UCLA
COMPUTATIONAL AND APPLIED MATHEMATICS

On Factorisations of the Hessenberg Matrices Arising from Polynomial Iterative Methods

Victor Eijkhout

October 1996

CAM Report 96-44

Department of Mathematics
University of California, Los Angeles
Los Angeles, CA. 90024-1555

On Factorisations of the Hessenberg matrices arising from polynomial iterative methods*

Victor Eijkhout

November 1, 1996

Abstract

We present an investigation of the Hessenberg matrices that arise from polynomial iterative methods such as the conjugate gradients method, in particular of their various factorisations. Application of the results to iterative methods, including a proof that (quasi-) minimising iterative methods can always be written as a two-term residual smoothing method, is given.

1 Introduction

Many studies of iterative methods mention Hessenberg matrices. Since these studies are mostly concerned with minimisation and orthogonalisation properties of the iterative methods, it is never quite clear how much the properties of the Hessenberg matrices depend on such properties of the sequences they relate to.

This paper brings together various facts on these Hessenberg matrices, in particular, various theorems on their factorisations. At first, the presentation is isolated from the iterative methods the matrices arise from, but the results are then applied to the iterative methods, thus stressing the fundamentality of the Hessenberg matrix itself.

Section 2 introduces polynomial iterative methods and residual sequences, and shows how they produce Hessenberg matrices with zero column sums. Additionally, some simple results are derived for future reference. Section 3 brings together a number of results on the LU and QR factorisations of Hessenberg

*This research was supported by the ARO under contract DAAL-03-9-C-0047 (Univ. of Tenn. subcontract ORA 4466.04, Amendment 1 and 2).

matrices with zero column sums, prepatory to the theorem in Section 4 that (quasi-) minimising iterative methods can always be written as a two-term residual smoothing method.

2 Motivation: polynomial iterative methods

In this section we will define polynomial iterative methods, and show how they give rise to a special type of Hessenberg matrix, which we will then study in the rest of this paper.

Polynomial iterative methods for the solution of a linear system $Ax = f$ use the following basic ingredients: a starting guess x_1, the corresponding residual $r_1 = Ax_1 - f$, and the Krylov sequence

$$k_1 = r_1, \qquad k_{i+1} = Ak_i. \tag{1}$$

We write this more compactly by introducing a matrix notation for sequences, e.g., $X = (x_1, x_2, \ldots)$, and using a shift operator matrix

$$J = (\delta_{i,j+1}) = \begin{pmatrix} 0 & & \emptyset & \\ 1 & 0 & & \\ & 1 & 0 & \\ & & \ddots & \ddots \\ \emptyset & & & \end{pmatrix}.$$

The Krylov sequence definition then becomes $AK = KJ$.

Definition 1 *A polynomial method is defined as a sequence X defined by*

$$x_{i+1} = x_1 + \sum_{j \le i} k_j u_{ji} = x_1 + \sum_{j \le i} A^{j-1} k_1 u_{ji}$$

where K is the Krylov sequence (1) and u_{ji} are coefficients with $u_{ii} \neq 0$.

An equivalent definition is

$$x_{i+1} = x_i + \sum_{j \le i} k_j u_{ji}$$

with different choices of u_{ji}. Multiplying this equation by A, and noting that

$$Ax_{i+1} - Ax_i = Ax_{i+1} - f - Ax_i + f = r_{i+1} - r_i,$$

we find for the residuals that

$$r_{i+1} = r_i + \sum_{j \le i} Ak_j u_{ji}.$$

In matrix form we can write these equations as

$$X\begin{pmatrix} -1 & & & \\ 1 & -1 & & \\ & \ddots & \ddots & \end{pmatrix} = X(J-I) = KU, \qquad R(J-I) = AKU \tag{2}$$

with U a non-singular upper triangular matrix. This matrix formalism for iterative methods was first used in [1].

