Matrix Spectrum Dichotomy in Boundary Value Problems for Differential Algebraic Equations

Kostin V.I.[*] and Shtykel T.L.[†]

Differential Algebraic Equations are equations with singular matrices of coefficients at the derivative. The simplest form of such equation looks like

$$Lx = A\frac{dx}{dt} - x = f(t) \tag{1}$$

where $\det A = 0$.

In classical theory (for exact data and exact computing) the matrix can be represented via Jordan decomposition

$$A = T^{-1}\begin{pmatrix} J_1 & 0 \\ 0 & N \end{pmatrix} T, \qquad \det T \neq 0, \quad \det J_1 \neq 0, \quad N^{k-1} \neq 0, \quad N^k = 0.$$

Here size of the nilpotent cell k is the so called *index* of the matrix A. Standard change of unknowns $Tx = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ allows to determine y_1 by usual way whereas for y_2 we have

$$y_2 = -(g_2(t) + Ng_2'(t) + \cdots + N^{k-1}g_2^{(k-1)}(t)).$$

---

[*]Institute of Geophysics, Novosibirsk. Participation of the first author was partly supported by RFBR (Grant 97-05-65280), Competition of Grants on fundamental research in the field of mathematics of the Russian Ministry of Gen. and Prof. Education and France Russian Center of Appl. Math. and Inf. (Grant "Reliable numerical methods of solving of nonlinear conditionally well posed problems in application to inverse problems of wave propagation theory").

[†]Novosibirsk State University, Novosibirsk.

Thus the component y_2 is uniquely defined by the right hand side and does not allow varying of the inital data for it. Besides, some components of the right hand side must posess aditional smoothnes in comparision with the classical case.

This approach can be criticized if it is used in applications where in a realistic situation matrix A is an approximation of some "ideal" A_{ideal} with probably known error estimation of $\|A - A_{\text{ideal}}\|$. High sensitivity of eigenvalues of a matrix to its perturbations has led many authors to the following definition of ε-spectrum of a matrix

$$\Lambda_\varepsilon(A) = \{\lambda \in C | \exists \Delta A : \det(A + \Delta A - \lambda I) = 0, \|\Delta A\| \leq \varepsilon \|A\|\}$$

which more adequately reflects situation than eigenvalues themselves.

1 Boundary value problem and criterion of its well-posedness

Because of specific properties mentioned above settings of boundary value problems for DAE require special investigation. The setting seems to be insensitive to perturbations in a matrix which may make it nondegenerate.
Definition W_σ *($\sigma > 0$) is a class of vector-valued functions $y(t) \in L_2$ which admit expansion for complex t as entire analytic functions with estimation of the type $|y(t)| \leq Ce^{\sigma|t|}$.*

Due to the Paley-Wiener theorem W_σ coincides with Fourier images of L_2-functions with supports on the interval $[-\sigma, \sigma]$.

We consider equation (1) as a boundary value problem in the class W_σ.

Theorem 1. *The boundary value problem is uniquely resolved in W_σ if and only if the matrix A has no eigenvalues in the union of rays $(-i\infty, -i/\sigma] \cup [i/\sigma, i\infty)$ of the imaginary axis.*

Similarly to the classical case one can introduce the *Green matrix* $G_\sigma(t, A)$ for the problem under consideration. $G_\sigma(t, A)$ is defined as a matrix that belongs to W_σ and satisfy to the equation

$$A \frac{d}{dt} G_\sigma(t, A) - G_\sigma(t, A) = \frac{\sin \sigma t}{\pi t} I.$$

It gives representation of the solution

$$x(t) = \int_{-\infty}^{\infty} G_\sigma(t - s, A) f(s) ds$$

of the boundary value problem under considearation.

Using the criterion of well-posedness given by the Theorem 1 we must have reliable answer to the question: *Does a given matrix A have any eigenvalues on the rays $(-i\infty, -i/\sigma] \cup [i/\sigma, i\infty)$ of the imaginary axis?*

We stress that we are not interested in placements of separate eigenvalues of a matrix or in its Jordan structure – only mutual disposition of the spectrum and the rays is crucial.

First we propose a criterion which is somewhat straightforward : *Matrix A has no eigenvalues on the rays if an only if the norm of the symmetric positive definite matrix*

$$H = \int_{-\infty}^{\infty} G_\sigma^*(t, A) G_\sigma(t, A) dt$$

is finite.

Similar criterion was used in works on dichotomy of a matrix spectrum by the imaginary axis [1]. Unfortunately, decay of the Green matrix at the infinity may be slow and the criterion is not useful.

Use of Jukowskii's function allows transform the rays into a circle and reduce the question on disposition of the eigenvalues with respect to the rays to the question on disposition with respect to the unit circle of the eigenvalues of a matrix being a function of the initial matrix. So, we come to

Theorem 2. *The matrix A has no eigenvalues on the rays $(-i\infty, -i/\sigma] \cup [i/\sigma, i\infty)$ if and only if the matrix equation*

$$i\sigma AB^2 - 2B + i\sigma A = 0$$

has unique solution B with the spectrum lying inside the unit cirle.

Remark. It can be shown that matrix B is a functon of matrix A and $B = i\sigma A(I + \sqrt{I + \sigma^2 A^2})^{-1}$.

The Theorem 2 gives us a new form of the criterion: *Matrix A has no eigenvalues on the rays $(-i\infty, -i/\sigma] \cup [i/\sigma, i\infty)$ if and only if the system of matrix equations*

$$\begin{gathered} i\sigma AB^2 - 2B + i\sigma A = 0 \\ B^* H B - H = -I \end{gathered} \tag{2}$$

is uniquely resolvable with respect to the pair of B, H and $H = H^* > 0$.

