Linearizations for Interpolation Bases

Piers W. Lawrence[1] Rob M. Corless[2]

[1]Department of Computer Science, KU Leuven

[2]Department of Applied Mathematics, The University of Western Ontario

Nonlinear Eigenvalue Problems, 24 April 2014

Outline

Introduction: Polynomial eigenproblems in interpolation bases

- We are interested in solving the polynomial eigenvalue problem

$$P(\lambda)x = \left(\sum_{i=0}^{n} F_i \phi_i(\lambda)\right) x = 0 \,,$$

  where $F_k \in \mathbb{C}^{m\times m}$, $\deg P(\lambda) = n$, and the $\phi_i(\lambda)$ form a basis for polynomials of degree at most n.
- In Lagrange form, the basis elements are $\phi_k(\lambda) = \ell_k(\lambda)$, defined for a set of $n+1$ distinct interpolation nodes $\{\sigma_0, \cdots, \sigma_n\}$ by

$$\ell_i(\lambda) = \frac{\prod_{k\neq i}(\lambda - \sigma_k)}{\prod_{k\neq i}(\sigma_i - \sigma_k)} \,.$$

Introduction: Barycentric Lagrange interpolation

- The Lagrange interpolation formula

$$P(\lambda) = \sum_{i=0}^{n} F_i \ell_i(\lambda)$$

- First form of the barycentric Lagrange formula

$$P(\lambda) = \prod_{i=0}^{n} (\lambda - \sigma_i) \sum_{j=0}^{n} \frac{w_j}{(\lambda - \sigma_j)} F_j \ ,$$

where the barycentric weights w_j are computed as

$$w_j^{-1} = \prod_{\substack{k=0 \\ k \neq j}}^{n} (\sigma_j - \sigma_k)$$

Introduction: Linearization of PEP's

- Basic concept:

$$P(\lambda) \Rightarrow L(\lambda) = \lambda B - A\,,$$

  turn a degree n polynomial eigenvalue problem into a linear one of larger dimension (usually of dimension mn).
- The linearization should be a strong linearization of the polynomial matrix: $L(\lambda)$ is a strong linearization of $P(\lambda)$ if it is a linearization for $P(\lambda)$ and $\mathrm{rev}L(\lambda)$ is a linearization for $\mathrm{rev}P(\lambda)$
- The left and right eigenvectors of $P(\lambda)$ should be easily recoverable from the eigenvectors of the linearization
- Infinite eigenvalues (where they exist) should be easily detectable

Introduction: Linearizations of PEP's

- Companion linearization for $\phi_i(\lambda) = \lambda^i$

$$\mathcal{M}(\lambda) = \begin{bmatrix} \lambda A_n - A_{n-1} & -A_{n-2} & \cdots & A_0 \\ I_m & \lambda I_m & & \\ & \ddots & \ddots & \\ & & I_m & \lambda I_m \end{bmatrix}$$

- Arrowhead linearization of Corless [1] $\phi_k(\lambda) = \ell_k(\lambda)$

$$\mathcal{L}(\lambda) = \lambda B - A = \begin{bmatrix} 0 & F_0 & \cdots & F_n \\ -w_0 I_m & (\lambda - \sigma_0) I_m & & \\ \vdots & & \ddots & \\ -w_n I_m & & & (\lambda - \sigma_n) I_m \end{bmatrix}$$

[1] R. M. Corless, *Generalized companion matrices in the Lagrange basis*, in Proceedings EACA, L. Gonzalez-Vega and T. Recio, eds., June 2004, pp. 317–322.

Introduction: Linearization of PEP's in the Lagrange basis

- $\mathcal{L}(\lambda)$ is a strong linearization of $\widehat{P}(\lambda) = 0 \cdot \lambda^{n+2} + 0 \cdot \lambda^{n-1} + P(\lambda)$
- The formulation of the linearization introduces $2m$ spurious infinite eigenvalues
- We will be disposed of these eigenvalues later
- A right eigenvector x of $P(\lambda)$ is related to a right eigenvector z of $\mathcal{L}(\lambda)$ via

$$\begin{bmatrix} 0 & F_0 & \cdots & F_n \\ -w_0 I_m & (\lambda - \sigma_0) I_m & & \\ \vdots & & \ddots & \\ -w_n I_m & & & (\lambda - \sigma_n) I_m \end{bmatrix} \begin{bmatrix} \ell(\lambda)x \\ \ell_0(\lambda)x \\ \vdots \\ \ell_n(\lambda)x \end{bmatrix} = \begin{bmatrix} P(\lambda)x \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Zero leading coefficients

