TR-1923 September 1987

Invariant Subspaces and Capital Punishment (A Participatory Paper)

G. W. Stewart*

ABSTRACT

The notion of invariant subspaces is useful in a number of theoretical and practical applications. In this paper we give an elementary treatment of invariant subspaces that stresses their connection with simple eigenvalues and their eigenvectors.

*Department of Computer Science and Institute for Physical Science and Technology, University of Maryland, College Park, MD 20742.

Invariant Subspaces
and
Capital Punishment
(A Participatory Paper)

G. W. Stewart*

1 Introduction

The notion of an invariant subspace of a matrix or a linear operator is useful in many branches of mathematics, both pure and applied. Unfortunately, many students have difficulty with the elementary facts about invariant subspaces even though they correspond exactly to equivalent facts about eigenvectors. The purpose of this note is to exhibit this correspondence by developing the theory of a simple eigenvalue and its eigenvectors in such a way that the generalization to invariant subspaces is obvious. We shall be concerned with two aspects of the subject: the constructive algebraic theory and first order perturbation expansions.

To make the relation between eigenvectors and invariant subspaces clear, we must be careful about our notation. In this note, scalars will be denoted by lower-case Greek letters, vectors by lower-case Latin letters, and matrices by upper-case Latin letters. The symbol $\|\cdot\|$ will denote the Euclidean norm; i.e.,

$$\|x\|^2 = x^{\mathrm{H}}x,$$

where x^{H} is the conjugate transpose of x. The same symbol will also denote the spectral norm of a matrix, which is defined by

$$\|A\| = \max_{\|x\|=1} \|Ax\|. \tag{1.1}$$

We shall have occasion to refer to linear operators that map matrices into matrices, and the norms of such operators will be defined in analogy with (1.1). For example if $\mathbf{T}$ is an operator, then

$$\|\mathbf{T}\| = \max_{\|P\|=1} \|\mathbf{T}P\|.$$

For more on norms see [4] or [12].

*Department of Computer Science and Institute for Physical Science and Technology, University of Maryland, College Park, 20742.

2 Eigenvectors and eigenvalues

Simply stated, an eigenvector of an $n \times n$ matrix A is a nonzero vector whose direction remains the same when it is multiplied by A. It may stretch, or shrink, or even change its orientation; but it always emerges a scalar multiple of itself. Thus if $x_1 \neq 0$ is an eigenvector of A, there is a unique scalar λ_1 such that

$$Ax_1 = \lambda_1 x_1.$$

The scalar λ_1 is called the eigenvalue of A corresponding to x_1.

I will suppose that the reader is familiar with the following elementary facts about eigenvalues and eigenvectors, which can be found in any respectable textbook on linear algebra. The eigenvalues of A are the zeros of the characteristic polynomial $\det(\lambda I - A)$, and hence a matrix of order n has n eigenvalues, counting multiplicities. If X is nonsingular, then the similarity transformation $X^{-1}AX$ leaves the eigenvalues of A unchanged. Moreover, the eigenvector x_1 is transformed into $X^{-1}x_1$. Finally, if A can be partitioned in the form

$$A = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix}, \tag{2.1}$$

where A_{11} is $m \times m$, then the m eigenvalues of A_{11} are a subset of those of A, while the $n-m$ eigenvalues of A_{22} form the complementary subset (here, as always, we count multiplicities).

We will be concerned with two questions about eigenvectors and their eigenvalues. First, there is the problem of left eigenvectors. If λ_1 is an eigenvalue of A then $\lambda_1 I - A$ is singular and has a left null vector. In other words, there is a vector y_1 such that

$$y_1^{\mathrm{H}} A = \lambda_1 y_1^{\mathrm{H}}.$$

The vector y_1 is called a *left eigenvector* of A, and it is natural to ask if there is any nice relation between left and right eigenvectors belonging to the same eigenvalue.

Another question concerns the behavior of an eigenvalue and its eigenvector when A is perturbed. Specifically, suppose A is replaced by $\tilde{A} = A+E$ for some small matrix E. Then we may ask how the eigenvalues and eigenvectors of $\tilde{A}$ behave as a function of E.

