A Class of Hessenberg Matrices
with Known Pseudoinverse and Drazin Inverse

By Inderjit Singh, George Poole and Thomas Boullion

Abstract. In this paper, a class of Hessenberg matrices is presented for adoption as test matrices. The Moore-Penrose inverse and the Drazin inverse for each member of this class are determined explicitly.

1. Introduction. Most numerical problems associated with solving a system of linear equations involve only rational numbers. However, square matrices over the real number field are considered in this paper.

Howell and Gregory [6] have shown how to avoid problems which arise in solving the matrix equation $Ax = b$ as a result of rounding errors in computer schemes. Specifically, they have shown how to use residue arithmetic to avoid ill-conditioned problems. Using a similar approach, Stallings and Boullion [12] have shown how to significantly reduce rounding errors in computer schemes which compute the Moore-Penrose inverse (pseudoinverse) for a given matrix. However, the rounding errors are not necessarily completely eliminated.

Chow [2] has presented a class of Hessenberg matrices which may be used as test matrices in checking the accuracy of matrix inversion programs. In this paper, a class of Hessenberg matrices is presented such that the pseudoinverse and Drazin inverse can be explicitly computed for each member. Furthermore, the eigenvalues and eigenvectors are known for the members of this class. Therefore, it appears reasonable that such a class of matrices may be useful as test matrices.

2. Definitions and Notation. One should distinguish between the class of matrices in [2] which are offered as test matrices and the class given below. Only square matrices over the real number field are considered.

Definition 2.1. The pseudoinverse of a matrix $A$ is the unique solution $A^+$ of the four matrix equations $AXA = A$, $XAX = X$, $(AX)^T = AX$ and $(XA)^T = XA$ where $(\cdot)^T$ denotes the matrix transpose.

Definition 2.2. The index of a matrix $A$ is the smallest nonnegative integer $\text{Ind}(A) = k$ such that $\text{rank}(A^k) = \text{rank}(A^{k+1})$.

Definition 2.3. The Drazin inverse of a matrix $A$ is the unique solution $A^D$ of the three matrix equations $AX = XA$, $XAX = X$, $A^{k+1}X = A^k$, where $\text{Ind}(A) = k$. 
H_n shall denote the Hessenberg matrix of order n, where

\[
H_n = \begin{bmatrix}
\alpha & 1 & \cdot & \cdot & \cdot \\
\alpha^2 & \alpha & 1 & \cdot & \cdot \\
\alpha^3 & \alpha^2 & \alpha & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & 1 \\
\alpha^n & \alpha^{n-1} & \cdot & \cdot & \alpha
\end{bmatrix},
\]

and \(\alpha\) is an arbitrary real number.

The eigenvalues of \(H_n\) were determined in [2]. \(H_n\) has \(k = [n/2]\) eigenvalues equal to 0 (where \([n/2]\) denotes the largest integer not exceeding \(n/2\)) and whose remaining eigenvalues are

\[4\alpha \cos^2\left(\frac{m\pi}{n+2}\right), \quad m = 1, 2, \ldots, n-k.\]

The reader can refer to [4] for the corresponding eigenvectors.

3. Pseudoinverse of \(H_n\). There are several algorithms available for computing \((H_n)^+\), [1], [3], [5], [10]. The general form for \((H_n)^+\) is presented in this section.

Case 1 \((n = 2)\). If \(H_2\) is the matrix

\[
H_2 = \begin{bmatrix}
\alpha & 1 \\
\alpha^2 & \alpha
\end{bmatrix}, \quad \text{then} \quad (H_2)^+ = \begin{bmatrix}
\frac{\alpha}{(\alpha^2 + 1)^2} & \frac{\alpha^2}{(\alpha^2 + 1)^2} \\
\frac{1}{(\alpha^2 + 1)^2} & \frac{\alpha}{(\alpha^2 + 1)^2}
\end{bmatrix}.
\]

This can be easily verified by direct substitution into the four defining equations.

Case 2 \((n \geq 3)\). If \(H_n\) is the Hessenberg matrix of order \(n \geq 3\), then

\[
(H_n)^+ = \begin{bmatrix}
\frac{\alpha}{\alpha^2 + 1} & 0 & 0 & \cdots & 0 \\
\frac{1}{\alpha^2 + 1} & -\alpha & 1 & 0 & \cdots & 0 \\
-\alpha & 1 & 0 & \ddots & \ddots & \ddots \\
\cdots & \cdots & \cdots & \ddots & \ddots & \ddots \\
\cdots & \cdots & \cdots & \cdots & \ddots & \ddots \\
-\alpha & 1 & 0 & \cdots & \cdots & \frac{\alpha}{\alpha^2 + 1}
\end{bmatrix}.
\]

