8C Consequences of Generalized Eigenspace Decomposition

Square Roots of Operators

Recall that a square root of an operator $T \in \mathcal{L}(V)$ is an operator $R \in \mathcal{L}(V)$ such that $R^2 = T$ (see 7.36). Every complex number has a square root, but not every operator on a complex vector space has a square root. For example, the operator on $\mathbf{C}^3$ defined by $T(z_1, z_2, z_3) = (z_2, z_3, 0)$ does not have a square root, as you are asked to show in Exercise 1. The noninvertibility of that operator is no accident, as we will soon see. We begin by showing that the identity plus any nilpotent operator has a square root.

8.39 *identity plus nilpotent has a square root*

Suppose $T \in \mathcal{L}(V)$ is nilpotent. Then $I + T$ has a square root.

Proof Consider the Taylor series for the function $\sqrt{1+x}$:

8.40 $$\sqrt{1+x} = 1 + a_1 x + a_2 x^2 + \cdots.$$

We do not find an explicit formula for the coefficients or worry about whether the infinite sum converges because we use this equation only as motivation.

Because $a_1 = \frac{1}{2}$, the formula above implies that $1 + \frac{x}{2}$ is a good estimate for $\sqrt{1+x}$ when x is small.

Because T is nilpotent, $T^m = 0$ for some positive integer m. In 8.40, suppose we replace x with T and 1 with I. Then the infinite sum on the right side becomes a finite sum (because $T^k = 0$ for all $k \geq m$). Thus we guess that there is a square root of $I + T$ of the form

$$I + a_1 T + a_2 T^2 + \cdots + a_{m-1} T^{m-1}.$$

Having made this guess, we can try to choose $a_1, a_2, \ldots, a_{m-1}$ such that the operator above has its square equal to $I + T$. Now

$$\begin{aligned}\bigl(I + a_1 T &+ a_2 T^2 + a_3 T^3 + \cdots + a_{m-1} T^{m-1}\bigr)^2 \\ &= I + 2a_1 T + \bigl(2a_2 + a_1^2\bigr) T^2 + (2a_3 + 2a_1 a_2) T^3 + \cdots \\ &\quad + (2a_{m-1} + \text{terms involving } a_1, \ldots, a_{m-2}) T^{m-1}.\end{aligned}$$

We want the right side of the equation above to equal $I + T$. Hence choose a_1 such that $2a_1 = 1$ (thus $a_1 = 1/2$). Next, choose a_2 such that $2a_2 + a_1^2 = 0$ (thus $a_2 = -1/8$). Then choose a_3 such that the coefficient of T^3 on the right side of the equation above equals 0 (thus $a_3 = 1/16$). Continue in this fashion for each $k = 4, \ldots, m-1$, at each step solving for a_k so that the coefficient of T^k on the right side of the equation above equals 0. Actually we do not care about the explicit formula for the a_k's. We only need to know that some choice of the a_k's gives a square root of $I + T$. ■

The previous lemma is valid on real and complex vector spaces. However, the result below holds only on complex vector spaces. For example, the operator of multiplication by -1 on the one-dimensional real vector space $\mathbf{R}$ has no square root.

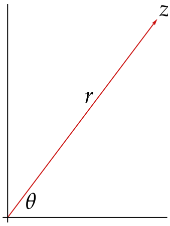
For the proof below, we need to know that every $z \in \mathbf{C}$ has a square root in $\mathbf{C}$. To show this, write

$$z = r(\cos\theta + i\sin\theta),$$

where r is the length of the line segment in the complex plane from the origin to z and θ is the angle of that line segment with the positive horizontal axis. Then

$$\sqrt{r}\left(\cos\tfrac{\theta}{2} + i\sin\tfrac{\theta}{2}\right)$$

is a square root of z, as you can verify by showing that the square of the complex number above equals z.



Representation of a complex number with polar coordinates.

8.41 *over $\mathbf{C}$, invertible operators have square roots*

Suppose V is a complex vector space and $T \in \mathcal{L}(V)$ is invertible. Then T has a square root.

Proof Let $\lambda_1, \ldots, \lambda_m$ be the distinct eigenvalues of T. For each k, there exists a nilpotent operator $T_k \in \mathcal{L}\big(G(\lambda_k, T)\big)$ such that $T|_{G(\lambda_k,T)} = \lambda_k I + T_k$ [see 8.22(b)]. Because T is invertible, none of the λ_k's equals 0, so we can write

$$T|_{G(\lambda_k,T)} = \lambda_k\left(I + \frac{T_k}{\lambda_k}\right)$$

for each k. Because T_k/λ_k is nilpotent, $I + T_k/\lambda_k$ has a square root (by 8.39). Multiplying a square root of the complex number λ_k by a square root of $I + T_k/\lambda_k$, we obtain a square root R_k of $T|_{G(\lambda_k,T)}$.

