Conjugation

John McCuan

April 17, 2022

Russell Newton asked a question about the significance of conjugation with respect to diagonalizable linear functions (operators) $L : V \to V$. I think I offered a fairly solid explanation from the point of view of the matrix framework for linear transformations on $\mathbb{R}^n$ or $\mathbb{C}^n$, but my attempt to adapt that discussion to the more abstract and general context of $L : V \to V$ was somewhat lacking, not to mention confusing. In the end, I am not able to see any particular compelling discussion that strictly parallels the discussion for matrices which I will outline below (as a review). From a somewhat different perspective, however, I think there is something interesting to say about conjugation for linear functions, so I will try to present that in a second section.

1 The Matrix Framework

As may be familiar to many of you, if you start with an $n \times n$ matrix $A = (a_{ij})$ with either real or complex entries, then one can view the matrix as corresponding to a linear function $L : \mathbb{R}^n \to \mathbb{R}^n$ (in the former case) or $L : \mathbb{C}^n \to \mathbb{C}^n$ (in either case) given by matrix multiplication

$$L\mathbf{x} = A\mathbf{x},$$

and the diagonalizability of the matrix A is essentially equivalent to what we have introduced as the definition of the digaonalizability of the function L, namely, the existence of a basis $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ of eigenvectors. In this setting, the basic objective is to find the diagonal matrix corresponding to L and thus related to A. The nature of the relation is nicely illustrated by a **matrix/mapping diagram**:

$$\begin{array}{ccccc} & & A & & \\ & V & \longrightarrow & V & \\ C & \downarrow & & \downarrow & C \\ & V & \longrightarrow & V & \\ & & D & & \end{array}$$

Each of the labels on the arrows in this diagram represents a matrix and accordingly a basis is required to make the initial connection between A and L as well as between each of the matrices C and D and some linear function. The usual interpretation is that the standard unit basis vectors are chosen as the basis for each of the spaces V in the diagram, but D is the matrix of the linear function L with respect to the basis of eigenvectors, while A is the matrix of L with respect to the standard unit basis vectors. This kind of "thinking of the vector space V with respect to, or attached to, a particular basis is not natural to and does not lend itself to the consideration of linear functions the **linear mapping diagram** considered below, but it works well in this context of matrices. More precisely, we take the matrix C^{-1} as the matrix with columns $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_j$ which may be viewed as corresponding to the linear function assigning the standard unit basis vector $\mathbf{e}_j$ to the vector $\mathbf{v}_j$ (with respect to the standard unit basis $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$. Then the matrix D which we "read off" from the matrix/mapping diagram is the **matrix conjugation**

$$D = CAC^{-1}.$$

Assuming each of the basis vectors $\mathbf{v}_j$ is an eigenvector of L with $L\mathbf{v}_j = \lambda_j \mathbf{v}_j$, we can compute the matrix D as follows:

$$D\mathbf{e}_j = CAC^{-1}\mathbf{e}_j = CA\mathbf{v}_j = C(\lambda_j v_j) = \lambda_j C v_j = \lambda_j \mathbf{e}_j.$$

Thus, D is the diagonal matrix with the eigenvalue λ_j of L in the j-th diagonal entry. This is pretty straightforward, and from the mapping point of view, we can simply say that the matrix of L with respect to the basis $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is the diagonal matrix D.

Exercise 1 *Given a diagonalizable matrix A corresponding to $L : \mathbb{R}^n \to \mathbb{R}^n$, or more generally a diagonalizable linear function $L : V \to V$ defined on an n-dimensional vector space V, show there can be no more than n distinct eigenvalues, and denoting the eigenspaces corresponding to distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_m$ by*

$$E_j = \{\mathbf{v} \in \mathbb{R}^n : A\mathbf{v} = \lambda_j \mathbf{v}\} \qquad \textit{for } j = 1, 2, \ldots, m\}$$

there holds

$$\sum_{j=1}^{m} \dim E_j = n.$$

2 Linear Mapping Framework

We recall that we can denote by $\mathcal{L}(V \to V)$ the finite dimensional vector space of linear functions defined on an n-dimensional vector space V. There is an operation of composition on this vector space, but the operation does not lead to a group structure. Not every linear function has an inverse with respect to composition, and typically not even the nonzero elements can be expected to have inverses in general.

There is a subset $\mathrm{Aut}(V \to V)$ of invertible elements in $\mathcal{L}(V \to V)$ which contains the identity mapping and is a group under composition. Note that $\mathrm{Aut}(V \to V)$ is not a vector space; it is neither closed under addition nor does it contain the zero mapping. Nevertheless, this subset has an interesting use with regard to the diagonalizable linear functions. Let

$$\Delta(V \to V) = \{L \in \mathcal{L}(V \to V) : L \text{ is diagonalizable}\}.$$

Exercise 2 *Let $\mathcal{I} = \Delta(V \to V) \cap \mathrm{Aut}(V \to V)$ denote the intersection of the diagonalizable linear functions with the invertible linear functions in $\mathcal{L}(V \to V)$.*

(a) *Characterize the intersection $\mathcal{I}$.*

(b) *Give an example to show the intersection $\mathcal{I}$ is not (typically) closed under addition.*

(c) *Under what circumstances if $\mathcal{I} \cup \{\zeta\}$, where $\zeta : V \to V$ by $\zeta v \equiv \mathbf{0}$ is the zero mapping, closed under addition?*

