UMIACS-TR-91-105 July 1991
CS-TR 2721 Revised March 1993

Perturbation Theory for Rectangular Matrix Pencils

G. W. Stewart*

ABSTRACT

The theory of eigenvalues and eigenvectors of rectangular matrix pencils is complicated by the fact that arbitrarily small perturbations of the pencil can cause them disappear. However, there are applications in which the properties of the pencil ensure the existence of eigenvalues and eigenvectors. In this paper it is shown how to develop a perturbation theory for such pencils.

*Department of Computer Science and Institute for Advanced Computer Studies, University of Maryland, College Park, MD 20742. This work was supported in part by the Air Force Office of Scientific Research under Contract AFOSR-87-0188. This report is available by anonymous ftp from `thales.cs.umd.edu` in the directory `pub/reports`.

Perturbation Theory for Rectangular Matrix Pencils

G. W. Stewart

ABSTRACT

The theory of eigenvalues and eigenvectors of rectangular matrix pencils is complicated by the fact that arbitrarily small perturbations of the pencil can cause them disappear. However, there are applications in which the properties of the pencil ensure the existence of eigenvalues and eigenvectors. In this paper it is shown how to develop a perturbation theory for such pencils.

In this note we will be concerned with the perturbation theory for the generalized eigenvalue problem[1]

$$\beta Ax = \alpha Bx, \tag{1}$$

where A and B are $m \times n$ matrices with $m > n$ and α and β are normalized so that

$$|\alpha|^2 + |\beta|^2 = 1.$$

Although rectangular matrix pencils have a long history and a well developed theory of canonical forms (e.g., see [1, 5, 6]), their eigenvalues and eigenvectors have been less well studied. There are two reasons.

In the first place, eigenvalues and eigenvectors may fail to exist. For example, if

$$A = \begin{pmatrix} 1 \\ \alpha \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 1 \\ \beta \end{pmatrix},$$

then (1) has no nontrivial solution unless $\alpha = \beta$. In the second place, even if (1) has a solution, an arbitrarily small perturbation can make it go away, as the above example also shows. It is not easy to construct a general perturbation theory for objects that may not exist and can vanish at the drop of a hat.

Nonetheless, in some applications the nature of the matrices A and B insure the existence of eigenvectors, and the conditions that do so are stable under small perturbations (for an example from game theory, see [3]). In these circumstances

[1]The homogeneous form given here seems preferable to the more conventional form $Ax = \lambda Bx$. We call the pair (α, β) an eigenvalue of the problem.

it is reasonable to try to develop a perturbation theory. In this note we shall show how to do so by reducing the problem to a square one.

We begin by describing the space of solutions of the unperturbed problem. First note that equation (1) says that Ax and Bx lie in the same one-dimensional subspace. Generalizing this observation, we say that a subspace $\mathcal{X}$ of dimension k is an eigenspace of the pencil $\beta A - \alpha B$ if there is a subspace $\mathcal{Y}$ of dimension k such that

$$A\mathcal{X} + B\mathcal{X} \subset \mathcal{Y}. \tag{2}$$

Here the sums and products are the usual Minkowski operations; e.g., $A\mathcal{X} = \{\, Ax : x \in \mathcal{X} \,\}$. Since the sum of two eigenspaces is an eigenspace, there is a unique maximal eigenspace, which we shall call the Ω-eigenspace of the pencil. In what follows $\mathcal{X}$ will be the Ω-eigenspace of the pencil $\beta A - \alpha B$ and $\mathcal{Y}$ will be the corresponding subspace in (2). We will assume that $\mathcal{X}$ has dimension $k > 0$.

The relation between eigenspaces and eigenvectors can be seen as follows. Let

$$X = (X_1\ X_2) \quad \text{and} \quad Y = (Y_1\ Y_2)$$

be unitary matrices such that $\mathcal{R}(X_1) = \mathcal{X}$ and $\mathcal{R}(Y_1) = \mathcal{Y}$. Then $Y^{\mathrm{H}}AX$ and $Y^{\mathrm{H}}BX$ have the forms

$$Y^{\mathrm{H}}AX = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix} \quad \text{and} \quad Y^{\mathrm{H}}BX = \begin{pmatrix} B_{11} & B_{12} \\ 0 & B_{22} \end{pmatrix}. \tag{3}$$

It follows that if x is an eigenvector of the pencil $\beta A_{11} - \alpha B_{11}$, then $X_1 x$ is an eigenvector of (1).

