SPECTRAL THEOREM FOR NORMAL MATRICES

K. Žitný, J. Kozánek and J. J. Koliha

Faculty of Mathematics and Physics, Charles University, Prague, Czech Republic
Institute of Thermomechanics, ASCR, Dolejškova 5, 182 00 Prague 8, Czech Republic
Department of Mathematics and Statistics, University of Melbourne, VIC 3010, Australia

Motto: *A good part of matrix theory is functional analytical in spirit.*
R. Bhatia

This classroom note is intended to give an alternative proof of the 'spectral theorem for normal matrices' by emphasizing the linear transformation aspect of linear algebra in the spirit of Halmos's well known book [1] or the monograph [3]. The proofs of three lemmas are suggested as exercises, designed to supplement routine problems, and to give a student the opportunity to see the norm more as a tool of analysis than just a simple tool of measurement.

Let H be a finite dimensional Hilbert space. The space of all linear operators from H to H is denoted by $\mathcal{L}(H)$, with the norm of $T \in \mathcal{L}(H)$ defined by $\|T\| = \sup\{\|Tx\| : x \in H, \|x\| = 1\}$. For each $T \in \mathcal{L}(H)$ there exists a unique *adjoint* operator $T^* \in \mathcal{L}(H)$, satisfying the celebrated B^*-identity:

$$\|T^*T\| = \|T\|^2.$$

An operator $T \in \mathcal{L}(H)$ is *normal* if $T^*T = TT^*$ and *Hermitian* if $T^* = T$. The *numerical range* of $T \in \mathcal{L}(H)$ is defined by

$$V(T) = \{\langle Tx, x\rangle : x \in H, \|x\| = 1\},$$

and the *numerical radius* of T is the nonnegative number

$$v(T) = \sup_{\lambda \in V(T)} |\lambda|.$$

We note that $|\langle Tx, x\rangle| \le v(T)\|x\|^2$ for all $x \in H$, and that $v(T) \le \|T\|$ by the Schwarz inequality.

The next three lemmas can be assigned as exercise; they are not completely trivial, and require some ingenuity on part of the student, but all have elementary solutions. Some hints may be supplied, in particular for Lemma 1 we may suggest the use of the parallelogram identity

$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2$$

to prove the first inequality.

Lemma 1. *For any* $T \in \mathcal{L}(H)$,

$$\tfrac{1}{2}\|T\| \le v(T) \quad \textit{and} \quad v(T^2) \le v(T)^2.$$

Lemma 2. *If* $T \in \mathcal{L}(H)$ *is a Hermitian operator, then* $v(T) = \|T\|$.

Lemma 3. *If* $T \in \mathcal{L}(H)$ *is a normal operator, then* $v(T) = \|T\|$.

All our results up to now hold for bounded linear operators on an arbitrary Hilbert space. The next result plays a key role in the eigenvalue problem for normal operators, with the proof specific to finite dimensional Hilbert spaces.

Lemma 4. *Every normal operator* $T \in \mathcal{L}(H)$ *on a finite dimensional Hilbert space* H *has an eigenvalue* $\lambda \in \mathbb{C}$ *satisfying* $|\lambda| = \|T\|$.

Proof. First we observe that the function $f(x) = \langle Tx, x\rangle$ is continuous on the closed unit ball $B = \{x \in H : \|x\| \le 1\}$. This follows from the inequality

$$|f(x) - f(y)| = |\langle T(x-y), x\rangle + \langle Ty, x-y\rangle| \le 2\|T\|\|x-y\|$$

valid for any $x, y \in B$. In our case B is compact, and thus the continuous function $x \mapsto |f(x)|$, $x \in B$, attains its maximum on B at some point $h \in B$. We show that $\|h\| = 1$. We have

$$\|T\| = v(T) = |\langle Th, h\rangle| \le \|Th\|\|h\| \le \|T\|\|h\|^2 \le \|T\|.$$

Hence $\|h\| = 1$, and $|\langle Th, h\rangle| = \|Th\|\|h\|$. Then Th and h are linearly dependent, and so $Th = \lambda h$ for some $\lambda \in \mathbb{C}$. Finally, $|\lambda| = |\langle \lambda h, h\rangle| = |\langle Th, h\rangle| = \|T\|$. □

We hope that the proof of the following main theorem (treated in any rigorous book on linear algebra) provides a good example of the advantages of the functional-analytical approach to matrix theory.

