Numerical methods for algebraic two-parameter eigenvalue problems

Bor Plestenjak
Department of Mathematics
University of Ljubljana
Slovenia

www-lp.fmf.uni-lj.si/plestenjak/bor.htm

Overview

- Two-parameter eigenvalue problem
  - origin
  - applications
  - tensor product approach
  - right definite problem
- Numerical methods
  - overview
  - continuation method
    - right definite case
    - weakly elliptic case
  - Jacobi–Davidson type method
    - right definite case
    - general case
- Conclusions

Two-parameter eigenvalue problem

- We consider two-parameter eigenvalue problem

$$\begin{aligned} A_1 x &= \lambda\, B_1 x + \mu\, C_1 x \\ A_2 y &= \lambda\, B_2 y + \mu\, C_2 y, \end{aligned} \tag{W}$$

  where A_i, B_i, C_i are $n \times n$ matrices, $\lambda, \mu \in \mathbb{C}$, and $x, y \in \mathbb{C}^n$

- Eigenvalue: a pair (λ, μ) that satisfies (W) for nonzero x and y

- Eigenvector: the tensor product $x \otimes y$

- **Problem: compute some (all) eigenvalues (λ, μ) and eigenvectors $x \otimes y$**

Separation of variables (s.o.v.)

$$\Delta u + \nu u = 0 \text{ on } \Omega, \qquad u|_{\delta\Omega} = 0$$

Rectangle: $\Omega = [0, a] \times [0, b]$, s.o.v. $\Longrightarrow$

$$x'' + \lambda x = 0, \qquad x(0) = x(a) = 0,$$
$$y'' + \mu y = 0, \qquad y(0) = y(b) = 0.$$

Circle: $\Omega = \{x^2 + y^2 \leq a^2\}$, polar coordinates, s.o.v. $\Longrightarrow$

$$\Phi'' + \lambda\Phi = 0, \qquad \Phi(0) = \Phi(2\pi) = 0,$$
$$r^{-1}(rR')' + (\nu - \lambda r^{-2})R = 0, \qquad R(0) < \infty, R(a) = 0.$$

Ellipse: $\Omega = \{(x_1/c_1)^2 + (x_2/c_2)^2 \leq 1\}$, elliptic coordinates, s.o.v. $\Longrightarrow$

$$v_1'' + (2\lambda \cosh(2y_1) - \mu)v_1 = 0$$
$$v_2'' - (2\lambda \cos(2y_1) - \mu)v_2 = 0.$$

Some two-parameter problems that appear in the algebraic form

Osborne (1963): The optimum value of the overrelaxation parameter ω in the SOR method for a separable elliptic partial differential equation in two independent variables can be obtained from the eigenvalues of a certain two-parameter eigenvalue problem.

Leiseutre, Mamishev, et al (2001): The estimation of material electrical properties from measurements of interdigital dielectrometry sensors. When the sensors are applied to the material that is composed of two layers, the properties of the individual layers are the eigenvalues of the appropriate two-parameter eigenvalue problem.

Cottin (2001): Dynamic model updating. We have a spring-mass model where the mass matrix is known and the stiffness parameter values of two springs have to be updated based on the outside measurements of the natural frequencies. The updated parameters are the eigenvalues of a two-parameter problem.

Tensor product approach

$$\begin{aligned} A_1 x &= \lambda B_1 x + \mu C_1 x \\ A_2 y &= \lambda B_2 y + \mu C_2 y \end{aligned} \tag{W}$$

- On $S := \mathbb{C}^n \otimes \mathbb{C}^n$ of the dimension n^2 we define

$$\begin{aligned} \Delta_0 &= B_1 \otimes C_2 - C_1 \otimes B_2 \\ \Delta_1 &= A_1 \otimes C_2 - C_1 \otimes A_2 \\ \Delta_2 &= B_1 \otimes A_2 - A_1 \otimes B_2. \end{aligned}$$

- Two-parameter problem (W) is equivalent to coupled GEP

$$\begin{aligned} \Delta_1 z &= \lambda \Delta_0 z \\ \Delta_2 z &= \mu \Delta_0 z \end{aligned} \tag{Δ}$$

where $z = x \otimes y$.

