Eigenvectors and Adjoints

P. R. HALMOS
Indiana University
Bloomington, Indiana

Communicated by Hans W. E. Schwerdtfeger

If λ is an eigenvalue of a linear transformation A on a finite dimensional, complex inner-product space H, then the complex conjugate λ^* is an eigenvalue of the adjoint transformation A^*. In other words, if there exists a nonzero vector f such that $Af = \lambda f$, then there exists a nonzero vector g such that $A^*g = \lambda^*g$. A generalization of the concept of eigenvector (an *eigensequence*) is a finite sequence $\langle f_1, \ldots, f_k\rangle$ of vectors such that each Af_j is a linear combination of the f_i's, that is, $Af_j = \sum_i \lambda_{ij} f_i$. The corresponding appropriate generalization of the requirement that f be not zero is that the set $\{f_1, \ldots, f_k\}$ be linearly independent. The span of the vectors in an eigensequence is an invariant subspace.

An elegant way to discuss eigensequences (of a fixed length k) is to consider the direct sum $\hat{H}$ of k copies of H. For each linear transformation A on H, write $\hat{A}$ for the linear transformation on $\hat{H}$ defined by $\hat{A}\langle f_1, \ldots, f_k\rangle = \langle Af_1, \ldots, Af_k\rangle$. For each $k \times k$ matrix λ, write $\hat{\lambda}$ for the linear transformation on $\hat{H}$ defined by

$$\hat{\lambda}\langle f_1, \ldots, f_k\rangle = \langle g_1, \ldots, g_k\rangle,$$

where

$$g_j = \sum_i \lambda_{ij} f_i.$$

The matrix λ deserves to be called an *eigenmatrix* of the transformation A in case there exists a linearly independent sequence $\hat{f} = \langle f_1, \ldots, f_k\rangle$ such that $\hat{A}\hat{f} = \hat{\lambda}\hat{f}$. The purpose of this note is to prove and discuss the following generalization of the introductory sentence.

Theorem 1. *If λ is an eigenmatrix of a linear transformation A on a finite dimensional, complex inner-product space H, then the conjugate transpose λ^* is an eigenmatrix of the adjoint transformation A^*.*

Proof. Coordinatize, and thus, in particular, replace the linear transformation A by a matrix A. Since A is similar to its transpose A', there exists an invertible matrix S such that $A = S^{-1}A'S$. Similarly, since λ is similar to its transpose, there exists an invertible $k \times k$ matrix σ such that $\lambda = \sigma^{-1}\lambda'\sigma$. Since the mappings $A \to \hat{A}$ and $\sigma \to \hat{\sigma}$ are multiplicative, it follows that, if $\hat{f}$ is an eigensequence of A (linearly independent) with eigenmatrix λ, i.e., if $\hat{A}\hat{f} = \hat{\lambda}\hat{f}$, then

$$\hat{S}^{-1}\hat{A}'\hat{S}\hat{f} = \hat{\sigma}^{-1}\hat{\lambda}'\hat{\sigma}\hat{f}.$$

Multiply on the left by $\hat{S}\hat{\sigma}$ and use the commutativity of numerical matrices such as $\hat{\sigma}$ with inflated transformations such as $\hat{S}$ to get

$$\hat{A}'(\hat{\sigma}\hat{S}\hat{f}) = \hat{\lambda}'(\hat{\sigma}\hat{S}\hat{f}).$$

Form the complex conjugate of both sides; the result is

$$\hat{A}^*\hat{g} = \hat{\lambda}^*\hat{g},$$

where $\hat{g}$ is the sequence $\langle g_1, \ldots, g_k \rangle$ such that the coordinates of the terms g_i are the complex conjugates of the corresponding coordinates of the corresponding terms of $\hat{\sigma}\hat{S}\hat{f}$. What remains to be proved is that $\hat{g}$ is linearly independent or, what comes to the same thing, that $\hat{\sigma}\hat{S}\hat{f}$ is such. Since S is invertible, it is clear that $\langle Sf_1, \ldots, Sf_k \rangle$ is linearly independent along with $\langle f_1, \ldots, f_k \rangle$, i.e., that $\hat{S}\hat{f}$ is independent along with $\hat{f}$. What is left to prove is that $\hat{\sigma}$ preserves independence, i.e., (with an unimportant change of notation), that, if $\hat{f}$ is independent, then so is $\hat{\sigma}\hat{f}$. This too is clear: the invertible matrix $\hat{\sigma}$ defines an invertible transformation on the span of the terms of $\hat{f}$.

