Journal of Computational Mathematics, Vol.8, No.1, 1990, 23–29.

THE DRAZIN INVERSE OF HESSENBERG MATRICES*

Miao Jian-ming
(*Shanghai Normal University, Shanghai, China*)

Abstract

The Drazin inverse of a lower Hessenberg matrix is considered. If A is a singular lower Hessenberg matrix, and $a_{i,i+1} \neq 0, i = 1, 2, \cdots, n-1$, then A^D can be given, and expressed explicitly by elements of A. The structure of the Drazin inverse of a lower Hessenberg matrix is also studied.

§1. Introduction

The Drazin inverse A^D of A can be characterized as the unique matrix satisfying the three equations

$$XAX = X, \tag{1.1a}$$

$$AX = XA, \tag{1.1b}$$

$$A^{k+1}X = A^k, \ \text{Index}(A) = k. \tag{1.1c}$$

If $k = 1$, then the Drazin inverse of A is called the group inverse of A and denoted by $A^{\#}$.

The Drazin inverse has been shown to have numerous applications[1]. In [2], it is used to give a closed form for solutions of systems of linear differential equations with singular coefficient matrices. In [5], $A^{\#}$ is used to study finite Markov chains. See [1] for an extensive discussion of the Drazin inverse.

Since an arbitrary square matrix can be reduced to Hessenberg form by means of unitary similarity transformations, and the Drazin inverse is well behaved with respect to similarity, that is $(P^{-1}AP)^D = P^{-1}A^DP$, the study of the Drazin inverse of a Hessenberg matrix is important.

Yasuhiko Ikebe[4] studied the structure of the inverse of a Hessenberg matrix. Cao Wei-lu and W. J. Stewart[3] generalized some results of [4]. Miao Jian-ming[6] presented the form of the group inverse of a Hessenberg matrix, and also generalized the results to more general matrices, called block upper (lower) s-diagonal matrices.

In this paper, we shall present the form of the Drazin inverse of a singular lower Hessenberg matrix with $a_{i,i+1} \neq 0, i = 1, 2, \cdots, n-1$.

§2. Notations and Preliminary Results

Throughout this paper, let $A = (a_{ij})_1^n$ be a lower Hessenberg matrix of order n, and $a_{i,i+1} \neq 0, i = 1, 2, \cdots, n-1$. Let A be partitioned as

$$A = \left(\begin{array}{c|cccc} a_{11} & a_{12} & & & \\ \vdots & \vdots & \ddots & & \\ a_{n-1,1} & a_{n-1,2} & \cdots & a_{n-1,n} \\ \hline a_{n1} & a_{n2} & \cdots & a_{n,n} \end{array} \right) \equiv \begin{pmatrix} c & B \\ \alpha & d^* \end{pmatrix}. \tag{2.1}$$

* Received April 8, 1987.

Clearly B is a lower triangular matrix, and B^{-1} exists. Let

$$T = \begin{pmatrix} 0 & 0 \\ B^{-1} & 0 \end{pmatrix}, \tag{2.2}$$

$$x = \begin{pmatrix} 1 \\ -B^{-1}c \end{pmatrix}, \tag{2.3}$$

$$y^* = (-d^* B^{-1}, 1). \tag{2.4}$$

Then we have

Lemma 1[3]. *If* $\alpha_1 = d^* B^{-1} c - \alpha \neq 0$, *then* A *is nonsingular, and*

$$A^{-1} = T - \alpha_1^{-1} x y^*. \tag{2.5}$$

If $\alpha_1 = 0$, *then* A *is singular.*

Lemma 2. *Let*

$$\alpha_1 = d^* B^{-1} c - \alpha, \quad \alpha_i = y^* T^{i-2} x, \quad i = 2, 3, \cdots. \tag{2.6}$$

If $\alpha_1 = \alpha_2 = \cdots = \alpha_k = 0, \alpha_{k+1} \neq 0$, *then* $\mathrm{Index}(A) = k$.

