DIAGONALIZING HERMITIAN MATRICES OF CONTINUOUS FUNCTIONS

JUSTIN CYR, JASON EKSTRAND, NATHAN MEYERS,
CRYSTAL PEOPLES, AND JUSTIN PETERS

ABSTRACT. The problem of diagonalizing hermitian matrices of continuous fiunctions was studied by Grove and Pederson in 1984. While diagonalization is not possible in general, in the presence of differentiability conditions we are able to obtain positive results in the case of 2×2 matrices. It remains open whether our results can be extended to $n \times n$ matrices.

1. INTRODUCTION

In this note we consider a special case of the problem of diagonalizing hermitian matrices of continuous functions. We say that a hermitian element $A \in M_n(C[a,b])$ is *diagonalizable* if there is a unitary element $U \in M_n(C[a,b])$ such that U^*AU is diagonal. Thus, for each $t \in [a,b]$, $U(t)^*A(t)U(t)$ is diagonal, and in particular each $U(t)$ is unitary. We do not know who first realized that hermitian elements of $M_n(C[a,b])$ cannot, in general, be diagonalized, but evidently this is folklore. In 1983 Kadison ([3]) proved that if $A \in M_n(L^\infty[a,b])$ is hermitian, then A is diagonalizable by a unitary in $M_n(L^\infty[a,b])$. In fact Kadison proved more generally that if $\mathcal{A}$ is an abelian von Neumann algebra, a hermitian element in $M_n(\mathcal{A})$ is diagonalizable by a unitary in $M_n(\mathcal{A})$.

However for $M_n(C[a,b])$ the analogous statement is false. (Example 2.1) A year after Kadison's result appeared, Grove and Pedersen proved that, for a large class of compact Hausdorff spaces X, if $A \in M_n(C(X))$ is hermitian with the property that the eigenvalues of $A(x)$ are distinct for all $x \in X$, then A is diagonalizable in $M_n(C(X))$. Their proof, however, is an existence proof, and the question of how

such a diagonalization can be achieved is not addressed. [1] also considers the question of which compact Hausdorff spaces X have the property that every hermitian element $A \in M_n(C(X))$ is diagonalizable by a unitary in $M_n(C(X))$.

Here we consider the case where $X = [a, b]$, a compact interval. Surprisingly, even for 2×2 matrices it is not completely trivial to show that that if $A \in M_2(C[a, b])$ is hermitian with distinct eigenvalues, then there is a unitary $U \in M_2(C[a, b])$ such that U^*AU is diagonal. In fact, for 2×2 matrices we provide an algorithm for diagonalization in case where the two eigenvalues are distinct.

Our main result, Theorem 2.6, is to obtain a diagonalization theorem for hermitian elements A of $M_2(C[a, b])$ in which we relax the condition that the eigenvalues of $A(t)$ must be distinct for all $t \in [a, b]$. To do this we suppose A is continuously differentiable, and at those points t where the the eigenvalues of $A(t)$ coincide, the eigenvalues of $A'(t)$ are distinct.

This result gives rise to a variety of related questions which, as far as we are aware, have not been explored. The most obvious question is whether there is an analogue of Theorem 2.6 for $M_n(C[a, b])$. Our proof uses the fact that one has an explicit formula for the eigenvalues in the 2×2 case; thus an entirely different approach would be required in the general case. Another question that could be asked is, if one assumes "sufficient" diffierentiability, then the conditions on the eigenvalues can be removed. Example 2.8 shows that even if the hermitian matrix has C^∞ entries, it may not be diagonalizable. Finally we raise the question whether more can be done if the entries are analytic (2.9)

2. Diagonalization

We begin with an example showing that continuous diagonalization is not always possible.

Example 2.1. We work on $M_2(C[-1, 1])$. Let χ be the characteristic function of the interval $[0, 1]$, and let

$$A(t) = t \begin{bmatrix} 1 & \chi \\ \chi & \chi \end{bmatrix}.$$

Now the unitary matrix $U(t) = I$ diagonalizes $A(t)$ for $t \in [-1, 0]$ and

$$U(t) = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

diagonalizes $A(t)$ for $t \in (0, 1]$. If $W(t)$ is another one parameter family of unitary matrices such that $W(t)^*A(t)W(t)$ is diagonal, $t \in [-1, 1]$,

then a calculation shows that $W(t) = U(t)D(t)$ where $D(t)$ is a diagonal unitary. Thus, there is no choice of $W(t)$ which is continuous at $t = 0$.

