) $||\alpha\underline{x}|| = |\alpha| \cdot ||\underline{x}||$.

(iii) $||\underline{x} + \underline{y}|| \leq ||\underline{x}|| + ||\underline{y}||$ (triangle inequality).

Common Norms for $\underline{x} \in (\Im^n, \Im)$:

(1) L_p norm: $||\underline{x}||_p = [\sum_{i=1}^{n} |x_i|^p]^{1/p}$

(2) L_1 norm: $||\underline{x}||_1 = \sum_{i=1}^{n} |x_i|$

(3) L_2 norm: $||\underline{x}||_2 = \left[\sum_{i=1}^{n} |x_i|^2\right]^{1/2}$ (Euclidian norm)

(4) L_∞ norm: $||\underline{x}||_\infty = \max_i |x_i|$

Note: LVS together with an appropriate norm="normed LVS".

Inner product $<\underline{x},\underline{y}>$

$<\underline{x},\underline{y}>$ must satisfy:

(i) $<\underline{x},\underline{y}>^*=<\underline{y},\underline{x}>$

(ii) $<\underline{x},\underline{x}>\ >0 \ \Leftrightarrow \underline{x}\neq \underline{0}$; $<\underline{x},\underline{x}>=0 \ \Leftrightarrow \underline{x}=\underline{0}$

(Note that $<\underline{x},\underline{x}>=<\underline{x},\underline{x}>^*$ is real)

(iii) $<\underline{x},\alpha_1\underline{y}_1+\alpha_2\underline{y}_2>=\alpha_1<\underline{x},\underline{y}_1>+\alpha_2<\underline{x},\underline{y}_2>$

Example: $<\underline{x},\underline{y}>=\underline{x}^{*T}\underline{y}=\sum_{i=1}^{n} x_i^* y_i$

(1) $<\underline{x},\underline{y}>^*=(\underline{x}^{*T}\underline{y})^*=\underline{x}^T\underline{y}^*=\underline{y}^{*T}\underline{x}=<\underline{y},\underline{x}>$

(2) $<\underline{x},\underline{x}>=\underline{x}^{*T}\underline{x}=\sum_{i=1}^{n}|x_i|^2 \begin{cases} >0, & \underline{x}\neq\underline{0} \\ 0, & \underline{x}=\underline{0} \end{cases}$

(3) $<\underline{x},\alpha_1\underline{y}_1+\alpha_2\underline{y}_2>=\underline{x}^{*T}(\alpha_1\underline{y}_1+\alpha_2\underline{y}_2)=\alpha_1\underline{x}^{*T}\underline{y}_1+\alpha_2\underline{x}^{*T}\underline{y}_2$

(i)-(iii) in general yield:

(a) $< \underline{x}, \alpha \underline{y} >= \alpha < \underline{x}, \underline{y} >$

Proof: Use (iii) with $\alpha_2 = 0$.

(b) $< \alpha \underline{x}, \underline{y} >= \alpha^* < \underline{x}, \underline{y} >$

Proof: $< \alpha \underline{x}, \underline{y} >=< \underline{y}, \alpha \underline{x} >^* = \alpha^* < \underline{y}, \underline{x} >^* = \alpha^* < \underline{x}, \underline{y} >$.

(c) $< \alpha_1 \underline{x}_1 + \alpha_2 \underline{x}_2, \underline{y} >= \alpha_1^* < \underline{x}_1, \underline{y} > + \alpha_2^* < \underline{x}_2, \underline{y} >$

Proof: Use the previous rule.

(d) $$< \alpha_1 \underline{x}_1 + \alpha_2 \underline{x}_2, \beta_1 \underline{y}_1 + \beta_2 \underline{y}_2 >= \alpha_1^* \beta_1 < \underline{x}_1, \underline{y}_1 > + \alpha_1^* \beta_2 < \underline{x}_1, \underline{y}_2 > + \alpha_2^* \beta_1 < \underline{x}_2, \underline{y}_1 > + \alpha_2^* \beta_2 < \underline{x}_2, \underline{y}_2 >$$

Proof: Homework!!

Schwarz Inequality: $|<\underline{x},\underline{y}>| \leq <\underline{x},\underline{x}>^{1/2}<\underline{y},\underline{y}>^{1/2}$

Proof: Let $\underline{y} \neq 0$.

Define $L =< \underline{x}+\lambda\underline{y},\underline{x}+\lambda\underline{y}> \geq 0$

$$=<\underline{x},\underline{x}> +\lambda <\underline{x},\underline{y}> +\lambda^* <\underline{y},\underline{x}> +\lambda\lambda^* <\underline{y},\underline{y}>$$

Let us choose $\lambda = - <\underline{y},\underline{x}> / <\underline{y},\underline{y}> = - <\underline{x},\underline{y}>^* / <\underline{y},\underline{y}>$

Then $L =<\underline{x},\underline{x}> \; - |<\underline{x},\underline{y}>|^2 / <\underline{y},\underline{y}> \geq \; 0$.

