Consider a system of linear first order differential equations $\mathbf{x}'(t) = A\mathbf{x}(t)$ with initial value $\mathbf{x}(0) = \mathbf{x}_0$, and suppose that $A$ is similar to a Hessenberg matrix $H = S^{-1}AS$. Let

$$\mathbf{y}(t) = S^{-1}\mathbf{x}(t) \quad \text{and} \quad \mathbf{y}_0 = S^{-1}\mathbf{x}(0) = \mathbf{y}(0);$$

then

$$\mathbf{y}'(t) = S^{-1}\mathbf{x}'(t) = S^{-1}A\mathbf{x}(t) = (S^{-1}AS)S^{-1}\mathbf{x}(t) = H\mathbf{y}(t).$$

Thus the change of variables $\mathbf{y}(t) = S^{-1}\mathbf{x}(t)$ transforms the given initial value problem (IVP) into the equivalent one

$$\mathbf{y}'(t) = H\mathbf{y}(t) \quad \text{with} \quad \mathbf{y}(0) = \mathbf{y}_0. \quad (1)$$

For example, the matrix $A = \begin{bmatrix} -1 & -8 & 1 \\ -1 & -3 & 2 \\ -4 & -16 & 7 \end{bmatrix}$ is similar to

$$H = \begin{bmatrix} -1 & -4 & 1 \\ -1 & 5 & 2 \\ 0 & -8 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -4 & 1 \end{bmatrix} \begin{bmatrix} -1 & -8 & 1 \\ -1 & -3 & 2 \\ -4 & -16 & 7 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 4 & 1 \end{bmatrix}. $$

Thus the change of variables above transforms the system

$$\begin{bmatrix} x'_1(t) \\ x'_2(t) \\ x'_3(t) \end{bmatrix} = \begin{bmatrix} -1 & -8 & 1 \\ -1 & -3 & 2 \\ -4 & -16 & 7 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} \quad \text{with} \quad \begin{bmatrix} x_1(0) \\ x_2(0) \\ x_3(0) \end{bmatrix} = \begin{bmatrix} 7 \\ -1 \\ 5 \end{bmatrix}$$

into

$$\begin{bmatrix} y'_1(t) \\ y'_2(t) \\ y'_3(t) \end{bmatrix} = \begin{bmatrix} -1 & -4 & 1 \\ -1 & 5 & 2 \\ 0 & -8 & -1 \end{bmatrix} \begin{bmatrix} y_1(t) \\ y_2(t) \\ y_3(t) \end{bmatrix} \quad \text{with} \quad \begin{bmatrix} y_1(0) \\ y_2(0) \\ y_3(0) \end{bmatrix} = \begin{bmatrix} 7 \\ -1 \\ 9 \end{bmatrix}. $$

When $H$ is reduced, Jordan Canonical Form (the final topic in this course) is required to solve the IVP in (1). When $H$ is unreduced, however, we can solve the IVP in (1) by following the steps outlined below. But first we need some important theoretical results.

**Theorem 1** Let $H = (h_{ij})$ be an unreduced $n \times n$ Hessenberg matrix. Then $\{e_1, He_1, \ldots, H^{n-1}e_1\}$ is linearly independent.

**Proof.** First note that

$$He_1 = \begin{bmatrix} h_{11} \\ h_{21} \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \text{and} \quad H^2e_1 = \begin{bmatrix} h_{11}^2 + h_{12}h_{21} \\ h_{21}h_{11} + h_{22}h_{21} \\ h_{32}h_{21} \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}.$$

since $H$ is unreduced, $h_{21} \neq 0$ and $b_3 = h_{32}h_{21} \neq 0$. Inductively, assume that $H^{i-1}e_1 = [c_1 \cdots c_i 0 \cdots 0]^T$ with $c_i \neq 0$. Then
if eigenspace, solve the system

the algebraic multiplicity of the eigenvalue

De…nition 3 Let A be an unreduced Hessenberg matrix. If \[ |a_0 \cdots a_{n-1}|^T \neq 0, \]

An usual eigenvector associated with \( \lambda \) is a generalized eigenvector of order 1.