The norm of the matrix H can be used as a numerical characteristic $\omega_\sigma = \|H\|$ of the quality of the desired property of matrix A – larger values of ω_σ worse the quality.

Remark. It can be shown that solution H to the equation (2) can be represented in the form

$$H = \frac{1}{2}I + \frac{1}{2\pi}\int_{-\sigma}^{\sigma}(i\xi A - I)^{-*}(I + \sqrt{I+\sigma^2A^2})^*(I+\sqrt{I+\sigma^2A^2})(i\xi A - I)^{-1}\frac{d\xi}{\sqrt{\sigma^2-\xi^2}}$$

where $^{-*}$ denotes superposition of inversion and conjugate transposition of a matrix.

2 Solution of matrix equations

In order to solve the system of matrix equations (2) we consider auxilliary boundary value problem for a system of difference equations with constant coefficients

$$\begin{aligned} &i\sigma A u_{j-1} - 2u_j + i\sigma A u_{j+1} = 0, \quad (j = 1, 2, \ldots, 2^k), \\ &u_0 = u, \qquad u_{2^k+1} = v. \end{aligned} \tag{3}$$

for some integer k. Boundary conditions are posed as given values u_0, u_{2^k+1} of the unknown discrete vector-valued function u_j.

In further considerations we assume that matrix A has no eigenvalues on the rays $(-i\infty, -i/\sigma] \cup [i/\sigma, i\infty)$ which is the necessary and sufficient condition for this boundary value problem to be resolvable .

Solution to the problem (3) can be represented in explicit form:

$$u_j = (B^j - B^{4n+2-j})(I - B^{4n+2})^{-1}u + (B^{2n+1-j} - B^{2n+1+j})(I - B^{4n+2})^{-1}v.$$

We need only two of these relations

$$u_1 = B_k u + C_k v, \qquad u_{2^k} = C_k u + B_k v,$$

where

$$\begin{aligned} B_k &= B(I - B^{2^{k+1}})(I - B^{2^{k+1}+2})^{-1}, \\ C_k &= B^{2^k}(I - B^2)(I - B^{2^{k+1}+2})^{-1}. \end{aligned}$$

Since the spectrum of B is supposed to lie inside the unit circle then $B_k \to B, C_k \to 0$ as k tends to infinity. Approximation to B will be obtained if a relation between u_0, u_{2^k}, u_{2^k+1} in the form $u_{2^k} = C_k u_0 + B_k u_{2^k+1}$ is deduced from the equations $i\sigma A u_{j-1} - 2u_j + i\sigma A u_{j+1} = 0$. To do this we

- Rewrite the initial 2^k equations of the second order (3 point) as 2^{k-1} equations of the first order ($j = 0, 1, 2, \dots, 2^{k-1} - 1$):

$$\begin{pmatrix} i\sigma A & -2I \\ 0 & i\sigma A \end{pmatrix} \begin{bmatrix} u_{2j} \\ u_{2j+1} \end{bmatrix} + \begin{pmatrix} i\sigma A & 0 \\ -2I & i\sigma A \end{pmatrix} \begin{bmatrix} u_{2(j+1)} \\ u_{2(j+1)+1} \end{bmatrix} = 0;$$

- For $j = 0, 2, 4, \dots$ exclude

$$\begin{bmatrix} u_{2(j+1)} \\ u_{2(j+1)+1} \end{bmatrix}$$

  from the pair of those equations which relate this vector.

A result will have a form of the system of 2 point difference equations with constant coefficients

$$\begin{pmatrix} D_1 & E_1 \\ F_1 & G_1 \end{pmatrix} \begin{bmatrix} u_{2j} \\ u_{2j+1} \end{bmatrix} + \begin{pmatrix} H_1 & I_1 \\ J_1 & K_1 \end{pmatrix} \begin{bmatrix} u_{2(j+2)} \\ u_{2(j+2)+1} \end{bmatrix} = 0.$$

Repeating the procedure $k-1$ times we come to equations

$$\begin{pmatrix} D_{k-1} & E_{k-1} \\ F_{k-1} & G_{k-1} \end{pmatrix} \begin{bmatrix} u_0 \\ u_1 \end{bmatrix} + \begin{pmatrix} H_{k-1} & I_{k-1} \\ J_{k-1} & K_{k-1} \end{pmatrix} \begin{bmatrix} u_{2^k} \\ u_{2^k+1} \end{bmatrix} = 0$$

Excluding u_1 we obtain necessary form of relations. This completes computations.

Remark that all the above exclusions can be made by use of numerically reliable QR-decomposition.

Rate of convergence of the described procedure can be expressed in terms of the parameter $\omega_\sigma = \|H\|$:

$$\|B_k - B\| \le 2\sqrt{\omega_\sigma}\, q^{2^k+1/2}\Big(1 + \sqrt{\omega_\sigma}\frac{q^{2^k+1}}{1 - q^{2^k+1}}\Big)^{-1}, \qquad q = \frac{\omega_\sigma}{\omega_\sigma + 1}.$$

Matrix H can be obtained by the power doubling method applied to the finite sum

$$H_k = \sum_{n=0}^{2^k-1} (B^*)^n B^n$$

with an error estimate

$$\|H_k - H\| \le \omega_\sigma^2 \left(\frac{\omega_\sigma}{\omega_\sigma + 1}\right)^{2^k-1} .$$

Provided $\omega_\sigma < \omega^*$ all the above estimates can be checked numerically. If any of them fail then the hypothesis $\omega_\sigma < \omega^*$ must be rejected.

References

[1] S.K. Godunov *Modern aspects of linear algebra.* Scientific book, Novosibirsk,1997.