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
Motivation

Some theoretical results.

Numerical Results.

Conclusions
Motivation: Two problems involving Companion matrices

- Initial value problems:

\[
\begin{align*}
\{ & y^{(m)}(t) + a_{m-1}y^{(m-1)}(t) + \cdots + a_1y'(t) + a_0y(t) = 0, \quad t \geq a \\
& y(a) = \alpha_0, \quad y'(a) = \alpha_1, \ldots, \quad y^{(m-1)}(a) = \alpha_m.
\end{align*}
\]

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

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

\( \Delta \) The solution of the above problems depends on the Jordan form of \( C \)!
Motivation: Two problems involving Companion matrices

- Initial value problems:

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

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

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

- 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 vice versa.

\( \triangle \) 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: $y(t) = e^{Ct}\alpha$, $\alpha = [\alpha_0, \ldots, \alpha_{m-1}]^T$ with

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

- 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 $\lambda_1, \ldots, \lambda_p$ denote the $p$ distinct eigenvalues of $C$ and $m_1, \ldots, m_p$ 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

$$\begin{bmatrix} J_{\lambda_1} & \cdots & \cdots & \cdots \\ \vdots & & & \vdots \\ \cdots & \cdots & & \cdots \\ J_{\lambda_p} & \cdots & \cdots & \cdots \end{bmatrix} = \begin{bmatrix} L_1 & \cdots & \cdots & \cdots \\ \vdots & & & \vdots \\ \cdots & \cdots & & \cdots \\ L_p & \cdots & \cdots & \cdots \end{bmatrix} C \begin{bmatrix} R_1 & \cdots & \cdots & \cdots \\ \vdots & & & \vdots \\ \cdots & \cdots & & \cdots \\ R_p & \cdots & \cdots & \cdots \end{bmatrix} \equiv LCR,$$

where for $i = 1, \ldots, p$, $J_{\lambda_i} = \begin{pmatrix} \lambda_i & 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & \lambda_i \end{pmatrix} \in \mathbb{C}^{m_i \times m_i}$, and

$$LR = RL = I \quad (m \times m \text{ identity matrix}).$$
Jordan Form of $C$

- Columns of $R_i = [r_1, r_2, \ldots, r_{m_i}]$ are the so-called generalized right eigenvectors of $C$ associated with $\lambda_i$. They satisfy
  
  $\begin{cases}
  Cr_1 = \lambda_i r_1 \\
  Cr_i = \lambda_i r_i + r_{i-1}, \ i = 2, \ldots, m_i.
  \end{cases}$

  $\{r_1, r_2, \ldots, r_{m_i}\}$: Right Jordan Chain of $C$
  $r_1$: Leading right generalized eigenvector

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

  $\{l_1, l_2, \ldots, l_{m_i}\}$: Left Jordan Chain of $C$
  $l_{m_i}$: Leading left generalized eigenvector
Proposition 1. Define $\Phi(\lambda) = [1, \lambda, \ldots, \lambda_{m-1}]^T$,

$$H = \begin{bmatrix} a_1 & a_2 & \cdots & a_{m-1} & 1 \\ a_2 & \cdots & a_{m-1} & 1 \\ \vdots & \vdots & \ddots & \vdots \\ a_{m-1} & 1 \\ 1 \end{bmatrix}, \ r_i = H \frac{\Phi^{(i-1)}(\lambda_l)}{(i-1)!}, \text{ and } \ \tilde{y}_i = \frac{\tilde{\Phi}^{(m_i-i)}(\lambda_l)}{(m_l-i)!}.$$

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

$$\tilde{L}_l R_l \equiv \begin{bmatrix} \tilde{y}_1^* \\ \vdots \\ \tilde{y}_{m_l}^* \end{bmatrix} \begin{bmatrix} r_1 & \cdots & r_{m_l} \end{bmatrix} = \begin{pmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_{m_l-1} & \alpha_{m_l} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \cdots & \cdots & \cdots & \alpha_{m_l-1} & \alpha_{m_l} \end{pmatrix} \equiv F_l \quad (2)$$

where $\alpha_i = \frac{\pi^{(m_l+i-1)}(\lambda_l)}{(m_l+i-1)!}$.
**Proposition 2.** Define $L^*_i = [l_1, l_2, \cdots, l_{m_i}] = \tilde{L}^*_i F_i^{-*}$. The set $\{l_1, l_2, \cdots, l_{m_i}\}$ is a left Jordan chain of $C$ associated with the eigenvalue $\lambda_i$, $l_{m_i}$ being the leading left eigenvector. The left and right Jordan chains are normalized so that

$$L_i R_i \equiv \begin{bmatrix} l_1^* \\ \vdots \\ l_{m_i}^* \end{bmatrix} \begin{bmatrix} r_1 & \cdots & r_{m_i} \end{bmatrix} = I \in \mathbb{R}^{m_i \times m_i}. \tag{3}$$

Similarly, we define $\tilde{R}_i = [\tilde{r}_1, \tilde{r}_2, \cdots, \tilde{r}_{m_i}] = [r_1, \cdots, r_{m_i}] F_i^{-1}$. The set $\{\tilde{r}_1, \tilde{r}_2, \cdots, \tilde{r}_{m_i}\}$ is a right Jordan chain of $C$ associated with the eigenvalue $\lambda_i$, and

$$\tilde{L}_i \tilde{R}_i = I \in \mathbb{R}^{m_i \times m_i}.$$
Confluent Vandermonde Matrices: An inversion formula