- Lagrange basis polynomials $\phi_k(\lambda) = \ell_k(\lambda)$ are not degree graded
- How do we answer the simple question: *What is the degree?*
- Leading coefficient of $P(\lambda)$ is

$$[\lambda^n](P(\lambda)) = \sum_{i=0}^{n} w_i F_i$$

- If the leading coefficient is singular, there are additional infinite eigenvalues to be deflated

Zero Leading Coefficients (scalar case)

Theorem

If the leading coefficients $\left[\lambda^{n-j}\right](p(\lambda)) = 0$ *for all* j, $0 \le j \le r-1$, *then* $\left[\lambda^{n-r}\right](p(\lambda)) = \mathbf{f}^T \mathbf{\Sigma}^r \mathbf{w}$. *That is, if all of the first* r *leading coefficients of* $p(\lambda)$ *are zero, then the* $(r+1)$*st leading coefficient is* $\mathbf{f}^T \mathbf{\Sigma}^r \mathbf{w}$.

Proof.

If the first r leading coefficients of $p(\lambda)$ are zero, then we can remove up to r interpolation nodes, and update the barycentric formula. Let $K_j = \{k_1, \ldots, k_j\}$ be a set of j unique integers, and $\{\sigma_k | k \in K_j\}$ be the set of nodes to be removed. The updated barycentric formula is

$$\prod_{\substack{i=0 \\ i \notin K_j}}^{n} (\lambda - \sigma_i) \sum_{\ell=0}^{n} \left(\prod_{k \in K_j} (\sigma_\ell - \sigma_k) \frac{w_\ell f_\ell}{\lambda - \sigma_\ell} \right).$$

Zero Leading Coefficients (scalar case)

Theorem

If the leading coefficients $\left[\lambda^{n-j}\right](p(\lambda)) = 0$ *for all* j, $0 \le j \le r-1$, *then* $\left[\lambda^{n-r}\right](p(\lambda)) = \mathbf{f}^T \mathbf{\Sigma}^r \mathbf{w}$. *That is, if all of the first* r *leading coefficients of* $p(\lambda)$ *are zero, then the* $(r+1)$*st leading coefficient is* $\mathbf{f}^T \mathbf{\Sigma}^r \mathbf{w}$.

Proof (Continued).

The updated barycentric formula has leading coefficient

$$[\lambda^{n-j}](p(\lambda)) = \sum_{\ell=0}^{n} q_j(\sigma_\ell) w_\ell f_\ell = \mathbf{f}^T q_j(\mathbf{\Sigma}) \mathbf{w}\ ,$$

where $q_j(\lambda) = \prod_{k \in K_j} (\lambda - \sigma_k)$. An induction on r shows that

$$\left[\lambda^{n-r}\right](p(\lambda)) = \mathbf{f}^T \mathbf{\Sigma}^r \mathbf{w}\ .$$

□

Zero leading coefficients (matrix case)

- The previous theorem holds for each scalar entry of $P(\lambda)$, and thus if the first r leading coefficients of the polynomial are zero then the $(r+1)$st leading coefficient is given by

$$
\begin{aligned}
[\lambda^{n-r}](P(\lambda)) &= F^T(\mathbf{\Sigma}^r \mathbf{w} \otimes I_m) \\
&= \sum_{i=0}^{n} F_j \sigma_j^r w_j
\end{aligned}
$$

- Reduction process to Hessenberg form produces an orthonormal basis for

$$\mathcal{K}_n(\mathbf{\Sigma}, \mathbf{w})$$

Spurious infinite eigenvalues

- The linearization $\mathcal{L}(\lambda)$ has dimension $(n+2)m$
- $\mathcal{L}(\lambda)$ linearizes $\widehat{P}(\lambda) = 0 \cdot \lambda^{n+2} + 0 \cdot \lambda^{n-1} + P(\lambda)$
- If we do not deflate these spurious infinite eigenvalues initially, they may never be detected correctly
- An initial reduction to block Hessenberg allows us to deflate these exactly
- Simultaneously transform the scalar entries of $P(\lambda)$ to a degree graded basis, still based on the nodes
- We can detect other infinite eigenvalues at the same time