$\Rightarrow \; |<\underline{x},\underline{y}>| \leq <\underline{x},\underline{x}>^{1/2}<\underline{y},\underline{y}>^{1/2}$ Since $\underline{y} \neq 0$

Triangle Inequality: $||\underline{x}+\underline{y}|| \; \leq \; ||\underline{x}|| + ||\underline{y}||$

Proof: (for $||\underline{x}|| =<\underline{x},\underline{x}>^{1/2}$)

$$||\underline{x}+\underline{y}||^2 =<\underline{x}+\underline{y},\underline{x}+\underline{y}>=<\underline{x},\underline{x}> + <\underline{x},\underline{y}> + <\underline{y},\underline{x}> + <\underline{y},\underline{y}>$$

$$= ||\underline{x}||^2 + ||\underline{y}||^2 + 2\,\text{Re} <\underline{x},\underline{y}>$$

$(<\underline{x},\underline{y}> + <\underline{x},\underline{y}>= 2\,\text{Re} <\underline{x},\underline{y}>)$

$$\leq ||\underline{x}||^2 + ||\underline{y}||^2 + 2|<\underline{x},\underline{y}>| \quad (\text{Re}(a+jb) = a \leq |a+jb|)$$

$$\leq ||\underline{x}||^2 + ||\underline{y}||^2 + 2||\underline{x}|| \cdot ||\underline{y}|| = (||\underline{x}|| + ||\underline{y}||)^2$$

$\Rightarrow \; ||\underline{x}+\underline{y}|| \leq ||\underline{x}|| + \underline{y}||$ as required.

Matrix Norms: $||A||$

$||A||$ must have same properties as matrix norms:

(i) $||A|| \geq 0$ and real; $||A|| = 0 \;\Leftrightarrow\; A = [0]$
(ii) $||\alpha A|| = |\alpha| \cdot ||A||$
(iii) $||A + B|| \leq ||A|| + ||B||$

Also desirable for norms to be "consistent", i.e., $||AB|| \leq ||A|| \cdot ||B||$

Frobenious Norm: $||A||_F = || \sum_i \sum_j |a_{ij}|^2 ||^{1/2}$

Extension of L_2 norm to matrices.

If $A = [\underline{a}_1 \; \underline{a}_2 \cdots \underline{a}_n],\; ||A||_F^2 = \sum_{i=1}^m \sum_{j=1}^n a_{ij}^* a_{ij} = \sum_{j=1}^n \underline{a}_j^{*T} \underline{a}_j =$ Tr $\left[A^{*T} A\right]$

This norm is consistent.

Extension of L_∞ Norm: $||A||_\infty = \max_{i,j} |a_{ij}|$

This is not consistent and rarely used.

Example: $A = B = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, $||A||_\infty = ||B||_\infty = 1$

$AB = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}$, $||AB||_\infty = 2 > ||A||_\infty \cdot ||B||_\infty = 1$.

Induced Norms

Many matrix norms are said to be "induced" by a vector norm according to

$$||A|| = \max_{\underline{x} \neq \underline{0}} \frac{||A\underline{x}||}{||\underline{x}||} = \max_{\underline{x}} ||A\underline{x}|| \text{ with } ||\underline{x}|| = 1$$

This norm insures:

(1) $||A|| \geq \frac{||A\underline{x}||}{||\underline{x}||} \quad \Rightarrow \quad ||A\underline{x}|| \leq ||A|| \cdot ||\underline{x}||$

(2) $||A\underline{x}|| = ||A|| \cdot ||\underline{x}||$ for some $\underline{x} \in X$.

(3) $||AB|| = \max_{\underline{x}} \frac{||A(B\underline{x})||}{||\underline{x}||} \leq \max_{\underline{x}} \frac{||A|| \cdot ||B\underline{x}||}{||\underline{x}||} = ||A|| \cdot ||B||$, i.e., consistent

Projection Matrix, P

A projection matrix has following properties:

(1) $P = P^{*T}$ (2) $P^2 = P$

$I - P$ is also a projection matrix

Proof: $(I-P)^{*T} = I - P^{*T} = I - P$ and
$(I-P)^2 = I - 2P + P^2 = I - P$

Projection Theorem

Let $\underline{x}_1, \underline{x}_2, \cdots, \underline{x}_n$ be a basis for $\Im^n$.
A subset of the basis vector $\underline{x}_1, \underline{x}_2, \cdots, \underline{x}_m$, generates a LVS $M \subset \Im^n$.
Any vector $\underline{y} \in \Im^n$ can be expressed as $\underline{y} = \underline{y}_n + \underline{y}_p$, where $\underline{y}_p \in M$ and $\underline{y}_n = \underline{y} - \underline{y}_p$ is in a space normal to M, $M^{\perp}$; where $\Im^n = M \cup M^{\perp}$.
Furthermore, $\underline{y}_p = P\underline{y},\ P = A(A^{*T}A)^{-1}A^{*T},\ A = [\underline{x}_1\ \underline{x}_2 \cdots \underline{x}_m]$.
Also, $\underline{y}_n = (I-P)\underline{y}, \quad \underline{y}_p^{*T}\underline{y}_n = \underline{y}^{*T}P^{*T}(I-P)\underline{y} = 0$.
And, $\underline{y}_p + \underline{y}_n = P\underline{y} + (I-P)\underline{y} = \underline{y}$.

Special Case: $\underline{x}_1, \underline{x}_2, \cdots, \underline{x}$ "orthonormal"

Here $\underline{x}_i^{*T}\underline{x}_j = \delta_{ij} = \begin{cases} 1, & i=j \\ 0 & \text{else} \end{cases} \quad \Rightarrow \; A^{*T}A = [\underline{x}_i^{*T}\underline{x}_j] = I_m$

$\Rightarrow P = AA^{*T} = \sum_{i=1}^{m} \underline{x}_i \underline{x}_I^{*T}$.

