On Inverses of Hessenberg Matrices*

Yasuhiko Ikebe

Department of Computer Sciences
Northwestern University
Evanston, Illinois 60201

Submitted by Alston Householder

ABSTRACT

The lower half of the inverse of a lower Hessenberg matrix is shown to have a simple structure. The result is applied to find an algorithm for finding the inverse of a tridiagonal matrix. With minor modifications, the technique applies to block Hessenberg matrices.

1. INTRODUCTION

A square matrix $A=(a_{ij})$, $i,j=1,\dots,n$, is called a *lower(upper) Hessenberg matrix* if $a_{ij}=0$ for all pairs (i,j) such that $i+1<j$ $(j+1<i)$. We shall prove the following theorem on the structure of the inverse of a lower Hessenberg matrix:

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

By taking the transpose, we see that a similar theorem holds for an upper Hessenberg matrix.

By the symbol $A=\{a_i,b_i,c_i\}_1^n$, we denote the tridiagonal matrix

$$A=\begin{bmatrix} b_1 & c_1 & & & 0 \\ a_2 & b_2 & c_2 & & \\ & \ddots & \ddots & \ddots & \\ & & a_{n-1} & b_{n-1} & c_{n-1} \\ 0 & & & a_n & b_n \end{bmatrix}.$$

*This work was supported in part by Air Force grant AFOSR 77-3183.

0024-3795/79/020093+5 $01.75

A tridiagonal matrix is a square matrix which is both a lower and an upper Hessenberg matrix. As a direct consequence of Theorem 1 and of the remark that follows, we obtain the following theorem on the structure of the inverse of a tridiagonal matrix.

THEOREM 2. *Let a tridiagonal matrix* $A=\{a_i,b_i,c_i,\}_1^n$ *be given such that none of the* a_i*'s and* c_i*'s vanish. Let* $A^{-1}=(\alpha_{ij})$ *exist. Then four column vectors* $u=(u_i)$, $v=(v_i)$, $x=(x_i)$ *and* $y=(y_i)$, $i=1,\dots,n$, *exist such that*

$$\alpha_{ij} = \begin{cases} u_i v_j, & i \geqslant j, \\ x_i y_j, & i \leqslant j. \end{cases}$$

Examination of the proof (see the next section) reveals that Theorems 1 and 2 may be extended to the case where the matrix elements are themselves square matrices of the same order:

THEOREM 3. *Let* $A=(A_{ij})$, $i,j=1,\dots,n$, *be a block lower Hessenberg matrix* ($A_{ij}=0$ *for* $i+1<j$), *where the* A_{ij} *are square matrices of a fixed order, say* m. *Let the superdiagonal blocks* $A_{i,i+1}$, $i=1,\dots,n-1$, *have inverses. Let* A^{-1} *exist, and write* $A^{-1}=B=(B_{ij})$, *where* B_{ij} *is* $m\times m$, $i,j=1,\dots,n$. *Then*

$$B_{ij} = X_i Y_j, \qquad i \leqslant j,$$

for a set of $m\times m$ *matrices* $X_1,\dots,X_n$, $Y_1,\dots,Y_n$.

THEOREM 4. *Let* $A=\{A_i,B_i,C_i\}_{i=1}^n$ *be a block tridiagonal matrix where the* A_i, B_i *and* C_i *are* $m\times m$ *matrices. Let* $A_i^{-1}(i=2,\dots,n)$ *and* C_i^{-1} $(i=1,\dots,n-1)$ *exist. Let* A^{-1} *exist, and write* $A^{-1}=B=(B_{ij})$, *where* B_{ij} *is* $m\times m$, $i,j=1,\dots,n$. *Then*

$$B_{ij} = \begin{cases} U_i V_j, & i \geqslant j, \\ X_i Y_j, & i \leqslant j, \end{cases}$$

for a set of $m\times m$ *matrices* $U_1,\dots,V_1,\dots,X_1,\dots,Y_1,\dots,Y_n$.

Theorem 2 was proved by Bukhberger and Emel'yanenko [1] under the further assumption that A is symmetric. If A is symmetric, so is A^{-1}. In this case A depends only on $2n-1$ numbers and Theorem 2 indicates that A^{-1} is representable in terms of $2n-1$ numbers. For $x_i y_j=(kx_i)\cdot(k^{-1}y_j)$, $i,j=1,\dots,n$, for any fixed nonzero constant k, and hence any one nonzero x_i or y_j

may be assigned an arbitrary nonzero value, say 1 (note that $x_1 \neq 0$). This is interesting in view of the fact that A^{-1} is usually a full matrix (all or almost all elements of A^{-1} are nonzero). Theorem 1 represents a generalization of Bukhberger and Emel'yanenko's result in [1]. Their proof depends on a determinantal identity for the matrix inverse. In this note, we give a determinant-free proof of Theorem 1. (See Sec. 2.) The proof is quite elementary and paves a way for further generalization to block matrices, as indicated by Theorems 3 and 4. A simple recursive algorithm may be developed for computing the x_i and y_j in Theorem 1, as we show in Sec. 3.

