Linearizations for Interpolation Bases

Piers Lawrence[1,2], Marc Van Barel[1], and Paul Van Dooren[2]

[1]Department of Computer Science, KU Leuven

[2] Institute of Information and Communication Technologies, Electronics and Applied Mathematics (ICTEAM), UCLouvain

Matrix Polynomials

In this work we consider matrix polynomials of the form

$$P(\lambda) = \sum_{k=0}^{d} P_k \phi_k(\lambda)\,,$$

where the P_k's are constant n by n matrices, and the set of $\{\phi_0(\lambda), \dots, \phi_d(\lambda)\}$ form a basis for polynomials of degree at most d.

Furthermore, we are interested in solving polynomial eigenvalue problems, that is, we compute a pair (λ, x) satisfying $P(\lambda)x = 0$, where $\lambda \in \mathbb{C}$, and $x \in \mathbb{C}^n \setminus \{0\}$.

Barycentric Lagrange Form

A matrix polynomial $P(\lambda)$ of degree at most d can be uniquely determined by $d+1$ samples $P_k = P(\sigma_k)$, at a distinct set of nodes

$$\{\ \sigma_0 \quad \cdots \quad \sigma_d\ \}\,.$$

The polynomial interpolant can be written in Lagrange form

$$P(\lambda) = \sum_{k=0}^{d} P_k \ell_k(\lambda)\,, \qquad (1)$$

where the $\ell_k(\lambda)$'s are the Lagrange basis polynomials

$$\ell_k(\lambda) = \prod_{\substack{j=0\\ j\neq k}}^{d} \frac{(\lambda-\sigma_j)}{(\sigma_k-\sigma_j)}\,.$$

The barycentric formula, or modified Lagrange formula [1] is obtained from (1) by first defining the so called barycentric weights

$$w_k^{-1} = \prod_{\substack{j=0\\ j\neq k}}^{d} (\sigma_k - \sigma_j)\,,$$

each of the $\ell_k(\lambda)$'s can then be written as

$$\ell_k(\lambda) = \ell(\lambda)\frac{w_k}{(\lambda-\sigma_k)}\,,$$

where $\ell(\lambda) = \prod_{k=0}^{d}(\lambda-\sigma_k)$ is known as the node polynomial. We then take out the common factor of $\ell(\lambda)$ in the Lagrange formula (1), and this gives the barycentric form of the Lagrange formula

$$P(\lambda) = \ell(\lambda)\sum_{k=0}^{d}\frac{w_k}{(\lambda-\sigma_k)}P_k\,. \qquad (2)$$

Linearization of Polynomials

Linearization replaces the polynomial eigenvalue problem for $P(\lambda)$ by a larger linear eigenvalue problem $\mathcal{L}(\lambda)$. The linearization should, of course, have the following desirable properties [3]:

- the linearization $\mathcal{L}(\lambda)$ is immediately constructable from the data in $P(\lambda)$,
- eigenvectors of $P(\lambda)$ can be easily recovered from eigenvectors of $\mathcal{L}(\lambda)$,
- $\mathcal{L}(\lambda)$ is a strong linearization of $P(\lambda)$, partial multiplicities of the finite and infinite eigenvalues are preserved.

Barycentric Linearization

First proposed in [2], the following $(d+2)n$ by $(d+2)n$ linearization $\mathcal{L}(\lambda) = \lambda B - A$ can be formed directly from the barycentric formula (2)

$$\mathcal{L}(\lambda) = \begin{bmatrix} 0 & P_0 & \cdots & P_d \\ -w_0 I & (\lambda-\sigma_0)I & & \\ \vdots & & \ddots & \\ -w_d I & & & (\lambda-\sigma_d)I \end{bmatrix}. \qquad (3)$$

This linearization is a strong linearization of the matrix polynomial

$$\widehat{P}(\lambda) = 0\cdot\lambda^{d+2} + 0\cdot\lambda^{d+1} + P(\lambda)\,,$$

and thus has an additional $2n$ eigenvalues at infinity.

