A Review of Nonlinear Eigenvalue Problems

Françoise Tisseur
School of Mathematics
The University of Manchester

ftisseur@ma.man.ac.uk
http://www.ma.man.ac.uk/~ftisseur/

ILAS Conference, Providence, June 3-7, 2013.

Nonlinear Eigenvalue Problems (NEPs)

Let $F\colon \Omega \to \mathbb{C}^{n\times n}$ be a given matrix valued function.

The **nonlinear eigenvalue problem**: Find scalars λ and nonzero $x, y \in \mathbb{C}^n$ satisfying $F(\lambda)x = 0$ and $y^* F(\lambda) = 0$.

- λ is an e'val, x, y are corresponding right and left e'vecs.
- E'vals are solutions of $\det(F(\lambda)) = 0$.

In practice, elements of F most often **polynomial, rational or exponential functions of λ**.

Nonlinear Eigenvalue Problems (NEPs)

Let $F\colon \Omega \to \mathbb{C}^{n\times n}$ be a given matrix valued function.

The **nonlinear eigenvalue problem**: Find scalars λ and nonzero $x, y \in \mathbb{C}^n$ satisfying $F(\lambda)x = 0$ and $y^* F(\lambda) = 0$.

- λ is an e'val, x, y are corresponding right and left e'vecs.
- E'vals are solutions of $\det(F(\lambda)) = 0$.

In practice, elements of F most often **polynomial, rational or exponential functions of λ**.

When F depends on e'vecs, see Jarlebring's talk, 11:00.

Comments

- NEPs underpin many areas of computational science and engineering.
- Collection of Nonlinear Eigenvalue Problems **NLEVP**, Betcke, Higham, Mehrmann, Schröder, T., 2011.
- NEPs can be very difficult to solve:
  - nonlinear,
  - large problem size,
  - poor conditioning,
  - lack of “good" numerical methods.
- Common feature of recent progress: clever conversion/approx. of $F(\lambda)x = 0$ by $(\mathcal{A} - \lambda\mathcal{B})\xi = 0$.

Polynomial Eigenvalue Problems (PEP)

Consider $n \times n$ matrix polynomial

$$P(\lambda) = \lambda^\ell A_\ell + \cdots + \lambda A_1 + A_0 \quad \text{(monomial form)}$$
$$= \phi_\ell(\lambda) P_\ell + \cdots + \phi_1(\lambda) P_1 + P_0 \quad \text{(non-monomial form)},$$

where $\phi_i(\lambda)$ is a polynomial of degree i.

- Assume $P(\lambda)$ is **regular**, i.e., $\det(P(\lambda)) \not\equiv 0$.
- $P(\lambda)$ has ℓn **e'vals**. Finites e'vals are roots of $\det(P(\lambda)) = 0$.
- Zero e'vals when $\det(A_0) = 0$ and ∞ e'vals when $\det(A_\ell) = 0$.

For theory, see Gohberg, Lancaster, Rodman, **Matrix Polynomials**, 1982 (Academic Press), 2009 (SIAM).

Schur's Theorem for Complex Matrices

Matrix version: if $A \in \mathbb{C}^{n\times n}$ then there exists a unitary U such that $U^*AU = T$ is a triangular matrix.

Subspaces version:

Theorem

Let $A \in \mathbb{C}^{n\times n}$. There are subspaces $\mathcal{V}_1, \dots, \mathcal{V}_n$ of $\mathbb{C}^n$ satisfying

(i) $\mathbb{C}^n = \mathcal{V}_1 \oplus \mathcal{V}_2 \oplus \cdots \oplus \mathcal{V}_n$,

(ii) *for $k = 1: n$, $\mathcal{V}_1 \oplus \cdots \oplus \mathcal{V}_k$, is A-invariant,*

(iii) *for $k = 1: n$, $\mathcal{V}_k = \langle u_k \rangle$, where $u_1, \dots u_n$ form an orthonormal system of vectors of $\mathbb{C}^n$.*

Schur-like Theorem for $Q(\lambda)$

Let $Q(\lambda) = \lambda^2 M + \lambda D + K$, $\det M \neq 0$.

