An Explicit Jordan Decomposition of Companion Matrices

Fermín S. V. Bazán
e-mail: fermin@mtm.ufsc.br
http://www.mtm.ufsc.br/˜fermin
CFM-UFSC

Universidade Federal de Santa Catarina

September – 2005

Content

- Motivation
- Some theoretical results.
- Numerical Results.
- Conclusions

Motivation: Two problems involving Companion matrices

- Initial value problems:

$$\begin{cases} y^{(m)}(t) + a_{m-1}y^{(m-1)}(t) + \cdots + a_1 y'(t) + a_0 y(t) = 0, \ t \geq a \\ y(a) = \alpha_0, \ y'(a) = \alpha_1, \cdots, y^{(m-1)}(a) = \alpha_m. \end{cases}$$

have solutions of the form: $y(t) = e^{Ct}\alpha, \quad \alpha = [\alpha_0, \ldots, \alpha_{m-1}]^T$ with

$$C = \begin{bmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \cdots & 0 & -a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_{m-1} \end{bmatrix} \tag{1}$$

- Computation of roots of polynomials

$$\pi(t) = t^m + a_{m-1}t^{m-1} + \cdots + a_1 t + a_0,$$

can be done by extracting the eigenvalues of C and viceversa

△ The solution of the above problems depends on the **Jordan form** of C !

Motivation: Two problems involving Companion matrices

- Initial value problems:

$$\begin{cases} y^{(m)}(t) + a_{m-1} y^{(m-1)}(t) + \cdots + a_1 y'(t) + a_0 y(t) = 0, \ t \geq a \\ y(a) = \alpha_0, \ y'(a) = \alpha_1, \cdots, y^{(m-1)}(a) = \alpha_m. \end{cases}$$

have solutions of the form: $y(t) = e^{Ct}\alpha, \quad \alpha = [\alpha_0, \ldots, \alpha_{m-1}]^T$ with

$$C = \begin{bmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \cdots & 0 & -a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_{m-1} \end{bmatrix} \tag{1}$$

- Computation of roots of polynomials

$$\pi(t) = t^m + a_{m-1} t^{m-1} + \cdots + a_1 t + a_0,$$

can be done by extracting the eigenvalues of C and viceversa

△ The solution of the above problems depends on the **Jordan form** of C !

Motivation: Two problems involving Companion matrices

- Initial value problems:

$$\begin{cases} y^{(m)}(t) + a_{m-1}y^{(m-1)}(t) + \cdots + a_1 y'(t) + a_0 y(t) = 0, \ t \geq a \\ y(a) = \alpha_0, \ y'(a) = \alpha_1, \cdots, y^{(m-1)}(a) = \alpha_m. \end{cases}$$

have solutions of the form: $y(t) = e^{Ct}\alpha, \quad \alpha = [\alpha_0, \ldots, \alpha_{m-1}]^T$ with

$$C = \begin{bmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \cdots & 0 & -a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_{m-1} \end{bmatrix} \tag{1}$$

- Computation of roots of polynomials

$$\pi(t) = t^m + a_{m-1}t^{m-1} + \cdots + a_1 t + a_0,$$

can be done by extracting the eigenvalues of C and viceversa

△ The solution of the above problems depends on the **Jordan form** of C !

Jordan Form of C

If $\lambda_1, \dots, \lambda_p$ denote the p distinct eigenvalues of C and $m_1, \dots, m_p$ the respective algebraic multiplicities, i.e.:

$$\pi(t) = (t-\lambda_1)^{m_1}(t-\lambda_2)^{m_2}\cdots(t-\lambda_p)^{m_p} \ \text{with} \ m_1 + \cdots + m_p = m,$$

a Jordan form of C can be given as

$$\begin{bmatrix} J_{\lambda_1} & & \\ & \ddots & \\ & & J_{\lambda_p} \end{bmatrix} = \begin{bmatrix} L_1 \\ \vdots \\ L_p \end{bmatrix} C \begin{bmatrix} R_1 & \dots & R_p \end{bmatrix} \equiv L\,C\,R,$$

$$\text{where for} \quad i = 1, \dots p,\ J_{\lambda_i} = \begin{pmatrix} \lambda_i & 1 & & & \\ & . & . & & \\ & & . & . & \\ & & & . & 1 \\ & & & & \lambda_i \end{pmatrix} \in \mathbb{C}^{m_i \times m_i}, \ \text{and}$$

$$LR = RL = I \quad (m \times m \ \text{identitymatrix}).$$

Jordan Form of C

- Columns of $R_l = [r_1, r_2, \dots, r_{m_l}]$ are the so-called generalized right eigenvectors of C associated with λ_l. They satisfy

$$\begin{cases} C\,r_1 = \lambda_l r_1 \\ \\ C\,r_i = \lambda_l r_i + r_{i-1},\ i = 2, \dots m_l. \end{cases}$$