Example: $\underline{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \; \underline{x}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \; \underline{y} = \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix}$

$M =$ Space generated by $\underline{x}_1, \underline{x}_2 = R(A)$ where

$A = [\underline{x}_1 \; \underline{x}_2] = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}$.

Hence $A^TA = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}, \; A(A^TA)^{-1}A^T = P = \begin{bmatrix} 2/3 & 1/3 & -1/3 \\ 1/3 & 2/3 & 1/3 \\ -1/3 & 1/3 & 2/3 \end{bmatrix}$

$\underline{y}_p = P\underline{y} = A(A^TA)^{-1}A^T\underline{y} = \begin{bmatrix} 4/3 \\ 8/3 \\ 4/3 \end{bmatrix}; \;\; \underline{y}_n = \underline{y} - \underline{y}_p = \begin{bmatrix} 2/3 \\ -2/3 \\ 2/3 \end{bmatrix}$

$\underline{y}_p^T \underline{y}_n = 0$ as required.

Forming an Orthogonal Basis – Gram Schmidt Procedure

Given a basis $\{\underline{u}_1, \underline{u}_2, \cdots, \underline{u}_n\}$ for $\Im^n$, this algorithm constructs an alternative basis $\{\underline{w}_1, \underline{w}_2, \cdots, \underline{w}_n\}$ where $\underline{w}_i^{*T}\underline{w}_j = \delta_{ij}$. The steps are

(1) Let $\underline{w}_1 = \underline{u}_1/||\underline{u}_1||, \;\;$ where $||\underline{x}|| = (\underline{x}^{*T}\underline{x})^{1/2}$.

(2) Set $k = 1$.

(3) Find $\alpha_{ki} = \underline{w}_i^{*T}\underline{u}_{k+1}, \;\; i = 1, 2, \cdots, k$

(4) Get $\underline{v}_{k+1} = \underline{u}_{k+1} - \sum_{i=1}^{k} \alpha_{ki}\underline{w}_i$.

Note: $\underline{v}_{k+1} = \underline{u}_{k+1} - \sum_{i=1}^{k} \alpha_{ki}\underline{w}_i = \left[I - \sum_{i=1}^{k} \underline{w}_i\underline{w}_i^{*T}\right]\underline{u}_{k+1}$

$\Rightarrow \; \underline{v}_{k+1} = (I - P_{w_{(k)}})\underline{u}_{k+1} =$

$\underline{u}_{k+1}$ – projection of $\underline{u}_{k+1}$ on $\{\underline{w}_1, \underline{w}_2, \cdots, \underline{w}_k\}$.

(5) Finally $\underline{w}_{k+1} = \underline{v}_{k+1}/||\underline{v}_{k+1}||$.

(6) Set $k = k + 1$.

(7) If $k = n$ done; else goto (3).

Example: $\underline{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, $\underline{u}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$, $\underline{u}_3 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$

$$\underline{w}_1 = \underline{u}_1 / (\underline{u}_1^T \underline{u}_1)^{1/2} = \frac{1}{\sqrt{2}} \underline{u}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

$$\alpha_{11} = \underline{w}_1^{*T} \underline{u}_2 = \frac{1}{\sqrt{2}} \Rightarrow \underline{v}_2 = \underline{u}_2 - \frac{1}{\sqrt{2}} \underline{w}_1 = \begin{bmatrix} -1/2 \\ 1/2 \\ 1 \end{bmatrix}.$$

$$\underline{w}_2 = \underline{v}_2 / ||\underline{v}_2|| = \frac{1}{\sqrt{6}} \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}.$$

$\alpha_{21} = \underline{w}_1^{*T} \underline{u}_3 = \frac{1}{\sqrt{2}}$ and $\alpha_{22} = \underline{w}_2^{*T} \underline{u}_3 = \frac{1}{\sqrt{6}}$

$$\underline{v}_3 = \underline{u}_3 - \alpha_{21} \underline{w}_1 - \alpha_{22} \underline{w}_2 = \begin{bmatrix} -2/3 \\ -2/3 \\ 2/3 \end{bmatrix} \Rightarrow \underline{w}_3 = \underline{v}_3 / ||\underline{v}_3|| = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}.$$

Approximate Solution to $A_{m\times n}\underline{x}=\underline{y}$ when No Solution Exists and $\rho(A)=n$

Consider real $\underline{x},\underline{y}$. The criterion to choose $\underline{x}$ is to minimize $||A\underline{x}-\underline{y}||^2$. Let

$J=||A\underline{x}-\underline{y}||^2=(A\underline{x}-\underline{y})^T(A\underline{x}-\underline{y})=\underline{x}^TA^TA\underline{x}-\underline{x}^TA^T\underline{y}-\underline{y}^TA\underline{x}+\underline{y}^T\underline{y}$.

$\nabla_{\underline{x}}J=2A^TA\underline{x}-2A\underline{y}=0$ (at minima or maxima) since

$\nabla_{\underline{x}}(\underline{x}^TP\underline{x})=(P+P^T)\underline{x}$ and $\nabla_{\underline{x}}(\underline{x}^T\underline{c})=\underline{c}$.

As $\nabla^2_{\underline{x}}J=A^TA$ is positive definite so the corresponding solution $\underline{x}=(A^TA)^{-1}A^T\underline{y}$ is a minima.