Theorem (Zaballa, T., 2013)

Let $\lambda I - A \in \mathbb{C}[\lambda]^{2n\times 2n}$ be a linearization of $n \times n$ $Q(\lambda)$. There are subspaces $\mathcal{V}_1, \dots, \mathcal{V}_n$ of $\mathbb{C}^{2n}$ satisfying

(i) $\mathbb{C}^{2n} = \mathcal{V}_1 \oplus \mathcal{V}_2 \oplus \cdots \oplus \mathcal{V}_n$,

(ii) *for* $k = 1\colon n$, $\mathcal{V}_1 \oplus \cdots \oplus \mathcal{V}_k$ *is* A*-invariant,*

(iii) *for* $k = 1\colon n$, $\dim \mathcal{V}_k = 2$ *and* $\mathcal{V}_k = \langle u_k, Au_k \rangle$, *where* $u_1, \dots u_n$ *form an orthogonal system of vectors of* $\mathbb{C}^{2n}$.

- $\mathcal{V}_k = \langle u_k, Au_k \rangle$ is a **Krylov subspace** of dimension 2.
- The u_j are **generating vectors**.
- If $U = [u_1 \dots u_n]$ then $\mathcal{U} = [U \; AU]$ is nonsingular.

Schur-like Theorem: Matrix Form

Let $Q(\lambda) = \lambda^2 M + \lambda D + K \in \mathbb{C}[\lambda]^{n\times n}$ with $\det(M) \neq 0$.

Theorem (Zaballa, T., 2012)

For any linearization $\lambda I - A$ of $Q(\lambda)$, there exists $U \in \mathbb{C}^{2n\times n}$ with ***orthonormal columns*** *s.t. $[U\ AU]$ is nonsingular and*

$$[U\ AU]^{-1}A[U\ AU] = \begin{bmatrix} 0 & -T_0 \\ I_n & -T_1 \end{bmatrix},$$

where $I_n\lambda^2 + T_1\lambda + T_0$ is triangular and equivalent to $Q(\lambda)$.

- The columns of U are generating vectors for A.
- Extends to **arbitrary degree matrix polynomials**.
- **Issue: how to compute U?** No efficient way yet!

Standard Solution Process ($\ell = 2$)

Find all λ and x satisfying $Q(\lambda)x = (\lambda^2 M + \lambda D + K)x = 0$.

- Commonly solved by **linearization**:
  - **Convert** $Q(\lambda)x = 0$ into $(\mathcal{A} - \lambda\mathcal{B})\xi = 0$, e.g.,

    $$\mathcal{A} - \lambda\mathcal{B} = \begin{bmatrix} K & 0 \\ 0 & I \end{bmatrix} - \lambda \begin{bmatrix} -D & -M \\ I & 0 \end{bmatrix}, \quad \xi = \begin{bmatrix} x \\ \lambda x \end{bmatrix}.$$

  - **Solve** $(\mathcal{A} - \lambda\mathcal{B})\xi = 0$ with an eigensolver for generalized eigenproblem (e.g., QZ algorithm).
  - **Recover** eigenvectors of $Q(\lambda)$ from those of $\mathcal{A} - \lambda\mathcal{B}$.
- Solution process extend to degree $\ell > 2$.
- Numerical issues with this process.

Numerical Issues

- **Infinitely many linearizations** to choose from.
  - Vector spaces of pencils $\mathbb{L}_1(P), \mathbb{L}_2(P), \mathbb{DL}(P)$, [$M^4$, 06], [Townsend, Nakatsukasa, Noferini, 12] (see Noferini's talk, today at 5:45).
  - Fiedler linearizations [Antoniou, Vologiannidis 04].
  - Linearizations for P expressed in polynomial bases [Amiraslani, Corless, Lancaster,09].

