Nonlinear Eigenvalue Problems: An Introduction

Cedric Effenberger

Seminar for Applied Mathematics
ETH Zurich

Pro*Doc Workshop
Disentis, August 18–21, 2010
Overview

1. Definition
2. Applications
3. Solution
4. Solution of polynomial / rational EVPs
Linear Eigenvalue Problems

Standard eigenvalue problem
Given $A \in \mathbb{C}^{n \times n}$, we seek
- eigenvalues $\lambda \in \mathbb{C}$, and
- eigenvectors $x \in \mathbb{C}^n \setminus \{0\}$
such that $Ax = \lambda x$.

Generalized eigenvalue problem
Given $A, B \in \mathbb{C}^{n \times n}$, we seek
- eigenvalues $\lambda \in \mathbb{C}$, and
- eigenvectors $x \in \mathbb{C}^n \setminus \{0\}$
such that $Ax = \lambda Bx$. 
Generalization to nonlinear EVPs

Both problems can be reformulated as

\[ T(\lambda)x = 0, \]

where

\[ T(\lambda) = \begin{cases} A - \lambda I, & \text{for the standard EVP,} \\ A - \lambda B, & \text{for the generalized EVP.} \end{cases} \]

- \[ T(\lambda) = \lambda^d A_d + \cdots + \lambda A_1 + A_0 \quad \sim \quad \text{polynomial EVP} \]
- \[ T(\lambda) = f_1(\lambda)A_1 + \cdots + f_r(\lambda)A_r \quad \sim \quad \text{nonlinear EVP} \]

\[ A_j \in \mathbb{C}^{n \times n} \text{ constant matrices, } f_j : D \to \mathbb{C} \text{ analytic,} \]
\[ D \subset \mathbb{C} \text{ open, connected} \]
Free vibrations of mechanical systems with damping

Equations of motion

\[ M\ddot{u} + C\dot{u} + Ku = 0 \]

- \( M \): mass matrix
- \( C \): damping matrix
- \( K \): stiffness matrix

- System of second-order, linear, homogeneous ODEs

- Eigenfrequencies \( \omega \) of the mechanical system are eigenvalues of

\[ (\omega^2 M + \omega C + K)u = 0. \]

\( \Rightarrow \) Quadratic eigenvalue problem
Free vibrations of string with elastically attached mass

- vibrating string \( \Rightarrow -u'' = \lambda u \) in \( \Omega = (0, 1) \)
- left end clamped \( \Rightarrow u(0) = 0 \)
- mass elastically attached to right end \( \Rightarrow u'(1) + \frac{\alpha \lambda}{\lambda - \alpha} u(1) = 0 \)

\( \lambda \)-dependent boundary condition!

Discretization with \( n \) piecewise linear FE leads to

\[
\left[ K + \frac{\alpha \lambda}{\lambda - \alpha} e_n e_n^T - \lambda M \right] x = 0.
\]
### Delay eigenvalue problems

**Linear delay differential equation with \( r \) discrete delays**

\[
\dot{u}(t) = A_0 u(t) + A_1 u(t - \tau_1) + \cdots + A_r u(t - \tau_r)
\]

- models influences which take effect only after some time
- famous example: Hot shower problem
- stability analysis involves nonlinear EVPs

**Delay eigenvalue problem**

\[
(-\lambda I + A_0 + e^{-\tau_1 \lambda} A_1 + \cdots + e^{-\tau_r \lambda} A_r)x = 0
\]
Electronic bandstructure computations

- **photonic crystal**: lattice of mixed dielectric media
- control electromagnetic waves by designing the crystal such that it inhibits their propagation
- **complete photonic band gap**: frequency range with no propagation of electromagnetic waves of *any* polarization travelling in *any* direction
Electronic bandstructure computations II

- 2D crystal: periodic in $x$- and $y$-direction; homogeneous in $z$-direction
- consider only electromagnetic waves propagating in the $xy$-plane
Time-harmonic modes of electromagnetic wave \((E, H)\) can be decomposed into

- transverse electric (TE) polarized modes \((E_x, E_y, 0, 0, 0, H_z)\),
- transverse magnetic (TM) polarized modes \((0, 0, E_z, H_x, H_y, 0)\).