Theorem 2.2. *Let $A \in M_2(C[a,b])$ such that $A(t)$ is Hermitian and $A(t)$ has distinct eigenvalues for every $t \in [a,b]$. Then there exists $U \in M_2(C[a,b])$ such that $U(t)$ is unitary and $U(t)^* A(t)U(t)$ is diagonal for all $t \in [a,b]$. Furthermore, if A consists of real-valued functions then U can be chosen to be real-valued.*

It will be useful to introduce notation for $A(t)$.

$$(1) \qquad A(t) = \begin{bmatrix} f(t) & h(t) \\ \bar{h}(t) & g(t) \end{bmatrix} \in M_2(C[a,b])$$

where f, g are real-valued functions in $C[a,b]$, and h is complex-valued. From the characteristic polynomial for $A(t)$

$$p_{A(t)}(x) = \det[xI - A(t)]$$
$$= x^2 - (f(t) + g(t))x + f(t)g(t) - |h(t)|^2$$

we obtain the two eigenfunctions

$$(2) \qquad \lambda^+(t) = \frac{f(t) + g(t) + \sqrt{(f(t) - g(t))^2 + 4|h(t)|^2}}{2}$$

$$(3) \qquad \lambda^-(t) = \frac{f(t) + g(t) - \sqrt{(f(t) - g(t))^2 + 4|h(t)|^2}}{2}$$

Note that $\lambda^+(t) = \lambda^-(t)$ if and only if $A(t) = \lambda^+(t)I$.

Lemma 2.3. *Let $A(t)$ be as in (1) and assume that $A(t)$ has distinct eigenvalues for all $t \in [a,b]$. Let $\lambda(t)$ be an eigenvalue of $A(t)$ for every $t \in [a,b]$. For $t \in [a,b]$, the following are equivalent:*

    (1) $h(t) = 0$,
    (2) $f(t) - \lambda(t) = 0$ or $g(t) - \lambda(t) = 0$,
    (3) either $f(t) - \lambda(t) = 0$ or $g(t) - \lambda(t) = 0$, but not both.

Proof. If $h(t) = 0$ then $A(t)$ is diagonal, and since $\lambda(t)$ is an eigenvalue for $A(t)$, $f(t) - \lambda(t) = 0$ or $g(t) - \lambda(t) = 0$ but not both since $A(t)$ has distinct eigenvalues. If $f(t) - \lambda(t) = 0$ or $g(t) - \lambda(t) = 0$ then by equations (2) and (3), $|h(t)|^2 = 0$ and so $h = 0$. $\square$

Let λ refer either to λ^+ or λ^-. We wish to find $\mathbf{v}(t) = \begin{bmatrix} v_1(t) \\ v_2(t) \end{bmatrix}$ continuous and non-vanishing such that $A(t)\mathbf{v}(t) = \lambda(t)\mathbf{v}(t)$. Equivalently,

$$(4) \qquad (f(t) - \lambda(t))v_1(t) + h(t)v_2(t) = 0$$

$$(5) \qquad \bar{h}(t)v_1(t) + (g(t) - \lambda(t))v_2(t) = 0$$

along with the condition that $v_1(t)$ and $v_2(t)$ are never simultaneously zero. The following construction allows us to do this explicitly.

Construction 2.4. Let $A \in M_2(C[a, b]$ have distinct eigenvalues for all $t \in [a, b]$, and fix λ to be either λ^+ or λ^- throughout the construction. Input: $n \in \mathbb{N}$, $t_n \in [a, b)$, and $\mathbf{v}(t_n) \in \mathbb{C}^2$ such that $\mathbf{v}(t_n) \neq 0$ and $A(t_n)\mathbf{v}(t_n) = \lambda(t_n)\mathbf{v}(t_n)$.

