Numerical methods for the tridiagonal hyperbolic quadratic eigenvalue problem

Bor Plestenjak
Department of Mathematics
University of Ljubljana
Slovenia

Quadratic eigenvalue problem (QEP)

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

We consider a special case when

- M, C, K are real symmetric tridiagonal matrices and
- QEP is hyperbolic (overdamped).

An example is the overdamped mass-spring system.

Our goal is to compute all the eigenvalues, i.e. scalars λ such that $\det(Q(\lambda)) = 0$. The eigenvectors can be later obtained by inverse iteration.

If M is nonsingular $n \times n$ matrix then there are $2n$ eigenvalues.

Overview

- Hyperbolic and overdamped QEP
- Inertia of a hyperbolic QEP
- $\det(Q(\lambda))$ and its derivatives
- Rank two divide and conquer approach
- Laguerre method and bisection
- Ehrlich–Aberth method
- Durand–Kerner method
- Numerical experiments
- Conclusion
- Diagonal homotopy

Hyperbolic QEP

$Q(\lambda) = \lambda^2 M + \lambda C + K$ is hyperbolic if $M > 0$ and

$$(x^T C x)^2 - 4(x^T M x)(x^T K x) > 0$$

for all $x \neq 0$.

Properties:

- all eigenvalues and eigenvectors are real
- eigenvalues are semisimple
- a gap between n largest (primary) and n smallest (secondary) eigenvalues
- n linearly independent vectors associated with the primary and the secondary eigenvalues, respectively
- Q is hyperbolic iff there exists λ_0 such that $Q(\lambda_0) < 0$.

Overdamped QEP

For each $x \neq 0$ the equation

$$\mu^2 x^T M x + \mu x^T C x + x^T K x$$

has two real solutions $\mu_1(x) < \mu_2(x)$.

Minimax principle: if $\lambda_{2n} \leq \cdots \leq \lambda_1$ are eigenvalues of a hyperbolic QEP $Q(\lambda)$ then

$$\lambda_i = \max_{\substack{S \subset \mathbb{R}^n \\ \dim(S)=i}} \min_{0 \neq x \in S} \mu_2(x)$$

and

$$\lambda_{n+i} = \max_{\substack{S \subset \mathbb{R}^n \\ \dim(S)=i}} \min_{0 \neq x \in S} \mu_1(x)$$

for $i = 1, \ldots, n$.

A hyperbolic QEP is overdamped if $C > 0$ and $K \geq 0$.

Additional properties:

- all eigenvalues are negative
- overdamped QEP are shifted hyperbolic QEP:

$$Q(\lambda + \theta) = \lambda^2 M + \lambda(C + 2\theta M) + K + \theta C + \theta^2 M.$$

Inertia of a hyperbolic QEP

Inertia of a symmetric matrix A is a triple $(\nu(A), \zeta(A), \pi(A))$, where ν, ζ and π are the numbers of negative, zero and positive eigenvalues of A, respectively.

Theorem: Let γ be such that $Q(\gamma) < 0$.

a) If $\gamma \leq \mu$ then $\nu(Q(\mu))$ equals the number of eigenvalues of Q that are greater than μ.

b) If $\mu \leq \gamma$ then $\nu(Q(\mu))$ equals the number of eigenvalues of Q that are smaller than μ.

Proof: $Q(\mu)$ is a symmetric matrix with ordered eigenvalues

$$\sigma_n(\mu) \leq \cdots \leq \sigma_1(\mu)$$

for each μ. σ_i are continuos functions of μ. λ is an eigenvalue of Q iff one of $\sigma_1(\lambda), \ldots, \sigma_n(\lambda)$ is 0,

At $\mu = \gamma$ all σ_i are negative and at $\mu = \pm\infty$ all σ_i are positive. Each σ_i crosses the x-axis exactly twice, once left and once right of γ.

$$\begin{array}{ccccccccccccc} 0 & 1 & 2 & \cdots & n-1 & n & \cdots & n & n-1 & \cdots & 2 & 1 & 0 \\ -\infty & & & \longleftarrow & & \gamma & & \longrightarrow & & & & & \infty \end{array}$$

Bisection and other methods

Based on the inertia we can use bisection to obtain the k-th eigenvalue, but the convergence is slow. Therefore we apply methods that were successfully applied to tridiagonal generalized eigenproblems.