- (W) is nonsingular $\Longleftrightarrow$ Δ_0 is invertible.
- $\Delta_0^{-1}\Delta_1$ and $\Delta_0^{-1}\Delta_2$ commute.

Right definite problem

Problem

$$\begin{aligned} A_1 x &= \lambda B_1 x + \mu C_1 x, \\ A_2 y &= \lambda B_2 y + \mu C_2 y. \end{aligned} \qquad \text{(W)}$$

is right definite when

- A_i, B_i, C_i real symmetric
- $\begin{vmatrix} x^T B_1 x & x^T C_1 x \\ y^T B_2 y & y^T C_2 y \end{vmatrix} > 0$ for nonzero x, y (equivalent to Δ_0 s.p.d.)

If (W) is right definite then

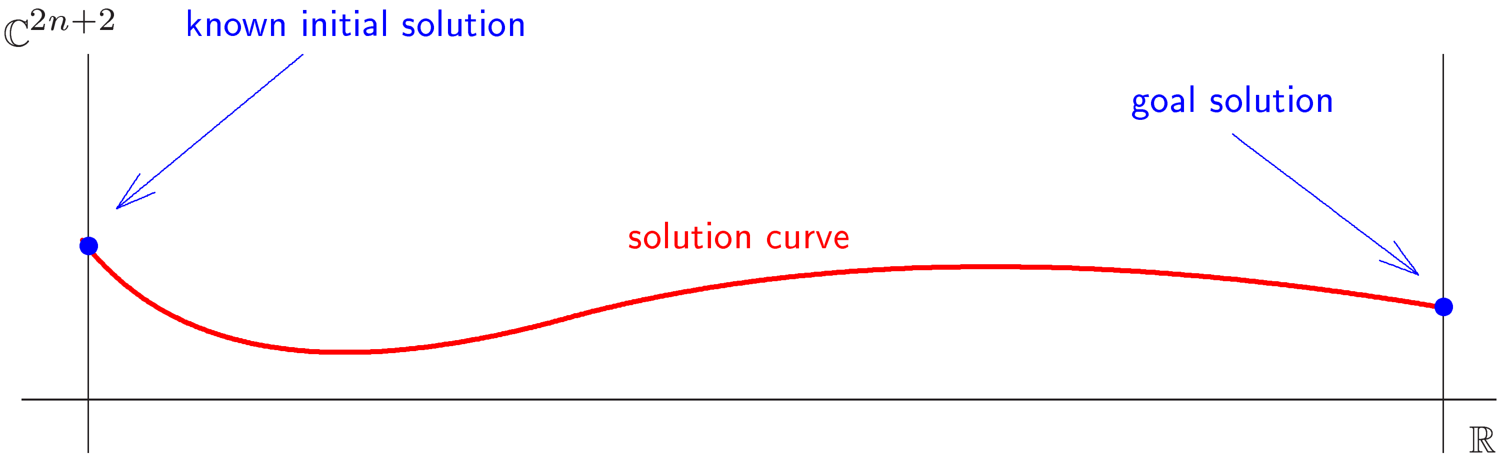
- eigenpairs are real
- there exist n^2 linearly independent eigenvectors
- eigenvectors of distinct eigenvalues are Δ_0-orthogonal, i.e. $(x_1 \otimes y_1)^T \Delta_0 (x_2 \otimes y_2) = 0$

Some available numerical methods

- Blum, Curtis, Geltner (1978) and Browne, Sleeman (1982): gradient method,
- Blum, Chang (1978): Minimum Residual Quotient Iteration (MRQI) for the problem $Ax = \lambda Bx + \mu Cx$ and conditions $\|x\| = 1$, $f(x) = 0$, where f is a real functional,
- Bohte (1980): Newton's method for eigenvalues,
- Müller (1982): continuation method for one two-parameter equation,
- Slivnik, Tomšič (1986): solving (Δ) with standard numerical methods,
- Ji, Jiang, Lee (1992): generalization of MRQI for (W): Generalized Rayleigh Quotient Iteration (GRQI).
- Shimasaki (1995): continuation method for a special class of RD problems.