With all its coordinates and complex conjugates brazenly flaunted, this proof is colossally ugly. Some of the ugliness can be removed, but I have not been able to find a completely clean functorial proof, and the proof above has, at least, the virtue of being brisk. The complex conjugation, i.e., the treatment of the sesquilinear theory associated with unitary geometry rather than the bilinear theory associated with duality in general, is motivated by the intended applications and generalizations;

it plays a relatively small role in the unnaturality of the proof. The statement of Theorem 1, unlike the proof, is quite easy to clean up, and, in view of the intrinsic interest of the theorem, it is worthwhile to do so. Here is an equivalent formulation.

THEOREM 2. *To each subspace M invariant under a linear transformation A on a finite dimensional, complex inner-product space there corresponds a subspace N invariant under A^* such that $A^*|N$ is similar to $(A|M)^*$.*

Proof. Given M, let $\hat{f} = \langle f_1, \ldots, f_k \rangle$ be a basis for M, and let λ be the matrix of $A|M$ with respect to that basis; then, by definition, $\hat{A}\hat{f} = \hat{\lambda}\hat{f}$. Note that the matrix of $(A|M)^*$ is λ^*. Apply Theorem 1 to get an independent sequence $\hat{g} = \langle g_1, \ldots, g_k \rangle$ such that $\hat{A}^*\hat{g} = \hat{\lambda}^*\hat{g}$. This says that, if N is the span of $\{g_1, \ldots, g_k\}$, then the matrix of $A^*|N$ with respect to the basis $\{g_1, \ldots, g_k\}$ is λ^*. Since $(A|M)^*$ and $A^*|N$ have the same matrix (with respect to two suitable bases), the asserted similarity follows. It is easy to see that the argument is reversible, and that, thus, Theorem 1 can be derived from Theorem 2.

It should be remarked that, although concepts of unitary geometry (adjoint) are present in Theorem 2, the similarity it asserts cannot be replaced by unitary equivalence. For a counterexample let A be a linear transformation on a three-dimensional space that has the matrix

$$\begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

with respect to some orthonormal basis. If M is the span of the eigenvectors of A, then M has the orthonormal basis $\langle\langle 1, 0, 0\rangle, \langle 0, 0, 1\rangle\rangle$. The transformation $A|M$ has two distinct eigenvalues; the corresponding eigenvectors are not orthogonal. The only two-dimensional invariant subspace of A^* on which A^* has two distinct eigenvalues (and therefore the only possible candidate for N) has the orthonormal basis $\langle\langle 0, 1, 0\rangle, \langle 0, 0, 1\rangle\rangle$. Since $A^*|N$ is Hermitian, but $A|M$ is not, it follows that $A^*|N$ cannot be unitarily equivalent to $(A|M)^*$.

The eigenmatrix theorem has a tempting generalization. The transformation $\hat{A}$ on $\hat{H}$ can be expressed as a $k \times k$ matrix whose entries are transformations on H, namely, the matrix whose diagonal entries are A and whose off-diagonal entries are 0. The transformation $\hat{\lambda}$ on $\hat{H}$ can also be expressed as a $k \times k$ matrix whose entries are transformations on H; the $\langle i, j\rangle$ entry is the scalar transformation λ_{ij}. It follows that $\hat{A} - \hat{\lambda}$ is (or, better, has) a $k \times k$ matrix (block matrix) whose entries are pairwise commutative linear transformations on H. Tempting generalization: if T is a $k \times k$ block matrix whose entries are pairwise commutative linear transformations on a finite dimensional, complex inner-product space, and if there exists an independent sequence in its kernel, then the same is true for its adjoint. Verdict: false. Counterexample, with $k = 2$:

$$T = \begin{pmatrix} A & 1 \\ 0 & 0 \end{pmatrix}, \quad \text{where} \quad A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

(and the remaining entries in T are the indicated 2×2 scalar matrices). The independent sequence $\langle\langle 1, 0\rangle, \langle 0, -1\rangle\rangle$ is in the kernel of T, but every sequence $\langle f, g\rangle$ in the kernel of T^* has $f = 0$, and, consequently, no such sequence can be independent.