Unfortunately, there is no easy answer to these questions. Part of the problem is that an eigenvalue may have two or more linearly independent eigenvectors. Moreover, the eigenvalues of $\tilde{A}$ need not be nice, smooth functions of E. For example consider the matrix

$$\tilde{A} = \begin{pmatrix} 1 & 1 \\ \epsilon & 1 \end{pmatrix}. \tag{2.2}$$

The eigenvalues of $\tilde{A}$ are easily seen to be $1 \pm \sqrt{\epsilon}$, neither of which is differentiable at $\epsilon = 0$.

There is, however, a case where these questions have straightforward answers. When λ_1, regarded as a zero of the characteristic polynomial, has multiplicity one, it is called a *simple* eigenvalue. Its eigenvector x_1, which is unique up to a scalar multiple, is called a simple eigenvector. In the next three sections, we will show that there is a pretty algebraic theory relating a simple eigenvector with its corresponding left eigenvector. Moreover, simple eigenvectors and eigenvalues behave nicely under perturbations. Throughout these three sections, we will assume that the pair λ_1 and x_1 are simple. Anticipating some later generalizations, we will express this fact as follows. Let $\Lambda_1 = \{\lambda_1\}$ and let Λ_2 be the set consisting of the remaining $n-1$ eigenvalues of A. Then λ_1 is simple if and only if $\Lambda_1 \cap \Lambda_2 = \emptyset$.

One final assumption. Since x_1 remains an eigenvector when it is multiplied by a nonzero constant, we may assume without loss of generality that

$$x_1^{\mathrm{H}} x_1 = 1.$$

3 The constructive algebraic theory[1,2]

The principal result of this section is that to λ_1 and x_1 there corresponds a y_1 of the same dimensions as x_1 such that

$$y_1^{\mathrm{H}} A = \lambda_1 y_1^{\mathrm{H}}.$$

Moreover $y_1^{\mathrm{H}} x_1 = 1$, so that x_1 and y_1 are not orthogonal. However, the way we achieve this result is as important as the result itself. We shall show how to construct a similarity transformation that reduces A to the block diagonal form

$$\begin{pmatrix} \lambda_1 & 0 \\ 0 & L_2 \end{pmatrix}.$$

[1] Do not read the footnotes the first time through the next three sections: they will only confuse you.

[2] Do not read the footnotes the second time through: they will only distract you

The formulas that define this transformation are themselves of theoretical interest.

The first step is to construct a matrix Y_2 with orthonormal columns such that $(x_1\ Y_2)$ is unitary. This can be done constructively by any of several numerical techniques.[3]

Next consider the matrix

$$\begin{pmatrix} x_1^{\mathrm{H}} \\ Y_2^{\mathrm{H}} \end{pmatrix} A(x_1\ Y_2) = \begin{pmatrix} x_1^{\mathrm{H}} A x_1 & x_1^{\mathrm{H}} A Y_2 \\ Y_2^{\mathrm{H}} A x_1 & Y_2^{\mathrm{H}} A Y_2 \end{pmatrix}.$$

Since $Ax_1 = x_1\lambda_1$, $x_1^{\mathrm{H}} x_1 = 1$, and $Y_2^{\mathrm{H}} x_1 = 0$, we have

$$\begin{pmatrix} x_1^{\mathrm{H}} \\ Y_2^{\mathrm{H}} \end{pmatrix} A(x_1\ Y_2) = \begin{pmatrix} \lambda_1 & b^{\mathrm{H}} \\ 0 & L_2 \end{pmatrix},$$

where $b^{\mathrm{H}} = x_1^{\mathrm{H}} A Y_2$ and

$$L_2 = Y_2^{\mathrm{H}} A Y_2.$$

Thus this unitary similarity transformation reduces the problem of finding the eigenvalues of A to that of finding those of L_2, since by the comments concerning equation (2.1) the eigenvalues of L_2 form the set Λ_2.[4]