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

\[
\tilde{L}_i^* = \\
\begin{bmatrix}
0 & \cdots & 0 & 0 & 1 \\
0 & \cdots & 0 & 1 & \lambda_i \\
0 & \cdots & 2 & 2\lambda_i & \lambda_i^2 \\
0 & \cdots & 6\lambda_i & 3\lambda_i^2 & \lambda_i^3 \\
0 & \cdots & \vdots & \vdots & \vdots \\
(m_i - 1)! & \vdots & \vdots & \vdots & \vdots \\
(m - m_i + 1)\lambda_i^{m-m_i} & \cdots & (m-1)(m-2)\lambda_i^{m-3} & (m-1)\lambda_i^{m-2} & \lambda_i^{m-1}
\end{bmatrix},
\]

a consequence of Proposition 2 is

\[
\tilde{L}^{-1} = [R_1 \cdots R_p]F^{-1} \quad \text{with} \quad F = \text{diag}(F_1, \ldots, F_p).
\]
Objective: Estimate the sensitivity of roots of
\[ p(t) = t^m + a_{m-1} t^{m-1} + \cdots + a_1 t + a_0 \]
to small perturbations in the coefficients \( a_j \).

\( \lambda \): root of \( p(t) \) of multiplicity \( d \).

\( \tilde{p}(t) \): monic polynomial with coeff. \( \tilde{a}_j = a_j + \Delta a_j \).

\( \tilde{\lambda}_k, k = 1, \ldots, d \): roots of \( \tilde{p}(t) \) that approximate \( \lambda \)

\[ |\Delta \lambda| = \max_{1 \leq k \leq d} |\lambda - \tilde{\lambda}_k| =? \]

Definition. [Chatelin 1996] Assuma the \( \Delta a_j \)’s satisfy the componentwise inequalities
\[ |\Delta a_j| \leq \varepsilon \alpha_j, \ j = 1, \ldots, m - 1, \]

where \( \alpha_j \) are arbitrary non negative real numbers. The componentwise relative condition number of the root \( \lambda \) of multiplicity \( d \) is defined by
\[ \kappa^C(\lambda) = \lim_{\varepsilon \to 0} \sup_{|\Delta a_j| \leq \varepsilon \alpha_j} \frac{|\Delta \lambda|}{|\lambda| \varepsilon^{1/d}}. \]
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 \( a_j \).

\( \lambda \): root of \( \pi(t) \) of multiplicity \( d \).

\( \tilde{\pi}(t) \): monic polynomial with coeff. \( \tilde{a}_j = a_j + \Delta a_j \).

\( \tilde{\lambda}_k, k = 1, \ldots, d \): roots of \( \tilde{\pi}(t) \) that approximate \( \lambda \)

\[ |\Delta \lambda| = \max_{1 \leq k \leq d} |\lambda - \tilde{\lambda}_k| =? \]

Definition. [Chatelin 1996] Assume the \( \Delta a_j \)'s satisfy the componentwise inequalities

\[ |\Delta a_j| \leq \varepsilon \alpha_j, \ j = 1, \ldots, m-1, \]  (5)

where \( \alpha_j \) are arbitrary non negative real numbers. The componentwise relative condition number of the root \( \lambda \) of multiplicity \( d \) is defined by

\[ \kappa^C(\lambda) = \lim_{\varepsilon \to 0} \sup_{|\Delta a_j| \leq \varepsilon \alpha_j} \frac{|\Delta \lambda|}{|\lambda|\varepsilon^{1/d}}. \]  (6)
Definition. Assuma the $\Delta a_j$’s satisfy the normwise inequality

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

The normwise relative condition number of the root $\lambda$ of multiplicity $d$ is defined by

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

Results

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

$$\kappa(\lambda) = \frac{1}{|\lambda|} \left( \frac{d! \| \phi(\lambda) \| \alpha}{|\pi^{(d)}(\lambda)|} \right)^{1/d}.$$  \hspace{1cm} (10)
**Definition.** Assuma the $\Delta a_j$’s satisfy the normwise inequality

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

The *normwise relative* condition number of the root $\lambda$ of multiplicity $d$ is defined by

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

**Results**

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

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

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

(11)

Consequences: For \( \varepsilon \) 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):
  \[
  |\Delta \lambda| \approx \varepsilon^{1/d} \kappa_a(\lambda)
  \]
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

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

(11)

**Consequences:** For $\varepsilon$ 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):
  $$|\Delta \lambda| \approx \varepsilon^{1/d} \kappa_a(\lambda)$$
**Theorem.** Consider $\pi(t)$ and $q(t)$ so that

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

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

$$
\kappa(\lambda_l) = \frac{1}{|\lambda_l|} \tilde{\kappa}(\lambda_l)^{1/m_l} \|a\|^{1/m_l} \rho^{1/m_l}, \quad (12)
$$

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

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

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

**Table:** Condition numbers, ratio \(\rho\), and theoretical predicted error in \(\lambda\) 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) $\lambda = 15 + 1.5i$. $\circ$: Exact eigenvalue, $\ast$: Approximate eigenvalue.
b) $\lambda = 10 + i$. $\circ$: 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


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