Eigenvector Relations

To relate the left and right eigenvectors of $\mathcal{L}(\lambda)$ to those of $P(\lambda)$, we form left and right sided factorizations

$$G(\lambda)\mathcal{L}(\lambda) = g^T \otimes P(\lambda)\,,$$

$$\mathcal{L}(\lambda)H(\lambda) = h \otimes P(\lambda)\,,$$

where the matrix polynomials $G(\lambda)$ and $H(\lambda)$ are given by

$$G(\lambda) = \ell(\lambda)\begin{bmatrix} I & -\frac{P_0}{\lambda-\sigma_0} & \cdots & -\frac{P_d}{\lambda-\sigma_d}\end{bmatrix}, \qquad (4)$$

$$H(\lambda) = \ell(\lambda)\begin{bmatrix} I \\ \frac{w_0}{\lambda-\sigma_0}I \\ \vdots \\ \frac{w_d}{\lambda-\sigma_d}I \end{bmatrix}, \qquad (5)$$

and additionally, both g and h are equal to e_1, the first unit vector.

From these factorizations we may recover the left and right eigenvectors of $P(\lambda)$ from those of $\mathcal{L}(\lambda)$ as follows. Suppose that (λ, v) is a right eigenpair of $\mathcal{L}(\lambda)$ then

$$G(\lambda)\mathcal{L}(\lambda)v = (e_1^T \otimes P(\lambda))v = P(\lambda)(e_1^T \otimes I)v = P(\lambda)x\,,$$

where x is extracted from the first n rows of v. Similarly, suppose that (λ, u^H) is a left eigenpair of $\mathcal{L}(\lambda)$, then

$$u^H\mathcal{L}(\lambda)H(\lambda) = u^H(e_1 \otimes P(\lambda)) = y^H P(\lambda)\,,$$

where y is extracted from the first n rows of u.

Backward Errors

The normwise backward errors of finite approximate left and right eigenpairs (λ, y) and (λ, x), respectively, of a matrix polynomial $P(\lambda)$ are given by

$$\eta_P(\lambda,x) = \min\{\varepsilon : (P(\lambda)+\Delta P(\lambda))x = 0,\ \|\Delta P_j\| \le \varepsilon\|P_j\|,\ 0\le j\le d\}\,,$$

$$\eta_P(\lambda,y^H) = \min\{\varepsilon : y^H(P(\lambda)+\Delta P(\lambda)) = 0,\ \|\Delta P_j\| \le \varepsilon\|P_j\|,\ 0\le j\le d\}\,,$$

where the coefficients P_j of the matrix polynomial are perturbed by ΔP_j, and both are expressed in the Lagrange basis.

Backward Error Bounds

Theorem 1. *Let (λ, v) be an approximate right eigenpair of $\mathcal{L}(\lambda)$. Then for $x = (e_1^T \otimes I)v$, we have*

$$\frac{\eta_P(\lambda,x)}{\eta_{\mathcal{L}}(\lambda,v)} \le \frac{(|\lambda|\|B\|_2 + \|A\|_2)\|G(\lambda)\|_2}{\sum_{i=0}^{d}\|P_i\|_2|\ell_i(\lambda)|}\cdot\frac{\|v\|_2}{\|x\|_2}$$
$$\le \frac{(|\lambda| + \|A\|_2)\left(1+\sum_{i=0}^{d}\frac{\|P_i\|_2}{|\lambda-\sigma_i|}\right)|\ell(\lambda)|}{\sum_{i=0}^{d}\|P_i\|_2|\ell_i(\lambda)|}\cdot\frac{\|v\|_2}{\|x\|_2}\,.$$

Similarly, let (λ, u^H) be an approximate left eigenpair of $\mathcal{L}(\lambda)$. Then for $y^H = u^H(e_1\otimes I)$ we have

$$\frac{\eta_P(\lambda,y^H)}{\eta_{\mathcal{L}}(\lambda,u^H)} \le \frac{(|\lambda|\|B\|_2 + \|A\|_2)\|H(\lambda)\|_2}{\sum_{i=0}^{d}\|P_i\|_2|\ell_i(\lambda)|}\cdot\frac{\|u\|_2}{\|y\|_2}$$
$$\le \frac{(|\lambda| + \|A\|_2)\left(1+\sum_{i=0}^{d}\frac{|w_i|}{|\lambda-\sigma_i|}\right)|\ell(\lambda)|}{\sum_{i=0}^{d}\|P_i\|_2|\ell_i(\lambda)|}\cdot\frac{\|u\|_2}{\|y\|_2}\,.$$

Error Bound Consequences

We may interpret the above backward error bounds in the following way:

- The interpolation nodes should be chosen to be close to the eigenvalues of interest.
- The contribution of the conditioning of the basis defined by the nodes is significant:
  - well conditioned node distributions are essential for small backward errors.
- The norms of the coefficient matrices $\|P_i\|_2$ should have similar magnitudes.