The converse is also true: all eigenvectors of (1) have the form $X_1 x$. This is a consequence of the fact that the pencil $\beta A_{22} - \alpha B_{22}$ does not have an eigenspace. For if it did, we could reduce A_{22} and B_{22} as above, so that the transformed pencil assumes the form

$$\beta \begin{pmatrix} A_{11} & \hat{A}_{12} & \hat{A}_{13} \\ 0 & \hat{A}_{22} & \hat{A}_{23} \\ 0 & 0 & \hat{A}_{33} \end{pmatrix} - \alpha \begin{pmatrix} B_{11} & \hat{B}_{12} & \hat{B}_{13} \\ 0 & \hat{B}_{22} & \hat{B}_{23} \\ 0 & 0 & \hat{B}_{33} \end{pmatrix}.$$

Then the original pencil has an eigenspace of dimension greater than k corresponding to the pencil

$$\beta \begin{pmatrix} A_{11} & \hat{A}_{12} \\ 0 & \hat{A}_{22} \end{pmatrix} - \alpha \begin{pmatrix} B_{11} & \hat{B}_{12} \\ 0 & \hat{B}_{22} \end{pmatrix},$$

which contradicts the maximality of the Ω-subspace.

In what follows, we will assume that the matrices of the pencil $\beta A - \alpha B$ are in the *reduced form* (3). Note that since the reduction is a unitary equivalence any perturbation in the original pencil corresponds (in a unitarily invariant norm) to a perturbation of the same size in the reduced form.

Let us now consider the perturbed matrices

$$\tilde{A} = \begin{pmatrix} \tilde{A}_{11} & \tilde{A}_{12} \\ E_{21} & \tilde{A}_{22} \end{pmatrix} \equiv \begin{pmatrix} A_{11} + E_{11} & A_{12} + E_{12} \\ E_{21} & A_{22} + E_{22} \end{pmatrix}$$

and

$$\tilde{B} = \begin{pmatrix} \tilde{B}_{11} & \tilde{B}_{12} \\ F_{21} & \tilde{B}_{22} \end{pmatrix} \equiv \begin{pmatrix} B_{11} + F_{11} & B_{12} + F_{12} \\ F_{21} & B_{22} + F_{22} \end{pmatrix},$$

where

$$\|E_{ij}\|, \|F_{ij}\| \le \epsilon, \qquad i, j = 1, 2. \tag{4}$$

Here $\|\cdot\|$ denotes both the Euclidean vector norm and the subordinate spectral matrix norm. Let $\tilde{x} = (\tilde{x}_1^{\mathrm{H}}\ \tilde{x}_2^{\mathrm{H}})^{\mathrm{H}}$ be a normalized eigenvector of the corresponding pencil:

$$\tilde{\beta}\tilde{A}\tilde{x} = \tilde{\alpha}\tilde{B}\tilde{x}, \qquad \|\tilde{x}\| = 1.$$

We are going to show that under a natural condition (namely, that $\tilde{\tau}$ defined below is not small) the second component $\tilde{x}_2$ of $\tilde{x}$ is small.

First note that since the pencil $\beta A_{22} - \alpha B_{22}$ has no eigenvectors, there is a number $\tau > 0$ such that

$$\|\tilde{\beta} A_{22}\tilde{x}_2 - \tilde{\alpha} B_{22}\tilde{x}_2\| \ge \tau\|\tilde{x}_2\|. \tag{5}$$

Since $|\tilde{\alpha}|^2 + |\tilde{\beta}|^2 = 1$, if

$$\tilde{\tau} \equiv \tau - \sqrt{2}\epsilon > 0, \tag{6}$$

then

$$\|\tilde{\beta}\tilde{A}_{22}\tilde{x}_2 - \tilde{\alpha}\tilde{B}_{22}\tilde{x}_2\| \ge \tilde{\tau}\|\tilde{x}_2\|.$$

Since $\tilde{x}$ is an eigenvector

$$0 = \|\tilde{\beta}(E_{21}\tilde{x}_1 - \tilde{A}_{22}\tilde{x}_2) - \tilde{\alpha}(F_{21}\tilde{x}_1 - \tilde{B}_{22}\tilde{x}_2)\| \ge \tilde{\tau}\|\tilde{x}_2\| - \|\tilde{\beta}E_{21}\tilde{x}_1 - \tilde{\alpha}F_{21}\tilde{x}_1\|.$$

