Jordan Transformations in two dimensions

John McCuan

April 17, 2022

Generally we consider a linear function $L : V \to V$ admitting a two-dimensional invariant subspace[1] $W \subset V$ containing a one-dimensional eigenspace

$$\mathrm{span}\{v\}$$

associated with an eigenvalue λ and such that no vector $w \in W \setminus \mathrm{span}\{v\}$ is an eigenvector. We focus attention on the restriction

$$L\Big|_W : W \to W,$$

and for most of the discussion the values of L at points outside of the subspace W will be of little or no consequence. Accordingly we refer to the restriction simply as $L : W \to W$. For any vector $w \in W \setminus \mathrm{span}\{v\}$ we have

$$W = \mathrm{span}\{v\} \oplus \mathrm{span}\{w\}$$

and we take as our first task to analyze the image vector

$$Lw = av + bw.$$

Note that we know $Lv = \lambda v$, so the first column of the matrix of L is $(\lambda, 0)^\perp$:

$$A = \begin{pmatrix} \lambda & * \\ 0 & * \end{pmatrix}.$$

[1]Note that V may be infinite dimensional here.

Lemma 1 *If* $w_0 \in W \setminus \operatorname{span}\{v\}$ *where* W *is a two-dimensional invariant subspace as above with exactly one one-dimensional eigensubspace spanned by an eigenvector* v*, and* w_0 *is* **not** *an eigenvector, then*

$$Lw_0 = av + \lambda w_0 \qquad \text{where} \qquad a \neq 0.$$

Proof: As mentioned above

$$W = \operatorname{span}\{v\} \oplus \operatorname{span}\{w_0\}$$

and $Lw_0 = av + bw_0$ for some unique a and b in the field. We know $a \neq 0$ because otherwise w_0 is an eigenvector. This of course holds even if there might be another eigenvector within $W \setminus \operatorname{span}\{v\}$.

Since $a \neq 0$, and $Lw_0 = av + bw_0$, it follows that

$$(L - \lambda\,\mathrm{id})w_0 = av + (b - \lambda)w_0 \neq \mathbf{0} \tag{1}$$

$$(L - \lambda\,\mathrm{id})(L - \lambda\,\mathrm{id})w_0 = (b - \lambda)(L - \lambda\,\mathrm{id})w_0. \tag{2}$$

The computation (2) follows quickly from the observation that $(L - \lambda\,\mathrm{id})v = \mathbf{0}$ because (λ, v) is an eigenvalue/eigenvector pair. We conclude from (1-2) that $(L - \lambda\,\mathrm{id})w_0$ is an eigenvector for $L - \lambda\,\mathrm{id} : W \to W$. In particular,

$$L(L - \lambda\,\mathrm{id})w_0 = \lambda(L - \lambda\,\mathrm{id})w_0 + (b - \lambda)(L - \lambda\,\mathrm{id})w_0 = b(L - \lambda\,\mathrm{id})w_0. \tag{3}$$

Exercise 1 *Compute* $L(L - \lambda\,\mathrm{id})w_0$ *directly using* (1) *to obtain* (3).

This means $(L - \lambda\,\mathrm{id})w_0$ is an eigenvector for L in W. The associated eigenvalue is b, but the only available eigenvalue for the restriction is λ. Therefore $b = \lambda$ and the assertion of the lemma is established. $\square$

In view of the lemma the matrix of the restriction of L to W with respect to the basis $\{v, w_0\}$ where v is any eigenvector in W and w_0 is any vector with $w_0 \in W \setminus \operatorname{span}\{v\}$ is

$$A = \begin{pmatrix} \lambda & a \\ 0 & \lambda \end{pmatrix} \qquad \text{with} \qquad a \neq 0. \tag{4}$$

It is traditional to make a specific choice of the vector $w = w_0$ in relation to the particular eigenvector v under consideration chosen to make the resulting matrix (4) both definite and, in some sense, "simplest."

Given a vector $w \in W \setminus \operatorname{span}\{v\}$ then with $Lw = av + \lambda w_0$ we take $w_1 = w/a$ so that

$$Lw_1 = \frac{1}{a} Lw = v + \frac{1}{a}\ \lambda w = v + \lambda w_1.$$

The matrix of L (restricted to W) with respect to the basis $\{v, w_1\}$ is

$$J = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix},$$

and this is called the **Jordan form matrix**.

There is one last observation I would like to make about this two-dimensional case: While the vector $w = w_1$ is **not an eigenvector**, it may be considered as an element in what is called a **cyclic basis**.[2] To see the cyclic form we use the operator $L - \lambda\,\mathrm{id}$ which played the important role of generating an eigenvector in the proof of the lemma. To be precise, $(L - \lambda\,\mathrm{id})w_1 = v$ so that the basis $\{v, w_1\}$ may be written as

$$\{(L - \lambda\,\mathrm{id})w_1, w_1\}.$$

Generally, the **cycle** in the cyclic basis starts from the end of the basis:

$$w,\ (L - \lambda\,\mathrm{id})w.$$

In this case, we can choose any vector $w \in W \setminus \operatorname{span}\{v\}$ to get a cyclic basis

$$\{(L - \lambda\,\mathrm{id})w = v, w\}$$

with respect to which the matrix of L has the form

$$A = \begin{pmatrix} \lambda & a \\ 0 & \lambda \end{pmatrix},$$

and we can choose $w_1 = w/a$ to get the particular cyclic/Jordan basis with respect to which we get

$$J = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}.$$

This situation for higher dimensional Jordan subspaces is not quite so simple.

[2] Some textbooks, including Axler's *Linear Algebra Done Right,* call the elements of the cyclic basis "generalized eigenvectors."