The next step is to remove the vector b^{H} by a similarity transformation of the form

$$\begin{pmatrix} 1 & -p^{\mathrm{H}} \\ 0 & I \end{pmatrix} \begin{pmatrix} \lambda_1 & b^{\mathrm{H}} \\ 0 & L_2 \end{pmatrix} \begin{pmatrix} 1 & p^{\mathrm{H}} \\ 0 & I \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & L_2 \end{pmatrix} \tag{3.1}$$

If we equate the $(1,2)$-blocks on each side of (3.1), we get the following equation for p^{H}:

$$\lambda_1 p^{\mathrm{H}} - p^{\mathrm{H}} L_2 = b^{\mathrm{H}}. \tag{3.2}$$

---

[3] For example, $(X_1\ Y_2)$ can be computed as a product of Householder transformations. See [4] or [12].

[4] The above reduction can be repeated on L_2 and so on. When an eigenvector is used at each stage, the result is that A is reduced to upper triangular form by a unitary similarity transformation—a very useful decomposition due to Schur [9]. When a sequence of invariant subspaces is used, one arrives at a block triangular form. A particularly important application occurs when A is real but λ_1 is complex. If we take the columns of X_1 to form a real basis for the two dimensional space spanned by the eigenvector of λ_1 and and its conjugate transpose, which is an eigenvector of $\bar{\lambda}_1$, then the transformation is real, and L_1 is a real 2×2 matrix whose eigenvalues are λ_1 and $\bar{\lambda}_1$. Continuing in the same manner, we get a variant of the Schur decomposition in which all the complex eigenvalues are contained in 2×2 blocks on the diagonal, a form computed by the widely used QR algorithm [4, Ch. 7].

We are not quite finished, since it is possible that (3.2) has no solution or more than one solution. However, (3.2) is equivalent to the linear operator equation $\mathbf{S}p^{\mathrm{H}} = b^{\mathrm{H}}$, where $\mathbf{S}$ is defined by

$$\mathbf{S}p^{\mathrm{H}} \stackrel{\mathrm{def}}{=} \lambda_1 p^{\mathrm{H}} - p^{\mathrm{H}} L_2.$$

Now it can be shown that a necessary and sufficient condition for $\mathbf{S}$ to be nonsingular is that $\Lambda_1 \cap \Lambda_2 = \emptyset$, which is true since x_1 is simple.[5] Thus p^{H} exists and is unique.

Let us now multiply out the two transformations by which we reduced A to the block diagonal form (3.1). On the right we have

$$(x_1\ Y_2)\begin{pmatrix} 1 & p^{\mathrm{H}} \\ 0 & I \end{pmatrix} = (x_1\ \ x_1 p^{\mathrm{H}} + Y_2) = (x_1\ X_2),$$

where

$$X_2 = x_1 p^{\mathrm{H}} + Y_2.$$

On the left we have

$$\begin{pmatrix} 1 & -p^{\mathrm{H}} \\ 0 & I \end{pmatrix}\begin{pmatrix} x_1^{\mathrm{H}} \\ Y_2^{\mathrm{H}} \end{pmatrix} = \begin{pmatrix} x_1^{\mathrm{H}} - p^{\mathrm{H}} Y_2^{\mathrm{H}} \\ Y_2^{\mathrm{H}} \end{pmatrix} = \begin{pmatrix} y_1^{\mathrm{H}} \\ Y_2^{\mathrm{H}} \end{pmatrix},$$

where

$$y_1^{\mathrm{H}} = x_1^{\mathrm{H}} - p^{\mathrm{H}} Y_2^{\mathrm{H}}.$$

Thus we have shown that we can replace the original unitary matrix $(x_1\ Y_2)$ by two matrices $(x_1\ X_2)$ and $(y_1\ Y_2)$ such that

$$\begin{pmatrix} y_1^{\mathrm{H}} \\ Y_2^{\mathrm{H}} \end{pmatrix}(x_1\ X_2) = \begin{pmatrix} 1 & 0 \\ 0 & I \end{pmatrix} \tag{3.3}$$

[i.e., $(x_1\ X_2)^{-1} = (y_1\ Y_2)^{\mathrm{H}}$] and

$$\begin{pmatrix} y_1^{\mathrm{H}} \\ Y_2^{\mathrm{H}} \end{pmatrix} A(x_1\ X_2) = \begin{pmatrix} \lambda_1 & 0 \\ 0 & L_2 \end{pmatrix}. \tag{3.4}$$

[5] This is obvious for eigenvectors, since the operator is represented by the matrix $\lambda_1 I - L_2$. The proof in the general case is more difficult. One approach is to use Schur's theorem to reduce L_1 and L_2 to upper triangular matrices [1]. This results in a constructive algorithm for solving (3.2).