Since $R(J-I) = AKU = KJU$, $r_1 = k_1$, and

$$\begin{pmatrix} 1 & J-I \\ \emptyset & \end{pmatrix} = (I - J^t) \qquad \text{and} \qquad \begin{pmatrix} 1 & JU \\ \emptyset & \end{pmatrix} = \begin{pmatrix} 1 & \emptyset \\ \emptyset & U \end{pmatrix} \equiv V,$$

we find that

$$R(I-J^t) = KV \Rightarrow R = KV(I-J^t)^{-1},$$

that is, $R = KU$ for some non-singular upper triangular matrix U. From this, it follows that $r_n = P_i(A)k_1$ where $P_n(\cdot)$ is an $n-1$-st degree polymial with its coefficients in the n-th column:

$$P_n(t) = u_{nn}t^{n-1} + \cdots + u_{1n}.$$

From equation 2 and $R = KU$ we find that for some non-singular upper triangular matrix U:

$$R(J-I) = ARU \qquad \text{or} \qquad AR = R(J-I)U^{-1}. \tag{3}$$

We see that $AR = RH$ with $H = (J-I)U^{-1}$ an upper Hessenberg matrix. Note that H has zero column sums because $J - I$ has zero column sums, and is irreducible, i.e., $h_{i+1,i} \neq 0$.

Definition 2 *A sequence R is called a 'residual sequence' for the system $Ax = f$ if there is a polynomial iterative method X such that $r_i = Ax_i - f$.*

We summarise the above results:

Lemma 3 *A residual sequence R for the system $Ax - f$ satisfies $AR = RH$ where H is an irreducible upper Hessenberg matrix with zero column sums.*

Most of the rest of this paper will be concerned with such Hessenberg matrices with zero column sums.

We also state the converse of lemma 3:

Lemma 4 *A R satisfying $AR = RH$ where H is an irreducible upper Hessenberg matrix with zero column sums, is a residual sequence for a system with A.*

Proof (sketch only): We will show later that H can be split as $(J-I)V$. Now V can always be written as $U_2^{-1}U_1$ where $U_2 = \left(\begin{smallmatrix}1 & \emptyset \\ \emptyset & U\end{smallmatrix}\right)(I-J^t)$ for some upper triangular U[1]. Let f and X be such that $Ax_i - f = r_i$, and define $K = RU_2^{-1}$, then $R(J-I) = KJU_1$, whence $X(J-I) = A^{-1}KJU_1$, and from $AKU_1 = R(J-I)$ it follows that $X(J-I) = KU_1$. Therefore $AK = KJ$. Qed

We prove a simple lemma characterising when linear combinations of a residual sequence are again a residual sequence.

Lemma 5 *Let U be a non-singular upper triangular matrix with $u_{11} = 1$, and let H be an irreducible upper Hessenberg matrix. If H has zero column sums, then $U^{-1}HU$ has zero column sums iff U has column sums identically 1.*

Proof: Introducing the vectors $e^t = (1, 1, \ldots)$ and $0^t = (0, 0, \ldots)$, we can formulate the zero column sums of matrix H as $e^tH = 0^t$. The statement that U has column sums identically 1 translates as $e^tU = e^t$. Clearly, also $e^tU^{-1} = e^t$. Now, $e^tU^{-1}HU = e^tHU = 0^tU = 0^t$. For the reverse implication, note that

$$e^tU^{-1}HU = 0^t \Rightarrow e^tU^{-1}H = 0^t \Rightarrow e^tU^{-1} = \alpha e^t \qquad \text{for some scalar } \alpha,$$

and $\alpha = 1$ follows from $u_{11} = 1$. •

This lemma has the following implication for residual sequences: if R is a residual sequence, and G is a transformation by means of an upper triangular matrix, i.e., $G = RU$, or

$$g_i = \sum_{j \le i} r_j u_{ji},$$

then G is a residual sequence iff the combinations are convex, that is, $\sum_j u_{ji} = 1$ for all i. Convex combinations of residual sequences occur in two places: deriving accelerated methods from stationary methods, and deriving minimising methods such as GMRES and QMR from full orthogonalisation and bi-conjugate gradients respectively.