Theorem 1. *If $T \in \mathcal{L}(H)$ is a normal operator on a finite dimensional Hilbert space H, then there exists an orthonormal basis of H consisting of eigenvectors of T.*

Proof. We recall that eigenvectors of a normal operator T corresponding to distinct eigenvalues are orthogonal. Hence in the finite dimensional space H there exists a maximal orthonormal set $\mathcal{B} = \{x_1, \dots, x_s\}$ consisting entirely of eigenvectors of T. We need to show that $\mathcal{B}$ is an orthonormal basis of H. Let M be the linear subspace of H spanned by $\mathcal{B}$. Let $\{\lambda_1, \dots, \lambda_s\}$ be the corresponding set of eigenvalues for $\mathcal{B}$, and let $Px = \sum_{j\in J}\langle x, x_j\rangle x_j$ be the orthogonal projection of H onto M, where $J = \{1, \dots, s\}$. Using known elementary properties of normal operators, we show that P commutes with T:

$$\begin{aligned} PTx &= \sum_{j\in J}\langle Tx, x_j\rangle x_j = \sum_{j\in J}\langle x, T^* x_j\rangle x_j = \sum_{j\in J}\langle x, \overline{\lambda_j} x_j\rangle x_j \\ &= \sum_{j\in J}\lambda_j\langle x, x_j\rangle x_j = T\Big(\sum_{j\in J}\langle x, x_j\rangle x_j\Big) = TPx, \quad x \in H. \end{aligned} \tag{1}$$

Suppose that $(I-P)T \neq 0$. Since T and P commute, the operator $(I-P)T$ is normal. By Lemma 4 there exists a unit vector x_0 and $\lambda \in \mathbb{C}$ such that $(I-P)Tx_0 = \lambda x_0$ and $|\lambda| = \|(I-P)T\| \neq 0$. Then $Px_0 = 0$ and $Tx_0 = \lambda x_0$; $x_0 \in N(P) = M^{\perp}$ contradicts the maximality of $\mathcal{B}$. This proves that $(I-P)T = 0$.

Assume that $P \neq I$. Since $T(I-P) = 0$, every unit vector $x_0 \in N(P) = M^{\perp}$ is an eigenvector for T, which again contradicts the maximality of $\mathcal{B}$. This proves $P = I$, that is, $M = H$. □

If $\mu_1, \dots, \mu_r$ are the distinct eigenvalues of T, write $K = \{1, \dots, r\}$.

Theorem 2. (Spectral decomposition for normal operators on a finite dimensional Hilbert space H.) *Let $T \in \mathcal{L}(H)$ be normal. If $\{\mu_k : k \in K\}$ are the distinct eigenvalues of T and $P_k \in \mathcal{L}(H)$ is the orthogonal projection of H onto the eigenspace $N(\mu_k I - T)$, $k \in K$, then*

$$Tx = \sum_{k\in K}\mu_k P_k x, \quad x \in H. \tag{2}$$

Proof. By Theorem 1 there exists an orthonormal basis $\mathcal{B} = \{x_j : j \in J\}$ of H consisting of eigenvectors of T. Partition $J = \bigcup_{k\in K} J_k$ so that the eigenvectors $\{x_j : j \in J_k\}$ correspond to the eigenvalue μ_k. Then

$$Tx = T\Big(\sum_{j\in J}\langle x, x_j\rangle x_j\Big) = \sum_{j\in J}\langle x, x_j\rangle Tx_j = \sum_{j\in J}\lambda_j\langle x, x_j\rangle x_j$$

$$= \sum_{k\in K} \sum_{j\in J_k} \lambda_k \langle x, x_j\rangle x_j = \sum_{k\in K} \lambda_k P_k x. \qquad \square$$

It is interesting to note that the preceding procedure can be extended to normal compact operators on an infinite dimensional space H. The key difference is that in the proof of Lemma 4 we use the weak continuity of the compact operator T, and the fact that the closed unit ball in H is weakly compact. Proofs of Theorem 1 and Theorem 2 have almost obvious adaptations to the infinite dimension (they are already written in a suggestive way to allow for this extension).

The reader may find it interesting to compare our approach to spectral theorem with a recent Korányi's paper [2] on the finite dimensional case, where the starting point is the singular value decomposition, and with Martin's note [4] treating the infinite dimensional case.

References

[1] P. Halmos, *Finite Dimensional Vector Spaces*, Van Nostrand, New York, 1958.

[2] Adam Korányi, Around the finite-dimensional spectral theorem, *Amer. Math. Monthly* **108** (2001), 120–125.

[3] I. Marek and K. Žitný, *Matrix Analysis for Applied Sciences* I, II, Teubner, Leipzig, 1983, 1985.

[4] A. D. Martin, On the spectral theorem for compact symmetric operators, *Amer. Math. Monthly* **72** (1965), 892–894.