By the generalized eigenspace decomposition (8.22), a typical vector $v \in V$ can be written uniquely in the form

$$v = u_1 + \cdots + u_m,$$

where each u_k is in $G(\lambda_k, T)$. Using this decomposition, define an operator $R \in \mathcal{L}(V)$ by

$$Rv = R_1 u_1 + \cdots + R_m u_m.$$

You should verify that this operator R is a square root of T, completing the proof. ■

By imitating the techniques in this subsection, you should be able to prove that if V is a complex vector space and $T \in \mathcal{L}(V)$ is invertible, then T has a k^{th} root for every positive integer k.

Jordan Form

We know that if V is a complex vector space, then for every $T \in \mathcal{L}(V)$ there is a basis of V with respect to which T has a nice upper-triangular matrix (see 8.37). In this subsection we will see that we can do even better—there is a basis of V with respect to which the matrix of T contains 0's everywhere except possibly on the diagonal and the line directly above the diagonal.

We begin by looking at two examples of nilpotent operators.

8.42 example: *nilpotent operator with nice matrix*

Let T be the operator on $\mathbf{C}^4$ defined by

$$T(z_1, z_2, z_3, z_4) = (0, z_1, z_2, z_3).$$

Then $T^4 = 0$; thus T is nilpotent. If $v = (1, 0, 0, 0)$, then T^3v, T^2v, Tv, v is a basis of $\mathbf{C}^4$. The matrix of T with respect to this basis is

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

The next example of a nilpotent operator has more complicated behavior than the example above.

8.43 example: *nilpotent operator with slightly more complicated matrix*

Let T be the operator on $\mathbf{C}^6$ defined by

$$T(z_1, z_2, z_3, z_4, z_5, z_6) = (0, z_1, z_2, 0, z_4, 0).$$

Then $T^3 = 0$; thus T is nilpotent. In contrast to the nice behavior of the nilpotent operator of the previous example, for this nilpotent operator there does not exist a vector $v \in \mathbf{C}^6$ such that $T^5v, T^4v, T^3v, T^2v, Tv, v$ is a basis of $\mathbf{C}^6$. However, if we take $v_1 = (1, 0, 0, 0, 0, 0)$, $v_2 = (0, 0, 0, 1, 0, 0)$, and $v_3 = (0, 0, 0, 0, 0, 1)$, then $T^2v_1, Tv_1, v_1, Tv_2, v_2, v_3$ is a basis of $\mathbf{C}^6$. The matrix of T with respect to this basis is

$$\left(\begin{array}{ccc} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} & \begin{matrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{matrix} & \begin{matrix} 0 \\ 0 \\ 0 \end{matrix} \\ \begin{matrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{matrix} & \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} & \begin{matrix} 0 \\ 0 \end{matrix} \\ \begin{matrix} 0 & 0 & 0 \end{matrix} & \begin{matrix} 0 & 0 \end{matrix} & \begin{pmatrix} 0 \end{pmatrix} \end{array}\right).$$

Here the inner matrices are blocked off to show that we can think of the 6-by-6 matrix above as a block diagonal matrix consisting of a 3-by-3 block with 1's on the line above the diagonal and 0's elsewhere, a 2-by-2 block with 1 above the diagonal and 0's elsewhere, and a 1-by-1 block containing 0.

Our next goal is to show that every nilpotent operator $T \in \mathcal{L}(V)$ behaves similarly to the operator in the previous example. Specifically, there is a finite collection of vectors $v_1, \dots, v_n \in V$ such that there is a basis of V consisting of the vectors of the form $T^j v_k$, as k varies from 1 to n and j varies (in reverse order) from 0 to the largest nonnegative integer m_k such that $T^{m_k} v_k \neq 0$. With respect to this basis, the matrix of T looks like the matrix in the previous example. More specifically, T has a block diagonal matrix with respect to this basis, with each block a square matrix that is 0 everywhere except on the line above the diagonal.

In the next definition, the diagonal of each A_k is filled with some eigenvalue λ_k of T, the line directly above the diagonal of A_k is filled with 1's, and all other entries in A_k are 0 (to understand why each λ_k is an eigenvalue of T, see 5.41). The λ_k's need not be distinct. Also, A_k may be a 1-by-1 matrix (λ_k) containing just an eigenvalue of T. If each λ_k is 0, then the next definition captures the behavior described in the paragraph above (recall that if T is nilpotent, then 0 is the only eigenvalue of T).

