Hessenberg Eigenvalue-Eigenmatrix Relations

Jens-Peter M. Zemke
Arbeitsbereich Mathematik
Technische Universität Hamburg-Harburg

25. Norddeutsches Kolloquium über
Angewandte Analysis und Numerische Mathematik
Universität Hamburg
14.–15. **Mai 2004**

Outline

Normal forms

Historical and theoretical approach:

Reduction to **Jordan normal form**

$$\widehat{V}^H A V = J, \qquad \widehat{V}^H V = I$$

$\Rightarrow$ eigenvalues + *eigenvectors* and *principal vectors*

Reduction to **Schur normal form**

$$U^H A U = R, \qquad U^H U = I$$

$\Rightarrow$ eigenvalues + *Schur vectors*

Jordan and Schur normal forms: explicit access to eigenvalues.

Drawbacks:

- frequently takes infinite number of steps (Abel)
- frequently numerically unstable (Wilkinson et al.)

Solution: seek “nearby” normal forms that behave stably, i.e.:

- computable in finite number of steps
- numerically stable
- not destroyed in GR iteration, e.g. QR iteration
- better tractable (e.g. divide and conquer)

Modern and numerically feasible approach:

Instead of Jordan normal form: reduction to **tridiagonal form**

$$\hat{Q}^H A Q = T, \qquad \hat{Q}^H Q = I$$

$\Rightarrow$ divide and conquer, RRR

Instead of Schur normal form: reduction to **Hessenberg form**

$$Q^H A Q = H, \qquad Q^H Q = I$$

$\Rightarrow$ QR iteration, "bulge chasing"

Summarized: Jordan $\rightarrow$ tridiagonal, Schur $\rightarrow$ Hessenberg.

Two conceptually distinct ways of computation:

50s: Lanczos (iterative computation of tridiagonal normal form)
Arnoldi (iterative computation of Hessenberg normal form)

60s: Householder & Givens (tridiagonal or Hessenberg matrix)

Afterwards invoke QR iteration (Francis & Kublanovskaya)

Methods are stable or cheap. First analysis of finite precision behaviour of symmetric Lanczos by Paige in his 1971 thesis.

Hessenberg structure

Hessenberg matrix:

$$H \in \mathbb{C}^{n \times n}, \qquad h_{ij} = 0 \quad \forall\ i+1 > j.$$

Unreduced := iff $h_{j+1,j} \neq 0$ for all $j = 1, \ldots, n-1$.

We consider only unreduced Hessenberg matrices, since

- natural occurence because of interest in eigenvalues
- all eigenvalues are nonderogatory

(first: obvious, latter: simple rank argument)

Question to be investigated is that of inheritance:

Matrix structure $\Rightarrow$ eigenmatrix structure?

Jordan: zeros in lower and upper part of Jordan matrix $\Rightarrow$ zeros in lower and upper part of eigenmatrix ($= I$).

Schur: zeros in lower part of Schur matrix $\Rightarrow$ zeros in lower part of eigenmatrix ($=$ upper triangular).

Replace zeros in matrix by "small" elements $\Rightarrow$ "small" elements in corresponding parts of eigenmatrix (modulo Jordan structure).

What about **Hessenberg and tridiagonal** matrices?

——— caution: real math starts here ———

Notation:

H	$H \in \mathbb{C}^{n\times n}$ unreduced Hessenberg
${}^{z}H$	family ${}^{z}H \equiv zI - H$
$P(z)$	adjugate $P(z) \equiv \mathrm{adj}({}^{z}H) = \mathrm{adj}(zI - H)$
$\chi(z)$	characteristic polynomial $\chi(z) \equiv \det({}^{z}H) = \det(zI - H)$

Results are obtained from Taylor expansion of adjugate of ${}^{z}H$.

Additional notation for specific eigenvalue necessary:

λ	eigenvalue of H with algebraic multiplicity α
J_λ	Jordan block to eigenvalue λ
V_λ	right invariant subspace corresponding to λ
$\widehat{V}_\lambda$	left invariant subspace corresponding to λ
$\omega(z)$	deflated polynomial $\chi(z) = \omega(z)(z - \lambda)^\alpha$

Relations between invariant subspaces, matrix and Jordan block:

$$HV_\lambda = V_\lambda J_\lambda, \quad \widehat{V}_\lambda^H H = J_\lambda \widehat{V}_\lambda^H, \quad \widehat{V}_\lambda^H V_\lambda = I_\alpha.$$

Natural restrictions (for all $\ell < \alpha$):

$$HV_\lambda^{[\ell]} = V_\lambda^{[\ell]} J_\lambda^{[\ell]}, \quad \widehat{V}_\lambda^{[\ell]H} H = J_\lambda^{[\ell]} \widehat{V}_\lambda^{[\ell]H}.$$