- P. (1999): continuation method for RD problem.
- P. (2000): continuation method for weakly elliptic problem.
- Hochstenbach, P. (2002): Jacobi-Davidson type method for RD problem.
- Hochstenbach, Košir, P. (2003): Jacobi-Davidson type method.

Continuation method



$t = 0$ $\longrightarrow$ $t = 1$

$$A_{10}x = \lambda B_{10}x + \mu C_{10}x$$
$$A_{20}y = \lambda B_{20}y + \mu C_{20}y$$

$$W_t = (1-t)W_0 + tW$$

$$A_1 x = \lambda B_1 x + \mu C_1 x$$
$$A_2 y = \lambda B_2 y + \mu C_2 y$$

initial problem W_0 with known solutions

homotopy

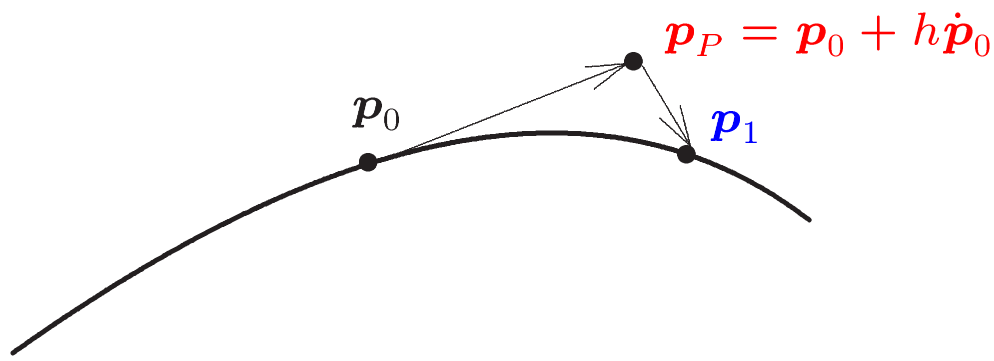
original problem W

Following the homotopy curve

- We numerically follow the homotopy curve by a prediction-correction scheme using arc length as the parameter.

- Euler's method is used as a predictor.

- Newton's method is used as a corrector.



Properties of a continuation method

- In tensor product space matrices Δ_i are of order n^2 and the time complexity is $\mathcal{O}(n^6)$.
- Matrices in the continuation method are of moderate size $\mathcal{O}(n)$.
  - One predictor-corrector step has time complexity $\mathcal{O}(n^3)$.
  - We have to multiply this number with n^2 as we are following n^2 curves.
  - We have to multiply it further with the number of P-C steps per curve.
- The continuation method allows an elegant parallel implementation.
- The continuation method works without approximations of eigenpairs.
- Even if we are interested in a small portion of eigenvalues, we have to compute all the eigenvalues.

Continuation method in the right definite case

We can assume that B_1 and C_2 are s.p.d.

Initial problem:

$$A_{10}x = \lambda B_1 x$$
$$A_{20}y = \mu C_2 y$$

W_t:

$$A_1 x = \lambda B_1 x + \mu t C_1 x,$$
$$A_2 y = \lambda t B_2 y + \mu C_2 y.$$

Eigenvalues and eigenvectors are real, we can use t as a parameter.

There are only finitely many singular points where we have multiple eigenvalues. As eigenvectors are Δ_0-orthogonal we can jump over.

Predictor: constant (the last approximation from the previous step)

Corrector: the tensor Rayleigh quotient iteration

Continuation method in the weakly elliptic case

Weakly elliptic: All matrices are symmetric and one of B_1, C_1, B_2, C_2 is definite, we can assume that B_1 and C_1 are positive definite.

For an eigenvalue (λ, μ) we have either $\lambda, \mu \in \mathbb{R}$ or $\lambda, \mu \notin \mathbb{R}$.

Initial problem: We can construct symmetric S_1 and S_2 in a way that

a) all eigenvalues of the two-parameter problem

$$\begin{aligned} S_1 x &= \lambda B_1 x + \mu C_1 x, \\ S_2 y &= \lambda B_2 y + \mu C_2 y. \end{aligned} \qquad (W_0)$$

are algebraically simple,

b) the construction reveals the solutions of (W_0).