  $\{r_1, r_2, \dots, r_{m_l}\}$: Right Jordan Chain of C
  r_1 : Leading rigth generalized eigenvector

- Rows of $L_l = \begin{bmatrix} l_1^* \\ \vdots \\ l_{m_l}^* \end{bmatrix}$ are the so-called generalized left eigenvectors of C

  $\{l_1, l_2, \dots, l_{m_l}\}$: Left Jordan Chain of C
  l_{m_l} : Leading left generalized eigenvector

Jordan Form: Main Results

Proposition 1. Define $\Phi(\lambda) = [1, \lambda, \ldots, \lambda_{m-1}]^T$,

$$H = \begin{bmatrix} a_1 & a_2 & \cdots & a_{m-1} & 1 \\ a_2 & \cdots & a_{m-1} & 1 & \\ \vdots & \vdots & 1 & & \\ a_{m-1} & 1 & & & \\ 1 & & & & \end{bmatrix}, \; r_i = H\,\frac{\phi^{(i-1)}(\lambda_l)}{(i-1)!}, \text{ and } \; \breve{l}_i = \frac{\bar{\phi}^{(m_l-i)}(\lambda_l)}{(m_l-i)!}.$$

The set $\{r_1, r_2, \cdots, r_{m_l}\}$ is a right Jordan chain of C associated with the eigenvalue λ_l and r_1 is the leading right eigenvector. The set $\{\breve{l}_1, \breve{l}_2, \cdots, \breve{l}_{m_l}\}$ is a left Jordan chain of C associated with the eigenvalue λ_l and $\breve{l}_{m_l}$ is the leading left eigenvector. The left and right Jordan chains satisfy

$$\breve{L}_l R_l \equiv \begin{bmatrix} \breve{l}_1^* \\ \vdots \\ \breve{l}_{m_l}^* \end{bmatrix} \begin{bmatrix} r_1 & \ldots & r_{m_l} \end{bmatrix} = \begin{pmatrix} \alpha_1 & \alpha_2 & . & \alpha_{m_{l-1}} & \alpha_{m_l} \\ & . & . & . & \alpha_{m_{l-1}} \\ & & . & . & . \\ & & & . & \alpha_2 \\ & & & & \alpha_1 \end{pmatrix} \equiv F_l \qquad (2)$$

where $\alpha_i = \frac{\pi^{(m_l+i-1)}(\lambda_l)}{(m_l+i-1)!}$.

Jordan Form: Main results

Proposition 2. Define $L_l^* = [l_1, l_2, \cdots, l_{m_l}] = \breve{L}_l^* F_l^{-*}$. The set $\{l_1, l_2, \cdots, l_{m_l}\}$ is a left Jordan chain of C associated with the eigenvalue λ_l, l_{m_l} being the leading left eigenvector. The left and right Jordan chains are normalized so that

$$L_l R_l \equiv \begin{bmatrix} l_1^* \\ \vdots \\ l_{m_l}^* \end{bmatrix} \begin{bmatrix} r_1 & \dots & r_{m_l} \end{bmatrix} = I \in \mathbb{R}^{m_l \times m_l}. \tag{3}$$

Similarly, we define $\breve{R}_l = [\breve{r}_1, \breve{r}_2, \dots, \breve{r}_{m_l}] = [r_1, \dots r_{m_l}] F_l^{-1}$. The set $\{\breve{r}_1, \breve{r}_2, \dots, \breve{r}_{m_l}\}$ is a right Jordan chain of C associated with the eigenvalue λ_l, and

$$\breve{L}_l \breve{R}_l = I \in \mathbb{R}^{m_l \times m_l}.$$

Confluent Vandermonde Matrices: An inversion formula

If $\breve{L}^*$ is a Confluent Vandermonde matrix: $\breve{L}^* = [\breve{L}_1^*, \breve{L}_2^*, \ldots, \breve{L}_p^*]$

$$\breve{L}_l^* = \begin{bmatrix} 0 & \cdots & 0 & 0 & 1 \\ 0 & \cdots & 0 & 1 & \lambda_l \\ 0 & & 2 & 2\lambda_l & \lambda_l^2 \\ 0 & & 6\lambda_l & 3\lambda_l^2 & \lambda_l^3 \\ \begin{matrix} 0 \\ (m_l-1)! \end{matrix} & \cdots & \vdots & \vdots & 0 \\ & & & & \\ \vdots & & & & \vdots \\ (m-m_l+1)\lambda^{m-m_l} & \cdots & (m-1)(m-2)\lambda_l^{m-3} & (m-1)\lambda_l^{m-2} & \lambda_l^{m-1} \end{bmatrix},$$

a consequence of Proposition 2 is

$$\breve{L}^{-1} = [R_1 \ \cdots R_p]F^{-1} \quad \text{with} \quad F = \operatorname{diag}(F_1, \ldots, F_p). \tag{4}$$