The eigenmatrix theorem has a pleasant effect on the structure of invariant subspaces. If A is a linear transformation on H, call a subspace M a *commuting range* (for A) in case there exists a linear transformation B on H such that B commutes with A and such that the range of B is equal to M. Similarly, call M a *commuting kernel* in case there exists a commuting C such that the kernel of C is equal to M. It is obvious that each commuting range and each commuting kernel are invariant under A; the pleasant surprise is the converse.

THEOREM 3. *Every subspace invariant under a linear transformation on a finite dimensional, complex inner-product space is both a commuting range and a commuting kernel.*

Remark. The assertion is that, for each A on H, and for each M invariant under A, there exists a B on H and there exists a C on H such that $AB = BA$, $AC = CA$, $\operatorname{ran} B = M$, and $\ker C = M$. It is not asserted that $B = C$.

Proof. Let $\hat{f} = \langle f_1, \ldots, f_k\rangle$ be a basis for M; the invariance of M implies that $\hat{f}$ is an eigensequence for A with eigenmatrix λ, say. By Theorem 1, there exists an independent eigensequence $\hat{g} = \langle g_1, \ldots, g_k\rangle$ for A^*, with eigenmatrix λ^*. Define a linear transformation B on H by

$$Bf = \sum_j (f, g_j) f_j.$$

The range of B is obviously included in M. If the g_j's were orthonormal, the reverse inclusion would be obvious too; since, however, they always are orthonormal with respect to some inner product equivalent to the given one, the reverse inclusion is true in any case. As for commutativity:

$$ABf = \sum_j (f, g_j)Af_j = \sum_j (f, g_j) \sum_i \lambda_{ij} f_i = \sum_i \sum_j (f, g_j)\lambda_{ij} f_i,$$

and

$$BAf = \sum_j (Af, g_j)f_j = \sum_j (f, A^*g_j)f_j = \sum_j \left(f, \sum_i \lambda_{ji}^* g_i\right) f_j$$

$$= \sum_j \sum_i (f, g_i)\lambda_{ji} f_j.$$

(What is really going on here is transparent in the case $k = 1$: if $Bf = (f, g_1)f_1$, then $ABf = (f, g_1)Af_1 = (f, g_1)\lambda f_1$ and $BAf = (Af, g_1)f_1 = (f, A^*g_1)f_1 = (f, \lambda^* g_1)f_1 = (f, g_1)\lambda f_1$. The general case could have been reported in a similarly condensed manner, but only at the cost of the consideration of an appropriately generalized inner product in the space $\hat{H}$; the machinery is too heavy to drag in for just one use.)

That takes care of commuting ranges. The assertion about commuting kernels follows by an application of the commuting range statement to A^* and $M^\perp$: if $A^*C^* = C^*A^*$ and ran $C^* = M^\perp$, then $AC = CA$ and ker $C = M$. The proof of Theorem 3 is complete.

How much of all this remains true with no essential change for infinite dimensional Hilbert spaces, bounded operators, and closed subspaces? Answer: none. Theorems 1 and 2 break down even for $k = 1$; the complex conjugate of an eigenvalue for A may fail to be an eigenvalue for A^*. Example: the adjoint of the unilateral shift. Theorem 3 also breaks down. Example: the bilateral shift on L^2 of the circle in the role of the operator and H^2 in the role of the subspace. D. E. Sarason has even constructed an irreducible example: the Hilbert space is H^2 of an annulus centered at 0, the operator is multiplication by the independent variable, and the subspace is the closure of the set of all polynomials.

Received January, 1970