If we use (3.3) to rewrite (3.4) in the form

$$\begin{pmatrix} y_1^{\mathrm{H}} \\ Y_2^{\mathrm{H}} \end{pmatrix} A = \begin{pmatrix} \lambda_1 & 0 \\ 0 & L_2 \end{pmatrix} \begin{pmatrix} y_1^{\mathrm{H}} \\ Y_2^{\mathrm{H}} \end{pmatrix},$$

it immediately follows that $y_1^{\mathrm{H}} A = \lambda_1 y_1^{\mathrm{H}}$; i.e. y_1 is invariant with respect to A^{H}, just as x_1 is invariant with respect to A. Moreover, from (3.2) it follows that $y_1^{\mathrm{H}} x_1 = 1$, which is what we set out to prove.[6]

The equations (3.3) and (3.4) can be cast in another useful form. Let

$$P_1 = x_1 y_1^{\mathrm{H}},$$

and let z be an arbitrary vector. From (3.3) we see that z can be written in the form $z = x_1\gamma_1 + X_2 g_2$, where $\gamma_1 = y_1^{\mathrm{H}} z$ and $g_2 = Y_2^{\mathrm{H}} z$. It then follows that

$$P_1 z = x_1 \gamma_1.$$

Geometrically this says that P_1 projects z onto the space spanned by x_1 along the direction of the space spanned by X_2. For this reason, P_1 is sometimes called the *spectral projection* associated with x_1.

Now let

$$P_2 = X_2 Y_2^{\mathrm{H}}$$

and

$$A_i = P_i A P_i \quad (i = 1, 2).$$

Then it is easily verified that $A_1 A_2 = A_2 A_1 = 0$ and

$$A = A_1 + A_2. \tag{3.5}$$

Equation (3.5) is sometimes called a *spectral decomposition* of A. It says that the matrix A, regarded as an operator, can be decomposed into the direct sum of two operators, one mapping the space spanned by x_1 onto itself and the other mapping the space spanned by X_2 onto itself.[7]

[6] We have actually established more. It is easily verified that $AX_2 = X_2 L_2$. Thus the space spanned by X_2 is also an invariant subspace of A. It is called the *complementary* invariant subspace.

[7] If we are given a sequence of independent invariant subspaces $X_1, X_2, \ldots, X_m$ that span $\mathcal{R}^n$, then we may construct corresponding left invariant subspaces $Y_1, Y_2, \ldots, Y_m$ and spectral projections $P_i = X_i Y_i^{\mathrm{H}}$. The analogue of (3.5) then becomes $A = \sum_{i=1}^{m} P_i A P_i$. When A is Hermitian, $m = n$, and we replace the sum by a Stieltje's integral, we obtain a finite dimensional analogue of the spectral representation theorem for self-adjoint operators in Hilbert space.

4 Perturbations of λ_1

In the next two sections we will consider what happens to λ_1 and x_1 when A is replaced by $\tilde{A} = A + E$, where E is a small error matrix. We will use an important technique, called *first order perturbation theory,* for computing approximations to the perturbations. As a general method, it begins with a nonlinear equation that implicitly defines the perturbed quantity as a function of the error. In our case, we shall be interested in the perturbed values $\tilde{\lambda}_1$ and $\tilde{x}_1$. Next we throw out of the equation all terms that are products of the error, either with itself or with the perturbation; i.e., all terms of order higher than the first. This gives us a *linear* equation which holds with increasing precision as the error approaches zero. If this linear equation has a unique solution, then the solution must approximate the perturbation with increasing accuracy as the error approaches zero.