We get $A\underline{x}=A(A^TA)^{-1}A^T\underline{y}$ or, $A\underline{x}$ is projection of $\underline{y}$ in a space generated by n columns of A.

Positive Definite Matrix

A matrix B is called positive definite if $\forall\ \underline{x}\quad \underline{x}^TB\underline{x}>0$ barring $\underline{x}=\underline{0}$.

Consider $C=A^TA$ for any $A_{m\times n}$. Let $\rho(A)=n$ then C is positive definite.

Proof: $\underline{x}^TC\underline{x}=\underline{x}^TA^TA\underline{x}=\underline{y}^T\underline{y}$ and given $\rho(A)=n$, $\underline{y}$ can only be $\underline{0}$ if $\underline{x}=\underline{0}$.

Which means $\underline{x}^TC\underline{x}=\sum_i y_i^2\neq 0$ for any $\underline{x}\neq 0$.

Solution to $A_{m\times n}\underline{x} = \underline{y}$ When Solution Exists But Is Not Unique

Here $m < n$, $\rho(A) = m$, $\gamma(A) = n - m$; $A, \underline{x}, \underline{y}$ real. The criterion is to choose $\underline{x}$ that minimize $\underline{x}^T\underline{x}$ subject to $A\underline{x} = \underline{y}$.

$J = \frac{1}{2}\underline{x}^T\underline{x} + \underline{\lambda}^T(A\underline{x} - \underline{y}) \;\; \Rightarrow$ At minima $\nabla_{\underline{x}} J = \underline{x} + A^T\underline{\lambda} = \underline{0} \;\; \Rightarrow \underline{x} = -A^T\underline{\lambda}$

Since $A\underline{x} = -AA^T\underline{\lambda} = \underline{y} \;\; \Rightarrow \; \underline{\lambda} = -(AA^T)^{-1}\underline{y} \;\; \Rightarrow \; \underline{x} = A^T(AA^T)^{-1}\underline{y}$.

Clearly $A\underline{x} = AA^T(AA^T)^{-1}\underline{y} = \underline{y}$.

$A^T(AA^T)^{-1}\underline{y} = A^T(AA^T)^{-1}A\underline{x}$ is the projection of $\underline{x}$ on the space generated by m columns of A^T.

Thus we can start with any solution of $A\underline{x} = \underline{y}$ and get its projection on the row space of A, this will give us the intended solution.

Eigenvalues, Eigenvectors of a Square Matrix

Definition: An eigenvector of $A_{n\times n}$ is any vector $\underline{x}$ for which $A\underline{x}$ yields a vector in the same direction as $\underline{x}$.
Mathematically If $\underline{x}$ is an eigenvector of A then $A\underline{x} = \lambda\underline{x}$, where λ is called eigenvalue associated with $\underline{x}$.
Eigenvectors are a special basis in the n dimensional space which play a central role in the “behavior” of a square matrix.
Example: (a) Block under compression

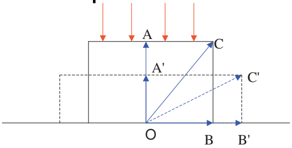

Fig. 2.1

Here $\overline{\text{OA}}' = k_a\overline{\text{OA}} = T[\overline{\text{OA}}]$ and $\overline{\text{OB}}' = k_b\overline{\text{OB}} = T[\overline{\text{OB}}]$ and Since
$\overline{\text{OC}} = \overline{\text{OA}} + \overline{\text{OB}}$
$\overline{\text{OC}}' = k_a\overline{\text{OA}} + k_b\overline{\text{OB}} = T[\overline{\text{OC}}]$. So $\overline{\text{OA}}$ and $\overline{\text{OB}}$ are orthogonal eigenvectors.

(b) Block under compression and shear

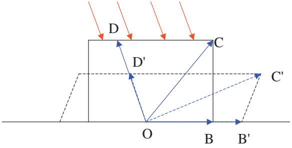

Fig. 2.2

Here $\overline{OD}$ and $\overline{OB}$ are skewed (non orthogonal) eigenvectors.

(c) Block under pure shear

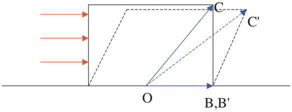

Fig. 2.3

Here only eigenvector is $\overline{OB}$. Must introduce a “generalized eigenvector” to obtain a basis including $\overline{OB}$.

(c) Iterative Impedance

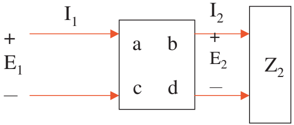

Fig. 2.4

Here $\begin{bmatrix} E_1 \\ I_1 \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} E_2 \\ I_2 \end{bmatrix} \stackrel{\triangle}{=} F \begin{bmatrix} E_2 \\ I_2 \end{bmatrix}$

In this case $E_1/I_1 = E_2/I_2 = Z_2$, so $F \begin{bmatrix} Z_2 \\ 1 \end{bmatrix} = \frac{I_1}{I_2} \begin{bmatrix} Z_2 \\ 1 \end{bmatrix}$

Hence $\begin{bmatrix} Z_2 \\ 1 \end{bmatrix}$ is an eigenvector of F with eigenvalue I_1/I_2.