- Laguerre's iteration
  - K. Li, T.Y. Li (1993) - symmetric tridiagonal eigenproblem
  - K. Li, T.Y. Li, Z. Zeng (1994) - generalized symmetric tridiagonal eigenproblem
- Ehrlich–Abert iteration
  - D.A. Bini, L. Gemignani, F. Tisseur (2003) - nonsymmetric tridiagonal eigenproblem
- Durand–Kerner method
  - K. Li (1999) - generalized symmetric tridiagonal eigenproblem

The above methods require stable and efficient computation of $\nu(Q(\lambda))$, $f(\lambda)$, $f'(\lambda)/f(\lambda)$ and $f''(\lambda)/f(\lambda)$, where $f(\lambda) = \det(Q(\lambda))$.

Three term recurrences

Let

$$Q(\lambda) = \begin{bmatrix} a_1 & b_1 & & & 0 \\ b_1 & a_2 & b_2 & & \\ & \ddots & \ddots & \ddots & \\ & & b_{n-2} & a_{n-1} & b_{n-1} \\ 0 & & & b_{n-1} & a_n \end{bmatrix},$$

where $a_i = a_i(\lambda)$ and $b_i = b_i(\lambda)$. Then

$$f_0 = 1,\ f_1 = a_1, \quad f_{r+1} = a_{r+1} f_r - b_r^2 f_{r-1}.$$

$$f_0' = 0,\ f_1' = a_1',$$
$$f_{r+1}' = a_{r+1}' f_r + a_{r+1} f_r' - 2 b_r b_r' f_{r-1} - b_r^2 f_{r-1}'.$$

$$f_0'' = 0,\ f_1'' = a_1'',$$
$$f_{r+1}'' = a_{r+1}'' f_r + 2 a_{r+1}' f_r' + a_{r+1} f_r'' - 2 b_r' 2^2 f_{r-1}$$
$$-2 b_r b_r'' f_{r-1} - 4 b_r b_r' f_{r-1}' - b_r^2 f_{r-1}''.$$

As the above recurrences may suffer from overflow–underflow problems, we define $d_i = \frac{f_i}{f_{i-1}}$, $g_i = \frac{f_i'}{f_i}$, $h_i = \frac{f_i''}{f_i}$ and obtain

$$d_1 = a_1, \quad d_{r+1} = a_{r+1} - \frac{b_r^2}{d_r}.$$

$$g_0 = 0,\ g_1 = \frac{a_1'}{a_1},$$
$$g_{r+1} = \frac{1}{d_{r+1}}(a_{r+1}' + a_{r+1} g_r - \frac{1}{d_r}(2 b_r b_r' + b_r^2 g_{r-1})).$$

$$h_0 = 0, h_1 = \frac{a_1''}{a_1},$$
$$h_{r+1} = \frac{1}{d_{r+1}}(a_{r+1}'' + 2 a_{r+1}' g_r + a_{r+1} h_r$$
$$- \frac{1}{d_r}(2 b_r'^2 + 2 b_r b_r'' + 4 b_r b_r' g_{r-1} + b_r^2 h_{r-1})).$$

It follows that $f_n = d_1 \cdots d_n$.

Time complexity for the computation of f, f' and f'' is approximately:

- f: 12 n
- f and f': 25n
- f, f' and f'': 44n

Remark: it is known that

$$f'(\lambda)/f(\lambda) = \mathrm{Tr}(Q(\lambda)^{-1}Q'(\lambda)).$$

Bini, Gemignani and Tisseur use this formula for a stable $\mathcal{O}(n)$ computation of f'/f via QR decomposition when $f(\lambda) = \det(A - \lambda I)$ and A is tridiagonal. Could this approach be generalized for the computation of $\mathrm{Tr}(A^{-1}B)$ where A and B are tridiagonal?