It is clear that if we are going to cast off products of errors and perturbations, we must first show that the perturbation approaches with the error. The easiest way to do this is to establish that the perturbed quantity is a differentiable function of the error.

For our problem, there is the complicating factor that x_1 needs some kind of normalization to make it unique. We shall require that

$$w^{\mathrm{H}}\tilde{x}_1 = 1 \tag{4.1}$$

for a fixed w which we will choose later.[8] It can then be shown that $\tilde{x}_1$ as a function of E has derivatives of all orders in some neighborhood of $E = 0$.[9] The same is *a fortiori* true of

$$\tilde{\lambda}_1 = (\tilde{x}_1^{\mathrm{H}}\tilde{x}_1)^{-1}\tilde{x}_1^{\mathrm{H}}A\tilde{x}_1.$$

All this justifies the following procedure for approximating $\tilde{\lambda}_1$ up to first order terms in E. Write

$$\tilde{\lambda}_1 = \lambda_1 + \mu$$

and

$$\tilde{x}_1 = x_1 + u.$$

Then

$$(A + E)(x_1 + u) = (x_1 + u)(\lambda_1 + \mu). \tag{4.2}$$

[8] Other normalizations in which w varies are possible. The most common is $w = \tilde{x}_1$, which amounts to imposing the condition $\|\tilde{x}_1\| = 1$. The role of normalization in the perturbation theory for eigenvectors is treated in detail in [7].

[9] This result is not trivial to establish. See for example [6].

Since $\mu, u = O(\|E\|)$, we may discard cross products to obtain[10]

$$Ax_1 + Ex_1 + Au = x_1\lambda_1 + u\lambda_1 + x_1\mu + O(\|E\|^2),$$

or since $Ax_1 = x_1\lambda_1$,

$$Ex_1 + Au = u\lambda_1 + x_1\mu + x_1\mu + O(\|E\|^2). \tag{4.3}$$

To solve (4.3) for μ we will choose $w^{\mathrm{H}} = y_1^{\mathrm{H}}$. It is easy to show that the normalization equation (4.1) will be satisfied if and only if $u = X_2 q$ for some q. Hence from (4.3)

$$x_1\mu = Ex_1 + X_2 q\lambda_1 + AX_2 q + O(\|E\|^2). \tag{4.4}$$

Now $y_1^{\mathrm{H}} X_2 q\lambda_1 = 0$ and $y_1^{\mathrm{H}} AX_2 q = \lambda_1 y_1^{\mathrm{H}} X_2 q = 0$. Hence on multiplying (4.4) by y^{H} we get

$$\mu = y_1^{\mathrm{H}} Ex_1 + O(\|E\|^2),$$

so that our approximation to $\tilde{\lambda}_1$ becomes

$$\tilde{\lambda}_1 \cong y_1^{\mathrm{H}}(A+E)x_1 = y_1^{\mathrm{H}}\tilde{A}x_1$$

The quantity $y_1^{\mathrm{H}}\tilde{A}x_1$ is sometimes called the generalized Rayleigh quotient because it generalizes the usual Rayleigh quotient for a Hermitian matrix.

Since $\|x_1\| = 1$, we have

$$\|\mu\| \le \|y_1\|\|E\| + O(\|E\|^2). \tag{4.5}$$

Since $1 \le \|y_1\|\|x_1\| = \|y_1\|$, the number $\|y_1\|$ in (4.5) can be interpreted as a factor which tells how $\|E\|$ is magnified in its effect on $\tilde{\lambda}_1$. Such a number is often called a *condition number* by numerical analysts.[11]

It is worth observing that $\|y_1\| = \|P_1\|$, where P_1 is the spectral projection introduced in §3. Thus the condition number of λ_1 is the norm of its spectral projector.[12]

---

[10] Actually, we are being a little heavy handed here, as was pointed out to me by B. N. Parlett. If we multiply by Y_2^{H} before we drop second order terms we obtain the equation

$$Y_2^{\mathrm{H}} EX_1 + Y_2^{\mathrm{H}} EU = M,$$

from which it is clear that the second order effects of the error will depend on how sensitive X_1 is to perturbations. As a general rule, one should try to get rid of higher order terms by algebraic means, before simply throwing them away.