Sensitivity of Roots: Condition Number

Objective: Estimate the sensitivity of roots of

$$\pi(t) = t^m + a_{m-1}t^{m-1} + \cdots + a_1 t + a_0$$

to small perturbations in the coefficients a_j.
λ : root of $\pi(t)$ of multiplicity d.
$\widetilde{\pi}(t)$: monic polynomial with coeff. $\widetilde{a}_j = a_j + \Delta a_j$.
$\widetilde{\lambda}_k,\ k = 1, \ldots, d$: roots of $\widetilde{\pi}(t)$ that approximate λ

$$|\Delta\lambda| = \max_{1 \le k \le d} |\lambda - \widetilde{\lambda}_k| = ?$$

Definition. [Chatelin 1996] Assuma the Δa_j's satisfy the componentwise inequalities

$$|\Delta a_j| \le \varepsilon \alpha_j,\ j = 1, \ldots, m-1, \tag{5}$$

where α_j are arbitrary non negative real numbers. The *componentwise relative* condition number of the root λ of multiplicity d is defined by

$$\kappa^{\mathcal{C}}(\lambda) = \lim_{\varepsilon \to 0} \sup_{|\Delta a_j| \le \varepsilon \alpha_j} \frac{|\Delta\lambda|}{|\lambda| \varepsilon^{1/d}}. \tag{6}$$

Sensitivity of Roots: Condition Number

Objective: Estimate the sensitivity of roots of

$$\pi(t) = t^m + a_{m-1}t^{m-1} + \cdots + a_1 t + a_0$$

to small perturbations in the coefficients a_j.
λ : root of $\pi(t)$ of multiplicity d.
$\widetilde{\pi}(t)$: monic polynomial with coeff. $\widetilde{a}_j = a_j + \Delta a_j$.
$\widetilde{\lambda}_k,\ k = 1, \ldots, d$: roots of $\widetilde{\pi}(t)$ that approximate λ

$$|\Delta\lambda| = \max_{1 \le k \le d} |\lambda - \widetilde{\lambda}_k| = ?$$

Definition. [Chatelin 1996] Assuma the Δa_j's satisfy the componentwise inequalities

$$|\Delta a_j| \le \varepsilon\alpha_j,\ j = 1, \ldots, m-1, \tag{5}$$

where α_j are arbitrary non negative real numbers. The *componentwise relative* condition number of the root λ of multiplicity d is defined by

$$\kappa^{\mathcal{C}}(\lambda) = \lim_{\varepsilon \to 0} \sup_{|\Delta a_j| \le \varepsilon\alpha_j} \frac{|\Delta\lambda|}{|\lambda|\varepsilon^{1/d}}. \tag{6}$$

Sensitivity of Roots: Condition Estimation

Definition. Assuma the Δa_j's satisfy the normwise inequality

$$||[\Delta a_0, \ldots \Delta a_{m-1}]|| \leq \varepsilon\alpha,\ \alpha > 0. \tag{7}$$

The *normwise relative* condition number of the root λ of multiplicity d is defined by

$$\kappa(\lambda) = \lim_{\varepsilon \to 0} \sup_{||[\Delta a]|| \leq \varepsilon\alpha} \frac{|\Delta\lambda|}{|\lambda|\varepsilon^{1/d}}. \tag{8}$$

- **Results**

$$\kappa^{C}(\lambda) = \frac{1}{|\lambda|}\left(\frac{d! \sum_{j=0}^{m-1} |\lambda^j|\alpha_j}{|\pi^{(d)}(\lambda)|}\right)^{1/d} \tag{9}$$

$$\kappa(\lambda) = \frac{1}{|\lambda|}\left(\frac{d!\ \|\ \phi(\lambda)\ \|\ \alpha}{|\pi^{(d)}(\lambda)|}\right)^{1/d} \tag{10}$$

Definition. Assuma the Δa_j's satisfy the normwise inequality

$$||[\Delta a_0, \ldots \Delta a_{m-1}]|| \leq \varepsilon\alpha,\ \alpha > 0. \tag{7}$$