8.44 definition: ***Jordan basis***

Suppose $T \in \mathcal{L}(V)$. A basis of V is called a *Jordan basis* for T if with respect to this basis T has a block diagonal matrix

$$\begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_p \end{pmatrix}$$

in which each A_k is an upper-triangular matrix of the form

$$A_k = \begin{pmatrix} \lambda_k & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_k \end{pmatrix}.$$

Most of the work in proving that every operator on a finite-dimensional complex vector space has a Jordan basis occurs in proving the special case below of nilpotent operators. This special case holds on real vector spaces as well as complex vector spaces.

8.45 ***every nilpotent operator has a Jordan basis***

Suppose $T \in \mathcal{L}(V)$ is nilpotent. Then there is a basis of V that is a Jordan basis for T.

Proof We will prove this result by induction on $\dim V$. To get started, note that the desired result holds if $\dim V = 1$ (because in that case, the only nilpotent operator is the 0 operator). Now assume that $\dim V > 1$ and that the desired result holds on all vector spaces of smaller dimension.

Let m be the smallest positive integer such that $T^m = 0$. Thus there exists $u \in V$ such that $T^{m-1}u \neq 0$. Let

$$U = \operatorname{span}\bigl(u, Tu, \ldots, T^{m-1}u\bigr).$$

The list $u, Tu, \ldots, T^{m-1}u$ is linearly independent (see Exercise 2 in Section 8A). If $U = V$, then writing this list in reverse order gives a Jordan basis for T and we are done. Thus we can assume that $U \neq V$.

Note that U is invariant under T. By our induction hypothesis, there is a basis of U that is a Jordan basis for $T|_U$. The strategy of our proof is that we will find a subspace W of V such that W is also invariant under T and $V = U \oplus W$. Again by our induction hypothesis, there will be a basis of W that is a Jordan basis for $T|_W$. Putting together the Jordan bases for $T|_U$ and $T|_W$, we will have a Jordan basis for T.

Let $\varphi \in V'$ be such that $\varphi\bigl(T^{m-1}u\bigr) \neq 0$. Let

$$W = \Bigl\{v \in V : \varphi\bigl(T^k v\bigr) = 0 \text{ for each } k = 0, \ldots, m-1\Bigr\}.$$

Then W is a subspace of V that is invariant under T $\Bigl($the invariance holds because if $v \in W$ then $\varphi\bigl(T^k(Tv)\bigr) = 0$ for $k = 0, \ldots, m-1$, where the case $k = m-1$ holds because $T^m = 0\Bigr)$. We will show that $V = U \oplus W$, which by the previous paragraph will complete the proof.

To show that $U + W$ is a direct sum, suppose $v \in U \cap W$ with $v \neq 0$. Because $v \in U$, there exist $c_0, \ldots, c_{m-1} \in \mathbf{F}$ such that

$$v = c_0 u + c_1 Tu + \cdots + c_{m-1} T^{m-1} u.$$

Let j be the smallest index such that $c_j \neq 0$. Apply T^{m-j-1} to both sides of the equation above, getting

$$T^{m-j-1} v = c_j T^{m-1} u,$$

where we have used the equation $T^m = 0$. Now apply φ to both sides of the equation above, getting

$$\varphi\bigl(T^{m-j-1} v\bigr) = c_j \varphi\bigl(T^{m-1} u\bigr) \neq 0.$$

The equation above shows that $v \notin W$. Hence we have proved that $U \cap W = \{0\}$, which implies that $U + W$ is a direct sum (see 1.46).

To show that $U \oplus W = V$, define $S\colon V \to \mathbf{F}^m$ by

$$Sv = \Bigl(\varphi(v), \varphi(Tv), \ldots, \varphi\bigl(T^{m-1}v\bigr)\Bigr).$$

Thus null $S = W$. Hence

$$\dim W = \dim \operatorname{null} S = \dim V - \dim \operatorname{range} S \geq \dim V - m,$$

where the second equality comes from the fundamental theorem of linear maps (3.21). Using the inequality above, we have

$$\dim(U \oplus W) = \dim U + \dim W \geq m + (\dim V - m) = \dim V.$$

Thus $U \oplus W = V$ (by 2.39), completing the proof. ■

Now the generalized eigenspace decomposition allows us to extend the previous result to operators that may not be nilpotent. Doing this requires that we deal with complex vector spaces.