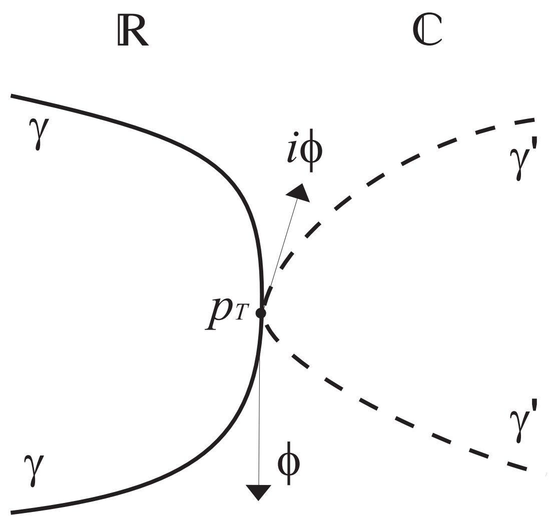
W_t:

$$(1-t)A_1 x + tS_1 x = \lambda B_1 x + \mu C_1 x,$$

$$(1-t)A_2 y + tS_2 x = \lambda B_2 y + \mu C_2 y.$$

Continuation method in the weakly elliptic case (cont.)

- We use arclength as a parameter.
- There are only finitely many values $t \in [0, 1]$ where (W_t) has a multiple eigenvalue.
- All bifurcations are turning points.
- We can not avoid the turning points.



Change from real to complex space in a quadratic turning point

JD-like method for the right definite case: extraction

Ritz–Galerkin conditions: search spaces = test spaces: $u_1 \in \mathcal{U}_{1k}$, $u_2 \in \mathcal{U}_{2k}$

$$(A_1 - \sigma B_1 - \tau C_1)u_1 \perp \mathcal{U}_{1k}$$
$$(A_2 - \sigma B_2 - \tau C_2)u_2 \perp \mathcal{U}_{2k}$$

$\Rightarrow$ projected right def. 2-parameter problem

$$\begin{aligned} U_{1k}^T A_1 U_{1k} c_1 &= \sigma U_{1k}^T B_1 U_{1k} c_1 + \tau U_{1k}^T C_1 U_{1k} c_1 \\ U_{2k}^T A_2 U_{2k} c_2 &= \sigma U_{2k}^T B_2 U_{2k} c_2 + \tau U_{2k}^T C_2 U_{2k} c_2 \end{aligned}$$

Ritz vectors: $u_1 = U_{1k} c_1$, $u_2 = U_{2k} c_2$, where $c_1, c_2 \in \mathbb{R}^k$

Ritz value: (σ, τ), Ritz pair: $((\sigma, \tau), u_1 \otimes u_2)$

JD-like method for the RD case: algorithm

1. $s_1 = u_1$ and $s_2 = u_2$ (starting vectors)

   for $k = 1, 2, \ldots$

2. $(U_{1,k-1}, s_1) \to U_{1k}$
   $(U_{2,k-1}, s_2) \to U_{2k}$

3. Extract appropriate Ritz pair $((\sigma, \tau), c_1 \otimes c_2)$ of

$$\begin{aligned} U_{1k}^T A_1 U_{1k} c_1 &= \sigma U_{1k}^T B_1 U_{1k} c_1 + \tau U_{1k}^T C_1 U_{1k} c_1 \\ U_{2k}^T A_2 U_{2k} c_2 &= \sigma U_{2k}^T B_2 U_{2k} c_2 + \tau U_{2k}^T C_2 U_{2k} c_2 \end{aligned}$$

4. $$\begin{aligned} r_1 &= (A_1 - \sigma B_1 - \tau C_1) u_1 \\ r_2 &= (A_2 - \sigma B_2 - \tau C_2) u_2 \end{aligned}$$

5. Stop if $(\|r_1\|^2 + \|r_2\|^2)^{1/2} \le \varepsilon$

6. Solve (approximately) an $s_1 \perp u_1$, $s_2 \perp u_2$ from corr. equation(s)

JD-like algorithm to find eigenpair of 2-parameter RD eigenvalue problem

JD-like method for the RD case: expansion, $s_1 \perp u_1$, $s_2 \perp u_2$

$$A_1(u_1 + s_1) = \lambda B_1(u_1 + s_1) + \mu C_1(u_1 + s_1)$$
$$A_2(u_2 + s_2) = \lambda B_2(u_2 + s_2) + \mu C_2(u_2 + s_2)$$