The *normwise relative* condition number of the root λ of multiplicity d is defined by

$$\kappa(\lambda) = \lim_{\varepsilon \to 0} \sup_{||[\Delta a]|| \leq \varepsilon\alpha} \frac{|\Delta\lambda|}{|\lambda|\varepsilon^{1/d}}. \tag{8}$$

- **Results**

$$\kappa^{\mathcal{C}}(\lambda) = \frac{1}{|\lambda|} \left(\frac{d! \sum_{j=0}^{m-1} |\lambda^j| \alpha_j}{|\pi^{(d)}(\lambda)|} \right)^{1/d} \tag{9}$$

$$\kappa(\lambda) = \frac{1}{|\lambda|} \left(\frac{d! \parallel \phi(\lambda) \parallel \alpha}{|\pi^{(d)}(\lambda)|} \right)^{1/d} \tag{10}$$

Condition Estimation: Results

When the perturbations are measured in an absolute sense, absolute conditions numbers are obtained: if $\|\Delta a\|_2 \leq \varepsilon$ (i.e, taking $\alpha = 1$), one obtains the absolute condition number

$$\kappa_a(\lambda) = \left(\frac{d!\|\phi(\lambda)\|_2}{|\pi^{(d)}(\lambda)|} \right)^{1/d} \tag{11}$$

Consequences: For ε small enough we obtain

- Relative error estimate (componentwise) :

$$\frac{|\Delta\lambda|}{|\lambda|} \approx \varepsilon^{1/d} \kappa^C(\lambda)$$

- Relative error estimate (normwise):

$$\frac{|\Delta\lambda|}{|\lambda|} \approx \varepsilon^{1/d} \kappa(\lambda)$$

- Absolute error estimate (normwise):

$$|\Delta\lambda| \approx \varepsilon^{1/d} \kappa_a(\lambda)$$

Condition Estimation: Results

When the perturbations are measured in an absolute sense, absolute conditions numbers are obtained: if $\|\Delta a\|_2 \leq \varepsilon$ (i.e, taking $\alpha = 1$), one obtains the absolute condition number

$$\kappa_a(\lambda) = \left(\frac{d!\|\phi(\lambda)\|_2}{|\pi^{(d)}(\lambda)|} \right)^{1/d} \tag{11}$$

Consequences: For ε small enough we obtain

- Relative error estimate (componentwise) :

$$\frac{|\Delta\lambda|}{|\lambda|} \approx \varepsilon^{1/d} \kappa^{C}(\lambda)$$

- Relative error estimate (normwise):

$$\frac{|\Delta\lambda|}{|\lambda|} \approx \varepsilon^{1/d} \kappa(\lambda)$$

- Absolute error estimate (normwise):

$$|\Delta\lambda| \approx \varepsilon^{1/d} \kappa_a(\lambda)$$

Most Important Result

Theorem. Consider $\pi(t)$ and $q(t)$ so that

$$\pi(t) = (t-\lambda_1)^{m_1}\cdots(t-\lambda_p)^{m_p}, \text{ and } \quad q(t) = \pi(t)/(t-\lambda_l)^{m_l-1}.$$

That is, $q(t)$ is a deflated polynomial of degree $m - m_l - 1$ the roots of which are: λ_l as simple root, and the remaining roots of $\pi(t)$ diffferent from λ_l. Let $\kappa(\lambda_l)$ and $\check{\kappa}(\lambda_l)$ be the condition numbers of λ_l viewed as multiple root of $\pi(t)$ and simple root of $q(t)$, respectively. Then

$$\kappa(\lambda_l) = \frac{1}{|\lambda_l|}\check{\kappa}(\lambda_l)^{1/m_l}\|a\|_2^{1/m_l}\rho^{1/m_l}, \tag{12}$$

where $\rho = \dfrac{\|\Phi(\lambda_l)\|_2}{\|\Psi(\lambda_l)\|_2} = \dfrac{\|[1,\lambda_l,\ldots,\lambda_l^{m-1}\|_2}{\|[1,\lambda_l,\ldots,\lambda_l^{m-m_l-2}\|_2.}$

Conclusion:
If λ_l is a well-conditioned root of $q(t)$ and ρ is not large, then λ_l may be a relatively well-conditioned multiple root of $\pi(t)$ provided that the multiplicity is not very large

Numerical Illustration

We consider the polynomial ($m = 20$)

$$\pi(t) = (t - \lambda(s))^5 (1 + t + \cdots + t^{15})$$

with $\lambda(s) = (1 + 9s) + si$, $0 \leq s \leq 2$. We illustrate the sensitivity of the multiple root $\lambda(s)$ as a function of s. The deflated polynomial is

$$q(t) = (t - \lambda(s))(1 + t + \cdots + t^{15}).$$

Notice that $s \approx 0 \Rightarrow \lambda(s) \approx 1$ (a very well conditioned root, Gautschi 1984.)