Eigenvalue, Eigenvector Calculation

If λ is an eigenvalue of $A_{n\times n}$ then we have $A\underline{x} = \lambda\underline{x}$ for some $\underline{x} \neq \underline{0}$.
Then $(\lambda I - A)\underline{x} = \underline{0}$ has a non-trivial solution. In other words the null space of $\lambda I - A$ is nonempty ($\rho(\lambda I - A) < n$ and $\gamma(\lambda I - A) > 0$).
Hence for any λ which is an eigenvalue of A we have $|\lambda I - A| = 0$.
So if $\lambda_1, \lambda_2, \cdots, \lambda_n$ are the eigenvalues then
$|\lambda I - A| = \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0 = (\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_n)$.
The above equation is known as the "characteristic equation of A".
In general roots may have multiplicity greater than 1
$\Rightarrow |\lambda I - A| = (\lambda - \lambda_1)^{n_1}(\lambda - \lambda_2)^{n_2}\cdots(\lambda - \lambda_m)^{n_m}$, where $\sum_{i=1}^{m} n_i = n$.

Solution of $A_{n\times n}\underline{x} = \underline{y}$

Let $\underline{z}_1, \underline{z}_2, \cdots, \underline{z}_n$ are lin. indep. eigenvectors and $\underline{y} = \sum_{i=1}^n \alpha_i \underline{z}_i$ and $\underline{x} = \sum_{i=1}^n \beta_i \underline{z}_i$, β_i unknown.

Then $A\underline{x} = A\sum_{i=1}^n \beta_i \underline{z}_i = \sum_{i=1}^n \beta_i \lambda_i \underline{z}_i = \underline{y} = \sum_{i=1}^n \alpha_i \underline{z}_i$

$\Rightarrow \sum_{i=1}^n (\alpha_i - \beta_i\lambda_i)\underline{z}_i = \underline{0}$

Since $\{\underline{z}_1, \underline{z}_2, \cdots, \underline{z}_n\}$ are lin. indep. then $\alpha_i - \beta_i\lambda_i = 0$ for all $i \;\Rightarrow\; \beta_i = \alpha_i/\lambda_i$.

Theorem: Eigenvectors of A having distinct eigenvalues are linearly independent.

Proof: Let $\underline{x}_1, \underline{x}_2, \cdots, \underline{x}_k$ be eigenvectors of A corresponding to distinct eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_n$.

Suppose $\{\underline{x}_1, \underline{x}_2, \cdots, \underline{x}_k\}$ are linearly dependent $\Rightarrow \sum_{i=1}^k c_i \underline{x}_i = \underline{0}$ for some $\underline{c} \neq \underline{0}$ where $\underline{c} = [c_1 \cdots c_n]^T$.

Since λ_is are distinct $(A - \lambda_j I)\underline{x}_i = (\lambda_i - \lambda_j)\underline{x}_i = \begin{cases} \underline{0}, & i = j \\ \neq \underline{0}, & i \neq j \end{cases}$

Thus

$$(A - \lambda_1 I) \sum_{i=1}^{k} c_i \underline{x}_i = \sum_{i=2}^{k} c_i (\lambda_i - \lambda_1) \underline{x}_i = \underline{0}$$

$$(A - \lambda_2 I)(A - \lambda_1 I) \sum_{i=1}^{k} c_i \underline{x}_i = \sum_{i=3}^{k} c_i (\lambda_i - \lambda_1)(\lambda_i - \lambda_2) \underline{x}_i = \underline{0}$$

$$\vdots$$

$$\prod_{j=1}^{k-1} (A - \lambda_j I) \sum_{i=1}^{k} c_i \underline{x}_i = c_k \underline{x}_k \prod_{j=1}^{k-1} (\lambda_k - \lambda_j) = \underline{0}$$

Since $\underline{x}_k \neq \underline{0}, \prod_{j=1}^{k-1} (\lambda_k - \lambda_j) \neq 0$; hence $c_k = 0$.
Since eigenvectors can be ordered in any way, we have $c_i = 0 \; \forall i$.
Contradiction, so the theorem follows.

Diagonalization of $A_{n\times n}$ with Distinct Eigenvalues

Suppose A has distinct eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_n$ with corresponding eigenvectors $\underline{u}_1, \underline{u}_2, \cdots, \underline{u}_n$.

Let $M \stackrel{\triangle}{=} [\underline{u}_1\ \underline{u}_2 \cdots \underline{u}_n]$, this matrix is called modal matrix.

Then M^{-1} exists since $\underline{u}_i$ are linearly independent.

Furthermore $AM = [A\underline{u}_1\ A\underline{u}_2 \cdots A\underline{u}_n] = [\lambda_1\underline{u}_1\ \lambda_2\underline{u}_2 \cdots \lambda_n\underline{u}_n] = M\text{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$

$\Rightarrow\ M^{-1}AM = M^{-1}M\text{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n) = \text{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$

$M^{-1}AM$ is a "similarity" transformation of A.

Note:

$|\lambda I - T^{-1}AT| = |T^{-1}(\lambda I - A)T| = |T^{-1}| \cdot |\lambda I - A| \cdot |T| = |\lambda I - A|$.

Hence $A,\ T^{-1}AT$ have same eigenvalues, same characteristic equation.

If $T^{-1}AT$ is diagonal, its diagonal elements must be eigenvalues of A.

Generalized Eigenvectors of $A_{n\times n}$

Suppose λ_i has multiplicity n_i. Then $n - n_i \leq \rho(\lambda_i I - A) \leq n - 1$ or $1 \leq \gamma(\lambda_i I - A) \leq n_i$.