Divide and conquer

Let

$$Q(\lambda) = \begin{bmatrix} a_1 & b_1 & & & 0 \\ b_1 & a_2 & b_2 & & \\ & \ddots & \ddots & \ddots & \\ & & b_{n-2} & a_{n-1} & b_{n-1} \\ 0 & & & b_{n-1} & a_n \end{bmatrix},$$

where $a_i = a_i(\lambda)$ and $b_i = b_i(\lambda)$. We choose $m \approx n/2$ and set $b_m = 0$. We obtain

$$Q_0(\lambda) = \begin{bmatrix} Q_1(\lambda) & 0 \\ 0 & Q_2(\lambda) \end{bmatrix}.$$

$Q_1(\lambda)$, $Q_2(\lambda)$ and therefore $Q_0(\lambda)$ are hyperbolic QEP. The eigenvalues $\widetilde{\lambda}_{2n} \leq \cdots \leq \widetilde{\lambda}_1$ of Q_0 are approximations for the eigenvalues $\lambda_{2n} \leq \cdots \leq \lambda_1$ of Q.

Let $Q(\lambda, t) = tQ(\lambda) + (1-t)Q_0(\lambda)$. If the eigenvalues of $Q(\lambda, t)$ are real for $t \in [0, 1]$ then we can show that they interlace.

Theorem: If $\lambda_n(K)^2 > 4\lambda_1(M)\lambda_1(C)$ then $Q(\lambda, t)$ is hyperbolic for $t \in [0, 1]$.

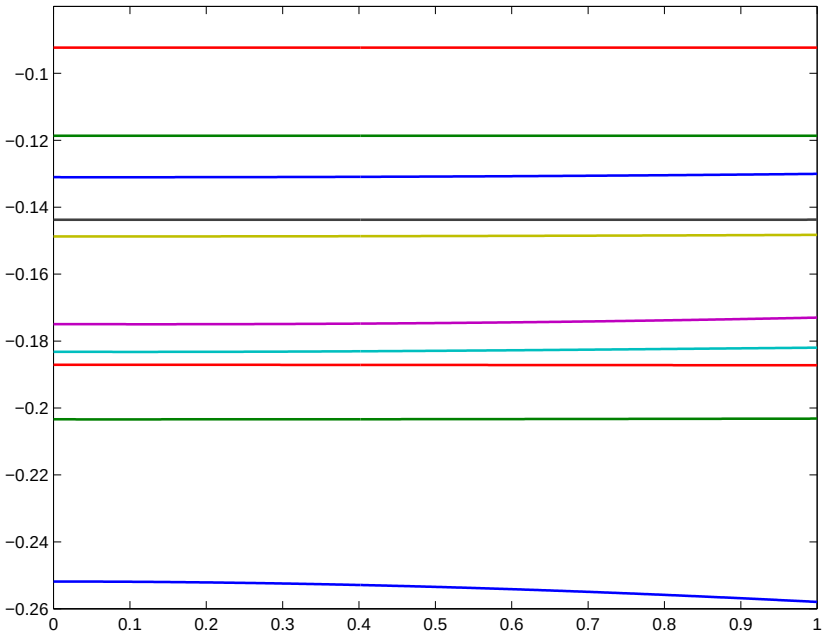
Interlacing property

Theorem: Let eigenvalues of $Q(\lambda, t)$ be real for $t \in [0, 1]$ and let $\widetilde{\lambda}_{2n} \leq \cdots \leq \widetilde{\lambda}_1$ be the eigenvalues of $Q_0(\lambda)$ and $\lambda_{2n} \leq \cdots \leq \lambda_1$ the eigenvalues of Q. Then:

a) $\widetilde{\lambda}_1 \leq \lambda_1$ and $\lambda_{2n} \leq \widetilde{\lambda}_{2n}$,

b) $\widetilde{\lambda}_{i+1} \leq \lambda_i \leq \widetilde{\lambda}_{i-1}$, for $i = 2, \ldots, n-1$ and $i = n+2, \ldots, 2n-1$,

c) $\widetilde{\lambda}_{n+1} \leq \lambda_{n+1} < \lambda_n \leq \widetilde{\lambda}_n$.