For TM polarized modes, the macroscopic Maxwell Equations reduce to a scalar equation for \(E_z\),

\[-\Delta E_z = \omega^2 \varepsilon(r, \omega) E_z.\]

- \(r\): spatial variable
- \(\omega\): frequency
- \(\varepsilon\): relative permittivity

\(\omega\)-dependent material parameter!
Electronic bandstructure computations IV

- **Bloch ansatz:** \( E_z = e^{ik \cdot r} u(r) \)

\[-(\nabla + ik) \cdot (\nabla + ik) u(r) = \omega^2 \varepsilon(r, \omega) u(r)\]

- **k:** wave vector
- **u:** periodic function on lattice

- **assumption:** lattice consists of 2 materials, one of which is air

\[\Omega = \Omega_1 \cup \Omega_2, \quad \varepsilon(r, \omega) = \begin{cases} 
\varepsilon_1 = 1, & r \in \Omega_1 \\
\varepsilon_2(\omega), & r \in \Omega_2 
\end{cases}\]

- **discretization** using discontinuous Galerkin with \( p \)-enhancement (Engström & Wang, 2010)

\[\left[ G - \omega^2 M_1 - \omega^2 \varepsilon_2(\omega) M_2 \right] u = 0\]
Electronic bandstructure computations V

- nonlinearity caused by $\omega$-dependency of $\varepsilon_2$
- popular model for permittivity: Lorentz model

$$\varepsilon_2(\omega) = \alpha + \sum_{k=1}^{K} \frac{\xi_k}{\eta_k - \omega^2 - i\gamma_k \omega} \rightsquigarrow \text{rational EVP}$$
Definition

Let $T : D \to \mathbb{C}^{n \times n}$ be holomorphic. We call

- $\rho(T) := \{\lambda \in D : T(\lambda) \text{ is invertible}\}$ the resolvent set of $T$,
- $\sigma(T) := D \setminus \rho(T)$ the spectrum of $T$.

Theorem

Either the resolvent set of $T$ is empty or the spectrum of $T$ consists of isolated eigenvalues.

The number of eigenvalues in $\sigma(T)$ may be infinite!
Newton-based methods

Two possibilities to apply Newton’s method:

1. \( \det T(\lambda) = 0 \) (scalar equation in \( \mathbb{C} \))
2. \( T(\lambda)x = 0 \) (vector equation in \( \mathbb{C}^n \))

We focus on the second case.

- \( x, \lambda \mapsto n + 1 \) unknowns
- \( T(\lambda)x = 0 \mapsto n \) constraints

We have to add one additional constraint.

\[ v^H x = 1 \]
Nonlinear inverse iteration

\[ F(x, \lambda) := \begin{bmatrix} T(\lambda)x \\ v^H x - 1 \end{bmatrix} \]

Application of Newton’s method

\[
0 \overset{!}{=} F(x, \lambda) + DF(x, \lambda) \begin{bmatrix} \hat{x} - x \\ \hat{\lambda} - \lambda \end{bmatrix} = \begin{bmatrix} T(\lambda)\hat{x} + (\hat{\lambda} - \lambda)T'(\lambda)x \\ v^H \hat{x} - 1 \end{bmatrix}
\]

yields

\[
\hat{x} = - (\hat{\lambda} - \lambda) T(\lambda)^{-1} T'(\lambda)x
\]

\[
\hat{\lambda} = \lambda - \frac{1}{v^H T(\lambda)^{-1} T'(\lambda)x}.
\]
Algorithm