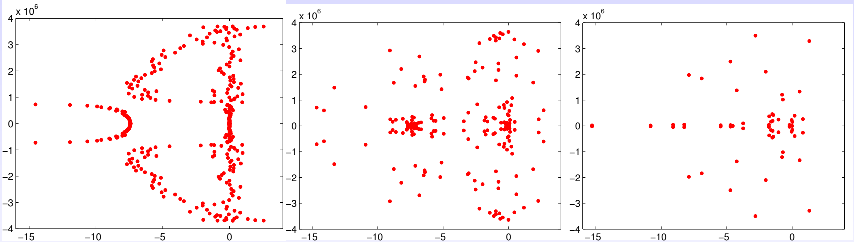
Numerical Issues

- **Infinitely many linearizations** to choose from.
  - Vector spaces of pencils $\mathbb{L}_1(P), \mathbb{L}_2(P), \mathbb{DL}(P)$, [M[4], 06], [Townsend, Nakatsukasa, Noferini, 12] (see Noferini's talk, today at 5:45).
  - Fiedler linearizations [Antoniou, Vologiannidis 04].
  - Linearizations for P expressed in polynomial bases [Amiraslani, Corless, Lancaster,09].
- Linearizations can have **widely varying eigenvalue condition numbers**.
- Solving PEPs with a backward stable alg. (e.g., QZ alg.) applied to a linearization can be **backward unstable** for the PEP.

Computed Spectra of C_1, L_1 and L_2

$$C_1(\lambda) = \lambda \begin{bmatrix} M & 0 \\ 0 & I \end{bmatrix} + \begin{bmatrix} D & K \\ -I & 0 \end{bmatrix},$$

$$L_1(\lambda) = \lambda \begin{bmatrix} M & 0 \\ 0 & -K \end{bmatrix} + \begin{bmatrix} D & K \\ K & 0 \end{bmatrix}, \quad L_2(\lambda) = \lambda \begin{bmatrix} 0 & M \\ M & D \end{bmatrix} + \begin{bmatrix} -M & 0 \\ 0 & K \end{bmatrix}$$



- Developed theory concerning the sensitivity and stability of linearizations [Higham, Mackey, T. 06, Higham, Li, T. 07, Grammont, Higham, T., 11].

Two Sided Factorizations

Suppose linearization L of P satisfies

$$L(\lambda)F(\lambda) = g \otimes P(\lambda), \qquad E(\lambda)L(\lambda) = h^T \otimes P(\lambda) \qquad (*)$$

with F, G of full rank and $g, h \in \mathbb{C}^\ell$.

$(*)$ Satisfied by most linearizations.

$(*)$ Yields relations between e'vecs for L & e'vecs for P:

- $z = F(\lambda)x \in \mathbb{C}^{\ell n}$ is a right e'vec of L with e'val λ iff $x \in \mathbb{C}^n$ is a right e'vec of P with e'val λ.
- If $w \in \mathbb{C}^{\ell n}$ is a left e'vec of L with e'val λ then $y = (g^* \otimes I_n)w \in \mathbb{C}^n$ is a left e'vec of P (if it's nonzero).

$(*)$ Implies $w^* L'(\lambda) z = y^* P'(\lambda) x$.

$(*)$ Relates residuals: $P(\lambda)x = E(\lambda)L(\lambda)z$.

Sensitivity and Stability of Linearizations

- Two sided factorizations allow analysis of growth factors $\kappa_L(\lambda)/\kappa_P(\lambda)$ (**conditioning**), $\eta_P(z_i, \lambda)/\eta_L(z, \lambda)$ (**backward error**).
- Importance of **scaling** QEPs/PEPs before computing e'vals via linearization.
  - Eigenvalue parameter scaling:

$$\lambda = \gamma\mu, \quad \widetilde{P}(\mu) := \delta P(\gamma\mu).$$

  - Does not affect sparsity of matrix coeffs.
  - γ, δ chosen to improve growth factors for conditioning and backward error.

`quadeig` for $Q(\lambda) = \lambda^2 M + \lambda D + K$.

- Eigensolver for dense (small to medium size) quadratics—**`quadeig`**.
- Incorporates:
  - Appropriate choice of linearization: uses $\begin{bmatrix} D & -I \\ K & 0 \end{bmatrix} - \lambda \begin{bmatrix} -M & 0 \\ 0 & -I \end{bmatrix}$.
  - Deflation of 0 and ∞ eigenvalues.
  - Eigenvalue parameter scaling (FLV/tropical). (see Sharify and Hook's talk, Tues., 4:45)
  - Careful recovery of the eigenvectors.
  - Backward stable when $\|D\| \lesssim (\|M\|\|K\|)^{1/2}$.
- MATLAB and Fortran implementations (NAG, LAPACK) [Hammarling, Munro, T. 2013].