Laguerre's method

Let $f(\lambda) = \det(Q(\lambda))$. Laguerre's iteration is

$$L_{\pm}(x) = x + \frac{2n}{\left(\frac{-f'(x)}{f(x)} \pm \sqrt{(2n-1)\left((2n-1)\left(\frac{-f'(x)}{f(x)}\right)^2 - 2n\frac{f''(x)}{f(x)}\right)}\right)}.$$

The method is globally convergent with cubic convergence in a neighbourhood of the corresponding simple eigenvalue.

If we add $\lambda_{2n+1} = -\infty$ and $\lambda_0 = \infty$ then for $x \in (\lambda_{i+1}, \lambda_i)$ we have

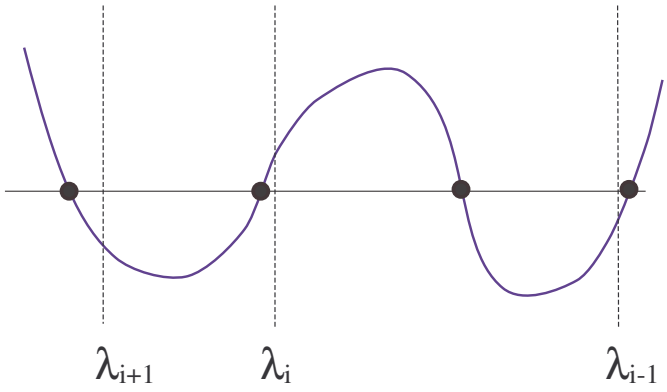
$$\lambda_{i+1} < L_-(x) < x < L_+(x) < \lambda_i.$$

Divide and conquer: eigenvalues $\widetilde{\lambda}_{2n} \leq \cdots \leq \widetilde{\lambda}_1$ of $Q_0(\lambda)$ are initial approximations for $\lambda_{2n} \leq \cdots \leq \lambda_1$.

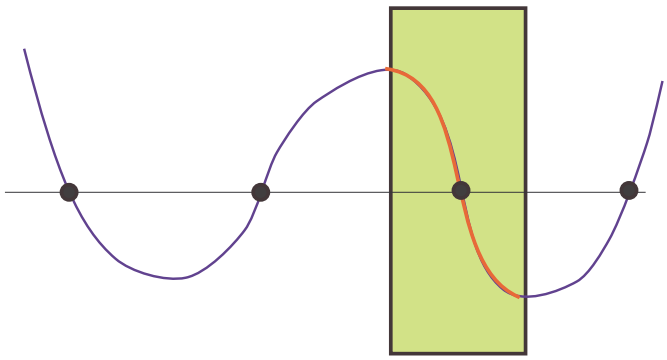
As $\widetilde{\lambda}_{i+1} \leq \lambda_i \leq \widetilde{\lambda}_{i-1}$ we can always use $\widetilde{\lambda}_i$ as an initial approximation for λ_i. From $\nu(Q(\widetilde{\lambda}_i))$ we see if $\lambda_i > \widetilde{\lambda}_i$ or $\widetilde{\lambda}_i < \lambda_i$ and then use L_+ or L_- sequence.

Bisection and Laguerre's method

If $\widetilde{\lambda}_i$ is closer to λ_{i-1} or λ_{i+1}, then the convergence can be very slow.



The necessary condition for the cubic convergence near a single eigenvalue λ is that the sign of $-f(x)'/f(x)$ agrees with the sign of $\lambda - x$. To improve the convergence we first use bisection on interval $[\lambda_i, \lambda_{i+1}]$ (or $[\lambda_i, \lambda_{i-1}]$) until the condition is achieved.