**Input:** approximate eigenpair \((x_0, \lambda_0)\) with \(v^H x_0 = 1\)

for \(j = 0, 1, \ldots\) until convergence do

solve \(T(\lambda_j) \tilde{x}_{j+1} = T'(\lambda_j) x_j\) for \(\tilde{x}_{j+1}\)

update \(\lambda_{j+1} := \lambda_j - \frac{v^H x_j}{v^H \tilde{x}_{j+1}}\)

normalize \(x_{j+1} := \frac{1}{v^H u_{j+1}} \tilde{x}_{j+1}\)

end for

- local quadratic convergence to simple eigenvalues
- main computational work lies in the solution of the linear system
Reducing the computational work

Replacing $T(\lambda_j)^{-1}$ by $T(\sigma)^{-1}$ for a fixed $\sigma$ leads to misconvergence!

But ...

\[
\hat{x} = - (\hat{\lambda} - \lambda) T(\lambda)^{-1} T'(\lambda)x \\
= x - T(\lambda)^{-1} T(\lambda)x - (\hat{\lambda} - \lambda) T(\lambda)^{-1} T'(\lambda)x \\
= x - T(\lambda)^{-1} [T(\lambda) + (\hat{\lambda} - \lambda) T'(\lambda)]x \\
= x - T(\lambda)^{-1} T(\hat{\lambda})x + O(|\hat{\lambda} - \lambda|^2)
\]

- Now $T(\sigma)^{-1}$ can be used in place of $T(\lambda)^{-1}$ safely.
- Even inexact solution of the linear system is possible.
Residual inverse iteration

**Algorithm**

**Input:** approximate eigenpair \((x_0, \sigma)\)

for \(j = 0, 1, \ldots\) until convergence do

solve \(v^H T(\sigma)^{-1} T(\lambda_{j+1}) x_j = 0\) for \(\lambda_{j+1}\)

solve \(T(\sigma) \Delta x = -T(\lambda_{j+1}) x_j\) for \(\Delta x\)

update \(\tilde{x}_{j+1} := x_j + \Delta x\)

normalize \(x_{j+1} := \frac{1}{v^H \tilde{x}_{j+1}} \tilde{x}_{j+1}\)

end for

- local linear convergence to simple eigenvalues
- convergence speed dependent on distance from \(\sigma\) to eigenvalue
Convergence

- string with elastically attached mass
- smallest magnitude eigenvalue using residual inverse iteration
- initial guess: $\sigma = 0$, $x$ discrete version of function $u(z) = z$

\begin{table}
\begin{tabular}{|l|c|}
\hline
it. & $\lambda$ \\
\hline
1 & 0.750000000000000 \\
2 & 0.6119289133137 \\
3 & 0.4992160868485 \\
4 & 0.4602962607221 \\
5 & 0.4573333718955 \\
6 & 0.4573184889546 \\
7 & 0.4573184889542 \\
8 & 0.4573184889542 \\
9 & 0.4573184889542 \\
\hline
\end{tabular}
\end{table}
Difficulties

- no global convergence
- can only handle one eigenvalue at a time
- cannot handle multiple eigenvalues
- consecutive runs may converge to the same eigenvalue

Linear eigensolvers (ARPACK, Jacobi-Davidson) exploit linear independence of eigenvectors to prevent reconvergence.

Example (loss of linear independence for NLEVPs)

The eigenvalues 1 and 2 of

\[
\begin{pmatrix}
1 & 1 \\
2 & 0
\end{pmatrix} + \lambda \begin{pmatrix}
-1 & -2 \\
0 & -3
\end{pmatrix} + \lambda^2 \begin{pmatrix}
1 & 0 \\
0 & 1
\end{pmatrix}
\]

share the same eigenvector \(\begin{pmatrix} 1 \\ 1 \end{pmatrix}\).
Several eigenvalues in the linear case

\[ Ax = \lambda x \]

Let \((x_1, \lambda_1), (x_2, \lambda_2)\) be two eigenpairs:

\[ Ax_1 = \lambda_1 x_1, \]
\[ Ax_2 = \lambda_2 x_2. \]