Eigensolvers for Large Sparse QEPs

For $Q(\lambda) = \lambda^2 M + \lambda D + I$, they include:

- Q-Arnoldi method [Meerbergen, 08], $S = \begin{bmatrix} 0 & I \\ -M & -D \end{bmatrix}$.
  - Krylov subspace $\mathcal{K}_p(S, v) = \mathrm{span}\{v, Sv, \dots, S^p v\}$.
  - Use structure in Arnoldi iter. to save memory.
  - Can do implicit restart.
- Second Order Arnoldi (SOAR) [Bai & Su, 05].
  - 2nd order Krylov: $\mathcal{G}_p(D, M, b) = \mathrm{span}\{r_0, r_1, \dots, r_p\}$, $r_0 = b, r_1 = -Dr_0,\ r_j = -Dr_{j-1} - Mr_{j-2}$.
    If $v = \begin{bmatrix} b \\ 0 \end{bmatrix}$ then $\mathcal{G}_p$ defines $\mathcal{K}_p$.
  - $X_p^* Q(\lambda) X_p = \lambda^2 M_p + \lambda D_p + I_p$, $\mathrm{span} X_p = \mathcal{G}_p$.
  - Cannot do implicit restart.
- Two-level orthogonal Arnoldi (TOAR) [Su & Lu, 12].

Rational Eigenvalue Problems (REPs)

- Arise in a variety of physical applications, e.g.,
  - acoustic emissions of high speed trains,
  - band structure calculations for photonic crystals.

- Matrix function takes the form

$$R(\lambda) = P(\lambda) + \sum_{i=1}^{m} \frac{s_i(\lambda)}{q_i(\lambda)} E_i,$$

  - $P(\lambda)$ is a matrix polynomial,
  - s_i and q_i are scalar polynomials,
  - E_i are constant matrices.

In practice, m is small and $\text{rank}(E_i)$ is low.

Numerical Methods

$$R(\lambda) = P(\lambda) + \sum_{i=1}^{m} \frac{s_i(\lambda)}{q_i(\lambda)} E_i.$$

- Brute-force approach: turn the REP into a PEP.
  - Employed by practitioners.
  - Not practical when $q_i(\lambda)$ has several poles and $m \gg 1$.
- Treat the REP as a general nonlinear eigenproblem.
  - Limits the exploitation of the underlying structure and properties of the REP.
- **Linearizations** for rational problems .

Linearizations for Rational Problems

[Su & Bai, 11]

- Rewrite $R(\lambda) = P(\lambda) + \sum_{i=1}^{m} \frac{s_i(\lambda)}{q_i(\lambda)} E_i$ as

$$R(\lambda) = P(\lambda) + U(C - \lambda D)^{-1} V^*, \qquad (*)$$

  where U, V are $n \times r$. In practice $r \ll n \cdot \sum_{i=1}^{m} deg(q_i)$.

- Convert $(*)$ into linear pencil of dim. $nk + r$.

Linearizations for Rational Problems

[Su & Bai, 11]

- Rewrite $R(\lambda) = P(\lambda) + \sum_{i=1}^{m} \frac{s_i(\lambda)}{q_i(\lambda)} E_i$ as

$$R(\lambda) = P(\lambda) + U(C - \lambda D)^{-1} V^*, \qquad (*)$$

where U, V are $n \times r$. In practice $r \ll n \cdot \sum_{i=1}^{m} deg(q_i)$.

- Convert $(*)$ into linear pencil of dim. $nk + r$.

Example: if $R(\lambda) = \lambda A + B + U(C - \lambda D)^{-1} V^*$ then $R(\lambda)x = 0$ is rewritten as a **linear eigenproblem**

$$\left(\lambda \begin{bmatrix} A & 0 \\ 0 & D \end{bmatrix} + \begin{bmatrix} B & U \\ V^* & -C \end{bmatrix}\right) \begin{bmatrix} x \\ (C - \lambda D)^{-1} V^* x \end{bmatrix} = 0.$$

Genuine Nonlinear Eigenproblems

$F(\lambda)x = 0$ with $F\colon \Omega \to \mathbb{C}^{n\times n}$, $\Omega \subseteq \mathbb{C}$.