As in Case 1, this can be easily verified by direct substitution after noting
A CLASS OF HESSENBERG MATRICES

\[ H_n(H_n)^+ = \begin{bmatrix} I_{n-2} & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & I_{n-2} & 0 \\ \end{bmatrix}, \quad (H_n)^+H_n = \begin{bmatrix} \frac{\alpha^2}{\alpha^2+1} & \frac{\alpha}{\alpha^2+1} & \cdots & \frac{\alpha}{\alpha^2+1} \\ \frac{\alpha}{\alpha^2+1} & \frac{\alpha^2}{\alpha^2+1} & \cdots & \frac{\alpha^2}{\alpha^2+1} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\alpha}{\alpha^2+1} & \frac{\alpha^2}{\alpha^2+1} & \cdots & I_{n-2} \\ \end{bmatrix}, \]

where \( I_{n-2} \) is the identity matrix of order \( n-2 \).

4. Drazin Inverse of \( H_n \). The index of \( H_n \) is first determined. If \( \alpha = 0 \), then \( H_n \) is nilpotent and \( \text{Ind}(H_n) = n \).

**Proposition.** If \( \alpha \neq 0 \), then \( \text{Ind}(H_n) = \lceil n/2 \rceil \).

**Proof.** Meyer [7] has shown that \( k \) is the index of \( H_n \) if \( k \) is the smallest integer such that \( \lim_{\epsilon \to 0} e^k(H_n + \epsilon I_n)^{-1} \) exists. From [2] it is known that, if \( (h_{ij}) = (H_n + \epsilon I_n)^{-1} \), then

\[
h_{ij} = \frac{(-1)^{i+j}\Delta_{i-1}^j \Delta_{n-j+1}^n}{\epsilon \Delta_n^i}, \quad \text{if } i \leq j,
\]

and

\[
h_{ij} = -\frac{\alpha(\epsilon\alpha)^{i-j}\Delta_{i-2}^j \Delta_{n-i}^n}{\Delta_n^i}, \quad \text{if } i > j,
\]

where

\[
\Delta_0 = 1, \quad \Delta_0' = 1, \quad \Delta_{-1}' = 1/\epsilon,
\]

\[
\Delta_1 = \epsilon, \quad \Delta_1' = \alpha + \epsilon,
\]

\[
\Delta_t = \epsilon \Delta_{t-1} + \alpha \epsilon \Delta_{t-2}, \quad \Delta_t' = \Delta_t + \alpha \Delta_{t-1}.
\]

Each \( \Delta_t \) is a polynomial of degree \( i \) in \( \epsilon \).

First, in computing the

\[
\lim_{\epsilon \to 0} e^k h_{ij} = \lim_{\epsilon \to 0} e^k \frac{(-1)^{i+j} \Delta_{i-1}^j \Delta_{n-j+1}^n}{\epsilon \Delta_n^i}
\]

(when \( i \leq j \)) attention is directed to the terms of smallest power in \( \epsilon \) of \( \Delta_t \) and \( \Delta_t' \).

Observe that

1. The exponent of \( \epsilon \) in the term of smallest degree in \( \Delta_t \) is \( t/2 \) when \( t \) is even and \((t+1)/2\) otherwise,

2. the exponent of \( \epsilon \) in the term of smallest degree in \( \Delta_t' \) is \([t/2]\).

Since, for a given \( n \), \( \epsilon \Delta_t' \) is fixed, the value of \( k \) depends on \( \Delta_{i-1}' \Delta_{n-j+1} \). In this polynomial, the exponent of \( \epsilon \) is minimum when \( i \) is smallest and \( j \) is largest.

Therefore, the integer \( k \) for which \( \lim_{\epsilon \to 0} e^k h_{1n} \) exists is also an integer for which \( \lim_{\epsilon \to 0} e^k h_{ij} \) exists \((i \leq j)\). Now

\[
\lim_{\epsilon \to 0} e^k h_{1n} = \lim_{\epsilon \to 0} e^k \frac{(-1)^{1+n} \Delta_0' \Delta_1}{\epsilon \Delta_n'} = \lim_{\epsilon \to 0} \frac{(-1)^{1+n} \epsilon^{k+1}}{\epsilon \Delta_n'}
\]

exists if \( k \geq \lceil n/2 \rceil \).
Second, when \( i > j \), a similar argument will show that \( \lim_{\varepsilon \to 0} e^{kH_{ij}} \) exists whenever \( \lim_{\varepsilon \to 0} e^{kH_{nj}} \) exists, which is true when \( k \geq [n/2] \).

Therefore, \( \lim_{\varepsilon \to 0} e^{k(H_n + \varepsilon I_n)^{-1}} \) exists whenever \( k \geq [n/2] \). To see that \( k \) is the smallest such integer observe that

\[
\lim_{\varepsilon \to 0} e^{q/2} - 1^p h_n = \lim_{\varepsilon \to 0} \frac{e{(n/2) - 1}(-1)^{+1} n \varepsilon}{\varepsilon A_n'}
\]

does not exist since the exponent of \( \varepsilon \) in the term of smallest degree in \( \Delta_n' \) is \([n/2]\).