The above equations can be merged:

\[ A \begin{bmatrix} x_1 & x_2 \end{bmatrix} = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} =: X \begin{bmatrix} \Lambda \end{bmatrix} \]

Hence, we have

\[ AX = X \Lambda. \]

\[ \leadsto \text{X spans an invariant subspace of } A. \]
Generalization to the nonlinear case

\[ f_1(\lambda)A_1x + \cdots + f_r(\lambda)A_rx = 0 \]

Assume \( r = 2 \) and let \( (x_1, \lambda_1), (x_2, \lambda_2) \) be two eigenpairs.

\[
A_1x_1f_1(\lambda_1) + A_2x_1f_2(\lambda_1) = 0, \\
A_1x_2f_1(\lambda_2) + A_2x_2f_2(\lambda_2) = 0.
\]

These equations can again be merged

\[
A_1 \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} f_1(\lambda_1) \\ f_1(\lambda_2) \end{bmatrix} + A_2 \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} f_2(\lambda_1) \\ f_2(\lambda_2) \end{bmatrix} = 0
\]

With \( X = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \) and \( \Lambda = \text{diag} (\lambda_1, \lambda_2) \) as before

\[
A_1Xf_1(\Lambda) + A_2Xf_2(\Lambda) = 0.
\]
Invariant pairs

Definition

\((X, \Lambda) \in \mathbb{C}^{n \times m} \times \mathbb{C}^{m \times m}\) is called an invariant pair if

\[ A_1Xf_1(\Lambda) + \cdots + A_rXf_r(\Lambda) = 0. \]

- need to exclude degenerate situations, such as \(X = 0\)
- linear case: require \(X\) to have full column rank
- not suitable for the nonlinear case:

\[
X = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad \Lambda = \begin{bmatrix} 1 \\ 2 \end{bmatrix}
\]

is a perfectly reasonable invariant pair of

\[
\left( \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix} - \lambda \begin{bmatrix} -1 & -2 \\ 0 & -3 \end{bmatrix} + \lambda^2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right)x = 0.
\]
An invariant pair \((X, \Lambda)\) is called **minimal** if
\[
\begin{bmatrix}
X \\
X \Lambda \\
\vdots \\
X \Lambda^{\ell-1}
\end{bmatrix}
\]
has full column rank for some integer \(\ell\).

The smallest such \(\ell\) is called the **minimality index** of \((X, \Lambda)\).

\[
X = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad \Lambda = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad \begin{array}{c}
\leadsto \\
\implies
\end{array} \\
\begin{bmatrix}
X \\
X \Lambda
\end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 2 \\ 1 & 2 \end{bmatrix}
\]
Minimal invariant pairs

\( (X, \Lambda) \) is a minimal invariant pair with minimality index 2.

**Theorem**

Let \((X, \Lambda)\) be a minimal invariant pair of a nonlinear EVP

\[ T(\lambda)x = 0. \]

Then every eigenvalue of \( \Lambda \) is an eigenvalue of \( T \).
Computation of invariant pairs

Apply Newton’s method to

\[ T(X, \Lambda) := A_1 X f_1(\Lambda) + \cdots + A_r X f_r(\Lambda) = 0. \]

- \( X, \Lambda \leadsto n \cdot m + m^2 \) unknowns
- \( T(X, \Lambda) = 0 \leadsto n \cdot m \) constraints

We have to impose \( m^2 \) additional constraints.

\[ N(X, \Lambda) := V^H \begin{bmatrix} X & X \Lambda & \cdots & X \Lambda^{\ell-1} \end{bmatrix} - I = 0, \quad V \text{ suitably chosen} \]

**Theorem**

Let \((X, \Lambda)\) be a minimal invariant pair of \( T \). The Fréchet derivative \( DF \) of \( F := [T_N] \) at \((X, \Lambda)\) is invertible iff the multiplicities of \( \Lambda \)'s eigenvalues match those of \( T \).
Ongoing work: Continuation of invariant pairs