Due to numerical errors, the condition $-f(x)'/f(x)(\lambda - x) > 0$ can also be achieved near λ_{i-1} or λ_{i+1}. Additional criteria is that near λ_i the sign of $f'(x)$ has to agree with $(-1)^{i+1}$.

Ehrlich–Aberth's method

The method simultaneously approximates all the zeros of a polynomial $f(\lambda) = \det(Q(\lambda))$. From an initial approximation $x^{(0)} \in \mathbb{C}^{2n}$ the method generates a sequence $x^{(j)} \in \mathbb{C}^{2n}$ which locally converges to the eigenvalues of $Q(\lambda)$. The equation is

$$x_j^{(k+1)} = x_j^{(k)} - \frac{\dfrac{f(x_j^{(k)})}{f'(x_j^{(k)})}}{1 - \dfrac{f(x_j^{(k)})}{f'(x_j^{(k)})} \displaystyle\sum_{\substack{l=1 \\ l \neq j}}^{2n} \frac{1}{x_j^{(k)} - x_l^{(k)}}}$$

for $j = 1, \ldots, 2n$.

If we implement the method in the Gauss–Seidel style then the convergence for simple roots is cubical and linear for multiple roots. We iterate only those eigenvalues that have not converged yet.

For an efficient use of the method we need good initial approximations, and fast and stable computation of $p(x)/p'(x)$.

For initial approximations we again use divide and conquer with rank two modifications. To eliminate multiple values in the initial approximation we slightly perturb the eigenvalues of $Q_0(\lambda)$.

Durand–Kerner's method

Another method that simultaneously approximates all the zeros of a polynomial $f(\lambda) = \det(Q(\lambda))$ is Durand–Kerner's method. As the method requires that the leading coefficient of the polynomial is one we apply it on

$$p(\lambda) = \frac{1}{\det(M)} \det(Q(\lambda)).$$

Similar to Ehrlich–Aberth the method generates a sequence $x^{(j)} \in \mathbb{R}^{2n}$ which locally converges to the eigenvalues of $Q(\lambda)$. The equation is

$$x_j^{(k+1)} = x_j^{(k)} - \frac{p(x_j^{(k)})}{\prod\limits_{\substack{l=1 \\ l \neq j}}^{2n} x_j^{(k)} - x_l^{(k)}}$$

for $j = 1, \ldots, 2n$.

If we implement the method in the Gauss–Seidel style then the convergence for simple roots is superquadratical and linear for multiple roots. As before, we iterate only the eigenvalues that have not converged yet.

Remark: K. Yi reports that although the method always converges in practice, this has not been proved yet.

Multiple eigenvalues

When we have multiple eigenvalue or close eigenvalues, we can expect problems and linear convergence in all three presented methods. Some solutions suggested in the literature are:

In Laguerre's method we may estimate the multiplicity or the number of eigenvalues in the cluster using the inertia. If the multiplicity of the eigenvalue is r we apply the modified Laguerre's iteration

$$L_{r\pm}(x) = x + \frac{2n}{\left(\frac{-f'(x)}{f(x)} \pm \sqrt{(2n-1)\left(\frac{2n-r}{r}\left(\frac{-f'(x)}{f(x)}\right)^2 - 2n\frac{f''(x)}{f(x)}\right)}\right)}.$$

which converges cubically.

In Ehrlich–Aberth's and Durand–Kerner's method, when we obtain the eigenvalue of multiplicity r, we fix all r instances and continue with the remaining eigenvalue approximations. In this way we prevent zero values in the denominator.

Comparison and numerical results

In Matlab we tested all three methods on a limited set of tridiagonal overdamped QEPs. We compared the average number of iterations. In all three methods one step (for one eigenvalue approximation) has linear time complexity.

- One step of Durand–Kerner's method is the cheapest as it requires only values of f.
- One step of Ehrlich–Aberth's method requires f and f' and is roughly equivalent to 2.4 Durand–Kerner steps.
- One step of Laguerre's method is the most expensive. It requires f, f' and f'', and is roughly equivalent to 4.7 Durand–Kerner steps.