[11] J. H. Wilkinson [13] appears to be the first to explicitly point out the role of this number, or rather its reciprocal.

[12] When it is a matter of eigenvalues, $\|y_1\|$ can be regarded as the secant of the angle between x_1 and y_1. Thus a simple eigenvalue becomes increasingly ill conditioned as its left and right eigenvectors approach orthogonality. Observe that for the matrix (2.2) the left and right eigenvectors are exactly orthogonal when $\epsilon = 0$.

5 Perturbations of x_1

The first step toward approximating $\tilde{x}_1$ is to derive an equation, like (4.2), from which we may cast out second order terms. To do this we make the following observation. Let $\tilde{Z}$ be a matrix whose columns span the orthogonal complement of the space spanned by $\tilde{x}_1$. Since $\tilde{A}\tilde{x}_1 = \tilde{x}_1\tilde{\lambda}_1$, we must have

$$\tilde{Z}^{\mathrm{H}}\tilde{A}\tilde{x}_1 = \tilde{Z}^{\mathrm{H}}\tilde{x}_1\tilde{\lambda}_1 = 0.$$

Conversely if

$$\tilde{Z}^{\mathrm{H}}\tilde{A}\tilde{x}_1 = 0, \tag{5.1}$$

then $\tilde{A}\tilde{x}_1$ must lie in the orthogonal complement of $\tilde{Z}$; that is, in the span of $\tilde{x}_1$. Consequently (5.1) characterizes the space spanned by $\tilde{x}_1$.

To use (5.1) we require an explicit representation of $\tilde{x}_1$. In this case, it is most convenient to choose $w = x_1$ as the normalizing factor in (4.1),[13] which means that $\tilde{x}_1$ can be written in the form

$$\tilde{x}_1 = x_1 + Y_2 r,$$

for some r to be determined. If we now take $\tilde{Z} = Y_2 - x_1 r^{\mathrm{H}}$, then it is easily verified that $\tilde{Z}^{\mathrm{H}}\tilde{x}_1 = 0$, so that $\tilde{Z}$ is orthogonal to $\tilde{x}_1$. Moreover, since $Y_2^{\mathrm{H}}\tilde{Z} = Y_2^{\mathrm{H}}Y_2 = I$, the columns of $\tilde{Z}$ are independent for any value of r, and hence they form a basis for the orthogonal complement of $\tilde{x}_1$.

With these definitions, it follows from (3.4) that (5.1) can be written

$$-r\lambda_1 + Y_2^{\mathrm{H}}Ex_1 - rx_1^{\mathrm{H}}Ex_1 + L_2 r - rb^{\mathrm{H}}r + Y_2^{\mathrm{H}}EY_2 r - rx_1^{\mathrm{H}}EY_2 r = 0.$$

Discarding second order terms, we get

$$r\lambda_1 - L_2 r = Y_2^{\mathrm{H}}Ex_1 + O(\|E\|^2).$$

Because $\Lambda_1 \cap \Lambda_2 = \emptyset$, the operator defined by

$$\mathbf{T}r \stackrel{\text{def}}{=} r\lambda_1 - L_2 r$$

is nonsingular. Hence

$$r = \mathbf{T}^{-1}Y_2^{\mathrm{H}}Ex_1 + O(\|E\|^2),$$

[13] The careful reader will note that the normalizing factor is different from the one we chose in the last section. This means that the perturbed representation $(\tilde{X}_1^{\mathrm{H}}\tilde{X}_1)^{-1}\tilde{X}_1^{\mathrm{H}}A\tilde{X}_1$ will be different from the representation $\tilde{L}_1$ of §4, although the two will be similar. It is an interesting exercise to verify that the approximate representations in this and the last section are similar up to terms of order $\|E\|^2$.

and

$$\|\tilde{x}_1 - x_1\| \le \|\mathbf{T}^{-1}\|\|E\| + O(\|E\|^2).$$

Thus $\mathbf{T}^{-1}$ is a condition number for $\tilde{x}_1$.