Related Linearization

Recently, other linearizations in the Lagrange basis have been proposed, such as a $n(d+1)$ dimension linearization of $P(\lambda)$ proposed in [4]:

$$\widehat{\mathcal{L}}(\lambda) = \begin{bmatrix} P_0 & P_1 & \cdots & & P_d \\ (\lambda-\sigma_0)I & (\sigma_1-\lambda)\theta_1 I & & & \\ & \ddots & \ddots & & \\ & & & (\lambda-\sigma_{d-1})I & (\sigma_d-\lambda)\theta_d I \end{bmatrix}$$

where $\theta_i = w_{i-1}/w_i$.

The following transformation E_θ, applied to the right of $\mathcal{L}(\lambda)$ decouples n of the spurious infinite eigenvalues from the linearization (3)

$$E_\theta = \begin{bmatrix} 0 & -I & & & \\ I & 0 & & & \\ & I & -\theta_1 I & & \\ & & \ddots & \ddots & \\ & & & I & -\theta_d I \end{bmatrix}.$$

Applying this on the right of $\mathcal{L}(\lambda)$ yields

$$E_\theta\mathcal{L}(\lambda) = \begin{bmatrix} -w_0 I & (\lambda-\sigma_0)(e_1^T\otimes I) \\ 0 & \widehat{\mathcal{L}}(\lambda)\end{bmatrix}.$$

Thus, we may easily apply the backward error analysis to the reduced linearization $\widehat{\mathcal{L}}(\lambda)$. The right sided factorization is almost identical, and the left sided factorization can be computed from $G(\lambda)E_\theta^{-1}$.

Damped gyroscopic system [5]

The matrix polynomial is constructed as follows: let N denote the 10×10 nilpotent matrix having ones on the subdiagonal and zeros elsewhere, and define $\widehat{M} = (4I_{10} + N + N^T)/6$, $\widehat{G} = N - N^T$, and $\widehat{K} = N + N^T - 2I_{10}$. Then define the matrices M, G, and K, using the Kronecker product $\otimes$, by

$$M = I_{10}\otimes\widehat{M} + 1.3\widehat{M}\otimes I_{10}\,,$$
$$G = 1.35 I_{10}\otimes\widehat{G} + 1.1\widehat{G}\otimes I_{10}\,,$$
$$K = I_{10}\otimes\widehat{K} + 1.2\widehat{K}\otimes I_{10}\,,$$

and the damping matrix by $D = \text{tridiag}\{-0.1, 0.2, -0.1\}$. The quadratic matrix polynomial we examine is defined by

$$P(\lambda) = \lambda^2 M + \lambda(G+D) + K\,.$$

We interpolate $P(\lambda)$ at $\{\sigma_0,\sigma_1,\sigma_2\} = \{-1.8, 0, 1.8\}$, and compute the eigenvalues of $P(\lambda)$ from the linearization $\mathcal{L}(\lambda)$. Figure 1 shows the eigenvalues, the pseudospectra of $P(\lambda)$, and the contour where the conditioning of the monomial basis is equal to that of the Lagrange basis.