Some applications:

- PDE models with λ-dependent boundary conditions or material conditions. E.g.,

$$F(\lambda) = K - \lambda M + i(\lambda - \sigma_1^2)^{\frac{1}{2}} W_1 + i(\lambda - \sigma_2^2)^{\frac{1}{2}} W_2 = 0$$

(model radio-frequency gun cavity, [Liao, 07]).

Genuine Nonlinear Eigenproblems

$F(\lambda)x = 0$ with $F\colon \Omega \to \mathbb{C}^{n\times n}$, $\Omega \subseteq \mathbb{C}$.

Some applications:

- PDE models with λ-dependent boundary conditions or material conditions. E.g.,

  $$F(\lambda) = K - \lambda M + i(\lambda - \sigma_1^2)^{\frac{1}{2}} W_1 + i(\lambda - \sigma_2^2)^{\frac{1}{2}} W_2 = 0$$

  (model radio-frequency gun cavity, [Liao, 07]).

- Characteristic function of delay differential eqns. E.g.,

  $$F(\lambda) = -\lambda I + A_1 + A_2 e^{-\tau\lambda} + A_3 \int_{-1}^{0} e^{\lambda s}(e^{(s+\frac{1}{2})^2} - e^{\frac{1}{4}})\, ds$$

  (delay-diff. eqn with distributed delays, [Jarlebring et al, 12]).

Genuine Nonlinear Eigenproblems

$F(\lambda)x = 0$ with $F\colon \Omega \to \mathbb{C}^{n\times n}$, $\Omega \subseteq \mathbb{C}$.

Some applications:

- PDE models with λ-dependent boundary conditions or material conditions. E.g.,

  $$F(\lambda) = K - \lambda M + i(\lambda - \sigma_1^2)^{\frac{1}{2}} W_1 + i(\lambda - \sigma_2^2)^{\frac{1}{2}} W_2 = 0$$

  (model radio-frequency gun cavity, [Liao, 07]).

- Characteristic function of delay differential eqns. E.g.,

  $$F(\lambda) = -\lambda I + A_1 + A_2 e^{-\tau\lambda} + A_3 \int_{-1}^{0} e^{\lambda s}(e^{(s+\frac{1}{2})^2} - e^{\frac{1}{4}})\, ds$$

  (delay-diff. eqn with distributed delays, [Jarlebring et al, 12]).

- Use of special basis functions in the discretization.

Methods for Nonlinear Eigenproblems

- Typically interested in only a few e'vals closest to a target point, a line or small region of complex plane.

- Numerical methods include
  - Newton methods, inverse iteration,. . . [Ruhe, 73].
  - Nonlinear Rayleigh-Ritz iterative procedure, e.g., nonlinear Arnoldi [Voss, 04], NRRIT, [Liao, Bai, Lee, Ko, 06], rational Krylov [Jarlebring, Voss, 05], Jacobi-Davidson [Voss, 04], ...
  - Contour integration [Beyn, Effenberger, Kressner, 11; Sakurai et al, 10]. (see Yamamoto's talk, Tues, 5:45)

Numerical Methods (Cont.)

- Approximate $F(\lambda)$ by a polynomial

$$F(\lambda) \approx P_d(\lambda) = a_0(\lambda)A_0 + a_1(\lambda)A_1 + \cdots + a_d(\lambda)A_d,$$

  with $a_j(\lambda)$, e.g., Taylor, Chebyshev, Newton, Hermite.

- Solve $P_d(\lambda)x = 0$ (see Karl Meebergen's talk).

Numerical Methods (Cont.)

- Approximate $F(\lambda)$ by a polynomial

  $$F(\lambda) \approx P_d(\lambda) = a_0(\lambda)A_0 + a_1(\lambda)A_1 + \cdots + a_d(\lambda)A_d,$$

  with $a_j(\lambda)$, e.g., Taylor, Chebyshev, Newton, Hermite.