Definition 4 Let A be an \( n \times n \) matrix with characteristic polynomial \( p(t) = (t - \lambda_1)^{r_1} \cdots (t - \lambda_k)^{r_k} \), where \( r_1 + \cdots + r_k = n \) and \( \lambda_i \in \mathbb{C} \). Then \( r_i \) is the algebraic multiplicity of \( \lambda_i \), and the dimension of the eigenspace associated with \( \lambda_i \) is the geometric multiplicity of \( \lambda_i \).

The geometric multiplicity of an eigenvalue is at least 1 and at most its algebraic multiplicity.

Definition 5 An \( n \times n \) matrix A is defective if A has an eigenvalue whose geometric multiplicity is less than its algebraic multiplicity.

Example 6 The characteristic polynomial of the unreduced Hessenberg matrix

\[
H = \begin{bmatrix}
1 & -1 \\
1 & 3
\end{bmatrix}
\]

is \( p(t) = \det \begin{bmatrix} 1-t & -1 \\
1 & 3-t \end{bmatrix} = (t-2)^2 \);

the algebraic multiplicity of the eigenvalue \( \lambda = 2 \) is therefore 2. To find a basis for the corresponding eigenspace, solve the system

\[
(H - 2I)x = 0
\]

then \( \{ \text{basis} \} \) is a basis and the geometric multiplicity of the eigenvalue \( \lambda = 2 \) is 1. Consequently the matrix is defective and is not diagonalizable.

Exercise 7 Show that the following matrices are defective:

a. \[
\begin{bmatrix}
1 & 4 \\
0 & 1
\end{bmatrix}
\]
b. \[
\begin{bmatrix}
-4 & 1 & 1 & 1 \\
-16 & 3 & 4 & 4 \\
-7 & 2 & 2 & 1 \\
-11 & 1 & 3 & 4
\end{bmatrix}
\]
**Theorem 8** Let $\lambda$ be an eigenvalue of an unreduced $n \times n$ Hessenberg matrix $H$. If the algebraic multiplicity of $\lambda$ is $m$, then $H$ has a generalized eigenvector $v_r$ of order $r$ associated with $\lambda$ for each $r = 1, 2, \ldots, m$.

**Proof.** Let $p(t)$ be the characteristic polynomial of $H$. Repeatedly divide $p(t)$ by $t - \lambda$ and until

$$p(t) = (t - \lambda)^m q(t) \text{ with } q(\lambda) \neq 0.$$ 

Consider the polynomial

$$(t - \lambda)^{m-1} q(t) = a_0 + a_1 t + \cdots + a_{n-2} t^{n-2} + t^{n-1}$$

whose vector of coefficients $[a_0 \cdots a_{n-2}]^T \neq 0$. Thus

$$(H - \lambda I)^{m-1} q(H) e_1 = a_0 e_1 + a_1 H e_1 + \cdots + a_{n-2} H^{n-2} e_1 + H^{n-1} e_1 \neq 0$$

by Corollary 2. Set $v_m = q(H) e_1$, then

$$(H - \lambda I)^{m-1} v_m \neq 0.$$ 

On the other hand, $p(H) = (H - \lambda I)^m q(H)$ implies

$$(H - \lambda I)^m v_m = (H - \lambda I)^m q(H) e_1 = p(H) e_1.$$ 

But $p(H) = 0$ by the Cayley-Hamilton Theorem, and we conclude that

$$(H - \lambda I)^m v_m = 0.$$ 

Therefore $v_m$ is a generalized eigenvector of order $m$ associated with $\lambda$. Finally, since $(H - \lambda I)^{m-1} q(H) e_1 \neq 0$ and $(H - \lambda I)^m q(H) e_1 = 0$, let

$$v_r = (H - \lambda I)^{m-r} q(H) e_1,$$

then $(H - \lambda I)^{r-1} v_r \neq 0$ and $(H - \lambda I)^r v_r = 0$. Thus $H$ has a generalized eigenvector of order $r$ associated with $\lambda$ for each $r = 1, 2, \ldots, m$. ■