Hence

$$\|\tilde{x}_2\| \le \frac{\sqrt{2}\epsilon}{\tilde{\tau}}.$$

Now let us perform the reduction procedure describe above on the perturbed pencil, but using only the one-dimensional space spanned by the eigenvector $\tilde{x}$. Specifically, let

$$X = \begin{pmatrix} R & Q \\ P & S \end{pmatrix}$$

be a unitary matrix whose first column is $\tilde{x}$ and for which

$$\|P\|, \|Q\| \le \|\tilde{x}_2\| \le \frac{\sqrt{2}\epsilon}{\tilde{\tau}}. \tag{7}$$

(The existence of of such a matrix is established in the appendix.) Let

$$Y = \begin{pmatrix} \hat{R} & \hat{Q} \\ \hat{P} & \hat{S} \end{pmatrix}$$

be a unitary matrix whose first column is proportional to $\tilde{A}x$ (or $\tilde{B}x$ if $\tilde{A}x = 0$). Then the first k columns of $Y^{\mathrm{H}}\tilde{A}X$ and $Y^{\mathrm{H}}\tilde{B}X$ have the forms

$$\begin{pmatrix} \tilde{\alpha} & \tilde{a}^{\mathrm{H}} \\ 0 & \tilde{A}_* \\ 0 & E_* \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \tilde{\beta} & \tilde{b}^{\mathrm{H}} \\ 0 & \tilde{B}_* \\ 0 & F_* \end{pmatrix}. \tag{8}$$

Thus $(\tilde{\alpha}, \tilde{\beta})$ is an eigenvalue of the pencil

$$\beta \begin{pmatrix} \tilde{\alpha} & \tilde{a}^{\mathrm{H}} \\ 0 & \tilde{A}_* \end{pmatrix} - \alpha \begin{pmatrix} \tilde{\beta} & \tilde{b}^{\mathrm{H}} \\ 0 & \tilde{B}_* \end{pmatrix}$$

But by direct computation, this pencil is

$$\begin{aligned} &\beta(\hat{R}^{\mathrm{H}}A_{11}R + \hat{R}^{\mathrm{H}}E_{11}R + \hat{R}^{\mathrm{H}}\tilde{A}_{12}P + \hat{Q}^{\mathrm{H}}E_{21}R + \hat{Q}^{\mathrm{H}}\tilde{A}_{22}P) - \\ &\quad \alpha(\hat{R}^{\mathrm{H}}B_{11}R + \hat{R}^{\mathrm{H}}F_{11}R + \hat{R}^{\mathrm{H}}\tilde{A}_{12}P + \hat{Q}^{\mathrm{H}}F_{21}R + \hat{Q}^{\mathrm{H}}\tilde{A}_{22}P) \end{aligned}$$

or

$$\beta(\hat{R}^{\mathrm{H}}A_{11}R + G) - \alpha(\hat{R}^{\mathrm{H}}B_{11}R + H),$$

where from (4) and (7)

$$\|G\|, \|H\| \le \left(2 + 2\sqrt{2}\frac{\max\{\|\tilde{A}\|, \|\tilde{B}\|\}}{\tilde{\tau}}\right)\epsilon.$$

We have thus shown that the eigenvalue $(\tilde{\alpha}, \tilde{\beta})$ is an eigenvalue of a perturbation of the pencil $\beta\hat{R}^{\mathrm{H}}A_{11}R - \alpha\hat{R}^{\mathrm{H}}B_{11}R$, a pencil which corresponds to the Ω-eigenspace of the original problem. Since the pencil and its perturbation are square, we may use standard perturbation theory to bound the perturbation in the eigenvalue and its eigenvector (for a survey of the perturbation theory of the generalized eigenvalue problem, see [4, Ch. VI]).

The dependence of the bounds on τ defined by (5) is entirely natural. If τ is small, then $(\tilde{\alpha}, \tilde{\beta})$ is almost an eigenvalue of the pencil $\beta A_{22} - \alpha B_{22}$ and could have come from the perturbed pencil $\beta\tilde{A}_{22} - \alpha\tilde{B}_{22}$. Such an eigenvalue need not be near an eigenvalue associated with the Ω-eigenspace of the original problem. Seen in this light, the condition (6) insures that the perturbed eigenvalue truly comes from the Ω-eigenspace of the original problem.