Reduction to block Hessenberg form

- We first construct a unitary Q such that

$$Q^*\mathbf{w} = \alpha\mathbf{e}_1, \quad Q\mathbf{\Sigma} = HQ\ ,$$

  with H upper Hessenberg.
- The process generates an orthonormal basis for

$$\mathcal{K}_n(\mathbf{\Sigma}, \mathbf{w})\ .$$

  Zero leading coefficients will also be detected.
- Generate a block upper Hessenberg matrix by applying this blockwise to the linearization $\mathcal{L}(\lambda)$

$$(Q^* \otimes I_m)\mathcal{L}(\lambda)(Q \otimes I_m)$$

Deflation of spurious infinite eigenvalues

- The block reduction has the form

$$\begin{bmatrix} 0 & \widehat{F}_0 & \widehat{F}_1 & \cdots & \widehat{F}_n \\ \alpha I_m & (\lambda - h_{0,0})I_m & \cdots & \cdots & -h_{0,n}I_m \\ & -h_{1,0}I_m & \ddots & & \vdots \\ & & \ddots & \ddots & \\ & & & -h_{n,n-1}I_m & (\lambda - h_{n,n})I_m \end{bmatrix}$$

- Since the top leftmost block is identically zero, we may always perform a block permutation on the first two block rows to deflate the $2m$ spurious infinite eigenvalues
- If the leading coefficient of $P(\lambda)$ is zero, then we have other additional eigenvalues at infinity. We take the SVD of the leading block then apply permutations to deflate these additional eigenvalues.

Deflation of spurious infinite eigenvalues

- The top left block ensures that the permutation introduces a further m zeros on the diagonal of B
- If a number of the leading coefficients are zero, then these immediately appear in the first block row.
- We can then transform the generalized eigenvalue problem to a standard one of dimension $mn \times mn$

One sided factorizations

- We look for left and right sided factorizations linking $\mathcal{L}(\lambda)$ to $P(\lambda)$
- $G(\lambda)\mathcal{L}(\lambda) = e_1^T \otimes P(\lambda)$, where

$$G(\lambda) = \begin{bmatrix} \ell(\lambda)I_m & -F_0\dfrac{\ell(\lambda)}{\lambda - \sigma_0} & \cdots & -F_n\dfrac{\ell(\lambda)}{\lambda - \sigma_n} \end{bmatrix}$$

- $\mathcal{L}(\lambda)H(\lambda) = e_1 \otimes P(\lambda)$, where

$$H(\lambda) = \begin{bmatrix} \ell(\lambda)I_m \\ \ell_0(\lambda)I_m \\ \vdots \\ \ell_n(\lambda)I_m \end{bmatrix}$$

Backward errors of PEP's solved by linearization

- Define

$$P(\lambda) + \Delta P(\lambda) = \sum_{j=0}^{n} (C + \Delta C_j)\phi_j(\lambda)$$

- Backward error for an approximate right eigenpair (λ, x) is given by

$$\eta_P(\lambda, x) = \min\{\varepsilon : (P(\lambda) + \Delta P(\lambda))x = 0, \\ \|\Delta C_j\|_2 \leq \varepsilon \|C_j\|_2,\ 0 \leq j \leq n\}$$

- Backward error for an approximate left eigenpair (λ, y^*) is given by

$$\eta_P(\lambda, y^*) = \min\{\varepsilon : y^*(P(\lambda) + \Delta P(\lambda)) = 0, \\ \|\Delta C_j\|_2 \leq \varepsilon \|C_j\|_2,\ 0 \leq j \leq n\}$$

Backward error relationships

- Backward error of a right eigenpair (λ, z) of $\mathcal{L}(\lambda)$:

$$\frac{\eta_P(\lambda, x)}{\eta_{\mathcal{L}}(\lambda, z)} \leq \frac{|\lambda|\|B\|_2 + \|A\|_2}{\sum_{i=0}^{n} \|F_i\|_2 |\ell_i(\lambda)|} \cdot \frac{\|G(\lambda)\|_2 \|z\|_2}{\|x\|_2} .$$