**Theorem 9** An unreduced $n \times n$ Hessenberg matrix has $n$ linearly independent generalized eigenvectors.

**Proof.** Let $\lambda_1, \lambda_2, \ldots, \lambda_k$ be the eigenvalues of $H$ and let $m_i$ be the algebraic multiplicity of $\lambda_i$. Then $p(t) = (t - \lambda_1)^{m_1}(t - \lambda_2)^{m_2} \cdots (t - \lambda_k)^{m_k}$ is the characteristic polynomial of $H$. By Theorem 8, each $\lambda_i$ has an associated generalized eigenvector $v_r^j$ of order $r$ for each $r = 1, 2, \ldots, m_i$. Denote these by

$$\left\{v_1^{\lambda_1}, \ldots, v_{m_1}^{\lambda_1}, v_2^{\lambda_2}, \ldots, v_{m_2}^{\lambda_2}, \ldots, v_k^{\lambda_k}, \ldots, v_{m_k}^{\lambda_k}\right\}$$

and suppose that

$$a_1^1 v_1^1 + \cdots + a_1^{m_1} v_{m_1}^1 + a_2^1 v_2^1 + \cdots + a_2^{m_2} v_{m_2}^1 + \cdots + a_k^1 v_k^1 + \cdots + a_k^{m_k} v_{m_k}^1 = 0. \quad (2)$$

Let $q(t) = (t - \lambda_2)^{m_2} \cdots (t - \lambda_k)^{m_k}$; since the factors of $q(t)$ commute, $q(H) v_i^j = 0$ for all $i$ and all $j \geq 2$. Now multiply both sides of (2) by $q(H)$; then

$$q(H) \left(a_1^1 v_1^1 + \cdots + a_1^{m_1} v_{m_1}^1\right) + q(H) \left(a_2^1 v_2^1 + \cdots + a_2^{m_2} v_{m_2}^1 + \cdots + a_k^1 v_k^1 + \cdots + a_k^{m_k} v_{m_k}^1\right) = 0$$

reduces to

$$a_1^1 q(H) v_1^1 + \cdots + a_1^{m_1} q(H) v_{m_1}^1 = 0. \quad (3)$$

Keeping in mind that $(t - \lambda_1)^p q(t) = q(t) (t - \lambda_1)^p$, multiply both sides of (3) by $(H - \lambda_1 I)^{m_1 - 1}$; then

$$a_1^{m_1} (H - \lambda_1 I)^{m_1 - 1} q(H) v_{m_1}^1 = 0. \quad (4)$$

3
implies $a_{1}^{m_{1}} = 0$. Now multiply both sides of (4) by $(H - \lambda_{1}I)^{m_{1} - 2}$; then
\[ a_{m_{1} - 1}^{1}(H - \lambda_{1}I)^{m_{1} - 2} q(H)v_{1}^{m_{1} - 1} = 0 \]
implies $a_{m_{1} - 1}^{1} = 0$, and so on inductively. Do this for each $\lambda_{i}$ and conclude that all coefficients vanish. ■

Example 10 Continuing Example 6, Let’s find a generalized eigenvector $v_{2}$ associated with $\lambda = 2$ such that \{v_{1}, v_{2}\} is linearly independent. Solve the system
\[(H - 2I)x = v_{1}\]

\[
\begin{bmatrix}
-1 & -1 & 1 \\
1 & 1 & 1
\end{bmatrix}
\rightarrow
\begin{bmatrix}
1 & 1 & :& -1 \\
0 & 0 & :& 0
\end{bmatrix};
\]
the general solution is
\[ t \begin{bmatrix}
-1 \\
1
\end{bmatrix} + \begin{bmatrix}
-1 \\
0
\end{bmatrix}. \]
Set $t = 0$ and obtain the particular solution $v_{2} = \begin{bmatrix}
-1 \\
0
\end{bmatrix}$. Note that
\[(H - 2I)^{2}v_{2} = (H - 2I)v_{1} = 0; \]
then $v_{2}$ is a generalized eigenvector of order 2 associated with $\lambda = 2$. Thus we obtain two linearly independent generalized eigenvectors associated with $\lambda = 2$:
\[
\{v_{1} = \begin{bmatrix}
-1 \\
1
\end{bmatrix}, v_{2} = \begin{bmatrix}
-1 \\
0
\end{bmatrix}\}. \]