Camille Jordan (1838–1922) published a proof of 8.46 in 1870.

8.46 *Jordan form*

Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. Then there is a basis of V that is a Jordan basis for T.

Proof Let $\lambda_1, \ldots, \lambda_m$ be the distinct eigenvalues of T. The generalized eigenspace decomposition states that

$$V = G(\lambda_1, T) \oplus \cdots \oplus G(\lambda_m, T),$$

where each $(T - \lambda_k I)|_{G(\lambda_k, T)}$ is nilpotent (see 8.22). Thus 8.45 implies that some basis of each $G(\lambda_k, T)$ is a Jordan basis for $(T - \lambda_k I)|_{G(\lambda_k, T)}$. Put these bases together to get a basis of V that is a Jordan basis for T. ■

Exercises 8C

1 Suppose $T \in \mathcal{L}(\mathbf{C}^3)$ is the operator defined by $T(z_1, z_2, z_3) = (z_2, z_3, 0)$. Prove that T does not have a square root.

2 Define $T \in \mathcal{L}(\mathbf{F}^5)$ by $T(x_1, x_2, x_3, x_4, x_5) = (2x_2, 3x_3, -x_4, 4x_5, 0)$.

(a) Show that T is nilpotent.

(b) Find a square root of $I + T$.

3 Suppose V is a complex vector space. Prove that every invertible operator on V has a cube root.

4 Suppose V is a real vector space. Prove that the operator $-I$ on V has a square root if and only if $\dim V$ is an even number.

5 Suppose $T \in \mathcal{L}(\mathbf{C}^2)$ is the operator defined by $T(w, z) = (-w - z, 9w + 5z)$. Find a Jordan basis for T.

6 Find a basis of $\mathcal{P}_4(\mathbf{R})$ that is a Jordan basis for the differentiation operator D on $\mathcal{P}_4(\mathbf{R})$ defined by $Dp = p'$.

7 Suppose $T \in \mathcal{L}(V)$ is nilpotent and $v_1, \ldots, v_n$ is a Jordan basis for T. Prove that the minimal polynomial of T is z^{m+1}, where m is the length of the longest consecutive string of 1's that appears on the line directly above the diagonal in the matrix of T with respect to $v_1, \ldots, v_n$.

8 Suppose $T \in \mathcal{L}(V)$ and $v_1, \ldots, v_n$ is a basis of V that is a Jordan basis for T. Describe the matrix of T^2 with respect to this basis.

9 Suppose $T \in \mathcal{L}(V)$ is nilpotent. Explain why there exist $v_1, \dots, v_n \in V$ and nonnegative integers $m_1, \dots, m_n$ such that (a) and (b) below both hold.
(a) $T^{m_1}v_1, \dots, Tv_1, v_1, \dots, T^{m_n}v_n, \dots, Tv_n, v_n$ is a basis of V.
(b) $T^{m_1+1}v_1 = \cdots = T^{m_n+1}v_n = 0$.

10 Suppose $T \in \mathcal{L}(V)$ and $v_1, \dots, v_n$ is a basis of V that is a Jordan basis for T. Describe the matrix of T with respect to the basis $v_n, \dots, v_1$ obtained by reversing the order of the v's.

11 Suppose $T \in \mathcal{L}(V)$. Explain why every vector in each Jordan basis for T is a generalized eigenvector of T.

12 Suppose $T \in \mathcal{L}(V)$ is diagonalizable. Show that $\mathcal{M}(T)$ is a diagonal matrix with respect to every Jordan basis for T.

13 Suppose $T \in \mathcal{L}(V)$ is nilpotent. Prove that if $v_1, \dots, v_n$ are vectors in V and $m_1, \dots, m_n$ are nonnegative integers such that

$$T^{m_1}v_1, \dots, Tv_1, v_1, \dots, T^{m_n}v_n, \dots, Tv_n, v_n \text{ is a basis of } V$$

and

$$T^{m_1+1}v_1 = \cdots = T^{m_n+1}v_n = 0,$$

then $T^{m_1}v_1, \dots, T^{m_n}v_n$ is a basis of $\operatorname{null} T$.

This exercise shows that $n = \dim \operatorname{null} T$. Thus the positive integer n that appears above depends only on T and not on the specific Jordan basis chosen for T.

14 Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. Prove that there does not exist a direct sum decomposition of V into two nonzero subspaces invariant under T if and only if the minimal polynomial of T is of the form $(z - \lambda)^{\dim V}$ for some $\lambda \in \mathbf{C}$.