- Solve $P_d(\lambda)x = 0$ (see Karl Meebergen's talk).
  - Linearization $\mathcal{A} - \lambda\mathcal{B}$ for non monomial basis [Amiraslani, Corless, Lancaster, 09].
  - For “clever" choice of $\mathcal{A} - \lambda\mathcal{B}$ can implement Arnoldi/rational Krylov iter. without fixing d.
    - Infinite Arnoldi [Jarlebring, Michiels, Meerbergen, 12].
    - Rational Krylov for NEP [Van Beeumen, Michiels, Meerbergen, 13].

  (See R. Van Beeumen's talk, Thur, 11:40)

Summary and Concluding Remarks

- Existence of Schur-like theorem for matrix polynomials.
- Better understanding of linearization process.
- Need to develop a backward stable algorithm for PEPs of degree larger than 2. Use of max-plus algebra.
- Rational eigenproblems with low rank matrices can be solved via linearization.
- New linearizations techniques for NEPs allowing a dynamic implementation of Arnoldi/Rational Krylov.
- Difficulties remain when the nonlinearity is difficult/impossible to approximate by a polynomial.

NLA Group in Manchester

References I

- A. Amiraslani, R. M. Corless, and P. Lancaster.
  Linearization of matrix polynomials expressed in polynomial bases.
  IMA J. Numer. Anal., 29:141–157, 2009.
- Z. Bai and Y. Su.
  SOAR: a second-order Arnoldi method for the solution of the quadratic eigenvalue problem.
  SIAM J. Matrix Anal. Appl., 26(3):640–659, 2005.
- T. Betcke, N. J. Higham, V. Mehrmann, C. Schröder, and F. Tisseur.
  NLEVP: A collection of nonlinear eigenvalue problems.
  ACM Trans. Math. Software, 39(2):7:1–7:28, Feb. 2013.

References II

- I. Gohberg, P. Lancaster, and L. Rodman.
  Matrix Polynomials.
  Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2009.
  ISBN 0-898716-81-8.
  xxiv+409 pp.
  Unabridged republication of book first published by Academic Press in 1982.
- L. Grammont, N. J. Higham, and F. Tisseur.
  A framework for analyzing nonlinear eigenproblems and parametrized linear systems.
  Linear Algebra Appl., 435(3):623–640, 2011.

References III

S. Hammarling, C. J. Munro, and F. Tisseur.
An algorithm for the complete solution of quadratic eigenvalue problems.
ACM Trans. Math. Software, 39(3):18:1–18:19, Apr. 2013.

N. J. Higham, R.-C. Li, and F. Tisseur.
Backward error of polynomial eigenproblems solved by linearization.
SIAM J. Matrix Anal. Appl., 29(4):1218–1241, 2007.

N. J. Higham, D. S. Mackey, and F. Tisseur.
The conditioning of linearizations of matrix polynomials.
SIAM J. Matrix Anal. Appl., 28(4):1005–1028, 2006.

References IV

N. J. Higham, D. S. Mackey, F. Tisseur, and S. D. Garvey.
Scaling, sensitivity and stability in the numerical solution of quadratic eigenvalue problems.
Internat. J. Numer. Methods Eng., 73(3):344–360, 2008.

E. Jarlebring, W. Michiels, and K. Meerbergen.
The infinite Arnoldi method and an application to time-delay systems with distributed delay.
Lecture Notes in Control and Information Sciences, 423, 2012.

References V

E. Jarlebring, W. Michiels, and K. Meerbergen.
A linear eigenvalue algorithm for the nonlinear eigenvalue problem.
Numer. Math., 122(1):169–195, 2012.

K. Meerbergen.
The quadratic Arnoldi method for the solution of the quadratic eigenvalue problem.
SIAM J. Matrix Anal. Appl., 30(4):1463–1482, 2008.

Y. Su and Z. Bai.
Solving rational eigenvalue problems via linearization.
SIAM J. Matrix Anal. Appl., 32(1):201–216, 2011.

References VI

F. Tisseur and K. Meerbergen.
The quadratic eigenvalue problem.
SIAM Rev., 43(2):235–286, 2001.

R. Van Beeumen, K. Meerbergen, and W. Michiels.
A rational krylov method based on hermite interpolation for nonlinear eigenvalue problems.
SIAM J. Sci. Comput., 35(1):A327–A350, 2013.

H. Voss.
An arnoldi method for nonlinear eigenvalue problems.
BIT, 44:387–401, 2004.