Here, ℓ denotes the number of principal vectors involved:

$J_\lambda^{[\ell]}$ Jordan block to eigenvalue λ of dimension $\ell + 1$

$V_\lambda^{[\ell]}$ *first* $\ell + 1$ columns of V_λ

$\widehat{V}_\lambda^{[\ell]}$ *last* $\ell + 1$ columns of $\widehat{V}_\lambda$

Well-known example of a natural restriction ($\ell = 0$):

$$Hv = v\lambda \quad \text{and} \quad \widehat{v}^H H = \lambda \widehat{v}^H.$$

Corollary: Let λ be nonderogatory. Then for all $\ell < \alpha$

$$\frac{P^{(\ell)}(\lambda)}{\ell\,!} = V_\lambda^{[\ell]} \omega\left(J_\lambda^{[\ell]}\right)\widehat{V}_\lambda^{[\ell]H}. \tag{1}$$

Enumerate right eigenbasis and flipped left eigenbasis:

$$V_\lambda \equiv \left(v_1, \ldots, v_\alpha\right) \qquad \text{and} \qquad \widehat{V}_\lambda F \equiv W_\lambda \equiv \left(w_1, \ldots, w_\alpha\right). \tag{2}$$

Then, in other words, for all $\ell < \alpha$

$$\frac{P^{(\ell)}(\lambda)}{\ell\,!} = \frac{\omega^{(\ell)}(\lambda)}{\ell\,!} v_1 w_1^H + \frac{\omega^{(\ell-1)}(\lambda)}{(\ell-1)!}\left(v_1 w_2^H + v_2 w_1^H\right) + \cdots$$
$$+\,\omega(\lambda)\left(\sum_{k=1}^{\ell+1} v_k w_{(\ell+1)-k+1}^H\right). \tag{3}$$

Example ($\alpha > 0$, $\ell = 0$):

$$P(\lambda) = v_1 \omega(\lambda) \widehat{v}_\alpha^H = v_1 \omega(\lambda) w_1^H$$

Example ($\alpha > 1$, $\ell = 1$):

$$\begin{aligned} P'(\lambda) &= \begin{pmatrix} v_1 & v_2 \end{pmatrix} \begin{pmatrix} \omega(\lambda) & \omega'(\lambda) \\ & \omega(\lambda) \end{pmatrix} \begin{pmatrix} \widehat{v}_{\alpha-1}^H \\ \widehat{v}_\alpha^H \end{pmatrix} \\ &= \begin{pmatrix} v_1 & v_2 \end{pmatrix} \begin{pmatrix} \omega(\lambda) & \omega'(\lambda) \\ & \omega(\lambda) \end{pmatrix} \begin{pmatrix} w_2^H \\ w_1^H \end{pmatrix} \end{aligned}$$

$P(z)$ linked to eigeninformation. How to use this knowledge?
$\Rightarrow$ Express adjugate in another way than in terms of cofactors.

The trick: express adjugate of ${}^zH$ in terms of **determinants of principal submatrices** of ${}^zH$, short principal subdeterminants.

In two stages: **lower** and **upper** part. Notation for lower part:

Polynomial vectors $\nu(z)$ and $\widehat{\nu}(z)$:

$$\nu(z) \equiv \left(\frac{\chi_{i+1:n}(z)}{\prod_{l=i+1}^{n} h_{l,l-1}}\right)_{i=1}^{n} \quad \text{and} \quad \widehat{\nu}(z) \equiv \left(\frac{\chi_{1:j-1}(z)}{\prod_{l=1}^{j-1} h_{l+1,l}}\right)_{j=1}^{n}. \tag{4}$$

Empty product is one, $\chi_{i:j}(z)$ defined by

$$\chi_{i:j}(z) \equiv \begin{cases} \det\left({}^{z}H_{i:j}\right), & 1 \leqslant i \leqslant j \leqslant n, \\ 1, & 1 + i = j. \end{cases} \tag{5}$$

${}^{z}H_{i:j}$ principal submatrix of ${}^{z}H$,

$$\det\left({}^{z}H_{i:j}\right) \equiv \det\left(zI - H_{i:j}\right).$$

Theorem: *Let* $H \in \mathbb{C}^{n\times n}$ *be unreduced upper Hessenberg. Let* $P(z)$ *denote the adjugate of* ${}^zH \equiv zI - H$*. Let* $\nu(z)$*,* $\hat{\nu}(z)$ *be as denoted above. Furthermore, let* tril(A) *denote the restriction of* A *to its* triangular lower part*.*

Then

$$\operatorname{tril}(P(z)) = \left(\prod_{l=1}^{n-1} h_{l+1,l}\right) \cdot \operatorname{tril}\left(\nu(z)\hat{\nu}(z)^T\right). \tag{6}$$

Especially, we have validity of relations

$$({}^zH)\nu(z) = \frac{\chi(z)}{\prod_{l=1}^{n-1} h_{l+1,l}} e_1, \qquad \hat{\nu}(z)^T({}^zH) = \frac{\chi(z)}{\prod_{l=1}^{n-1} h_{l+1,l}} e_n^T. \tag{7}$$

Gain: Expression for eigenvectors and principal vectors.