Finally, we note that the general technique we have outlined here is at least as important as the particular bounds. For example, it can be extended to obtain perturbations bounds on eigenspaces. Again, if we work with nonorthogonal diagonalizing transformations, so that $A_{12} = B_{12} = 0$ in (3) and $\tilde{a}^{\mathrm{H}} = \tilde{b}^{\mathrm{H}} = 0$ in (8), we obtain the result that

$$\tilde{\alpha} \propto \alpha + y^{\mathrm{H}}Ex + O(\max\{\|E\|, \|F\|\}^2)$$

and

$$\tilde{\beta} \propto \beta + y^{\mathrm{H}}Fx + O(\max\{\|E\|, \|F\|\}^2),$$

where y^{H} is the left eigenvector corresponding to (α, β). The details are left to the reader.

Appendix

The following extension theorem is more general than we need but may be of independent interest.

Let $p \geq q$ and let the matrix

$$X_1 = \begin{array}{cc} & q \\ p & \\ n-p & \end{array}\!\!\!\begin{pmatrix} X_{11} \\ X_{21} \end{pmatrix}$$

have orthonormal columns. Then there is a unitary matrix

$$X = \begin{array}{c} \\ p \\ n-p \end{array} \begin{array}{c} \begin{array}{ccc} q & p-q & n-p \end{array} \\ \begin{pmatrix} X_{11} & X_{12} & X_{13} \\ X_{21} & X_{22} & X_{23} \end{pmatrix} \end{array}$$

such that

$$\|(X_{21}\ \ X_{22})\| = \|X_{21}\|.$$

To establish the theorem, first assume that $p+q \le n$. By the CS decomposition (e.g., see [2, p.77]) we may assume that

$$X_1 = \begin{array}{c} \\ q \\ p-q \\ q \\ n-p-q \end{array} \begin{array}{c} q \\ \begin{pmatrix} C \\ 0 \\ S \\ 0 \end{pmatrix} \end{array},$$

where C and S are nonnegative diagonal matrices. Then the required extension is

$$X = \begin{array}{c} \\ q \\ p-q \\ q \\ n-p-q \end{array} \begin{array}{c} \begin{array}{cccc} q & p-q & q & n-p-q \end{array} \\ \begin{pmatrix} C & 0 & -S & 0 \\ 0 & I & 0 & 0 \\ S & 0 & C & 0 \\ 0 & 0 & 0 & I \end{pmatrix} \end{array},$$

in which we take

$$(X_{21}\ \ X_{22}) = \begin{pmatrix} S & 0 \\ 0 & 0 \end{pmatrix}.$$

On the other hand, if $p+q > n$, we take X_1 in the form

$$X_1 = \begin{array}{c} \\ p+q-n \\ n-p \\ p-q \\ n-p \end{array} \begin{array}{c} \begin{array}{cc} p+q-n & n-p \end{array} \\ \begin{pmatrix} I & 0 \\ 0 & C \\ 0 & 0 \\ 0 & S \end{pmatrix} \end{array}$$

In this case the required extension is

$$X = \begin{array}{c} \\ p+q-n \\ n-p \\ p-q \\ n-p \end{array} \begin{array}{c} \begin{array}{cccc} p+q-n & n-p & p-q & n-p \end{array} \\ \begin{pmatrix} I & 0 & 0 & 0 \\ 0 & C & 0 & -S \\ 0 & 0 & I & 0 \\ 0 & S & 0 & C \end{pmatrix} \end{array},$$

in which

$$(X_{21}\ \ X_{22}) = (0\ \ S\ \ 0).$$

References

[1] F. R. Gantmacher (1959). *The Theory of Matrices, Vols. I, II.* Chelsea Publishing Company, New York.

[2] G. H. Golub and C. F. Van Loan (1989). *Matrix Computations.* Johns Hopkins University Press, Baltimore, Maryland, 2nd edition.

[3] E. Marchi, J. A. Oviedo, and J. E. Cohen (1991). "Pertrubation Theory of a Nonlinear Game of von Neuman." *SIAM Journal on Matrix Analysis and Applications*, **12**, 592–596.

[4] G. W. Stewart and G.-J. Sun (1990). *Matrix Perturbation Theory.* Academic Press, Boston.

[5] P. Van Dooren (1979). "The Computation of Kronecker's Canonical Form of a Singular Pencil." *Linear Algebra and Its Applications*, **27**, 103–140.

[6] J. H. Wilkinson (1979). "Kronecker's Canonical Form and the QZ Algorithm." *Linear Algebra and Its Applications*, **28**, 285–303.