By Lemma 2.3, we cannot have $\lambda(t_n) = f(t_n) = g(t_n)$. Suppose $f(t_n) \neq \lambda(t_n)$. Then, by continuity, $f - \lambda$ is nonzero in a neighborhood of t_n. Let

$$t_{n+1} = \min(\{t \in [t_n, b] : f(t) = \lambda(t)\} \cup \{b\}).$$

On $[t_n, t_{n+1})$, define

$$(6) \qquad \mathbf{u}(t) = \begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix} = \begin{bmatrix} \frac{h(t)}{\lambda(t) - f(t)} \\ 1 \end{bmatrix}.$$

Then u_1 and u_2 satisfy (4) on $[t_n, t_{n+1})$. Since $A(t) - \lambda(t)I$ is singular, (4) and (5) form a dependent set of equations, and so (5) is satisfied. Therefore, $\mathbf{u}(t)$ is an eigenvector of $A(t)$ for $\lambda(t)$ for all $t \in [t_n, t_{n+1})$. Since $A(t)$ has distinct eigenvalues, the geometric multiplicity is one and there is some constant $c_1 \in \mathbb{C}$ with $c_1 \neq 0$ such that $\mathbf{v}(t_n) = c_1\mathbf{u}(t_n)$.

If $\lambda \neq f$ on all of $[t_n, b]$, then $t_{n+1} = b$ and we can extend $\mathbf{u}(t)$ to $[t_n, t_{n+1}]$ using (6). Then let $\mathbf{v}(t) = c_1\mathbf{u}(t)$ on $[t_n, b]$ and we are done.

Otherwise, $\lambda(t_{n+1}) = f(t_{n+1})$. Then, by Lemma 2.3, $\lambda(t_{n+1}) \neq g(t_{n+1})$ and by continuity there is some $t_n < \alpha < t_{n+1}$ such that $\lambda \neq f$ on $[\alpha, t_{n+1}]$. On $[\alpha, t_{n+1}]$, define

$$(7) \qquad \mathbf{w}(t) = \begin{bmatrix} w_1(t) \\ w_2(t) \end{bmatrix} = \begin{bmatrix} 1 \\ \frac{\bar{h}(t)}{\lambda(t) - g(t)} \end{bmatrix}.$$

By the same argument as above, $\mathbf{w}(t)$ is an eigenvector of $A(t)$ for $\lambda(t)$ for all $t \in [\alpha, t_{n+1}]$. Furthermore, since $\lambda(\alpha)$ has multiplicity one, there is some nonzero $c_2 \in \mathbb{C}$ such that $\mathbf{u}(\alpha) = c_2\mathbf{w}(\alpha)$. Then, let $\mathbf{v} = c_1\mathbf{u}$ on $[t_n, \alpha]$ and let $\mathbf{v} = c_1c_2\mathbf{w}$ on $[\alpha, t_{n+1}]$. Then we have that $\mathbf{v}$ is continuous and non-vanishing on $[t_n, t_{n+1}]$ and that $A(t)\mathbf{v}(t) = \lambda(t)\mathbf{v}(t)$ for all $t \in [t_n, t_{n+1}]$.

Finally, in the case where $f(t_n) = \lambda(t_n)$, we reverse the roles of (6) and (7).

Lemma 2.5. *For $A \in M_2(C[a, b])$, construction 2.4 terminates after finitely many steps.*

Proof. Let $A \in M_2(C[a,b])$ and let $t_0 = a$. For all iterations of the construction we make the same choice of λ as either λ^+ or λ^-. Suppose that construction 2.4 never terminates and yields the sequence $\{t_n\}$. Since $\{t_n\}$ is increasing and $t_n < b$ for every n, t_n has a limit, c, $a < c \leq b$. Observe that, for each $n > 0$, $f(t_n) = \lambda(t_n)$ or $g(t_n) = \lambda(t_n)$ by our choice of the t_n. Further observe that if $f(t_n) = \lambda(t_n)$ then $g(t_{n+1}) = \lambda(t_{n+1})$ and vice-versa because of the way we choose the t_n and the assumption that $t_n < b$. Therefore, since f, g, and λ are continuous, $\lim_{n\to\infty} f(t_n) - \lambda(t_n) = 0$ and $\lim_{n\to\infty} g(t_n) - \lambda(t_n) = 0$, and so $f(c) = g(c) = \lambda(c)$. However, this contradicts Lemma 2.3, and so construction 2.4 must terminate after only finitely many steps. $\square$

Finally, we can prove Theorem 2.2.