This completes the proof.

The Drazin inverse is now determined for \( H_n \), using elementary divisor theory (see [8]) and a technique described in [9]. Consider the characteristic matrix \((H_n - \lambda I_n)\). Let \( P_n(\lambda) \) denote the characteristic polynomial of \( H_n \) where \( P_n(\lambda) = \det(H_n - \lambda I_n) \). Also,

\[
P_n(\lambda) = (\alpha - \lambda)P_{n-1}(\lambda) + \sum_{j=1}^{n-1} (-1)^{n-j} \alpha^{n-1-j} P_{j-1}(\lambda),
\]

where \( P_0(\lambda) = 1, P_1(\lambda) = \alpha - \lambda, \) and \( P_n(\lambda) + \lambda P_{n-1}(\lambda) + \alpha P_{n-2}(\lambda) = 0 \) (expanding by rows). The solution of this last equation [4] is

\[
P_n(\lambda) = (-\alpha)^n 2^{n-1}(\lambda/4\alpha)^{(n-1)/2} \frac{\sin[(n + 2)\cos^{-1}(\lambda/4\alpha)^{1/2}]}{\sin[\cos^{-1}(\lambda/4\alpha)^{1/2}]}.\]

Since all determinantal divisors \( d_i \) corresponding to \( H_n \) are equal to one except \( d_n = \det(H_n - \lambda I_n) \), the minimum polynomial of \( H_n \) is \( P_n(\lambda) \). Suppose

\[
H_n = p^{-1} \begin{bmatrix} B & 0 \\ 0 & N \end{bmatrix} p,
\]

where \( B \) is nonsingular and \( N \) is nilpotent. Since \( \text{Ind}(H_n) = [n/2] \), the minimum polynomial of \( B \) is \( f(\lambda) = P_n(\lambda)/\lambda^{[n/2]} \) or

\[
f(\lambda) = \begin{cases} 
-\alpha^{(n+1)/2} \frac{\sin[(n + 2)\cos^{-1}(\lambda/4\alpha)^{1/2}]}{\sin[\cos^{-1}(\lambda/4\alpha)^{1/2}]} , & \text{if } n \text{ is odd}, \\
\alpha^{(n+1)/2} \frac{\lambda^{-1/2} \sin[(n + 2)\cos^{-1}(\lambda/4\alpha)^{1/2}]}{\sin[\cos^{-1}(\lambda/4\alpha)^{1/2}]} , & \text{if } n \text{ is even}.
\end{cases}
\]

The degree of \( f(\lambda) \) is \((n + 1)/2 \) or \( n/2 \) depending on whether \( n \) is odd or even. If \( f(\lambda) = a_n \lambda^n + a_{n-1} \lambda^{n-1} + \cdots + a_1 \lambda + a_0 = 0 \), then set

\[
g(\lambda) = \lambda^{-1} = -\frac{1}{a_0} (a_n \lambda^{n-1} + \cdots + a_2 \lambda + a_1).
\]

Therefore [9], if \( h(\lambda) = \lambda^{[n/2]} g^{[n/2]} + 1(\lambda) \), then \((H_n)^D = h(H_n) \). If \( \alpha = 0 \), \( (H_n)^D = 0 \).

Example. Consider

\[
H = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix},
\]
where \( \alpha = 1 \) and \( n = 3 \).

The index of \( H \) is 1, so that

\[
 f(\lambda) = -\frac{\sin \left[ 5 \cos^{-1} \left( \frac{\lambda}{4} \right)^{1/2} \right]}{\sin \left[ \cos^{-1} \left( \frac{\lambda}{4} \right)^{1/2} \right]}. 
\]

If \( \lambda = 4 \cos^2(\theta) \), then

\[
 f(\lambda) = -\frac{\sin(5\theta)}{\sin(\theta)} = -\frac{1}{\sin(\theta)} \left[ 8 \cos^2(\theta)\sin(\theta) - 16 \sin^3(\theta) \right] 
= -(8 \cos^2(\theta) - 16 \sin^2(\theta) \cos^2(\theta) - 3 + 4 \sin^2(\theta)).
\]

Upon substitution of \( \cos(\theta) = (\lambda/4)^{1/2}, \sin(\theta) = ((4 - \lambda)/4)^{1/2} \), \( f(\lambda) = -(\lambda^2 - 3\lambda + 1) \) is the minimum polynomial for \( B \). Therefore, \( h(\lambda) = \lambda^2(\lambda) = \lambda(3 - \lambda)^2 = 9\lambda - 6\lambda^2 + \lambda^3 \) and

\[
 H^D = h(H) = 9H - 6H^2 + H^3 = \begin{bmatrix} 2 & 2 & -3 \\ -1 & -1 & 2 \\ -1 & -1 & 2 \end{bmatrix}. 
\]