Let us focus only on the eigenvalues with multiplicity as an **unordered conditional**. More precisely, given an unordered set

$$\gamma = \{(\lambda_1, k_1), (\lambda_2, k_2), \ldots, (\lambda_m, k_m)\} \tag{1}$$

with $\lambda_1, \ldots, \lambda_m$ distinct elements of F and $k_j \in \mathbb{N} = \{1, 2, 3, \ldots\}$ with

$$\sum_{j=1}^{m} k_j = n = \dim V,$$

we can consider

$$P_\gamma = \{L \in \Delta(V \to V) : \lambda_j \text{ is an eigenvalue of } L \text{ with } \dim E_j = k_j,\ j = 1, 2, \ldots, m\}$$

and

$$E_j = \{v \in V : Lv = \lambda_j v\} \qquad \text{for} \qquad j = 1, 2, \ldots, m.$$

Each diagonalizable linear function $L \in \Delta(V \to V)$ is in precisely one P_γ. That is, denoting by Γ the indexing set of all sets γ as described in (1)

$$\Delta(V \to V) = \cup_{\gamma \in \Gamma} P_\gamma$$

is a **partition** of $\Delta(V \to V)$; each of the sets P_γ is nonempty and each pair P_γ and P_η with γ and η distinct are disjoint:

$$P_\gamma \cap P_\eta = \phi, \qquad \gamma \neq \eta.$$

The partition sets P_γ do not typically all have the same number of elements.

Exercise 3 *Give examples showing partition sets in $\Delta(V \to V)$ with different numbers of elements.*

The partition sets P_γ described above determined by an unordered collection of γ of eigenvalues with multiplicities has a particular structure:

Theorem 1 *(similarity theorem) If γ is an index set as described in (1) and L_0 is any fixed element of P_γ, then*

$$P_\gamma = \{T L_0 T^{-1} : T \in \mathrm{Aut}(V \to V)\}.$$

As per this theorem partition sets P_γ of the kind described above are called **conjugacy classes** or similarity classes and correspond to a mapping diagram:

$$\begin{array}{ccccc} & & L_0 & & \\ & V & \longrightarrow & V & \\ T & \downarrow & & \downarrow & T \\ & V & \longrightarrow & V & \\ & & L = T L_0 T^{-1} & & \end{array}$$

As mentioned above no notion of the vector space V with any particular basis should be attached to this diagram; that only leads to confusion. I won't give the full proof of the partition/conjugacy result stated above, but I will mention some sets and functions that are useful in organizing and carrying out the details of that proof. First of all, we can let

$$\mathcal{B}_0 = \{v_1, v_2, \dots, v_n\}$$

be a basis of V of eigenvectors for L_0. In order to determine a relation between L_0 and another linear function L with the same unordered eigenvalues and corresponding multiplicities, it is useful to organize and order the eigenvalues associated with each $v_j \in \mathcal{B}_0$. Let

$$A_\ell = \{v_j \in \mathcal{B}_0 : (\lambda_\ell, v_j) \text{ is an eigenvalue/eigenvector pair for } L_0\}$$

where $\lambda_1, \lambda_2, \dots, \lambda_m$ are the distinct eigenvalues of L_0. Similarly, if L is in the same partition P_γ containing L_0 and/or $L = TL_0T^{-1}$ for some $T \in \text{Aut}(V \to V)$, then the distinct eigenvalues of L are also $\lambda_1, \lambda_2, \dots, \lambda_m$, and we can similarly define

$$B_\ell = \{w_j \in \mathcal{B} : (\lambda_\ell, w_j) \text{ is an eigenvalue/eigenvector pair for } L\}$$

and $\mathcal{B} = \{w_1, w_2, \dots, w_n\}$ is a basis of V consisting of eigenvectors of L.

Notice that we can first find a **permutation**

$$\sigma : \{1, 2, 3, \dots, n\} \to \{1, 2, 3, \dots n\}$$

which is just a bijection such that

$$\mu_{\sigma(1)} \le \mu_{\sigma(2)} \le \mu_{\sigma(3)} \le \cdots \le \mu_{\sigma(n)}$$

where $L_0 v_j = \mu_j v_j$. Similarly, there is a permutation

$$\tau : \{1, 2, 3, \dots, n\} \to \{1, 2, 3, \dots n\}$$

for which

$$\nu_{\sigma(1)} \le \nu_{\sigma(2)} \le \nu_{\sigma(3)} \le \cdots \le \nu_{\sigma(n)}$$

with $Lw_j = \nu_j v_j$.

In order to show

$$P_\gamma \subset \{TL_0T^{-1} : T \in \text{Aut}(V \to V)\}$$

one must identify[1] a linear automorphism T for which $L = TL_0T^{-1}$. This can be accomplished by determining the appropriate correspondence between $\mathcal{B}$ and $\mathcal{B}_0$ using the permutations σ and τ above.

[1]This automorphism will not be unique in general because the ordering of the corresponding eigenvectors in eigenspaces of dimension greater than one can be switched.

In order to show

$$\{TL_0T^{-1} : T \in \mathrm{Aut}(V \to V)\} \subset P_\gamma$$

one needs to show the linear function $L = TL_0T^{-1} : V \to V$ has the same eigenvalues with the same multiplicities as L_0. Again, it should be emphasized that L is a different linear function from L_0; what the two mappings share in common are the collection of eigenvalues with corresponding multiplicities; they are not somehow "the same linear functions with respect to different bases." That doesn't make any sense. It can be said that the matrix of L with respect to the basis $\mathcal{B}_0 = \{v_1, v_2, \ldots, v_n\}$ is the same as the matrix of L_0 with respect to the basis $\mathcal{B} = \{w_1, w_2, \ldots, w_n\}$, but I'm not really seeing much useful information in that.[2] From the higher level point of view of classing distinct linear functions together on the basis of sharing the same eigenvalues (with corresponding multiplicities), the notion of conjugate linear mappings seems to have some significance.

[2]Maybe there is useful information in this assertion, but I don't know what it is.