- Backward error of a left eigenpair (λ, w^H) of $\mathcal{L}(\lambda)$:

$$\frac{\eta_P(\lambda, y^*)}{\eta_{\mathcal{L}}(\lambda, w^*)} \leq \frac{|\lambda|\|B\|_2 + \|A\|_2}{\sum_{i=0}^{n} \|F_i\|_2 |\ell_i(\lambda)|} \cdot \frac{\|H(\lambda)\|_2 \|w\|_2}{\|y\|_2} .$$

Backward Error Bound

Theorem

We may bound the ratio of the backward error of an approximate right eigenpair $(\lambda, \mathbf{x})$ *of* $P(z)$ *to the backward error of an approximate right eigenpair* $(\lambda, \mathbf{z})$ *of* $(\mathbf{A}, \mathbf{B})$ *by*

$$\frac{\eta_P(\lambda, \mathbf{x})}{\eta_{(\mathbf{A},\mathbf{B})}(\lambda, \mathbf{z})} \leq \sqrt{m}(|\lambda|+1) \max\left(1, \|\mathbf{A}\|_2\right) \left(\frac{|\ell(\lambda)|}{\min_j \|\mathbf{F}_j\|_2} + \frac{\sqrt{m}}{\min_k |w_k|} \right) \cdot \frac{\|\mathbf{z}\|_2}{\|\mathbf{x}\|_2}.$$

Similarly, for an approximate left eigenpair,

$$\frac{\eta_P(\lambda, \mathbf{y}^*)}{\eta_{(\mathbf{A},\mathbf{B})}(\lambda, \mathbf{w}^*)} \leq (|\lambda|+1) \max\left(1, \|\mathbf{A}\|_2\right) \frac{(|\ell(\lambda)|+1)}{\min_j \|\mathbf{F}_j\|_2} \cdot \frac{\|\mathbf{w}\|_2}{\|\mathbf{y}\|_2}.$$

Example: Damped Gyroscopic System

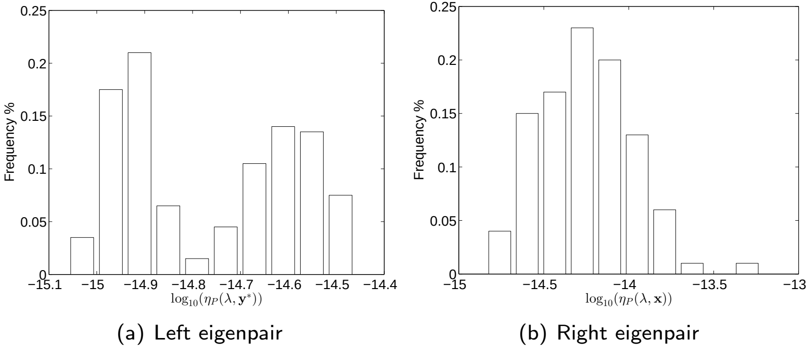

Figure: Damped gyroscopic system, backward error distributions.

Example: Damped Gyroscopic System

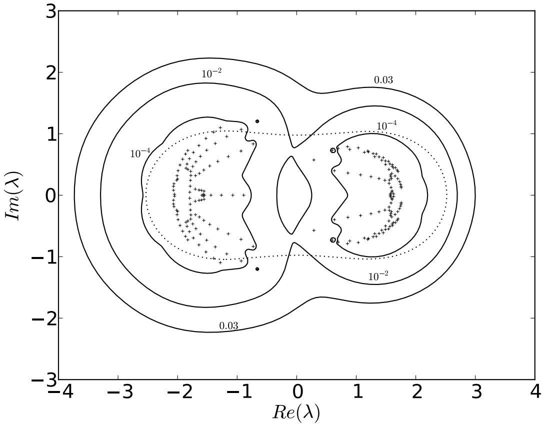

Figure: Damped gyroscopic system, distributions of eigenvalues and pseudospectra. The dotted line represents the level curve where $B_M(z) = B_L(z)$.

Conclusions

- Linearizations for PEP's in the Lagrange basis
- Detection of zero leading coefficients
- Deflation of spurious and non-spurious infinite eigenvalues
- Bounds on the backward error

Conclusions

- Linearizations for PEP's in the Lagrange basis
- Detection of zero leading coefficients
- Deflation of spurious and non-spurious infinite eigenvalues
- Bounds on the backward error

Thank you!