Proof. By [1, p.138], $\mathrm{Index}(A) = k$ is equivalent to saying k is the smallest nonnegative integer such that the limit

$$\lim_{\lambda \to 0} \lambda^k (A + \lambda I)^{-1} \tag{2.7}$$

exists. Now using Lemma 1 gives

$$(A + \lambda I)^{-1} = T(\lambda) - \alpha_1^{-1}(\lambda) x(\lambda) y^*(\lambda). \tag{2.8}$$

where $\lim_{\lambda \to 0} T(\lambda) = T$, $\lim_{\lambda \to 0} x(\lambda) = x$, $\lim_{\lambda \to 0} y^*(\lambda) = y^*$. Hence k is the smallest nonnegative integer such that the limit

$$\lim_{\lambda \to 0} \lambda^k \alpha_1^{-1}(\lambda) \tag{2.9}$$

exists. Let $J_0 \in C^{(n-1)\times(n-1)}$, and $f_1, f_{n-1} \in C^{(n-1)\times 1}$ denote the matrices

$$J_0 = \begin{pmatrix} 0 & & 0 \\ 1 & 0 & \\ & \ddots & \ddots & \\ 0 & 1 & 0 \end{pmatrix}, \quad f_1 = (1, 0, \cdots, 0)^T, f_{n-1} = (0, \cdots, 0, 1)^T.$$

Then

$$\alpha_1(\lambda) = (d^* + \lambda f_{n-1}^T)(B + \lambda J_0)^{-1}(c + \lambda f_1) - \alpha. \tag{2.10}$$

Let $N = B^{-1} J_0$. Then N is nilpotent of index $n-1$. Hence

$$(B + \lambda J_0)^{-1} = (I + \lambda N)^{-1} B^{-1} = [I - \lambda N + \lambda^2 N^2 - \cdots + (-1)^{n-2} \lambda^{n-2} N^{n-2}] B^{-1}. \tag{2.11}$$

By use of (2.11), (2.10) becomes

$$\alpha_1(\lambda) = \alpha_1 - \beta_2 \lambda + \beta_3 \lambda^2 - \cdots + (-1)^n \beta_{n+1} \lambda^n, \tag{2.12}$$

where

$$\beta_2 = (d^* B^{-1} J_0 - f_{n-1}^T) B^{-1} c - d^* B^{-1} f_1, \tag{2.13a}$$

$$\beta_i = (d^* B^{-1} J_0 - f_{n-1}^T) N^{i-3} B^{-1} (J_0 B^{-1} c - f_1), \quad i = 3, 4, \cdots, n+1. \tag{2.13b}$$

By definitions of x, y, J_0, f_1 and f_{n-1}, we have

$$d^*B^{-1}J_0 - f_{n-1}^T = -(y_2^*, y_3^*, \cdots, y_n^*), \tag{2.14}$$

$$J_0B^{-1}c - f_1 = -(x_1, x_2, \cdots, x_{n-1})^T. \tag{2.15}$$

Then from (2.13a), we have

$$\beta_2 = y^*x = \alpha_2. \tag{2.16}$$

Since from (2.6), we have

$$\begin{aligned}\alpha_i &= y^*\begin{pmatrix} 0 & 0 \\ B^{-1} & 0 \end{pmatrix}^{i-3}\begin{pmatrix} 0 & 0 \\ B^{-1} & 0 \end{pmatrix}x = y^*\begin{pmatrix} 0 & 0 \\ 0 & N \end{pmatrix}^{i-3}\begin{pmatrix} 0 & 0 \\ B^{-1} & 0 \end{pmatrix}x \\ &= [0, (y_2^*, \cdots, y_n^*)N^{i-3}]Tx = [(y_2^*, \cdots, y_n^*)N^{i-3}B^{i}, 0]x \\ &= (y_2^*, \cdots, y_n^*)N^{i-3}B^{-1}(x_1, \cdots, x_{n-1})^T, \quad i = 3, 4, \cdots,\end{aligned}$$

therefore

$$\beta_i = \alpha_i = y^*T^{i-2}x, \quad i = 3, 4, \cdots, n+1. \tag{2.17}$$

Hence by assumptions, we have

$$\lim_{\lambda \to 0} \lambda^k \alpha_1^{-1}(\lambda) = (-1)^k/\alpha_{k+1}.$$

Thus we conclude that $\mathrm{Index}(A) = k$.