3 Factorisations of the Hessenberg matrix

In this section we will consider the LU and QR factorisations of the Hessenberg matrices of residual sequences. The statements derived will be applied in the next section.

Above we saw that such Hessenberg matrices are given in the form $H = (J-I)U$ (equation 3). For future use we prove a simple lemma, showing that such a factorisation always exists for Hessenberg matrices with zero column sums.

[1]The proof is elementary but tedious.

Lemma 6 *Let H be an upper Hessenberg matrix with zero column sums. If H has maximum rank, it can be factored as $H = (J - I)U$.*

Proof: The zero column sums imply that $h_{21} = -h_{11}$. Since the matrix has maximum rank, the elements in the first column are nonzero. The first elimination step then entails adding the first row to the second. This does not change the column sums of any column, nor does it change the (column) rank of the remaining block. Hence we can inductively repeat this argument. •

Sometimes we are interested in a factorisation of the form $U(J - I)$ rather than $(J - I)U$.

Lemma 7 *A matrix $(I-J)B$ with B upper triangular can be factored as $C(I-J)$ with C upper triangular if B has constant row sums; then C has constant column sums, equal to the row sums of B.*

Conversely, a matrix $C(I - J)$ with C upper triangular can be factored as $(I - J)B$ with B upper triangular if C has constant column sums; then B has constant row sums, equal to the column sums of C.

Proof: Assume we have a matrix that can be factored as both $(I - J)B$ and $C(I - J)$ with B and C upper triangular. First we show that both B has constant row sums, and C constant column sums, and that the two constants are related. Define row sums $\beta_i = \sum_j b_{ij}$ and column sums $\gamma_i = \sum_j c_{ji}$. Then

$$C(I-J)e = Ce_1 = c_{11}e_1 \Rightarrow (I-J)Be = c_{11}e_1 \Rightarrow Be = c_{11}e \Rightarrow \beta_i \equiv c_{11} = \gamma_1.$$

Conversely,

$$e^t(I-J)B = e_n^t B = b_{nn}e_n^t \Rightarrow e^t C(I-J) = b_{nn}e_n^t$$
$$\Rightarrow e^t C = b_{nn}e^t \Rightarrow \gamma_i \equiv b_{nn} = \beta_n.$$

Now we show that B and C can be derived from each other. The matrix C can be derived by columns. First of all, $c_{11} = b_{11}$ follows directly from the above results. Suppose inductively that for some n the column c_{in} has been solved, then for $i \leq n$ the $n+1$-st column, excluding the diagonal element, follows from

$$c_{in} - c_{in+1} = b_{in} - b_{i-1n}$$

where we define $b_{0n} = 0$. The diagonal element follows from the constant column sums:

$$c_{n+1n+1} = \gamma_{n+1} - \sum_{i \leq n+1} c_{in+1}.$$

Conversely, we can derive B by rows. First, $b_{1j} = c_{1j} - c_{1j+1}$ can be stated immediately. Assuming that the n-th row coefficients b_{nj} have been solved, then the $n+1$-st row follows from

$$b_{n+1j} - b_{nj} = c_{n+1j} - c_{n+1j+1},$$

where we define $c_{in+1} = 0$. This concludes the statements to be proved. •

The QR decomposition of the Hessenberg matrix associated with a residual sequence takes a remarkably simple form.

Lemma 8 *Let H be an upper Hessenberg matrix of maximum rank with zero column sums, and let $H = QR$ be a decomposition into an orthonormal matrix and an upper triangular matrix. Then there is a diagonal matrix $\sigma = diag(\pm 1, \pm 1, \ldots)$ such that $Q \leftarrow Q\sigma$ is given by*

$$q_{kn} = -\frac{1}{\sqrt{n(n+1)}} \quad k \le n; \qquad q_{n+1n} = \sqrt{\frac{n}{n+1}}.$$

Furthermore, $Q = (J - I)B^{-1}$, where B is an upper bidiagonal matrix.