$\lambda(s)$	$\kappa(\lambda)$	$\kappa_a(\lambda)$	ρ	$\lvert\Delta\lambda\rvert/\lvert\lambda\rvert$
$19 + 2\imath$	$1.3169e+1$	$1.0480e+1$	$1.3322e+5$	$1.3169e-1$
$15 + 1.5\imath$	$1.0724e+1$	$8.6469e+0$	$5.1642e+4$	$1.0724e-1$
$10 + \imath$	$7.4747e+0$	$6.2102e+0$	$1.0201e+4$	$7.4747e-2$
$5 + 0.5\imath$	$3.9220e+0$	$3.4955e+0$	$6.3756e+2$	$3.9220e-2$
$1.45 + 0.05\imath$	$1.5800e+0$	$1.1384e+0$	$4.4310e+0$	$1.5800e-2$
1	$1.2693e+0$	$7.7495e-1$	$1.1180e+0$	$1.2693e-2$

Table: Condition numbers, ratio ρ, and theoretical predicted error in λ corresponding to a normwise relative input error in a_j such that $\|\Delta a\|/\|a\| = \varepsilon = 10^{-10}$. In this case, the predicted error is : $|\Delta\lambda|/|\lambda| \approx \varepsilon^{1/5}\kappa(\lambda)$

Numerical Illustration

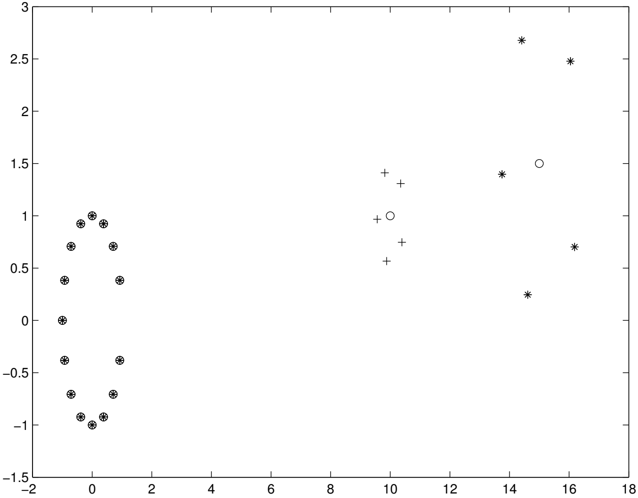

Figure: a) $\lambda = 15 + 1.5\imath$. $\circ$: Exact eigenvalue, $*$: Approximate eigenvalue.
b) $\lambda = 10 + \imath$. $\circ$: Exact eigenvalue, $+$: Approximate eigenvalue.

Conclusion

- The results are of theoretical interest: they serve to understand the sensitivity problem of multiple roots. Application of the results to problems from system theory (identification, modification) are the subject of ongoing work.
- The results can be extended to analyse the sensitivity problem of multiple eigenvalues of block companion matrices (roots of matrix polynomials) .

Some References

[1] F. Chaitin-Chatelin and V. Frayseé, *Lectures on Finite Precision computations*. SIAM, Philadelphia 1996.

[2] W. Gautschi, *Questions of numerical condition related to polynomials*, in MAAA Studies in Mathematics, Vol. 24, Studies in Numerical Analysis, G. H. Golub, ed., USA, 1984, The Mathematical Association of America, pp. 140-177.

[3] F. S. V. Bazán, *Matrix polynomials with partially prescribed eigenstrucutre: Eigenvalue sensitivity and condition estimation.* To appear in J. Comput. Appl. Math.

Conclusion

- The results are of theoretical interest: they serve to understand the sensitivity problem of multiple roots. Application of the results to problems from system theory (identification, modification) are the subject of ongoing work.
- The results can be extended to analyse the sensitivity problem of multiple eigenvalues of block companion matrices (roots of matrix polynomials) .

Some References

[1] F. Chaitin-Chatelin and V. Frayseé, *Lectures on Finite Precision computations*. SIAM, Philadelphia 1996.

[2] W. Gautschi, *Questions of numerical condition related to polynomials*, in MAAA Studies in Mathematics, Vol. 24, Studies in Numerical Analysis, G. H. Golub, ed., USA, 1984, The Mathematical Association of America, pp. 140-177.

[3] F. S. V. Bazán, *Matrix polynomials with partially prescribed eigenstrucutre: Eigenvalue sensitivity and condition estimation.* To appear in J. Comput. Appl. Math.