Two correction equations

$$\begin{aligned}(I - u_1u_1^T)(A_1 - \sigma B_1 - \tau C_1)(I - u_1u_1^T)s_1 &= -r_1 \\ (I - u_2u_2^T)(A_2 - \sigma B_2 - \tau C_2)(I - u_2u_2^T)s_2 &= -r_2\end{aligned}$$

- orthogonal projections preserve the symmetry
- we solve the equations only approximately with a Krylov subspace method with initial guess 0 (e.g., few steps of MINRES or GMRES).

JD-like method for RD case: computing more eigenpairs

Eigenvectors are Δ_0-orthogonal:

$$(x_1 \otimes y_1)^T \Delta_0 (x_2 \otimes y_2) = 0$$

Standard deflation techniques can not be applied:

- $(x \otimes y)^{\perp_{\Delta_0}}$ can not be written as $\mathcal{U} \otimes \mathcal{V}$, where $\mathcal{U} \subset \mathbb{R}^n$ and $\mathcal{V} \subset \mathbb{R}^n$.
- there can exist eigenvalues (λ, μ) and (λ', μ') with eigenvectors $x \otimes y$ and $x' \otimes y'$, respectively, such that $(\lambda, \mu) \neq (\lambda', \mu')$ and $x = x'$,

Our approach: In selection consider only Ritz vectors that are Δ_0-orthogonal to already computed eigenvectors.

Two-sided JD-like method for a general problem: extraction

Petrov–Galerkin conditions: search spaces $u_i \in \mathcal{U}_{ik}$, test spaces $v_i \in \mathcal{V}_{ik}$

$$\begin{aligned}(A_1 - \sigma B_1 - \tau C_1)u_1 &\perp \mathcal{V}_{1k},\\ (A_2 - \sigma B_2 - \tau C_2)u_2 &\perp \mathcal{V}_{2k},\end{aligned}$$

where $u_i \in \mathcal{U}_{ik} \backslash \{0\} \Rightarrow$ projected two-parameter problem

$$\begin{aligned}V_{1k}^* A_1 U_{1k} c_1 &= \sigma V_{1k}^* B_1 U_{1k} c_1 + \tau V_{1k}^* C_1 U_{1k} c_1,\\ V_{2k}^* A_2 U_{2k} c_2 &= \sigma V_{2k}^* B_2 U_{2k} c_2 + \tau V_{2k}^* C_2 U_{2k} c_2,\end{aligned}$$

where $u_i = U_{ik} c_i \neq 0$ for $i = 1, 2$ and $\sigma, \tau \in \mathbb{C}$.

Petrov vectors: $u_i = U_{ik} c_i$, $v_i = V_{ik} d_i$, $c_i, d_i \in \mathbb{C}^k$

Petrov value: (σ, τ), Petrov triple: $((\sigma, \tau), u_1 \otimes u_2, v_1 \otimes v_2)$

Two-sided JD-like method: algorithm

1. $s_i = u_i$ and $t_i = v_i$ (starting vectors)

   for $k = 1, 2, \ldots$

2. $(U_{i,k-1}, s_i) \to U_{ik}$
   $(V_{i,k-1}, t_i) \to V_{ik}$

3. Extract appropriate Petrov triple $((\sigma, \tau), c_1 \otimes c_2, d_1 \otimes d_2)$ of

$$\begin{aligned} V_{1k}^* A_1 U_{1k} c_1 &= \sigma V_{1k}^* B_1 U_{1k} c_1 + \tau V_{1k}^* C_1 U_{1k} c_1, \\ V_{2k}^* A_2 U_{2k} c_2 &= \sigma V_{2k}^* B_2 U_{2k} c_2 + \tau V_{2k}^* C_2 U_{2k} c_2, \end{aligned}$$

4. $$\begin{aligned} r_i^R &= (A_i - \sigma B_i - \tau C_i) u_i \\ r_i^L &= (A_i - \sigma B_i - \tau C_i)^* v_i \end{aligned}$$