Proof of Theorem 2.2. Let $A \in M_2(C[a,b])$ and let $t_0 = a$. Let $\mathbf{v}(\mathbf{t_0})$ be the eigenvector corresponding to $\lambda^+(t_0)$ for $A(t_0)$. Then applying construction 2.4 and Lemma 2.5 yields $\mathbf{v}_1(t) : [a,b] \to \mathbb{C}^2$ such that $A(t)\mathbf{v}_1(t) = \lambda^+(t)\mathbf{v}_1(t)$ for $t \in [a,b]$. Similarly, we can get $\mathbf{v}_2(t) : [a,b] \to \mathbb{C}^2$ such that $A(t)\mathbf{v}_2(t) = \lambda^-(t)\mathbf{v}_1(t)$ for $t \in [a,b]$.

Let $\mathbf{u}_j(t) = \frac{\mathbf{v}_j(t)}{\|\mathbf{v}_j(t)\|_2}$, $j = 1,\ 2$. Then, for all $t \in [a,b]$, we have that, since $A(t)$ is hermitian and $\lambda^+(t) \neq \lambda^-(t)$, $\mathbf{v}_1(t) \perp \mathbf{v}_2(t)$. Therefore, letting $U \in M_2(C[a,b])$ be defined by $U(t) = [\mathbf{u}_1(t)\ \mathbf{u}_2(t)]$, it follows that $U(t)$ is unitary and $U(t)^*A(t)U(t)$ is diagonal for all t.

Finally, it may be observed that when $h(t)$ is real-valued, and the initial $\mathbf{v}(t_0)$ is real, construction 2.4 yields a real-valued result. Therefore, if A is real-valued, U will be real-valued. $\square$

We now relax the assumption that the eigenvalues of $A(t)$ are distinct for all t.

Theorem 2.6. *Let $A \in M_2(C[a,b])$ be continuously differentiable and suppose that $A(t)$ has distinct eigenvalues except possibly at finitely many points t and at those points $A'(t)$ has distinct eigenvalues. Then there exists a unitary matrix $U \in M_2(C[a,b])$ such that U^*AU is diagonal.*

We begin with a lemma.

Lemma 2.7. *Let $\mathbf{v} : [a,b] \to \mathbb{R}^n$ be continuously differentiable. Suppose that $\mathbf{v}(t) = 0$ for only finitely many points t. Then there exists $\mu : [a,b] \to \mathbb{R}$ continuously differentiable such that $|\mu(t)| = \|\mathbf{v}(t)\|_2$.*

Proof. Suppose there is at least one point in the open interval at which $\mathbf{v}$ vanishes. Let $a < t_1 < t_2 < \ldots < t_s < b$ be those points in (a,b)

where $\mathbf{v}(t_i) = 0$ and, for each $1 \leq i \leq s+1$, define

$$\mu(t) = (-1)^i \|\mathbf{v}(t)\|_2$$

on $[t_{i-1}, t_i]$, with $t_0 = a$ and $t_{s+1} = b$.

Fix $1 \leq i \leq s$. Since $\mathbf{v}(t)$ is continuously differentiable, we may apply Taylor's theorem to each component of $\mathbf{v}(t)$ which yields

$$\mathbf{v}(t) = \mathbf{v}(t_i) + \mathbf{v}'(t_i)(t - t_i) + \mathbf{h}(t)(t - t_i)$$

where $\mathbf{h}(t) \to \mathbf{0}$ as $t \to t_i$. By assumption, $\mathbf{v}(t_i) = 0$, so

$$\mathbf{v}(t) = \mathbf{v}'(t_i)(t - t_i) + \mathbf{h}(t)(t - t_i) = (t - t_i)\big(\mathbf{v}'(t_i) + \mathbf{h}(t)\big).$$