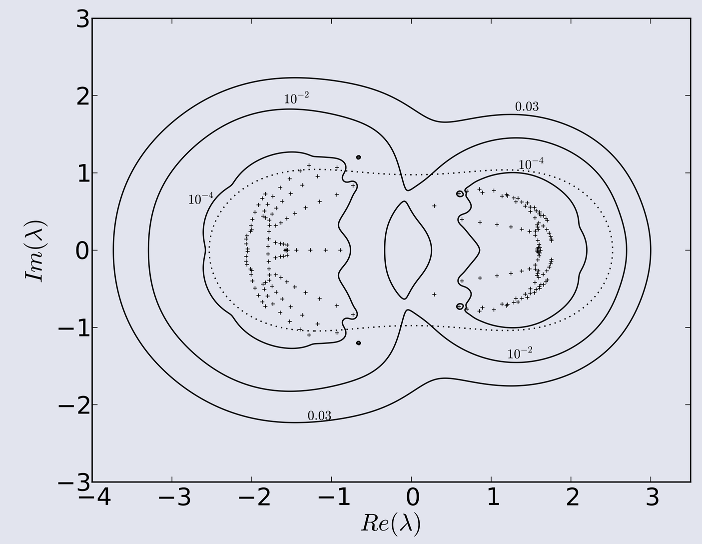

Figure 1: Eigenvalues and pseudospectra of a damped gyroscopic system.

Everywhere inside the dotted line represents where the Lagrange basis is better conditioned than the monomial basis. Figures 2 and 3 show the frequency of the obtained backward errors over all computed eigenpairs, as well as the pessimism index that indicates how many orders of magnitude larger the backward error bound is from the computed backward error.

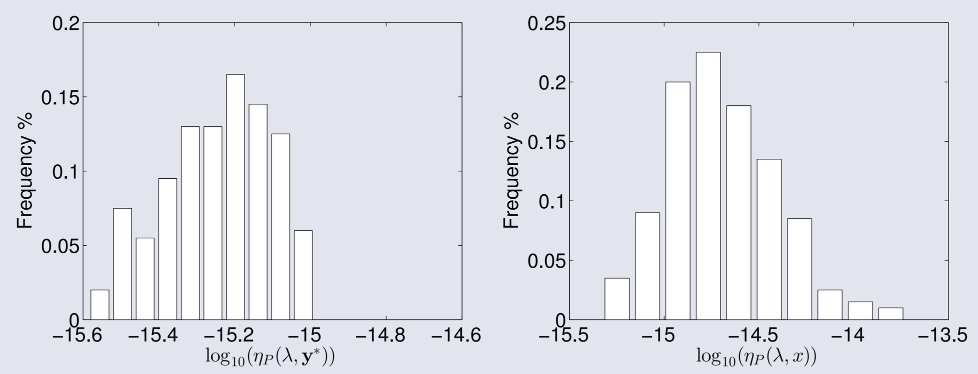

Figure 2: Damped gyroscopic system, backward error distributions for left and right eigenpairs.

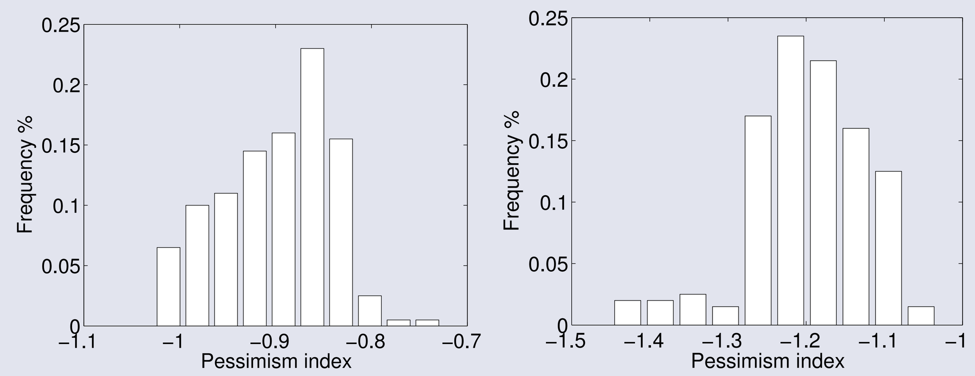

Figure 3: Damped gyroscopic system, pessimism index.

Speaker Enclosure [5]

We consider the quadratic

$$P(\lambda) = \lambda^2 M + \lambda C + K\,,$$

with $M, C, K \in \mathbb{C}^{107\times107}$, arising from a finite element model of a speaker enclosure. There is a large variation in the norms of the monomial basis coefficients:

$$\|M\|_2 = 1,\ \|C\|_2 = 5.7\times10^{-2},\ \|K\|_2 = 1\times10^{7}.$$

We interpolate the matrix polynomial at the nodes $\{\sigma_0,\sigma_1,\sigma_2\} = \{-i, 0, i\}$. At these nodes $\|P_j\|_2 \approx 1\times10^7$, and so we have already, in a sense, equalized the norms of the coefficients through interpolation.