Although there are no simple expressions for $\|\mathbf{T}^{-1}\|$, we can relate it to the eigenvalues of A as follows. It can be shown that the set of eigenvalues of $\mathbf{T}$ is the set $\Lambda_1 - \Lambda_2$; i.e., the set of pairwise differences of members of Λ_1 and Λ_2. Hence the eigenvalues of $\mathbf{T}^{-1}$ are the reciprocals of the members of $\Lambda_1 - \Lambda_2$. Since the norm of any operator is an upper bound for the magnitudes of the eigenvalues of the operator, we have

$$\|\mathbf{T}^{-1}\| \ge \frac{1}{\min|\Lambda_1 - \Lambda_2|}.$$

Thus if a member of Λ_1 is poorly separated from the members of Λ_2, then $\tilde{x}_1$ is necessarily sensitive to perturbations in A. Unfortunately, the converse is not true.[14]

6 Invariant Subspaces

We noted in §2 that the space spanned by an eigenvector does not change when it is multiplied by A. A natural generalization of this is to say that a k-dimensional subspace $\mathcal{X}_1$ is an *invariant subspace* of A if each vector in $\mathcal{X}_1$ is mapped by A back into $\mathcal{X}_1$; that is, if

$$A\mathcal{X}_1 \stackrel{\text{def}}{=} \{Ax : x \in \mathcal{X}_1\} \subset \mathcal{X}_1.$$

To establish some of the elementary properties of invariant subspaces, let the columns of X_1 form a basis for $\mathcal{X}_1$. Then the column space of AX_1 can be written as a linear combination of X_1; that is, there is a $k \times k$ matrix L_1 such that

$$AX_1 = X_1 L_1. \tag{6.1}$$

[14] As an exercise consider the sensitivity of the eigenvector $(1, 0, \ldots, 0)^{\mathrm{T}}$ of the matrix $\mathrm{diag}(0, W_{n-1})$, where W_k is the matrix illustrated below for $k = 4$:

$$\begin{pmatrix} 1 & -1 & -1 & -1 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Note that since the columns of X_1 are linearly independent, $X_1^{\mathrm{H}}X_1$ is nonsingular, and L_1 can be written

$$L_1 = (X_1^{\mathrm{H}}X_1)^{-1}X_1^{\mathrm{H}}AX_1.$$

There is a nice relation between the eigenvectors of A and those of L_1. First suppose that $L_1z = z\lambda$. Then from (6.1)

$$AX_1z = X_1L_1z = X_1z\lambda,$$

so that X_1z is an eigenvalue of A. Conversely if z is a vector with the property that $A(X_1z) = (X_1z)\lambda$, then $X_1L_1z = \lambda X_1z$, and it follows upon multiplying by $(X_1^{\mathrm{H}}X_1)^{-1}X_1^{\mathrm{H}}$ that $L_1z = z\lambda$. Thus to any eigenvector X_1z of A that lies in $\mathcal{X}_1$ there corresponds and eigenvector z of L_1 and vice versa.[15]

Even more is true. It can be shown (and you are going to show it in just a moment) that if λ is an eigenvalue of multiplicity m of L_1, then λ is an eigenvalue of multiplicity at least m of A. This allows us to partition the eigenvalues of A into two sets: the set Λ_1 of eigenvalues of L_1 and the set Λ_2 of the remaining eigenvalues of A. Since we count multiplicities, if λ is an eigenvalue of multiplicity m of L_1 and of multiplicity greater than m of A, the λ belongs to both Λ_1 and Λ_2.