Problem: Let $H$ be a complex $n \times n$ unreduced Hessenberg matrix. Solve the IVP $y' = Hy$ with $y(0) = y_{0}$.
Steps in the solution:
1. Using Krylov’s method, compute the characteristic polynomial of $H$. Then
\[ p(t) = (t - \lambda_{1})^{\alpha_{1}}(t - \lambda_{2})^{\alpha_{2}} \cdots (t - \lambda_{k})^{\alpha_{k}}; \]
the eigenvalue $\lambda_{i}$ has algebraic multiplicity $\alpha_{i}$.
2. Find a maximal linearly independent set of eigenvectors \{u_{1}, \ldots, u_{r}\}. Then each $u_{j}$ is associated with some $\lambda_{i}$. Set $y = e^{\lambda_{i}t}u_{j}$; then
\[ y' = e^{\lambda_{i}t}\lambda_{i}u_{j} = e^{\lambda_{i}t}Hu_{j} = H(e^{\lambda_{i}t}u_{j}) = Hy \]
so that $y$ is a particular solution.
3. Let $m_{i}$ be the geometric multiplicity of $\lambda_{i}$. If $m_{i} < \alpha_{i}$, find $q_{i} = \alpha_{i} - m_{i}$ generalized eigenvectors \{v_{1}, v_{2}, \ldots, v_{q_{i}}\} of $H$ in the following way: Set $v_{1} = u_{1}$ and solve the following successive sequence of linear equations in which $v_{r}$ is known and $v_{r+1}$ is unknown:
\[
(H - \lambda_{1}I)v_{2} = v_{1}
\]
\[
(H - \lambda_{1}I)v_{3} = v_{2}
\]
\[
(H - \lambda_{1}I)v_{4} = v_{3}
\]
\[ \vdots \]
\[
(H - \lambda_{1}I)v_{q_{i}} = v_{q_{i}-1}. \]
Then
\[ (H - \lambda_i I)^{q_i - 1} \mathbf{v}_{q_i} = (H - \lambda_i I)^{q_i - 2} \mathbf{v}_{q_i - 1} = \cdots = (H - \lambda_i I) \mathbf{v}_2 = \mathbf{v}_1. \]  
(5)

Since \((H - \lambda_i I) \mathbf{v}_1 = \mathbf{0}\), multiplying each expression in (5) by \(H - \lambda_i I\) gives
\[ (H - \lambda_i I)^q \mathbf{v}_{q_i} = (H - \lambda_i I)^{q_i - 1} \mathbf{v}_{q_i - 1} = \cdots = (H - \lambda_i I)^2 \mathbf{v}_2 = (H - \lambda_i I) \mathbf{v}_1 = \mathbf{0}. \]

Then \(\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{q_i}\}\) is linearly independent by Theorem 9 and \(\mathbf{v}_r\) is a generalized eigenvector of order \(r\) associated with \(\lambda_i\) for each \(r = 1, 2, \ldots, q_i\). Let
\[ \mathbf{y}_r = e^{\lambda_i t} \left( \mathbf{v}_r + t \mathbf{v}_{r-1} + \frac{t^2}{2!} \mathbf{v}_{r-2} + \cdots + \frac{t^{r-1}}{(r-1)!} \mathbf{v}_1 \right); \]
then
\[
\mathbf{y}_r' = \lambda_i e^{\lambda_i t} \left( \mathbf{v}_r + t \mathbf{v}_{r-1} + \frac{t^2}{2!} \mathbf{v}_{r-2} + \cdots + \frac{t^{r-1}}{(r-1)!} \mathbf{v}_1 \right) + e^{\lambda_i t} \left( \mathbf{v}_{r-1} + t \mathbf{v}_{r-2} + \cdots + \frac{t^{r-2}}{(r-2)!} \mathbf{v}_1 \right)
= e^{\lambda_i t} \left( (\mathbf{v}_{r-1} + \lambda_i \mathbf{v}_r) + t (\mathbf{v}_{r-2} + \lambda_i \mathbf{v}_{r-1}) + \frac{t^2}{2!} (\mathbf{v}_{r-3} + \lambda_i \mathbf{v}_{r-2}) + \cdots + \frac{t^{r-2}}{(r-2)!} (\mathbf{v}_1 + \lambda_i \mathbf{v}_2) + \frac{t^{r-1}}{(r-1)!} \lambda_i \mathbf{v}_1 \right)
= e^{\lambda_i t} \left( H \mathbf{v}_r + t H \mathbf{v}_{r-1} + \frac{t^2}{2!} H \mathbf{v}_{r-2} + \cdots + \frac{t^{r-1}}{(r-1)!} H \mathbf{v}_1 \right) = H \mathbf{y}_r
\]
so that \(\mathbf{y}_r\) is a particular solution.