Note that T is nilpotent of index n. Therefore $\alpha_i = 0$ for $i \geq n+2$.

Let e_1, e_n denote the first and the last column of the identity matrix I_n. Then we have

Lemma 3. *If* $\alpha_1 = \alpha_2 = \cdots = \alpha_k = 0, \alpha_{k+1} \neq 0$, *then*

(i) $Ax = 0, AT^ix = T^{i-1}x, i = 1, 2, \cdots, k-1, AT^kx = T^{k-1}x - \alpha_{k+1}e_n$;

(ii) $y^*A = 0, y^*T^iA = y^*T^{i-1}, i = 1, 2, \cdots, k-1, y^*T^kA = y^*T^{k-1} - \alpha_{k+1}e_1^T$;

(iii) $A^kT^ix = 0, i = 0, 1, \cdots, k-1$;

(iv) $y^*T^iA^k = 0, i = 0, 1, \cdots, k-1$.

Proof. Since

$$Ax = -\alpha_1 e_n,$$

$$AT^ix = \begin{pmatrix} I_{n-1} & 0 \\ d^*B^{-1} & 0 \end{pmatrix}T^{i-1}x = [I - \begin{pmatrix} 0 \\ y^* \end{pmatrix}]T^{i-1}x = T^{i-1}x - \alpha_{i+1}e_n, i = 1, 2, \cdots$$

by assumptions, we have (i).

The proof of (ii) is similar to that of (i), and so is omitted.

By using (i), we have

$$A^kT^ix = A^{k-1}T^{i-1}x = \cdots = A^{k-i}x = 0, \quad i = 0, 1, \cdots, k-1.$$

(iv) is similar to (iii), and is omitted.

§3. Representation of A^D

Theorem 1. *Let $A = (a_{ij})_1^n$ be a lower Hessenberg matrix, $a_{i,i+1} \neq 0, i = 1, 2, \cdots, n-1$, and A be partitioned as in (2.1). Let α_i be defined by (2.6). If $\alpha_1 = \alpha_2 = \cdots = \alpha_k = 0, \alpha_{k+1} \neq 0$, then* Index$(A) = k$, *and*

$$A^D = T + r_1 \sum_{i=0}^{k} T^i xy^* T^{k-i} + r_2 \sum_{i=0}^{k-1} T^i xy^* T^{k-1-i} + \cdots + r_{k+1} xy^*, \tag{3.1}$$

where

$$\begin{cases} r_1 = -1/\alpha_{k+1}, \\ r_p = -(r_1\alpha_{k+p} + r_2\alpha_{k+p-1} + \cdots + r_{p-1}\alpha_{k+2})/\alpha_{k+1}, \quad p = 2, 3, \cdots, k+1. \end{cases} \tag{3.2}$$

Proof. Let X denote the right-hand side of (3.1). We shall prove that X satisfies the three equations in (1.1).

By Lemma 3(i), we have

$$\begin{aligned} AX &= \begin{pmatrix} I_{n-1} & 0 \\ d^*B^{-1} & 0 \end{pmatrix} - r_1\alpha_{k+1}e_n y^* + r_1 \sum_{i=1}^{k} T^{i-1} xy^* T^{k-i} \\ &\quad + r_2 \sum_{i=1}^{k-1} T^{i-1} xy^* T^{k-1-i} + \cdots + r_k xy^* \\ &= I + r_1 \sum_{i=1}^{k} T^{i-1} xy^* T^{k-i} + r_2 \sum_{i=1}^{k-1} T^{i-1} xy^* T^{k-1-i} + \cdots + r_k xy^*. \end{aligned}$$

By Lemma 3(ii), we have

$$\begin{aligned} XA &= \begin{pmatrix} 0 & 0 \\ B^{-1}c & I_{n-1} \end{pmatrix} - r_1\alpha_{k+1} x e_1^T + r_1 \sum_{i=0}^{k-1} T^i xy^* T^{k-1-i} \\ &\quad + r_2 \sum_{i=0}^{k-2} T^i xy^* T^{k-2-i} + \cdots + r_k xy^* \\ &= I + r_1 \sum_{i=0}^{k-1} T^i xy^* T^{k-1-i} + r_2 \sum_{i=0}^{k-2} T^i xy^* T^{k-2-i} + \cdots + r_k xy^*. \end{aligned}$$