Proof: H has zero column sums, so Q has zero column sums. The values given satisfy this requirement plus orthonormality of the columns. Since QR decompositions are unique up to the sign of the columns of Q, there is a diagonal matrix $\sigma = diag(\pm 1, \pm 1, \ldots)$ such that $H = (Q\sigma)R$. Then, with $\alpha_n = \sqrt{n(n+1)}$:

$$\begin{aligned} Q &= \begin{pmatrix} -\alpha_1^{-1} & -\alpha_2^{-1} & & \\ \alpha_1^{-1} & -\alpha_2^{-1} & & \\ & 2\alpha_2^{-1} & \cdots & \\ & & & \ddots \end{pmatrix} \\ &= \begin{pmatrix} -1 & & & \\ 1 & -1 & & \\ & 1 & -1 & \\ & & \ddots & \ddots \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 & \cdots \\ & 2 & 2 & \cdots \\ & & 3 & \cdots \\ & & & \ddots \end{pmatrix} diag(\alpha_i^{-1}) \\ &= (J - I)\left[diag(\alpha_i)(I - J^t)diag(i^{-1})\right]^{-1} \end{aligned}$$

that is, $Q = (J - I)B^{-1}$ with

$$B = \begin{pmatrix} \sqrt{\frac{2}{1}} & -\sqrt{\frac{1}{2}} & & \\ & \sqrt{\frac{3}{2}} & -\sqrt{\frac{2}{3}} & \\ & & \ddots & \ddots \end{pmatrix}.$$

•

We regularly encounter sequences that are derived from a residual sequence by scaling, in particular normalised sequences. Let R satisfy $AR = RH$, and let $\Omega = diag(\omega_i)$ be a diagonal matrix such that $R = N\Omega$. Then $AN = N\tilde{H}$ with

$$\tilde{H} = \Omega H \Omega^{-1}. \qquad (4)$$

First we show any polynomial sequence can be scaled to a residual sequence.

Lemma 9 *If N is an arbitrary sequence satisfying $AN = N\tilde{H}$ with $\tilde{H}$ an upper Hessenberg matrix, then N can be scaled to a residual sequence R.*

Proof: First of all, the choice of ω_1 is immaterial. Let then ω_1 be given; the question is then how the other ω_i are to be computed. We have that

$$AR = R\Omega^{-1}\tilde{H}\Omega \equiv RH$$

where H has zero column sums.

For $n = 1, 2, \ldots$:

1. Let ω_k be given for $k \leq n$. This is true initially.
2. For $k < n$, $h_{kn} = \omega_k^{-1}\tilde{h}_{kn}\omega_n$ can be computed.
3. $h_{nn} = \tilde{h}_{nn}$.
4. From the zero column sum requirement, $h_{n+1n} = -\sum_{k\leq n} h_{kn}$, and from $h_{n+1n} = \omega_{n+1}^{-1}\tilde{h}_{n+1n}\omega_n$ we can compute ω_{n+1}. •

We are interested in relations between the QR decompositions of Hessenberg matrices that are equivalent through diagonal transformations as in equation (4).

Lemma 10 *Let $H_1 = Q_1U_1$ and $H_2 = Q_2U_2$ be QR decompositions of Hessenberg matrices that are related by $H_1 = \Omega H_2 \Omega^{-1}$ where Ω is a diagonal matrix. Then there is an upper triangular matrix T such that*

$$Q_2 = \Omega^{-1}Q_1T, \qquad U_2 = T^{-1}U_1\Omega.$$

If H_2 has zero column sums, the QR factorisations satisfy

$$\Omega^{-1}Q_1 = (J-I)B_1^{-1} \qquad \textit{and} \qquad Q_2 = (J-I)B_2^{-1},$$

where B_1 and B_2 are upper bi-diagonal matrices, and $T = B_1B_2^{-1}$.