Only $\gamma(\lambda_i I - A)$ lin. indep. solutions of $(\lambda_i I - A)\underline{x} = \underline{0}$ exist.

Thus if $\gamma(\lambda_i I - A) < n_i$, cannot find n_i conventional eigenvectors.

Example: $A = \begin{bmatrix} 2 & -1 & 2 \\ -1 & 2 & -2 \\ -1 & 1 & -1 \end{bmatrix}, \ |\lambda I - A| = (\lambda - 1)^3$

$$(A - \lambda I)\underline{x} = (A - I)\underline{x} = \begin{bmatrix} 1 & -1 & 2 \\ -1 & 1 & -2 \\ -1 & 1 & -2 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} a - b + 2c \\ -(a - b + 2c) \\ -(a - b + 2c) \end{bmatrix} = \underline{0}$$

Thus for $\underline{x}$ to be in the null space of $I - A$ we need $b = a + 2c$.

$\underline{x}_1 = [1\ 1\ 0]^T \ \ (a = 1, c = 0)$ and $\underline{x}_2 = [0\ 2\ 1]^T \ \ (a = 0, c = 1)$ are lin. indep. eigenvectors (not unique).

We can define a "generalized eigenvector" lin. indep. of $\underline{x}_1, \underline{x}_2$; say, $\underline{x}_3 = [0\ 1\ 0]^T$.

Generalized Eigenvectors – Single Chain Rule

Consider the case that eigenvalue λ_i with multiplicity k and where $B \stackrel{\triangle}{=} (A - \lambda_i I)$ has nullity 1.

Then $B\underline{x} = \underline{0}$ has one solution, say $\underline{v}_1$, i.e., $B\underline{v}_1 = \underline{0}$.

For the remaining “generalized eigenvectors” a reasonable choice is vectors $\underline{v}_2, \underline{v}_3, \cdots, \underline{v}_k$ that are “killed” by $B^2, B^3, \cdots, B^k$ respectively.

$$B\underline{v}_1 = \underline{0} \;\Rightarrow\; A\underline{v}_1 = \lambda_i \underline{v}_1$$
$$B\underline{v}_2 = \underline{v}_1 \;\Rightarrow\; B^2\underline{v}_2 = B\underline{v}_1 = \underline{0} \;\Rightarrow\; A\underline{v}_2 = \lambda_i \underline{v}_2 + \underline{v}_1$$
$$B\underline{v}_3 = \underline{v}_2 \;\Rightarrow\; B^3\underline{v}_3 = B^2\underline{v}_2 = \underline{0} \;\Rightarrow\; A\underline{v}_3 = \lambda_i \underline{v}_3 + \underline{v}_2$$
$$\vdots$$
$$B\underline{v}_k = \underline{v}_{k-1} \;\Rightarrow\; B^k\underline{v}_k = B^{k-1}\underline{v}_{k-1} = \underline{0} \;\Rightarrow\; A\underline{v}_k = \lambda_i \underline{v}_k + \underline{v}_{k-1}$$

Repeated use of the above yields:

$$\underline{v}_j = B\underline{v}_{j+1} = B^2\underline{v}_{j+2} = \cdots = B^{j-i}\underline{v}_i = \cdots = B^{k-j}\underline{v}_k$$
$$\underline{v}_1 = B\underline{v}_2 = B^2\underline{v}_3 = \cdots = B^{i-1}\underline{v}_i = \cdots = B^{k-1}\underline{v}_k$$

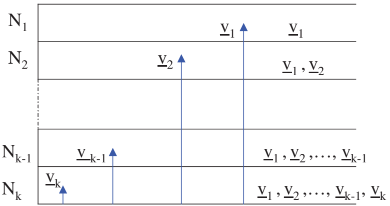

Fig. 2.5

Define N_k = null space of B^k.

Observe that $\underline{v}_1 \in N_1$, i.e., $B\underline{v}_1 = \underline{0}$ again $B^j \underline{v}_1 = \underline{0}$ for $j = 1, \cdots, k \quad \Rightarrow \underline{v}_1 \in N_j$.

Similarly $B^2 \underline{v}_2 = \underline{0}$ but $B\underline{v}_2 = \underline{v}_1 \neq \underline{0}$.

Hence $\underline{v}_2 \in N_2, N_3, \cdots, N_k$

$$\underline{v}_3 \in N_3, N_4, \cdots, N_k$$

$$\vdots$$

$$\underline{v}_k \in N_k$$

Vector $\underline{v}_k$ satisfying $B^k \underline{v}_k = \underline{0}$ is a "generalized eigenvector" of grade k.

"Chain" $\underline{v}_1, \underline{v}_2, \cdots, \underline{v}_k$ has length k.

Determining $\underline{v}_1, \underline{v}_2, \cdots, \underline{v}_k$ with $\gamma(B) = 1,\ n_i = k$

Step 1: Find $\underline{v}_k$ such that $B^k \underline{v}_k = (A - \lambda_i I)^k \underline{v}_k = \underline{0}$ but $B^{k-1}\underline{v}_k = \underline{v}_1 \neq \underline{0}$.

Step 2: Determine in turn $\underline{v}_j = B\underline{v}_{j+1}$ for $j = k-1, k-2, \cdots, 1$.