Joint project with Daniel Kressner (SAM, ETHZ) and Wolf-Jürgen Beyn (University of Bielefeld)

Parameter-dependent nonlinear EVP

\[ T(\lambda, s)x = 0 \]

- Goal: track several eigenvalues as the parameter \( s \) varies
- Idea:
  - use invariant pair \((X(s), \Lambda(s))\) at parameter value \( s \) as an initial guess for invariant pair at \( s + \Delta s \)
  - apply Newton to obtain invariant pair \((X(s + \Delta s), \Lambda(s + \Delta s))\) at parameter value \( s + \Delta s \)
  - repeat
Part II:
Solution of polynomial / rational EVPs
Linearization

Polynomial EVPs can be solved by transforming them into linear ones.

- Example: \((\lambda^2 M + \lambda C + K)x = 0\)
- introduce auxiliary variable \(y = \lambda x\)
- rewrite as \(\lambda My + Cy + Kx = 0\)

This can be written as a generalized linear EVP with the same eigenvalues:

\[
\begin{bmatrix}
- C & -K \\
I & 0
\end{bmatrix}
\begin{bmatrix}
y \\
x
\end{bmatrix}
= \lambda
\begin{bmatrix}
M & 0 \\
0 & I
\end{bmatrix}
\begin{bmatrix}
y \\
x
\end{bmatrix}.
\]

- linearizations not unique
- could also rewrite as \(\lambda My + \lambda Cx + Kx = 0\)

\[
\begin{bmatrix}
K & 0 \\
0 & I
\end{bmatrix}
\begin{bmatrix}
x \\
y
\end{bmatrix}
= \lambda
\begin{bmatrix}
- C & -M \\
I & 0
\end{bmatrix}
\begin{bmatrix}
x \\
y
\end{bmatrix}.
\]
There are entire vector spaces of linearizations (Mackey et al., 2006)!

- choice of linearization depends on underlying application
- structure preservation of interest

**Example**

The eigenvalues of an even polynomial eigenvalue problem

\[(\lambda^2 M + \lambda C + K)x = 0, \quad M = M^T, \; C = -C^T, \; K = K^T\]

occur in pairs \((\lambda, -\lambda)\).

- If \(M\) is invertible, the symmetric / skew-symmetric linearization

  \[
  \begin{bmatrix}
  K & 0 \\
  0 & M
  \end{bmatrix}
  \begin{bmatrix}
  x \\
  y
  \end{bmatrix}
  = \lambda
  \begin{bmatrix}
  -C & -M \\
  M & 0
  \end{bmatrix}
  \begin{bmatrix}
  x \\
  y
  \end{bmatrix}
  \]

  preserves this property.
Linearization III

- linearization of polynomial EVPs of degree $d > 2$ analogously by introducing the auxiliary vector

$$\begin{bmatrix}
x \\
\lambda x \\
\vdots \\
\lambda^{d-1} x
\end{bmatrix}$$

Conclusion 1:
- polynomial EVP of size $n$ $\xrightarrow{}$ linearized EVP of size $dn$
- polynomial EVP has exactly $dn$ eigenvalues (counting multiplicities)

Conclusion 2:
- The extended vectors above will always be linearly independent, even though the eigenvectors $x$ themselves may not.
Solution of rational EVPs

- most straightforward way: multiplication by the common denominator of all rational terms \(\leadsto\) polynomial EVP
- degree of resulting polynomial EVP may be very large
- sometimes smarter ways to solve rational EVPs:

Ongoing joint project with Daniel Kressner and Christian Engström (both SAM, ETHZ)

\[
\begin{align*}
G - \omega^2 M_1 - \omega^2 \left( \alpha + \frac{\xi}{\eta - \omega^2 - i\gamma\omega} \right) M_2 \right) u &= 0 \\
M_1 + M_2 &\text{ is a splitting of the total mass matrix of the problem.}
\end{align*}
\]
Rational EVPs in electronic bandstructure computations