Proof: We have $H_1\Omega = Q_1U_1\Omega = \Omega H_2 = \Omega Q_2U_2$, so $Q_1^t\Omega Q_2 = U_1\Omega U_2^{-1} \equiv T$, and T is clearly upper triangular. This proves the first statement of the lemma.

If H_2 has zero column sums, by lemma 8 its QR decomposition satisfies $Q_2 = (J-I)B_2^{-1}$, where B_2 is upper bi-diagonal. Since $Q_1^t\Omega Q_2$ is upper triangular, $Q_1^t\Omega(J-I)$ is also upper triangular, but it is also of lower Hessenberg form, hence it is of upper bi-diagonal form, say $B_1 \equiv Q_1^t\Omega(J-I)$.) From

$$Q_1^t\Omega(J-I) = U_1\Omega U_2^{-1}B_2$$

we find that $TB_2 = U_1\Omega U_2^{-1}B_2 = B_1$. This proves the second statement of the lemma. •

Corollary 11 *If Q is an orthonormal upper Hessenberg matrix (of size $n+1\times n$) and Ω is diagonal such that $\Omega^{-1}Q$ has zero column sums, then*

$$\Omega^{-1}Q = (J-I)B^{-1}$$

where B is of upper bidiagonal form.

Proof: Apply the previous lemma to $H_1 = Q$. •

4 Application to minimising methods and residual smoothing

The factorisation results in the previous section can be applied to minimising polynomial methods. We will show that minimising methods can be generated by two-term 'residual smoothing' recurrences.

We will first give a brief review of (quasi-) minimising methods. This restates results such as found in [2], but presented in a form more suited to the current discussion.

Minimising methods can be considered as convex combinations of an underlying residual sequence. Introducing the matrix $E = (\delta_{i1})$, we can state the following result on how convex combinations of a residual sequence can be computed.

Lemma 12 *Let R be a residual sequence, satisfying $AR = RH$, and let G derived by taking convex combinations of the R sequence. Then there is an upper triangular matrix V such that*

$$G(J-E) = -ARV \qquad \text{or} \qquad g_{i+1} = g_1 - \sum_{j\le i} Ar_j v_{ji}. \tag{5}$$

Proof: Since G consists of convex combinations of R, there is an upper triangular matrix V_1 with column sums identically 1 such that $G = RV_1$. Then $V_1(J-E)$ is an upper Hessenberg matrix with zero column sums. In lemma 6 we showed that such a matrix can be factored as $(J-I)V_2$. Noting that $H = (J-I)V_3$ for some upper triangular V_3, we now have

$$G(J-E) = RV_1(J-E) = R(J-I)V_2 = RHV_3^{-1}V_2 = RHV_4 = -ARV_5$$

where V_5 is an upper triangular matrix. $\bullet$

In minimising methods, the matrix V in equation (5) is chosen to satisfy some minimisation property. The method can be based on the sequence R, or on a normalised scaling $N = R\Omega^{-1}$ of it.

For $n \geq 1$, $g_{n+1} - g_1$ is the n-th column of ARV, that is,

$$g_{n+1} - g_1 = (ARV)_n = AR_n v_n = R_{n+1} H_n v_n. \tag{6}$$

Assume now that R is an orthogonal sequence, and note that with $e_1 = (1, 0, \ldots)^t$ we have $g_1 = Ge_1 = Re_1$. Then

$$\begin{aligned} \|g_{n+1}\|_{L^2(R^N)} &= \|Re_1 - (ARV)_n\|_{L^2(R^N)} \\ \text{(use 6):} &= \|R_{n+1}e_1 - R_{n+1}H_n v_n\|_{L^2(R^N)} \\ \text{(use } R = N\Omega \text{ and } N \text{ orthonormal:)} &= \|\Omega_{n+1}e_1 - \Omega_{n+1}H_n v_n\|_{L^2(R^{n+1})} \\ \text{(use 4:)} &= \left\| \|r_1\| e_1 - \tilde{H}_n \Omega_n v_n \right\|_{L^2(R^{n+1})} \end{aligned}$$

We see that $\|g_{n+1}\|$ can be minimised by minimising v_n. If R is not orthogonal, and we simply minimise v_n, this is referred to as 'quasi-minimisation'.