Example: $A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -3 & 3 \end{bmatrix}, \quad |\lambda I - A| = (\lambda - 1)^3$

Step 1: $A - I = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 1 & -3 & 2 \end{bmatrix};\ \rho(A_I) = 2,\ \gamma(A - I) = 1;$ single chain.

$$(A - I)^2 = \begin{bmatrix} 1 & -2 & 1 \\ 1 & -2 & 1 \\ 1 & -2 & 1 \end{bmatrix};\ \rho((I - A)^2) = 1,\ \gamma((I - A)^2) = 2.$$

Also $(A - I)^3 = [0]$, thus for any $\underline{v}_3$, $(I - A)^3 \underline{v}_3 = \underline{0}$.

Choose $\underline{v}_3$ such that

$$(A - I)^2 \begin{bmatrix} a \\ b \\ c \end{bmatrix} = (a - 2b + c) \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \underline{v}_1 \neq \underline{0}$$

Choose $\underline{v}_3$ such that $\underline{v}_1 \neq \underline{0}$, i.e., $a - 2b + c \neq 0$.

One such choice is $\underline{v}_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad \Rightarrow \underline{v}_1 = (A-I)^2 \underline{v}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

Step 2: $\underline{v}_2 = (A-I)\underline{v}_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$; $\underline{v}_1 = (A-I)\underline{v}_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ check!

How does A look like after similarity transform ($M^{-1}AM$)?

$M = [\underline{v}_1 \ \underline{v}_2 \ \underline{v}_3] = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \end{bmatrix}$ has $|M| = 1 \neq 0$, $\underline{v}_1, \underline{v}_2, \underline{v}_3$ are lin. indep.

$$AM = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 2 & 1 \end{bmatrix}, \quad M^{-1} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & -1 & 1 \\ 1 & -2 & 1 \end{bmatrix}, \quad M^{-1}AM = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}.$$

Multiple Generalized Eigenvector Chains for Same Eigenvalue

$\lambda_i =$ eigenvalue of A with multiplicity n_i.

$B \triangleq A - \lambda_i I,\ \rho_k \triangleq \rho\left(B^k\right),\ N_k \triangleq$ null space of B^k, $\gamma_k \triangleq$ dimension of N_k.

There are $\gamma(B)$ eigenvector chains of total length n_i.

Special Cases:

(a) $\gamma(B) = n_i$, there are n_i conventional eigenvectors, i.e., n_i chains of length 1.

(b) $\gamma(B) = 1$, there is one chain of length n_i.

Procedure for Finding $\gamma(B)$ Chains

(a) Determine integer g such that $\rho_g = n - n_i,\ \gamma_g = n_i$.

(g is the length of the longest chain)

(b) $r_g \triangleq \gamma_g - \gamma_{g-1} \geq 1$. Find r_g chains of length g, insuring all chains are lin. indep. (N_g contains r_g vectors not in N_{g-1}. Thus there are r_g chains of length g)

(c) Set $j = g - 1$.

(d) $r_j = \gamma_j - \gamma_{j-1} - \sum_{i=j+1}^{g} r_i$. Find r_j chains of length j.
(N_j contains $\gamma_j - \gamma_{j-1}$ vectors not in N_{j-1}. Of these, $r_{j+1} + r_{j+2} + \cdots + r_g$ are in chains longer than j. This leaves r_j chains of length j)
(e) Set $j = j - 1$. Stop if $j = 0$, else goto step (d).

Example: Let $n = 20,\ n_i = 6,\ \gamma(B) = \gamma(A - \lambda_i I) = 3$. Given $\rho_0 = \rho(B^0) = \rho(I) = 20,\ \gamma_0 = 0;\ \ \rho_1 = 17,\ \gamma_1 = 3;\ \ \rho_2 = 15,\ \gamma_2 = 5; \rho_3 = 14,\ \gamma_3 = 6$. Describe the procedure of finding generalized eigenvectors.

Solution:

(a) We have $\gamma_3 = 6 = n_i,\ g = 3$
(b) Here $r_3 = \gamma_3 - \gamma_2 = 1$. One chain of length 3. Find $\{\underline{v}_1, \underline{v}_2, \underline{v}_3\}$ with $B^3\underline{v}_3 = \underline{0}$, $B^2\underline{v}_3 = \underline{v}_1 \neq 0$ and $\underline{v}_2 = B\underline{v}_3$.
(c) $j = g - 1 = 2$.
(d) $r_2 = \gamma_2 - \gamma_1 - r_3 = 1$, this second chain has length 2. Find $\underline{u}$ with $B^2\underline{u} = \underline{0}$, $B\underline{u} \neq \underline{0}$ and ensure that $\underline{u}$ is lin. indep. to $\underline{v}_1, \underline{v}_2, \underline{v}_3$. Get $\underline{v}_5 = \underline{u}$ and $\underline{v}_4 = B\underline{u}$.

(e) $j = 1$
(d) $r_1 = \gamma_1 - \gamma_0 - r_2 - r_3 = 1$. Final length has length 1. Find $\underline{v}_6$ with $B\underline{v}_6 = \underline{0}$ and $\underline{v}_6$ is lin. indep. to $\underline{v}_1, \cdots, \underline{v}_5$.
(e) $j = 0$. Done.