Missing: Complete representation of adjugate $P(z)$.

Problem: The polynomials $p_{ij}(z)$ have (maximal) degree

$$\begin{aligned} \deg\left(p_{ij}(z)\right) &= n-1+j-i, \qquad && i \geqslant j, \\ \deg\left(p_{ij}(z)\right) &\leqslant n-2, && i < j. \end{aligned} \tag{8}$$

Pictorially (when all $h_{ij} \neq 0$, $i < j$):

$$\begin{pmatrix} n-1 & n-2 & \cdots & n-2 \\ n-2 & n-1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & n-2 \\ 0 & \cdots & n-2 & n-1 \end{pmatrix} \neq \begin{pmatrix} n-1 & n & \cdots & 2(n-1) \\ n-2 & n-1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & n \\ 0 & \cdots & n-2 & n-1 \end{pmatrix}$$

Solution for **upper** part: middle characteristic polynomials

Additional notation: regular **upper triangular** matrices

$H^{\triangle}(z)$ ${}^{z}H$ without first row and last column

$M^{\triangle}(z)$ inverse of $H^{\triangle}(z)$

Elements of inverse $M^{\triangle}(z)$ of regular $H^{\triangle}(z)$:

$$m_{ij}^{\triangle}(z) = \begin{cases} -\dfrac{\chi_{i+1:j}(z)}{\prod_{l=i}^{j} h_{l+1,l}}, & i \leqslant j, \\ 0, & i > j. \end{cases} \tag{9}$$

Theorem: *$H \in \mathbb{C}^{n \times n}$ unreduced upper Hessenberg. Polynomial vectors $\nu(z)$, $\hat{\nu}(z)$ as above. Strictly upper triangular $M(z)$:*

$$M(z) \equiv \begin{pmatrix} 0 & M^{\triangle}(z) & \\ \vdots & & \\ 0 & \cdots & 0 \end{pmatrix}. \tag{10}$$

Then

$$P(z) = \left(\prod_{l=1}^{n-1} h_{l+1,l} \right) \cdot \nu(z)\hat{\nu}(z)^T + \chi(z)M(z), \qquad \text{i.e.,} \tag{11}$$

$$p_{ij}(z) = \begin{cases} \chi_{1:j-1}(z) \left(\prod_{l=j}^{i-1} h_{l+1,l} \right) \chi_{i+1:n}(z) & i \geqslant j, \\[2ex] \dfrac{\chi_{1:j-1}(z)\chi_{i+1:n}(z) - \chi_{i+1:j-1}(z)\chi_{1:n}(z)}{\prod_{l=i}^{j-1} h_{l+1,l}} & i < j. \end{cases} \tag{12}$$

In short:

$$P(z) = \left(\prod_{l=1}^{n-1} h_{l+1,l} \right) \cdot \nu(z) \widehat{\nu}(z)^T \pmod{\chi(z)}. \tag{13}$$

Gluing parts together: for all $\ell < \alpha$

$$\left(\prod_{l=1}^{n-1} h_{l+1,l} \right) \cdot \left[\frac{\left(\nu \widehat{\nu}^T\right)^{(\ell)}}{\ell\,!} \right] (\lambda) = V_\lambda^{[\ell]} \omega \left(J_\lambda^{[\ell]} \right) \widehat{V}_\lambda^{[\ell]H}.$$

Can be used to define "natural" eigenmatrix.

General usage: switch between *analytic information* (principal subdeterminants) and *algebraic information* (eigeninformation).

Applications

- structured perturbation theory
  - Jacobi matrices [Hochstadt '74, Hald '76, Hochstadt '79, Elhay/Gladwell/Golub/Ram '99]
  - inverse eigenvalue problems [Moody T. Chu '98]
- stable methods for ODE
  - general linear methods with inherent Runge-Kutta stability [Butcher/Chartier '99, Butcher/Wright '03]
- error analysis of Krylov methods [Z. '03]
- QR (aggressive early deflation) [Braman/Byers/Mathias '02]