Higher dimensional Jordan subspaces

It is not difficult to imagine ordering additional cyclic basis vectors to obtain an invariant subspace containing precisely one one-dimensional eigenspace. Proceeding from the opposite direction this does occur and may be characterized as follows: Again in the setting where $L : V \to V$ is a linear function, if W is a k-dimensional invariant subspace in which v is an eigenvector and each $w \in W \setminus \operatorname{span}\{v\}$ is **not an eigenvector**, then there is some vector $w_1 \in W$ for which

$$\{w_1, (L - \lambda\,\mathrm{id})w_1, (L - \lambda\,\mathrm{id})^2 w_1, \ldots, (L - \lambda\,\mathrm{id})^{k-1} w_1\}$$

is a basis for W with

$$(L - \lambda\,\mathrm{id})^{k-1} w_1 = v.$$

Reversing the order we have the basis

$$\{v, (L - \lambda\,\mathrm{id})^{k-2} w_1, \ldots, (L - \lambda\,\mathrm{id})w_1, w_1\} \tag{5}$$

with respect to which the matrix of the restriction

$$L = L\big|_W$$

is

$$J = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & & 0 \\ 0 & 0 & \lambda & \ddots & \vdots \\ \vdots & & & \ddots & 1 \\ 0 & \cdots & & 0 & \lambda \end{pmatrix} \tag{6}$$

with

$$\begin{aligned} a_{jj} &= \lambda \qquad \text{for} \qquad j = 1, 2, \ldots, k, \\ a_{j,j+1} &= 1 \qquad \text{for} \qquad j = 1, 2, \ldots, k-1, \\ a_{ij} &= 0 \qquad \text{for} \qquad \text{otherwise.} \end{aligned}$$

Exercise 2 *Verify the matrix of the restriction L with respect to the basis* (5) *is given by the matrix* (6).

Exercise 3 *Show that in the case of a k-dimensional Jordan subspace described above, the subspace*

$$\mathrm{span}\{v, (L - \lambda\,\mathrm{id})^{k-2} w_1\} \tag{7}$$

is a two dimensional Jordan invariant subspace. Show that Jordan subspaces of all intermediate dimensions $\ell = 3, 4, \ldots, k-1$ can be generated by adjoining specific vectors w to the subspace given in (7). Characterize all possible choices of w at each step.

The last exercise illustrates the comment at the end of the last section suggesting the vector w_1 generating the cyclic basis must be chosen with some care. If one were to choose any $w \in W \setminus \mathrm{span}\{v\}$ in the higher dimensional case, one might choose, for example,

$$w = (L - \lambda\,\mathrm{id})^{k-2} w_1$$

and only generate the basis in (7).

Nothing we have said so far is restricted to the case of a real vector space V. In particular, it should be noted that the existence of non-trivial Jordan shear may be present in a linear function $L : V \to V$ defined on a complex vector space. Any such linear function is not diagonalizable. Indeed, we now state a kind of general result which is specifically restricted, not to only real vector spaces, but to **only complex vector spaces**.

Theorem 1 *(Jordan decomposition theorem) If V is a finite dimensional complex vector space and $L : V \to V$ is linear, then there exist invariant subspaces $W_1, W_2, \ldots, W_m$ with*

$$V = \bigoplus_{j=1}^{m} W_j$$

and associated with each W_j for $j = 1, 2, \ldots, m$ is an eigenvalue λ_j for which there is an eigenvector $v_j \in W_j$ with $Lv_j = \lambda_j v_j$ and each $w \in W_j \setminus \mathrm{span}\{v_j\}$ is not an eigenvector, but there is a vector $w_j \in W_j$ generating a cyclic basis

$$\{v_j, (L - \lambda\,\mathrm{id})^{k_j - 2} w_j, \ldots, (L - \lambda\,\mathrm{id}) w_j, w_j\}$$

where $k_j = \dim W_j$ so that the matrix of L with respect to the concatenation of these bases has the **block form**

$$\begin{pmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & J_m \end{pmatrix}$$

with each j_j a Jordan form matrix as indicated in (6) *with the eigenvalue λ_j on the diagonal. Note we are "allowing" here the degenerate case of a "Jordan subspace of dimension one" so that $J_j = (\lambda_j)$ is possible. Thus, the theorem includes the diagonalizable case for linear functions on finite dimensional complex vector spaces as well as diagonalizability on invariant subspaces of various dimensions. The situation as described by this theorem is as complicated as it can get, for example, for $L : \mathbb{C}^n \to \mathbb{C}^n$.*

The situation is actually a bit **more complicated** if V is allowed to be a real vector space. A finite dimensional real vector space can have a two-dimensional invariant subspace with **no real eigenvalue/eigenvector pair**.

Exercise 4 *Say $L : \mathbb{C}^2 \to \mathbb{C}^2$ satisfies the following*

(i) *$Lv \in \mathbb{R}^2$ for each $v \in \mathbb{R}^2 \subset \mathbb{C}^2$. Notice we are not saying $\mathbb{R}^2$ is an invariant subspace here because $\mathbb{R}^2$ is not a subspace of $\mathbb{C}^2$. We are simply saying $\mathbb{R}^2$ is an* **invariant subset**.

(ii) *There does not exist any vector $v \in \mathbb{R}^2 \backslash \{\mathbf{0}\}$ for which*

$$Lv \in \{cv : c \in \mathbb{R}\}.$$

Show the following:

(a) *The restriction*

$$T = L\Big|_{\mathbb{R}^2} : \mathbb{R}^2 \to \mathbb{R}^2$$

is a linear function on $\mathbb{R}^2$.

(b) *There are no (real) eigenvectors of T in $\mathbb{R}^2$.*

(b) *There are no (complex) eigenvectors of L in $\mathbb{R}^2$.*

Here is the main question: What is the Jordan block structure of L in this case?