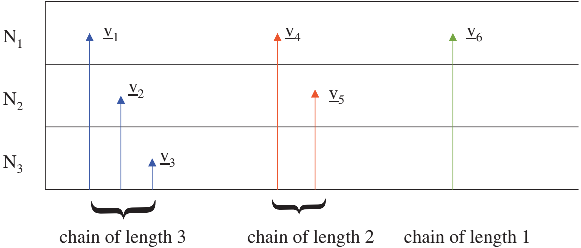

Fig. 2.6

Note: N_1 contains $\underline{v}_1, \underline{v}_4, \underline{v}_6$; $\gamma_1 = 3$.
N_2 contains $\underline{v}_1, \underline{v}_2, \underline{v}_4, \underline{v}_5, \underline{v}_6$; $\gamma_2 = 5$.
N_3 contains $\underline{v}_1, \underline{v}_2, \underline{v}_3, \underline{v}_4, \underline{v}_5, \underline{v}_6$; $\gamma_3 = 6$.

Example: $A = \begin{bmatrix} 2 & -1 & 2 \\ -1 & 2 & -2 \\ -1 & 1 & -1 \end{bmatrix}, \ |\lambda I - A| = (\lambda - 1)^3, \ n_i = n_1 = 3.$

Find a set of generalized eigenvectors.

Solution: $B = \begin{bmatrix} 1 & -1 & 2 \\ -1 & 1 & -2 \\ -1 & 1 & -2 \end{bmatrix}, \ \rho(B) = 1, \ \gamma(B) = 2 = \gamma_1$. Thus have 2 chains.

(Only possible choice one chain of length 2, another of length 1.)

(a) $(A - I)^2 = [0], \ \rho(B^2) = \rho_2 = 0, \ \gamma_2 = 3$. (maximum length of a chain is 2).

(b) $r_2 = \gamma_2 - \gamma_1 = 3 - 2 = 1 \ \ \Rightarrow$ one chain of length 2.

$$(A-I)\underline{v}_2 = (A-I)\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} a - b + 2c \\ -a + b - 2c \\ -a + b - 2c \end{bmatrix} = (a-b+2c)\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \underline{v}_1 \neq 0.$$

So we need $a - b + 2c \neq 0$.

One choice $\underline{v}_2 = [1 \ 0 \ 0]^T \ \ \Rightarrow \underline{v}_1 = (A - I)\underline{v}_2 = [1 \ -1 \ -1]^T$.

(c) $j = 2 - 1 = 1$

(d) $r_1 = \gamma_1 - \gamma_0 - r_2 = \gamma(A-I) - \gamma((A-I)^0) - r_2 = 2 - 0 - 1 = 1$
(one chain of length 1)

$(A-I)\underline{v}_3 = \underline{0} \;\Rightarrow\; (A-I)\begin{bmatrix} x \\ y \\ z \end{bmatrix} = (x-y+2z)\begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix}$ where

$\underline{v}_3 = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \neq 0$.

Thus we need $x - y + 2z = 0$. $\underline{v}_1$ satisfies this but want $\underline{v}_1, \underline{v}_2, \underline{v}_3$ lin. indep. A possible choice $\underline{v}_3 = [1\ 1\ 0]^T$. ($x = 1,\ y = 1\ z = 0$)

(e) $j = 1 - 1 = 0$. Done.

Here $M = [\underline{v}_1\ \underline{v}_2\ \underline{v}_3] = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 0 \end{bmatrix}$

Thus $M^{-1}AM = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_1 \end{bmatrix}$

Example: $\begin{bmatrix} 0 & 1 & 0 & j \\ 1 & 0 & -j & 0 \\ 0 & -j & 0 & 1 \\ j & 0 & 1 & 0 \end{bmatrix} = A^T$. Find a set of generalized eigenvectors.

Solution: col $3 = -j \times$ col 1 col $4 = j \times$ col 2 $\Rightarrow$ $\rho(A) = 2,\ \gamma(A) = 2$.

Again the eigenvalues are all zero as $|\lambda I - A| = \lambda^4 \Rightarrow A - \lambda_1 I = A$.

Since $\gamma(A) = 2$, we have two chains of total length 4. (either 2,2 or 3,1)

(a) $A^2 = [0]$. Hence $\rho(A^2) = 0,\ \gamma(A^2) = 4 \Rightarrow g = 2$. (Maximum chain len. is 2)

(b) $r_2 = \gamma_2 - \gamma_1 = 2 = \gamma(A^2) - \gamma(A) = 4 - 2 = 2$. (Two chains of length 2)

Let $\underline{v}_2 = \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$, we need $A^2\underline{v}_2 = \underline{0}$ and

$$A\underline{v}_2 \neq 0 \Rightarrow A\underline{v}_2 = \begin{bmatrix} b + jd \\ a - jc \\ -jb + d \\ ja + c \end{bmatrix} \neq 0 \text{ as } A^2 = [0].$$

Thus we need $a \neq jc$ or $b \neq -jd$. Two choices $\underline{v}_{21} = [0\ 0\ 1\ 0]^T$ and $\underline{v}_{22} = [0\ 0\ 0\ 1]^T$.

Then $\underline{v}_{11} = A\underline{v}_{21} = [0\ -j\ 0\ 1]^T$ and $\underline{v}_{12} = A\underline{v}_{22} = [j\ 0\ 1\ 0]^T$.

Here $M = [\underline{v}_{11}\ \underline{v}_{12}\ \underline{v}_{21}\ \underline{v}_{22}] = \begin{bmatrix} 0 & j & 0 & 0 \\ -j & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}$

Thus $M^{-1}AM = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & \lambda_1 & 1 \\ 0 & 0 & 0 & \lambda_1 \end{bmatrix}$