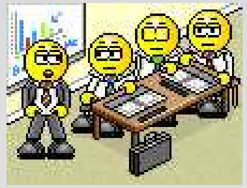

Transformed Companion Matrices as a Theoretical Tool in the Numerical Analysis of Differential Equations

Winfried Auzinger
Vienna University of Technology

(work in progress, with O. Koch and G. Schranz-Kirlinger)

CBC Workshop, Széchenyi István University, Györ
October 22, 2004

Contents

Companion matrix

- Companion (or Frobenius) matrix (nonderogatory):

$$C = \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ -\gamma_0 & -\gamma_1 & \cdots & -\gamma_{n-2} & -\gamma_{n-1} \end{pmatrix} \in \mathbb{C}^{n\times n}$$

- . . . associated with characteristic polynomial

$$\text{charpoly}(C) = p(\zeta) = \sum_{i=0}^{n} \gamma_i \, \zeta^i \quad (\gamma_n := 1)$$

in monomial (or 'Taylor') representation

Algorithmic relevance

- Solving the eigenvalue problem for C (lower Hessenberg) gives the zeros of p
- Monomial representation of p may be numerically unfavorable
- Use other representations (Lagrange, Newton, . . .)
  $\longrightarrow$ transformed companion matrix
- Literature: “Numerical Polynomial Algebra”, by Hans J. Stetter (SIAM Press, 2004)
  . . . multivariate polynomial systems
- We do not consider such algorithmic aspects here
- Rather: Companion matrices as a theoretical tool

Companion matrix as a theoretical tool

- Example 1: n - step Linear Multistep Method
  - characterized by a polynomial p of degree n (p not identical with stability function)
  - Companion matrix C represents equivalent one-step method in a higher dimensional space
- Example 2: Linear ODE of order n
  Companion matrix C represents equivalent 1st order system
- In such cases, the characteristic polynomial p is a ‘symbol’ for the method or the problem
- Stability estimates reduce to norm estimates for $\varphi(C)$, e.g.
  $\varphi(C) = C^{\nu}$ or $\varphi(C) = \exp(tC)$
  (= constant coefficient case)

Linear Multistep Methods

- ODE: $y'(t) = f(t, y(t))$
- n - step linear multistep method (stepsize h)

$$\sum_{k=0}^{n} \alpha_k \, y_{\nu+k} = h \sum_{k=0}^{n} \beta_k f(t_{\nu+k}, y_{\nu+k})$$

($y_{\nu+k} \approx y(t_{\nu+k})$)

- Backward Differentiation Formulas (BDF):
  $\beta_k = 0, \; k < n$, and $\beta_n = 1$
- simplest cases ($n = 1, \; n = 2$, A-stable):
  - $-y_\nu + y_{\nu+1} = hf(t_{\nu+1}, y_{\nu+1})$ (Backward Euler)
  - $\frac{1}{2} y_\nu - 2\, y_{\nu+1} + \frac{3}{2} y_{\nu+2} = hf(t_{\nu+2}, y_{\nu+2})$ (BDF 2)

Example: Stability of BDF 2 (i)

- To begin with: model problem $y' = \lambda y$, $\text{Re}\,\lambda < 0$, stiff
- Note: BDF 2 is A-stable $\Rightarrow$
  - $|y_\nu| \to 0$ for $\nu \to \infty$
  - But: Estimates for finite ν not directly available
  - ? $|y_\nu| \leq$?
- . . . This is not an open problem . . .
  - Write BDF 2 as a one-step method in $\mathbb{C}^2$ using companion matrix
  - Apply the Kreiss Matrix Theorem, or
  - Estimate using G-stability
- But: Both approaches have a very restricted scope

Example: Stability of BDF 2 (ii)

- Let $\mu := h\lambda$, $Y_\nu := (y_\nu, y_{\nu+1})^T$
- BDF 2 $\iff$ $Y_{\nu+1} = C(\mu)\, Y_\nu$, with companion matrix

$$C(\mu) = \begin{pmatrix} 0 & 1 \\ -\frac{1}{3-2\mu} & \frac{4}{3-2\mu} \end{pmatrix}, \quad |\text{EV}| \le 1 \ (\text{Re}\, \mu \le 0)$$