The testing overdamped QEPs have the following structure:

- M and K: diagonals are random values from $[0.5, 1]$, codiagonals are random values from $[0, 0.1]$.
- C: diagonals are random values from $[4, 5]$, codiagonals are random values from $[0, 0.5]$.

method	avg. iter.	avg. iter. in last step	× imp. factor	time
$n = 20$				
LB	6.8	4.0	14.4	2.5
AE	6.2	3.5	7.4	0.9
DK	13.1	8.5	8.5	0.8
$n = 40$				
LB	5.2	2.1	7.6	7.7
AE	5.0	2.1	4.4	2.2
DK	8.6	3.3	3.3	1.6
$n = 100$				
LB	4.5	1.9	6.8	43.6
AE	3.9	1.6	3.4	8.4
DK	6.8	2.5	2.5	4.5
$n = 200$				
LB	4.5	2.1	7.6	178.5
AE	3.4	1.4	2.9	27.3
DK	5.8	2.3	2.3	13.0
$n = 300$				
LB	3.9	1.6	5.8	365.6
AE	3.0	1.2	2.5	54.2
DK	4.6	1.6	1.6	19.9

Durand–Kerner's method is faster than `polyeig(K,C,M)` for $n \geq 100$. For example, `polyeig(K,C,M)` requires 39.6s for $n = 200$.

As the dimension of the matrices increases, the eigenvalues of $Q_0(\lambda)$ better approximate eigenvalues of $Q(\lambda)$ and the methods require fewer steps in the final divide and conquer phase.

Conclusions

Three eigensolvers for tridiagonal hyperbolic QEPs.

Generalizations:

- All methods can be easily parallelized.
- Similar approach might be generalized to:
  - nonsymmetric and non hyperbolic tridiagonal quadratic polynomial problems
  - tridiagonal polynomial problems

Future work:

- handling of multiple eigenvalues
- more numerical tests
- stable computation of f, f' and f''.
- can continuation and path following be as efficient as other three methods

Continuation method

A continuation method has been successfully applied to the symmetric and nonsymmetric eigenvalue problem, to the generalized eigenvalue problem and other eigenproblems. Chu, Li and Sauer (1988) presented a continuation method for a general λ-matrix problem that does not exploit the properties of an overdamped QEP as the following method.

Let M_0, C_0 and K_0 be diagonal matrices and let

$$Q(\lambda, t) = (1-t)Q_0(\lambda) + tQ_1(\lambda),$$

where

$$\begin{aligned} Q_0(\lambda) &= \lambda^2 M_0 + \lambda C_0 + K_0, \\ Q_1(\lambda) &= \lambda^2 M + \lambda C + K. \end{aligned}$$

We construct a homotopy $H : \mathbb{R}^n \times \mathbb{R} \times [0,1] \to \mathbb{R}^n \times \mathbb{R}$

$$H(x, \lambda, t) = \begin{bmatrix} Q(\lambda, t)x \\ \frac{1}{2}(x^T x - 1) \end{bmatrix}.$$

The solution of $H(x, \lambda, 1) = 0$ corresponds to the eigenpairs of Q. The solution of $H(x, \lambda, 0) = 0$ is trivial and we choose the initial matrices M_0, C_0, and K_0 in such a way that all eigenvalues of $Q_0(\lambda)$ are algebraically simple.

We follow the set $H^{-1}(0)$ from the known solution of $H(x, \lambda, 0) = 0$ at $t = 0$ to the solution of Q at $t = 1$.

Theorem: Let M_0, C_0, and K_0 be diagonal matrices such that

$$0 \leq (M_0)_{ii} < \lambda_{\min}(M), \qquad \lambda_{\max}(C) \leq (C_0)_{ii},$$

$$(K_0)_{ii} \leq \lambda_{\min}(K)$$

for $i = 1, \ldots, n$. Then $Q(\lambda, t)$ is a hyperbolic QEP for $t \in [0, 1]$.