5. Stop if $\left(\|r_1^R\|^2 + \|r_2^R\|^2 + \|r_1^L\|^2 + \|r_2^L\|^2\right)^{1/2} < \varepsilon$

6. Solve (approximately) subspace extensions s_i, t_i from corr. equation(s)

JD-like algorithm to find eigenpair of 2-parameter eigenvalue problem

Two-sided JD-like method: expansion, $s_i \perp a_i$, $t_i \perp b_i$

A correction equation for the vector u_i:

$$\left(I - \frac{c_i v_i^*}{v_i^* c_i}\right) (A_i - \sigma B_i - \tau C_i) \left(I - \frac{u_i a_i^*}{a_i^* u_i}\right) s_i = -r_i^R$$

for $i = 1, 2$, where $c_i \not\perp v_i$ and $a_i \not\perp u_i$.

Similarly, a correction equation for the vector v_i:

$$\left(I - \frac{d_i u_i^*}{u_i^* d_i}\right) (A_i - \sigma B_i - \tau C_i)^* \left(I - \frac{v_i b_i^*}{b_i^* v_i}\right) t_i = -r_i^L$$

for $i = 1, 2$, where $d_i \not\perp u_i$ and $b_i \not\perp v_i$.

Different choices of vectors a_i, b_i, c_i, d_i lead to different correction equations.

We suggest the preconditioner $M_i = A_i - \lambda_{\mathrm{T}} B_i - \mu_{\mathrm{T}} C_i$, where $(\lambda_{\mathrm{T}}, \mu_{\mathrm{T}})$ is the target.

Two-sided vs. one-sided

Statistics of the Jacobi–Davidson type method using the same set of 10 random initial vectors for computing 10 closest eigenvalues to the $(0, 0)$, matrices are of size 100.

For each eigenvalue we select the closest Petrov value to the origin until the residual becomes smaller than $\varepsilon_{\text{change}}$ and in the remaining steps we select Petrov triple with the minimum residual.

two-sided J-D												
	$\varepsilon_{\text{change}} = 10^{-1}$				$\varepsilon_{\text{change}} = 10^{-1.5}$				$\varepsilon_{\text{change}} = 10^{-2}$			
GMRES	In 10	Conv.	Avg.	Iter.	In 10	Conv.	Avg.	Iter.	In 10	Conv.	Avg.	Iter.
5	3.4	4.2	6.4	400.0	3.3	3.9	4.7	400.0	2.7	3.0	4.3	400.0
10	4.7	7.4	19.4	324.5	5.9	8.0	10.2	387.8	5.3	6.2	5.6	400.0
20	6.8	9.4	15.2	255.3	6.6	9.2	26.1	301.8	6.9	9.4	14.5	300.3
40	7.2	9.5	29.1	284.0	7.3	9.5	16.2	292.3	7.0	9.0	14.2	354.9

one-sided J-D												
	$\varepsilon_{\text{change}} = 10^{-1}$				$\varepsilon_{\text{change}} = 10^{-1.5}$				$\varepsilon_{\text{change}} = 10^{-2}$			
GMRES	In 10	Conv.	Avg.	Iter.	In 10	Conv.	Avg.	Iter.	In 10	Conv.	Avg.	Iter.
5	2.0	5.2	21.3	400.0	1.3	1.3	6.9	400.0	0.5	0.5	1.6	400.0
10	2.9	7.1	21.7	357.3	2.6	3.0	4.1	400.0	1.9	1.9	1.6	400.0
20	3.5	9.9	72.7	189.5	3.0	3.7	24.0	400.0	1.9	2.1	4.2	400.0
40	3.0	9.9	75.1	143.8	3.5	4.0	5.1	380.5	2.9	3.2	25.6	400.0

Conclusions

Continuation method:

- if we need all eigenvalues (eigenvectors)
- can be applied to right definite or weakly elliptic problems
- easy parallelization

Jacobi–Davidson type method:

- if we need selected eigenvalues (e.g., closest to the target)
- works for right definite and general nonsingular problems

Both methods:

- work with matrices of size $\mathcal{O}(n)$ and not $\mathcal{O}(n^2)$
- do not require initial approximations