- Wanted: Estimate $\| C(\mu) \| \le 1$, uniformly for $\text{Re}\, \mu \le 0$
  - For $\| \cdot \| = \| \cdot \|_2$ this does not hold (consider $\mu = 0$)
  - G-stability: Estimate O.K. for $\|Y\| = \langle G\,Y, Y\rangle_2$, $G = \dots$
  - But: $\nexists$ G-stable n - step LMM for $n > 2$
  - $\Rightarrow$ not generalizable
  - Kreiss Matrix Theorem: Restricted to const. coeff.

Example: Stability of BDF 2 (iii)

- . . . look for general, more flexible approach to stability
- Jordan canonical form for $C = C(\mu)$:
  - C is diagonalizable for $\mu \neq -1/2$, $C = X\,\Xi\,X^{-1}$
  - X = Vandermonde matrix $\begin{pmatrix} 1 & 1 \\ \xi_1 & \xi_2 \end{pmatrix}$, ξ_k = EV of C
  - X becomes singular (confluent) for $\mu \to -1/2$
  - Estimate $\|C\| \leq \|X\|\,\|\Xi\|\,\|X^{-1}\|$ useless near $\mu = -1/2$
- Note (charpoly$(C) = p(\zeta) = (\zeta - \xi_1)(\zeta - \xi_2)$):
  - Jordan form of C is discontinuous w.r.t. parameter μ
  - corresponds to Lagrange representation of p for $\xi_1 \neq \xi_2$
  - undefined for $\xi_1 = \xi_2$ ($\mu = -1/2$)

Example: Stability of BDF 2 (iv)

- Our approach:
  - Consider LU-decomposition of X, $X = L\,U$, with

$$L = \begin{pmatrix} 1 & 0 \\ \xi_1 & 1 \end{pmatrix}$$

  - Then, for all μ: $C = L\,J\,L^{-1}$, with

$$J = \begin{pmatrix} \xi_1 & 1 \\ 0 & \xi_2 \end{pmatrix}$$

  - … *"Bidiagonal canonical form"* of C
  - continuous w.r.t. parameter μ
  - cond(L) uniformly bounded for Re $\mu \le 0$

Example: Stability of BDF 2 (v)

- We can prove:
  - After appropriate diagonal scaling, $J \to \tilde{J} = D\, J\, D^{-1}$,

$$\| \tilde{J} \|_\infty \le 1 \quad \forall\, \mathrm{Re}\, \mu \le 0$$

  - $\Rightarrow$ uniform stability estimate

$$\| C^\nu \|_\infty \le K, \quad \forall\, \nu > 0\,, \quad \forall\, \mathrm{Re}\, \mu \le 0$$

  - but: ... does not work for $\| \cdot \| = \| \cdot \|_2$
  - Generalization: Transform C to *“Bidiagonal-Frobenius form”*, $C = L\, H\, L^{-1}$, with appropriate $\eta_1,\, \eta_2$ and

$$H = \begin{pmatrix} \eta_1 & 1 \\ -p[\eta_1] & -p[\eta_1, \eta_2] + \eta_2 \end{pmatrix}, \quad L = \begin{pmatrix} 1 & 0 \\ \eta_1 & 1 \end{pmatrix}$$

Example: Stability of BDF 2 (vi)

- We can prove:
  - With $\eta_1 = \eta_2 = \frac{1}{2}\,\text{trace}(C)$, and after appropriate diagonal scaling, $H \to \tilde{H} = D\,H\,D^{-1}$,

$$\|\,\tilde{H}\,\|_2 \le 1 \quad \forall\ \text{Re}\,\mu \le 0$$

    (and: scaled version $\tilde{L}$ of L uniformly well-conditioned)
  - Proof: Apply the Cohn-Schur-criterion to $\tilde{H}^T\tilde{H}$
- Interpretation in terms of $p = \text{charpoly}(C)$:
  - $\tilde{H}$ is associated with a (scaled) Newton-Taylor representation of p
  - = Newton representation allowed to be confluent (always well-defined)

Bidiagonal-Frobenius form (i)