Remark: If the QEP is hyperbolic but not overdamped, then we can not apply the above theorem. However, using an appropriate shift it is possible to transform each hyperbolic QEP into an overdamped QEP.

Theorem: There exists a subset S of $\mathbb{R}^n$ with full measure such that if $(k_1, \ldots, k_n) \in S$ and $K_0 = \mathrm{diag}(k_1, \ldots, k_n)$, then all eigenvalues of $Q(\lambda, t)$ are algebraically simple for $t \in [0, 1)$.

We have $2n$ disjoint smooth curves in $\mathbb{R}^n \times \mathbb{R}$, each leads from an eigenpair of $Q_0(\lambda)$ to the eigenpair of $Q_1(\lambda)$. The homotopy has the order-preserving property.

Following the eigencurves

We may follow the eigencurve through (x, λ, t) by the predictor-correction scheme. We have

$$\begin{bmatrix} \dot{x} \\ \dot{\lambda} \end{bmatrix} = \begin{bmatrix} Q(\lambda, t) & Q'(\lambda, t)x \\ x^T & 0 \end{bmatrix}^{-1} \begin{bmatrix} -(Q_1(\lambda) - Q_0(\lambda))x \\ 0 \end{bmatrix},$$

$$\dot{\lambda} = -\frac{x^T\Big(\lambda^2(A - A_0) + \lambda(B - B_0) + (C - C_0)\Big)x}{x^T\Big(2\lambda A(t) + B(t)\Big)x}.$$

Let us assume that the step size h was determined. Then we

a) use Hermite interpolation to predict $\lambda^{(0)}(t_{i+1})$ from $\lambda(t_{i-1})$, $\lambda'(t_{i-1})$, $\lambda(t_i)$, and $\lambda'(t_i)$.

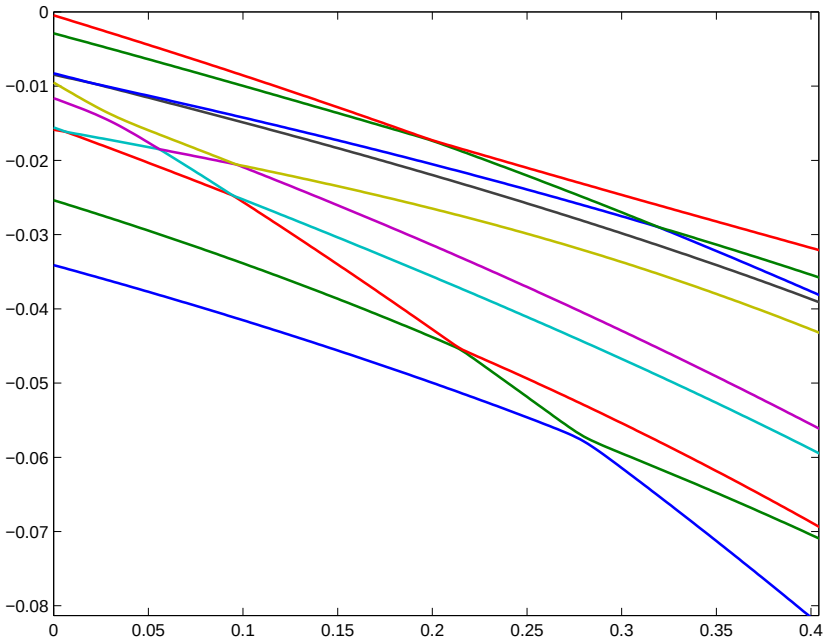
b) use one step of the inverse power method to predict the eigenvector:

$$Q(\lambda^{(0)}(t_{i+1}), t_{i+1})y = x^{0)},$$

$$x^{(1)} = \frac{y}{\|y\|}.$$

c) use a modified Rayleigh quotient method with the initial approximation $(\lambda^{(0)}, x^{(0)})$

Deflections



Eigenvalue curves do not cross each other but do come very close.