The eigenvalues are shown in Figure 4, and although we compute the eigenvalues from a complex valued linearization, all of the real parts are zero.

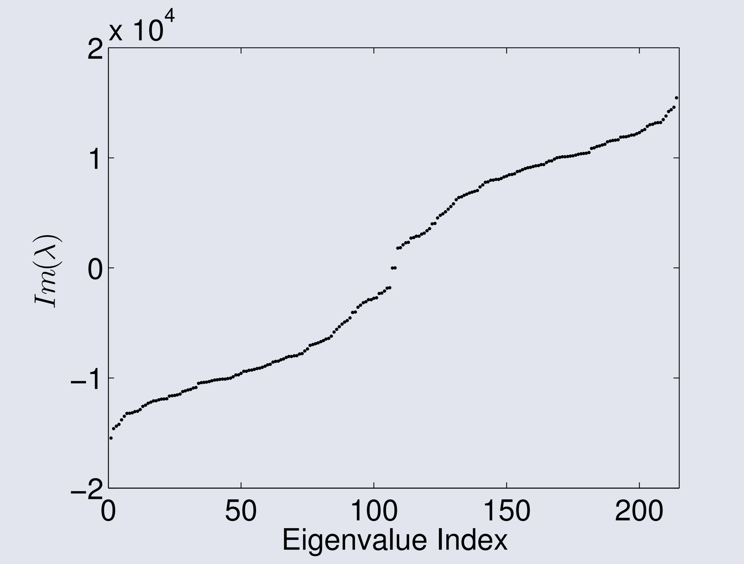

Figure 4: Computed eigenvalues

The backward error distributions are shown in Figure 5, and are remarkably small. The error bound in this case is around 4 orders of magnitude larger than the computed backward errors.

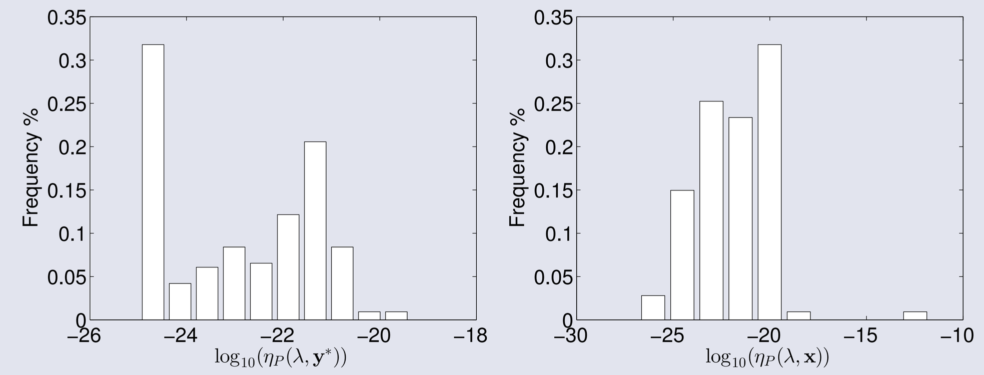

Figure 5: Speaker enclosure, backward error distribution for left and right eigenpairs.

References

[1] J.-P. Berrut and L. N. Trefethen, *Barycentric Lagrange interpolation*, SIAM Rev., 46 (2004), pp. 501–517.

[2] R. M. Corless, *Generalized companion matrices in the Lagrange basis.* In Proceedings EACA, (2004), pp. 317-322.

[3] D. Mackey, N. Mackey, C. Mehl and V. Mehrmann, *Vector spaces of linearizations for matrix polynomials*, SIAM J. Matrix Anal. Appl., 28(4) (2006), pp. 971-1004.

[4] R. Van Beeumen, W. Michiels and K. Meerbergen Linearization of Lagrange and Hermite interpolating matrix polynomials, to appear in IMA J. Numer. Anal.

[5] T. Betcke, N. J. Higham, V. Mehrmann, C. Schröder, F. Tisseur, *NLEVP: A collection of nonlinear eigenvalue problems*, ACM Trans. Math. Softw. 39(2) (2013) pp. 7:1-7:28.