Then, for $t \neq t_i$,

$$\begin{aligned}
\frac{d}{dt}\|\mathbf{v}(t)\|_2 &= \frac{d}{dt}\sqrt{\mathbf{v}(t) \cdot \mathbf{v}(t)} \\
&= \frac{\mathbf{v}(t) \cdot \mathbf{v}'(t)}{\|\mathbf{v(t)}\|_2} \\
&= \frac{(t - t_i)\big(\mathbf{v}'(t_i) + \mathbf{h}(t)\big) \cdot \mathbf{v}'(t)}{\|(t - t_i)\big(\mathbf{v}'(t_i) + \mathbf{h}(t)\big)\|_2}.
\end{aligned}$$

Since $\mathbf{v}'(t)$ is continuous, if $\mathbf{v}'(t_i) \neq 0$,

$$\lim_{t \to t_i^+} \frac{d}{dt}\|\mathbf{v}(t)\|_2 = \lim_{t \to t_i^+} \frac{(t - t_i)\mathbf{v}'(t_i) \cdot \mathbf{v}'(t)}{|(t - t_i)|\|\mathbf{v}'(t_i)\|_2} = \frac{\mathbf{v}'(t_i) \cdot \mathbf{v}'(t_i)}{\|\mathbf{v}'(t_i)\|_2} = \|\mathbf{v}'(t_i)\|_2$$

and if $\mathbf{v}'(t_i) = 0$,

$$\lim_{t \to t_i^+} \frac{d}{dt}\|\mathbf{v}(t)\|_2 = \lim_{t \to t_i^+} \frac{t - t_i}{|t - t_i|}\frac{\mathbf{h}(t)}{\|\mathbf{h}(t)\|} \cdot \mathbf{v}'(t) = 0 = \|\mathbf{v}'(t_i)\|_2$$

since $\mathbf{v}'(t) \to 0$ as $t \to t_i$.

Similarly,

$$\lim_{t \to t_i^-} \frac{d}{dt}\|\mathbf{v}(t)\|_2 = -\|\mathbf{v}'(t_i)\|_2$$

Therefore, by our definition of $\mu(t)$ and this difference of sign, $\mu'(t)$ is well defined and continuous at t_i.

If $\mathbf{v}$ never vanishes on (a, b), we can take $\mu(t) = ||\mathbf{v}(t)||$. Note that it is irrelevant whether $\mathbf{v}$ vanishes at one or both endpoints. $\square$

Proof of theorem 2.6. Let $h(t) = h_r(t) + ih_c(t)$ where h_r and h_c are real-valued. For $t \in [a, b]$, the eigenvalues of $A(t)$ are as in equations (2) and (3).

Let $Z \subseteq [a, b]$ be the set of points t at which $A(t)$ does not have distinct eigenvalues; then $|Z| < \infty$ by hypothesis. Let

$$\mathbf{v}(t) = [f(t) - g(t), 2h_r(t), 2h_c(t)]^T$$

and observe that $\mathbf{v}(t) = 0$ exactly on Z. Applying Lemma 2.7 to $\mathbf{v}(\mathbf{t})$, there is some continuously differentiable $\mu : [a, b] \to \mathbb{R}$ such that

$$|\mu(t)| = \|\mathbf{v}(t)\|_2 = \sqrt{(f(t) - g(t))^2 + 4(h_r^2(t) + h_c^2(t))}.$$