Hence we have

$$AX = XA = I + r_1 \sum_{i=0}^{k-1} T^i xy^* T^{k-1-i} + r_2 \sum_{i=0}^{k-2} T^i xy^* T^{k-2-i} + \cdots + r_k xy^*. \tag{3.3}$$

Now, using Lemma 3(iii) together with (3.3) gives

$$A^{k+1}X = A^k(AX) = A^k. \tag{3.4}$$

Finally,

$$XAX = X(AX) = X + r_1 \sum_{i=0}^{k-1} T^{i+1}xy^*T^{k-1-i} + r_2 \sum_{i=0}^{k-2} T^{i+1}xy^*T^{k-2-i} + \cdots + r_k Txy^*$$

$$+r_1^2\left(\alpha_{k+1}\sum_{i=0}^{k-1} T^{i+1}xy^*T^{k-1-i} + \alpha_{k+2}\sum_{i=0}^{k-1} T^{i}xy^*T^{k-1-i} + \alpha_{k+3}\sum_{i=0}^{k-2} T^{i}xy^*T^{k-2-i} + \cdots\right)$$

$$+r_2r_1\left(\alpha_{k+1}\sum_{i=0}^{k-1} T^{i}xy^*T^{k-1-i} + \alpha_{k+2}\sum_{i=0}^{k-2} T^{i}xy^*T^{k-2-i} + \cdots\right) + \cdots + r_{k+1}r_1\alpha_{k+1}xy^*$$

$$+r_1r_2\left(\alpha_{k+1}\sum_{i=0}^{k-2} T^{i+1}xy^*T^{k-2-i} + \alpha_{k+2}\sum_{i=0}^{k-2} T^{i}xy^*T^{k-2-i} + \alpha_{k+3}\sum_{i=0}^{k-3} T^{i}xy^*T^{k-3-i} + \cdots\right)$$

$$+r_2^2\left(\alpha_{k+1}\sum_{i=0}^{k-2} T^{i}xy^*T^{k-2-i} + \alpha_{k+2}\sum_{i=0}^{k-3} T^{i}xy^*T^{k-3-i} + \cdots\right) + \cdots + r_kr_2\alpha_{k+1}xy^*$$

$$+ \cdots + r_1r_k(\alpha_{k+1}Txy^* + \alpha_{k+2}xy^*) + r_2r_k\alpha_{k+1}xy^*$$

$$= X + (1 + r_1\alpha_{k+1})r_1 \sum_{i=0}^{k-1} T^{i+1}xy^*T^{k-1-i} + (r_1\alpha_{k+2} + r_2\alpha_{k+1})r_1 \sum_{i=0}^{k-1} T^{i}xy^*T^{k-1-i}$$

$$+ \cdots + (r_1\alpha_{2k+1} + r_2\alpha_{2k} + \cdots + r_{k+1}\alpha_{k+1})r_1xy^*$$

$$+(1 + r_1\alpha_{k+1})r_2 \sum_{i=0}^{k-2} T^{i+1}xy^*T^{k-2-i} + (r_1\alpha_{k+2} + r_2\alpha_{k+1})r_2 \sum_{i=0}^{k-2} T^{i}xy^*T^{k-2-i}$$

$$+ \cdots + (r_1\alpha_{2k} + r_2\alpha_{2k-1} + \cdots + r_k\alpha_{k+1})r_2xy^* + \cdots$$

$$+(1 + r_1\alpha_{k+1})r_kTxy^* + (r_1\alpha_{k+2} + r_2\alpha_{k+1})r_kxy^*.$$

Therefore, using (3.2) gives

$$XAX = X. \tag{3.5}$$

This completes the proof of Theorem 1.

Note that if we consider a singular upper Hessenberg matrix with $a_{i+1,i} \neq 0, i = 1, 2, \cdots, n-1$, and apply Theorem 1 to A^*, a similar result on the form of the Drazin inverse for an upper Hessenberg matrix is obtained.