We shall say that an invariant subspace is simple if the sets Λ_1 and Λ_2 are disjoint. We are going to verify that simple invariant subspaces have properties analogous to those of simple eigenvectors. Before we do this, note that without loss of generality we may take X_1 to be an orthonormal basis for $\mathcal{X}_1$, so that

$$X_1^{\mathrm{H}}X_1 = I$$

and

$$L_1 = X_1^{\mathrm{H}}AX_1.$$

7 Capital punishment

Return to §§3-5 with pencil and paper in hand. Write down each equation as it occurs, replacing each lower-case letter with the corresponding capital Latin letter; e.g., $\lambda \leftarrow L$, $x \leftarrow X$, etc. Wherever it is appropriate replace 1 with I. As you do this exercise reflect on the meaning of the equations in terms of invariant subspaces.

Now reread the same sections once more along the footnotes.

[15] By replacing $z\lambda$ by ZL, we see that the correspondence is actually between invariant subspaces of L_1 and invariant subspaces of A lying in $\mathcal{X}_1$.

8 Coda

If you have faithfully performed the exercise of the last section you will have learned some important facts about invariant subspaces. However, there is much more.

The algebraic theory of invariant subspaces becomes quite complicated when the subspace is not simple. For example, the subspace may fail to have a complementary invariant subspace. A treatment of some of the problems is given in [3].

A standard reference for perturbation theory is Kato's comprehensive book [6]. The passage from first order expansions to rigorous bounds has been treated by the author [10,11]. Davis and Kahan [2] have developed an elegant perturbation theory for Hermitian matrices (see also [8]).

People who calculate a invariant subspaces usually end up with matrices X_1 and L_1 for which $R = AX_1 - X_1L_1$ is not zero but is merely small. An important question, treated in the references cited in the last paragraph, is how near the space spanned by X_1 is to an invariant subspace of A. In many cases it is sufficient to know that there is a small matrix E such that the space spanned by X_1 is an invariant subspace of $A + E$. An exhaustive treatment of this question is given in [5].

9 Acknowledgements

I would like to thank my friends and colleagues at Fudan University in Shanghai for encouraging me to write this material down and especially Professor Jiang Erxiong, who made made my visit there possible. Berseford Parlett read the paper and gave me the benefit of his pointed observations. Finally, thanks to Gil Strang of MIT for the *bon mot* which gave the paper its title.

References

[1] R. H. Bartels and G. W. Stewart, "Algorithm 432: The Solution of the Matrix Equation $AX - BX = C$," *Communications of the ACM* **15** (1972) 820-826.

[2] C. Davis and W. M. Kahan, "The Rotation of Eigenvectors by a Perturbation, III," *SIAM J. Numer. Anal.* **7** (1970) 1-46.

[3] I. Gohberg, P. Lancaster, and L. Rodman, *Invariant Subspaces of Matrices with Applications*, John Wiley, New York, 1986.

[4] G. H. Golub and C. F. Van Loan, *Matrix Computations*, Johns Hopkins University Press, Baltimore, 1983.

[5] W. Kahan, B. N. Parlett, and E. Jiang, "Residual Bounds on Approximate Eigensystems of Nonnormal Matrices," *SIAM J. Numer. Anal.* **19** (1982) 470-484.

[6] T. Kato, *Perturbation Theory for Linear Operators*, Springer Verlag, New York, 1966.

[7] C. Meyer and G. W. Stewart, "Derivatives and Perturbations of Eigenvectors," to appear in *Lin. Alg. Appl.*

[8] B. N. Parlet *The Symmetric Eigenvalue Problem*, Prentice Hall, Englewood Cliffs, New Jersey, 1980.

[9] I. Schur, "Über die charakteristischen Würzeln einer linearen Substitution mit einer Anwenduug auf die Theorie der Integralgleichuugen," *Mathematische Annallen* **66** (1909) 448-510.

[10] G. W. Stewart "Error Bounds for Approximate Invariant Subspaces of Closed Linear Operators," *SIAM J. Numer. Anal.* **8** (1971) 796-808.

[11] G. W. Stewart "Error and Perturbation Bounds for Subspaces Associated with Certain Eigenvalue Problems," *SIAM Review* **15** (1973) 727-764.

[12] G. W. Stewart *Introduction to Matrix Computations*, Academic Press, New York, 1974.

[13] J. H. Wilkinson *The Algebraic Eigenvalue Problem*, Oxford University Press, Oxford, 1965.