Now let $\tilde{H}_n = Q_n U_n$ be a decomposition into an orthonormal matrix and an upper triangular matrix. Define

$$\bar{Q}_n = \begin{bmatrix} Q_n & \begin{matrix} * \\ \vdots \\ * \end{matrix} \end{bmatrix}, \qquad \bar{U}_n = \begin{bmatrix} & U_n & \\ 0 & \cdots & 0 \end{bmatrix}$$

where the final column of $\bar{Q}_n$ is chosen to make $\bar{Q}_n$ again an orthonormal matrix, then $\tilde{H}_n = \bar{Q}_n \bar{U}_n$. We now find

$$\begin{aligned} \|g_{n+1}\|_{L^2(R^N)} &= \left\| \|r_1\| e_1 - \tilde{H}_n \Omega_n v_n \right\|_{L^2(R^{n+1})} \\ &= \left\| \|r_1\| q^{(n+1)} - \bar{U}_n \Omega_n v_n \right\|_{L^2(R^{n+1})} \end{aligned}$$

where $q^{(n+1)^t} = (q_{11}, \ldots, q_{1n+1})$.

Obviously, $\|g_{n+1}\|$ is minimised if

$$U_n \Omega_n v_n = q^{(n)}, \tag{7}$$

and its value is $|q_{1n+1}|$. This fact was derived earlier, but differently, in [2].

Simple manipulation now gives the formula for updating G:

$$G(J - I) = R\Omega^{-1} Q D_Q \omega_1$$

where $D_Q = diag(q_{11}, q_{12}, \ldots)$.

Now for our main result.

Theorem 13 *If R is a residual sequence, and G a sequence of convex combinations derived by (quasi-) minimisation as in equation 5, then R and G are related by a two-term recurrence of the form*

$$\alpha_i g_{i+1} + (1 - \alpha_i) g_i = r_i.$$

Proof: Since H has zero column sums, so has $\Omega^{-1}Q$. Thus, by lemma 8 there is an upper bidiagonal matrix B such that

$$\Omega^{-1}Q = (J - I)B^{-1} \Rightarrow G(J - I) = R(J - I)B^{-1}D_Q\omega_1.$$

Now observe that

$$D_Q^{-1}Be = D_Q^{-1}Q^t\Omega(J - I)e = \begin{pmatrix} 1 & * & & \\ 1 & * & * & \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} \omega_1 \\ 0 \\ \vdots \end{pmatrix} = \omega_1 e.$$

Since $D_Q^{-1}B$ has constant column sums, in fact, identically ω_1, by lemma 7 there is an upper bidiagonal matrix $\tilde{B}$ with column sums identically 1 such that

$$G(J - I)D_Q^{-1}B\omega_1^{-1} = G\tilde{B}(J - I).$$

Factoring out the $J - I$ matrix, we find $G\tilde{B} = R$, that is, the G sequence can be derived from the R sequence by a two-term recurrence of the form

$$\alpha_i g_{i+1} + (1 - \alpha_i) g_i = r_i.$$

This is the basis for 'residual smoothing' methods [3]. •

5 Conclusion

We have shown how there is a close relationship between polynomial iterative methods and Hessenberg matrices with zero column sums. We have derived various facts about such matrices, and applied these to iterative methods, in particular showing that (quasi-) minimising methods can always be generated by two-term 'residual smoothing' methods.

References

[1] Alston S. Householder. *The theory of matrices in numerical analysis.* Blaisdell Publishing Company, New York, 1964. republished by Dover Publications, New York, 1975.

[2] Yousef Saad and Martin H. Schultz. GMRes: a generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM J. Sci. Stat. Comput.*, 7:856–869, 1986.

[3] Lu Zhou and Homer F. Walker. Residual smoothing techniques for iterative methods. *SIAM J. Sci. Stat. Comput.*, 15:297–312, 1992.