Theorem 2. *Let $A = (a_{ij})_1^n$ be an upper Hessenberg matrix with $a_{i+1,i} \neq 0, i = 1, 2, \cdots, n-1$, and A be partitioned as*

$$A = \left(\begin{array}{ccccc} a_{11} & a_{12} & \cdots & a_{1,n-1} & \vdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots\cdots & \cdots & \cdots \\ a_{21} & a_{22} & \cdots & a_{2,n-1} & \vdots & a_{2n} \\ & \ddots & \ddots & \vdots & \vdots & \vdots \\ & & & a_{n,n-1} & \vdots & a_{nn} \end{array}\right) = \begin{pmatrix} \tilde{d}^* & \tilde{\alpha} \\ \tilde{B} & \tilde{c} \end{pmatrix}. \tag{3.6}$$

Clearly, $\tilde{B}$ is an upper triangular matrix, and $\tilde{B}^{-1}$ exists. Let

$$\tilde{T} = \begin{pmatrix} 0 & \tilde{B}^{-1} \\ 0 & 0 \end{pmatrix}, \tag{3.7}$$

$$\tilde{x} = \begin{pmatrix} -\tilde{B}^{-1}\tilde{c} \\ 1 \end{pmatrix}, \tag{3.8}$$

$$\tilde{y}^* = (1, -\tilde{d}^*\tilde{B}^{-1}), \tag{3.9}$$

$$\tilde{\alpha}_1 = \tilde{d}^*\tilde{B}^{-1}\tilde{c} - \tilde{\alpha}, \tilde{\alpha}_i = \tilde{y}^*\tilde{T}^{i-2}\tilde{x}, \quad i = 2, 3, \cdots. \tag{3.10}$$

If $\tilde{\alpha}_1 = \tilde{\alpha}_2 = \cdots = \tilde{\alpha}_k = 0, \tilde{\alpha}_{k+1} \neq 0$, then $\text{Index}(A) = k$, and

$$A^D = \tilde{T} + r_1 \sum_{i=0}^{k} \tilde{T}^i \tilde{x}\tilde{y}^* \tilde{T}^{k-i} + r_2 \sum_{i=0}^{k-1} \tilde{T}^i \tilde{x}\tilde{y}^* \tilde{T}^{k-i-1} + \cdots + r_{k+1}\tilde{x}\tilde{y}^* \tag{3.11}$$

where

$$\begin{cases} r_1 = -1/\tilde{\alpha}_{k+1}, \\ r_p = -(r_1\tilde{\alpha}_{k+p} + r_2\tilde{\alpha}_{k+p-1} + \cdots + r_{p-1}\tilde{\alpha}_{k+2})/\tilde{\alpha}_{k+1}, \quad p = 2, 3, \cdots, k+1. \end{cases} \tag{3.12}$$

If A is a lower Hessenberg matrix, and $a_{i,i+1} = 0$ for some i, then A can be partitioned as a lower block-triangular matrix, and the following formular[1] is applicable:

$$\begin{pmatrix} C & 0 \\ B & D \end{pmatrix}^D = \begin{pmatrix} C^D & 0 \\ X & D^D \end{pmatrix}$$

where C, D are square, $\text{Index}(C) = 1$, $\text{Index}(D) = k$, and

$$X = (D^D)^2 \Big[\sum_{i=0}^{l-1} (D^D)^i BC^i\Big](I - CC^D) + (I - DD^D)\Big[\sum_{i=0}^{k-1} D^i B (C^D)^i\Big](C^D)^2 - D^D BC^D.$$

§4. The Structure of A^D

In [4], Yasuhiko Ikebe showed that the upper half of the inverse of a lower Hessenberg matrix has a simple structure. He proved the following theorem:

Theorem 3. *Let $A = (a_{ij})_1^n$ be a lower Hessenberg matrix and let $a_{i,i+1} \neq 0, i = 1, 2, \cdots, n-1$. Let $A^{-1} = (\alpha_{ij})$ exist. Then two column vectors $x = (x_1, \cdots, x_n)^T$ and $y = (y_1, \cdots, y_n)^T$ exist such that the upper half of A^{-1} equals the upper half of xy^T, i.e., $\alpha_{ij} = x_i y_j$ for $i \leq j$.*