- For each $n \times n$ companion matrix C: $C = L\,H\,L^{-1}$,

$$H = \begin{pmatrix} \eta_1 & 1 & & & \\ & \eta_2 & 1 & & \\ & & \ddots & \ddots & \\ & & & \eta_{n-1} & 1 \\ -p_{[1]} & -p_{[1\cdot\cdot 2]} & \cdots & -p_{[1\cdot\cdot n-1]} & -p_{[1\cdot\cdot n]}+\eta_n \end{pmatrix}$$

- $\eta_1, \ldots, \eta_n$ arbitrary (not necessarily distinct)
- $p_{[j\cdot\cdot k]}$ . . . [confluent] divided differences of p w.r.t. the η_j
- L from LU-decomposition of Vandermonde matrix $X = X(\eta_1, \ldots, \eta_n)$

Confluent divided differences

- $\eta_1, \ldots, \eta_n$ arbitrary (not necessarily distinct)
- [Confluent] divided differences (φ smooth enough, $j \leq k$):

$$\varphi[\eta_j, \ldots, \eta_k] :=$$

$$\begin{cases} \dfrac{\varphi[\eta_{j+1}, \ldots, \eta_k] - \varphi[\eta_j, \ldots, \eta_{k-1}]}{\eta_k - \eta_j}, & \eta_j \neq \eta_k \\ \lim\limits_{\varepsilon \to 0} \dfrac{\varphi[\eta_{j+1}, \ldots, \eta_k + \varepsilon] - \varphi[\eta_j, \ldots, \eta_{k-1}]}{\varepsilon}, & \eta_j = \eta_k \end{cases}$$

- Includes all possible combinations of confluent (‘derivative-like’) and nonconfluent cases

Bidiagonal-Frobenius form (ii)

- BF-form always well-defined, continuous w.r.t. varying C
- L (lower triangular):

$$L = \begin{pmatrix} | & | & & & | \\ x[\eta_1] & x[\eta_1, \eta_2] & \dots & \dots & x[\eta_1, \dots, \eta_n] \\ | & | & & & | \end{pmatrix}$$

  where $x[\eta] = (1, \eta, \dots, \eta^{n-1})^T$, $x[\eta_j]$ = j - th column of X
- Interpretation in terms of $p =$ charpoly(C) :
  - H associated with Newton-Taylor representation of p
  - = Newton representation allowed to be confluent

Bidiagonal-Frobenius form (iii)

- Special case: $\eta_j \equiv \xi_j$ = EV of C $\Rightarrow$ *Bidiagonal form*:

$$H = J = \begin{pmatrix} \xi_1 & 1 & & & \\ & \xi_2 & 1 & & \\ & & \ddots & \ddots & \\ & & & \xi_{n-1} & 1 \\ & & & & \xi_n \end{pmatrix}$$

- Remark: Matrix functions $\varphi(J)$ are upper tridiagonal, with

$$(\varphi(J))_{j,k} = \varphi[\xi_j, \ldots, \xi_k]$$

- Trivial case: $\eta_j \equiv 0 \Rightarrow H = C$
- BF-form can be generalized to companion matrix w.r.t. arbitrary basis polynomials

Stability of Linear Multistep Methods

- Approach much more versatile than G-stability or Kreiss Matrix Theorem
- Bidiagonal canonical form has been used for stability analysis of A(α)-stable BDF methods applied to stiff ODEs
  - $y' = \lambda(t)y$
  - $y' = A(t)y + \phi(t, y)$

    (A. Eder, G. Schranz-Kirlinger)
- To be done: Stability/convergence analysis for highly nonlinear problems

Further applications, Remarks

- Sharp growth estimates for higher order ODEs
- [Numerical] analysis of singular ODE systems:
  - Well-posedness of BVP depends on Jordan structure of boundary conditions
  - ... Use of bidiagonal form simplifies the analysis, e.g. for parameter-dependent problems
- General linear methods: [generalized] companion matrices play a role in the analysis – may be worth considering (Example: Butcher, Wright – ESIRK methods characterized by doubly companion matrices)
- General matrices: Frobenius canonical form $\longrightarrow$ can be transformed blockwise