- by suitable ordering of nodes:

\[
M_2 = \begin{bmatrix} 0 & \tilde{M}_2 \end{bmatrix} \in \mathbb{R}^{n \times n}
\]

with a positive definite \( \tilde{M}_2 \in \mathbb{R}^{m \times m}, m \ll n \)

- by Cholesky decomposition:

\[
M_2 = F^T F, \quad F \in \mathbb{R}^{m \times n}
\]

Idea (Bai & Su, 2008):
- write rational term(s) in transfer function form

\[
\frac{\omega^2 \xi}{\eta - \omega^2 - i\gamma \omega} = -\xi + b^T (A - \omega E)^{-1} b
\]
Conversion to transfer function form

- by partial fraction expansion

\[
\frac{p(\omega)}{q(\omega)} = \zeta + \frac{\sigma_1}{\omega - \rho_1} + \cdots + \frac{\sigma_k}{\omega - \rho_k}
\]

\[
= \zeta + \left(\frac{\omega}{\sigma_1} - \frac{\rho_1}{\sigma_1}\right)^{-1} + \cdots + \left(\frac{\omega}{\sigma_k} - \frac{\rho_k}{\sigma_k}\right)^{-1}
\]

\[
= \zeta + b^T (A - \omega E)^{-1} b,
\]

where

\[
b = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}, \quad A = \begin{bmatrix} -\frac{\rho_1}{\sigma_1} & \cdots & \cdots \\ & -\frac{\rho_k}{\sigma_k} \\ & & -\frac{1}{\sigma_k} \end{bmatrix}, \quad E = \begin{bmatrix} -\frac{1}{\sigma_1} & \cdots & \cdots \\ & \cdots & \cdots \\ & & -\frac{1}{\sigma_k} \end{bmatrix}
\]
Linearization of rational EVPs

- EVP with rational term in transfer function form:

\[
\hat{G} u - b^T (A - \omega E)^{-1} b F^T F u = \omega^2 \hat{M} u,
\]

where \( \hat{G} = G + \xi M_2 \), \( \hat{M} = M_1 + \alpha M_2 \)

- rearrange:

\[
b^T (A - \omega E) b F^T F = b^T (A - \omega E)^{-1} b \otimes F^T F
\]
\[
= B^T (A - \omega E)^{-1} B
\]

where \( B = b \otimes F \), \( A = A \otimes I \), \( E = E \otimes I \)
Linearization of rational EVPs II

\[ \hat{G} u - \mathcal{B}^T (A - \omega \mathcal{E})^{-1} \mathcal{B} u = \omega^2 \hat{M} u \]

Introducing the auxiliary variables \( v := -(A - \omega \mathcal{E})^{-1} \mathcal{B} u, \quad w := \omega u \), we obtain the linearized EVP

\[
\begin{bmatrix}
\hat{G} & \mathcal{B}^T \\
\mathcal{B} & A
\end{bmatrix}
\begin{bmatrix}
u \\
w
\end{bmatrix}
= \omega
\begin{bmatrix}
\hat{M} \\
\mathcal{E}
\end{bmatrix}
\begin{bmatrix}
u \\
w
\end{bmatrix}
\]

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>( p )</td>
<td>#DoF</td>
<td>size</td>
</tr>
<tr>
<td>2</td>
<td>288</td>
<td>1728</td>
</tr>
<tr>
<td>4</td>
<td>720</td>
<td>4320</td>
</tr>
<tr>
<td>6</td>
<td>1344</td>
<td>8064</td>
</tr>
<tr>
<td>8</td>
<td>2160</td>
<td>12960</td>
</tr>
</tbody>
</table>
introduced nonlinear EVPs
showed several sample applications
discussed Newton-based technique for determining one or several eigenpairs
demonstrated linearization techniques for polynomial and rational EVPs
outlined two ongoing projects