Using Theorem 1, we can obtain the following theorem on the structure of the Drazin inverse of a lower Hessenberg matrix, which is a generalization of Theorem 3. Let

$$x_i = T^i x, i = 0, 1, \cdots, k, \tag{3.1}$$

$$y_i^* = y^* T^i, i = 0, 1, \cdots, k. \tag{3.2}$$

Then we have

Theorem 4. *Let $A = (a_{ij})_1^n$ be a lower Hessenberg matrix and let $a_{i,i+1} \neq 0, i = 1, 2, \cdots, n-1$. Let $\text{Index}(A) = k$. Then there are two sets of column vectors $\{x_i\}_0^k$ and $\{y_i\}_0^k$ such that the upper half of A^D equals the upper half of*

$$r_1 \sum_{i=0}^{k} x_i y_{k-i}^* + r_2 \sum_{i=0}^{k-1} x_i y_{k-1-i}^* + \cdots + r_{k+1} x_0 y_0^*, \tag{3.3}$$

where r_i are defined as in (3.2).

Note. By the definition of T in (2.2), the first (last) i components of $x_i (y_i)$ are zeros.

For the special case that $\text{Index}(A) = 0, A^D$ becomes A^{-1}, and Theorem 3 is obtained from Theorem 4.

§5. Example

Let

$$A = \frac{1}{2}\begin{pmatrix} 2 & 2 & 0 \\ 2 & 2 & 2 \\ 2 & 2 & 1 \end{pmatrix}.$$

Then

$$c = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \quad \alpha = 1, \quad d^T = (1, \frac{1}{2}).$$

Thus

$$B^{-1} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}, \quad T = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ -1 & 1 & 0 \end{pmatrix}, \quad x = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad y^T = \frac{1}{2}(-1, -1, 2),$$

and

$$\alpha_1 = d^T B^{-1} c - \alpha = 0, \quad \alpha_2 = y^T x = 0, \quad \alpha_3 = y^T T x = -\frac{5}{2}.$$

Hence

$$k = 2, \quad x_0 = x, \quad x_1 = Tx = \begin{pmatrix} 0 \\ 1 \\ -2 \end{pmatrix}, \quad x_2 = T^2 x = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

$$y_0^T = y^T, \quad y_1^T = y^T T = (-\frac{3}{2}, 1, 0), \quad y_2^T = y^T T^2 = (1, 0, 0),$$

$$r_1 = -1/\alpha_3 = \frac{2}{5}, \quad r_2 = r_1^2 \alpha_4 = \frac{4}{25}, \quad r_3 = r_1(r_1\alpha_5 + r_2\alpha_4) = \frac{8}{125}.$$

Then we have

$$A^D = T + r_1(x_0 y_2^T + x_1 y_1^T + x_2 y_0^T) + r_2(x_0 y_1^T + x_1 y_0^T) + r_3 x_0 y_0^T = \frac{2}{125}\begin{pmatrix} 8 & 8 & 4 \\ 12 & 12 & 6 \\ 10 & 10 & 5 \end{pmatrix}.$$

The author wishes to thank his advisors Profs. Wang Guo-rong and Kuang Jiao-xun for many useful suggestions.

References

[1] S. L. Campbell and C. D. Meyer, Jr., Generalized Inverses of Linear Transformations, Pitman, London, 1979.

[2] S. L. Campbell, C. D. meyer, Jr. and N. J. Rose, Applications of the Drazin inverse to linear systems of differential equations, *SIAM J. Appl. Math.*, **31** (1976), 411–425.

[3] Cao Wei–lu and W. J. Stewart, A note on inverse of Hessenberg–like matrices, *Linear Algebra Appl.*, **76** (1986), 233–240.

[4] Yasuhiko Ikebe, On inverses of Hessenberg matrices, *Linear Algebra Appl.*, **24** (1979), 93–97.

[5] C. D. Meyer, Jr., The role of the group generalized inverse in the theory of finite Markov chains, *SIAM Rev.*, **17** (1975), 433–464.

[6] Miao Jian–ming, The group inverse and the Moore–Penrose inverse of Hessenberg matrices, submitted to *Linear Algebra Appl.*