2. PROOF OF THEOREM 1

We will prove Theorem 1. Let P_k $(k=1,2,\dots,n-1)$ denote the square submatrix of A formed from the first k rows and the second through the $(k+1)$st columns of A. Thus,

$$P_1 = (a_{12}), \qquad P_2 = \begin{pmatrix} a_{12} & 0 \\ a_{22} & a_{23} \end{pmatrix}, \qquad \dots .$$

The P_k are lower triangular matrices with nonvanishing diagonal elements. Hence P_k^{-1} exists $(n=1,\dots,n-1)$. Let $A^{-1}=(\alpha_{ij})$, $i,j=1,\dots,n$, and let the first row of A^{-1} be denoted by $(y_1,y_2,\dots,y_n)^T$, $\alpha_{1j}=y_j$, $j=1,\dots,n$. Let $c^{(k)}$, $k=1,\dots,n$, denote the column vector of dimension k composed of the first k components of the first column of A. Thus,

$$c^{(1)} = (a_{11}), \qquad c^{(2)} = (a_{11},a_{21})^T, \qquad \dots .$$

Equating the first $j-1$ components of the jth column $(j=2,\dots,n)$ in $AA^{-1}=I$, we obtain

$$c^{(j-1)}y_j + P_{j-1}(\alpha_{2j},\dots,\alpha_{jj})^T = 0$$

or

$$(\alpha_{2j},\dots,\alpha_{jj})^T = -P_{j-1}^{-1}c^{(j-1)}y_j.$$

Since P_{j-1} is the leading $(j-1)\times(j-1)$ submatrix of P_{n-1}, which is lower triangular, P_{j-1}^{-1} gives the leading $(j-1)\times(j-1)$ submatrix of P_{n-1}^{-1}, which is again lower triangular. It follows that $-P_{j-1}^{-1}c^{(j-1)}$ is identified with the leading $j-1$ components of $-P_{n-1}^{-1}c^{(n-1)}$. If we denote this vector by

$(x_2,\dots,x_n)^T$ and let x_1 denote 1, we find

$$(\alpha_{1j},\dots,\alpha_{jj})^T = (x_1,\dots,x_j)^T y_j, \qquad j = 1,\dots,n.$$

This completes the proof of Theorem 1.

REMARK. A completely similar proof applies to Theorem 3.

3. ALGORITHM

We will describe a recursive algorithm for computing the x_i and y_i in Theorem 1. Equating the last column in $AA^{-1}=I$, we obtain

$$A\cdot(x_1,\dots,x_n)^T = \left(0,\dots,0,y_n^{-1}\right). \tag{3.1}$$

Equating the first row in $A^{-1}A=I$, we find

$$(y_1,\dots,y_n)\cdot A = \left(x_1^{-1},0,\dots,0\right). \tag{3.2}$$

Since $x_1\neq 0$ and $y_n\neq 0$, we may choose either x_1 or y_n to be an arbitrary nonzero number. Equations (3.1) and (3.2) now give the following algorithm for finding $x_1,\dots,x_n,y_1,\dots,y_n$:

(1) Let a nonzero number be given as the value of x_1 (say $x_1=1$);

(2) $x_2=-a_{12}^{-1}a_{11}x_1$;

(3) $x_i=-a_{i-1,i}^{-1}\sum_{k=1}^{i-1}a_{i-1,k}x_k,\ i=3,\dots,n$;

(4) $y_n=\left(\sum_{k=1}^{n}a_{n,k}x_k\right)^{-1}$;

(5) $y_{n-1}=-y_n a_{n,n}a_{n-1,n}^{-1}$;

(6) $y_i=-\left(\sum_{k=i+1}^{n}y_k a_{k,i+1}\right)\cdot a_{i,i+1}^{-1},\ i=n-2,\dots,2,\ 1$;

(7) $x_1\sum_{k=1}^{n}y_k a_{k,1}=1$ (for check).

REMARK 1. The above algorithm may be used for computing the x_i and y_i in Theorem 3 if we read a_{ij} as A_{ij}, x_i as X_i and y_j as Y_j and if the order of operations is kept as it is indicated in (1)–(7).

REMARK 2. The inverse of a tridiagonal matrix $A=\{a_i,b_i,c_i\}_1^n$, where $a_i\neq 0$, $i=2,\dots,n$, $c_i\neq 0$, $i=1,\dots,n-1$, may be computed by applying the above algorithm twice, once for computing the upper half of A^{-1} and once for computing the lower half of A^{-1}.

REFERENCES

1 B. Bukhberger and G. A. Emel'yanenko, Methods of inverting tridiagonal matrices, *USSR Computational Math. and Math. Phys.* 13:10–20 (1973); MR 48 #2155.

Received 3 October 1977