Define $\lambda(t) = \frac{1}{2}[f(t) + g(t) + \mu(t)]$; then $\lambda(t)$ is continuously differentiable and $\lambda(t)$ is an eigenvalue for $A(t)$ for all $t \in [a, b]$. Let $B(t) = A(t) - \lambda(t)I$. Then $B(t)$ is continuously differentiable, $B \neq 0$ on $[a, b] \backslash Z$ and $B'(t) \neq 0$ for $t \in Z$. Applying Lemma 2.7 to $B(t)$, there exists $\tau : [a, b] \to \mathbb{R}$ continuously differentiable such that $|\tau(t)| = \|B(t)\|_2$ for all $t \in [a, b]$ (where $B(t)$ is treated as a vector with 4 entries). Next define

$$C(t) = \frac{1}{\tau(t)} B(t)$$

and observe that, by l'Hôpital's rule, we can define $C(t) = \frac{B'(t)}{\tau'(t)}$ for $t \in Z$ and $C(t)$ will be continuous on $[a, b]$. Since $\det : M_2(\mathbb{C}) \to \mathbb{C}$ is continuous and $\det(C(t)) = 0$ for all $t \notin Z$, $\det(C(t)) \equiv 0$ on $[a, b]$. Furthermore, for any $t \in [a, b]$, if $t \notin Z$ then $C(t)$ has distinct eigenvalues and if $t \in Z$, then $C(t) = \frac{B'(t)}{\tau'(t)}$ which has distinct eigenvalues. Therefore, $C(t)$ is well-defined with distinct eigenvalues and so by Theorem 2.2 we can find $U \in M_2(C[a, b])$ such that $U(t)$ is unitary and $U(t)^*C(t)U(t)$ is diagonal for all $t \in [a, b]$. Then

$$\begin{aligned}
U(t)^*A(t)U(t) &= U(t)^*(B(t) + \lambda(t)I)U(t) \\
&= U(t)^*B(t)U(t) + \lambda(t)I \\
&= \tau(t)U(t)^*C(t)U(t) + \lambda(t)I
\end{aligned}$$

is diagonal. $\qquad\square$

Example 2.8. Let

$$\varphi(t) = \begin{cases} \exp(-1/t^2) \text{ if } t \neq 0 \\ 0 \text{ if } t = 0 \end{cases}$$

and let

$$\psi(t) = \begin{cases} 0 \text{ if } t \leq 0 \\ \varphi(t) \text{ if } t > 0 \end{cases}.$$

Note that both φ and ψ are C^∞ functions, hence the matrix

$$A(t) = \begin{bmatrix} \varphi(t) & \psi(t) \\ \psi(t) & \psi(t) \end{bmatrix} = \varphi(t) \begin{bmatrix} 1 & \chi \\ \chi & \chi \end{bmatrix}$$

is C^∞ on $[-1, 1]$, where χ is the characteristic function of $[0, 1]$. Note that $A(t)$ has repeated eigenvalues only at $t = 0$.

Now 0 is an eigenvalue for $A(t)$ for every t. However, $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ is an eigenvector for 0 for $t \in [-1, 0)$ and $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ is an eigenvector for 0 for $t \in [0, 1]$. Therefore, $A(t)$ cannot be continuously diagonalized.

Since differentiability alone is not sufficient to guarantee diagonalizability of hermitian elements, one can ask if analyticity suffices.

Question 2.9. *Let $A \in M_n(C[a,b])$ be hermitian. Writing the entries of A as in (1), assume there is a domain $\mathcal{D} \subset \mathbb{C}$ whch contains $[a, b]$ and that f, g, h are analytic in $\mathcal{D}$ with f, g real on $[a, b]$. Is A diagonalizable over $[a, b]$?*

References

[1] K. Grove and G. K. Pedersen. "Diagonalizing Matrices over *C(X)*". *Journal of Functional Analysis* **59**:1 (1984), 65-89.
[2] R. Horn and C. Johnson. *Matrix Analysis*. Cambridge University Press, 1985.
[3] R. V. Kadison. "Diagonalizing Matrices over Operator Algebras". *Bulletin of the American Mathematical Society* **8**:1 (1983), 84-86.

Department of Mathematics, Iowa State University, Ames, IA 50011
Current address: Department of Mathematics, Syracuse University, 215 Carnegie Building, Syracuse, NY 13244-1150
E-mail address: `jrcyr@syr.edu`

E-mail address: `ekstrand@iastate.edu`

Current address: Department of Mathematics, Bowdoin College, Brunswick ME 04011
E-mail address: `nathanlmeyers@gmail.com`

E-mail address: `peoplesc@iastate.edu`

E-mail address: `peters@iastate.edu`