4. Finally, the general solution of the given system of linear differential equations consists of all linear combinations of the solutions found in steps 2 and 3. Given this, one can solve a system of linear equations and find the particular solution that solves the IVP.

Example 11: Solve the system
\[
\begin{align*}
y'_1 &= y_1 - y_2 \\
y'_2 &= y_1 + 3y_2
\end{align*}
\]
subject to the initial conditions \(y_1(0) = 5\), \(y_2(0) = -7\). Written in matrix form \(\mathbf{y}' = H \mathbf{y}\) the IVP becomes
\[
\begin{bmatrix} y'_1 \\ y'_2 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \quad \text{with} \quad \begin{bmatrix} y_1(0) \\ y_2(0) \end{bmatrix} = \begin{bmatrix} 5 \\ -7 \end{bmatrix}.
\]
(6)

In Examples 6 and 10 we found the linearly independent generalized eigenvectors
\[
\mathbf{v}_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \quad \text{and} \quad \mathbf{v}_2 = \begin{bmatrix} -1 \\ 0 \end{bmatrix}
\]
associated with the eigenvalue \(\lambda = 2\) of the unreduced Hessenberg coefficient matrix
\[
H = \begin{bmatrix} 1 & -1 \\ 1 & 3 \end{bmatrix}.
\]
Let
\[
\mathbf{y} = e^{2t} \mathbf{v}_1;
\]

differentiating we have
\[
\mathbf{y}' = e^{2t} (2 \mathbf{v}_1) = e^{2t} H \mathbf{v}_1 = H \mathbf{y}
\]
so that \(\mathbf{y} = e^{2t} \mathbf{v}_1\) is a particular solution. Now
\[
(H - 2I) \mathbf{v}_2 = \mathbf{v}_1 \quad \Rightarrow \quad H \mathbf{v}_2 = \mathbf{v}_1 + 2 \mathbf{v}_2.
\]
Let \( y = e^{2t} (v_2 + tv_1) \);

differentiating gives
\[
\begin{align*}
y' &= 2e^{2t} (v_2 + tv_1) + e^{2t} v_1 = e^{2t} [(v_1 + 2v_2) + t(2v_1)] \\
&= e^{2t} (Hv_2 + tv_1) = H e^{2t} (v_2 + tv_1) = Hy
\end{align*}
\]

so that \( y = e^{2t} (v_2 + tv_1) \) is a particular solution. Thus the general solution of system 6 is
\[
y = ae^{2t}v_1 + be^{2t} (v_2 + tv_1), \ a, b \in \mathbb{R}.
\]

To find the desired particular solution, solve \( av_1 + bv_2 = y(0) \):
\[
\begin{bmatrix}
-1 & 1 & 5 \\
1 & 0 & -7
\end{bmatrix}
\rightarrow
\begin{bmatrix}
1 & 0 & -7 \\
0 & 1 & -2
\end{bmatrix}.
\]

The solution of the IVP is therefore
\[
y = -7e^{2t}v_1 - 2e^{2t} (v_2 + tv_1).
\]

11-10-2010