Additional Topics in Linear Algebra Supplementary Material for Math 540

Joseph H. Silverman

E-mail address: jhs@math.brown.edu

Mathematics Department, Box 1917 Brown University, Providence, RI 02912 USA

Contents

CHAPTER 1

Characteristic Polynomials

As always, V is a finite dimensional vector space with field of scalars $\mathbb{F}$. But we do not assume that V is an inner product space.

1. The Characteristic Polynomial of a Linear Map

We begin with a fundamental definition.

Definition. Let $T \in \mathcal{L}(V)$ be a linear map. The *characteristic polynomial of* T is

$$P_T(z) = \det(zI - T).$$

We begin by showing that $P_T(z)$ may be computed using any matrix associated to T.

Proposition 1. *Let $\{\mathbf{v}_i\}$ be a basis for V, and let*

$$A = \mathcal{M}\big(T, \{\mathbf{v}_i\}\big) \in \mathrm{Mat}(n, \mathbb{F})$$

be the matrix for T associated to the chosen basis. Then

$$P_T(z) = \det(zI - A).$$

In particular, we can compute $P_T(z)$ using the matrix for T associated to any basis.

Proof. We proved that if $S \in \mathcal{L}(V)$ is any linear map, then the determinant of the matrix $B = \mathcal{M}\big(S, \{\mathbf{v}_i\}\big)$ doesn't depend on the choice of basis. That's what allows us to define $\det(S)$. In case you've forgotten, the proof uses the fact that if we choose some other basis $\{\mathbf{w}_i\}$, then the associated matrix $C = \mathcal{M}\big(S, \{\mathbf{w}_i\}\big)$ satisfies $C = G^{-1}BG$ for some invertible matrix G, so

$$\begin{aligned}\det(C) = \det(G^{-1}BG) &= \det(G^{-1})\det(B)\det(G) \\ &= \frac{1}{\det(G)}\det(B)\det(G) = \det(B).\end{aligned}$$

We now need merely observe that

$$\mathcal{M}\big(zI - T, \{\mathbf{v}_i\}\big) = z\mathcal{M}\big(I, \{\mathbf{v}_i\}\big) - \mathcal{M}\big(T, \{\mathbf{v}_i\}\big) = zI - A,$$

so by definition, $\det(zI - T) = \det(zI - A)$, and it doesn't matter what basis we use to compute A. $\square$

EXAMPLE 2. Let $T \in \mathcal{L}(\mathbb{R}^2)$ be given by

$$T(x, y) = (x + 2y, 3x + 4y).$$

Then the matrix of T for the standard basis of $\mathbb{R}^2$ is

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}.$$

So the characteristic polynomial of T is

$$\begin{aligned} P_T(z) = \det(zI - A) &= \det\left(\begin{pmatrix} z & 0 \\ 0 & z \end{pmatrix} - \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\right) \\ &= \det\begin{pmatrix} z-1 & -2 \\ -3 & z-4 \end{pmatrix} \\ &= (z-1)(z-4) - 6 \\ &= z^2 - 5z - 2. \end{aligned}$$

PROPOSITION 3. *Let $T \in \mathcal{L}(V)$. Then the roots of the characteristic polynomial $P_T(z)$ in $\mathbb{C}$ are exactly the eigenvalues of T.*

PROOF. We know that the eigenvalues of T are precisely the numbers $\lambda \in \mathbb{C}$ for which $\lambda I - T$ is not invertible. And we also proved that a linear map S is not invertible if and only if $\det(S) = 0$. Hence the eigenvalues of T are the numbers $\lambda \in \mathbb{C}$ such that $\det(\lambda I - T) = 0$, which are exactly the roots of $P_T(z)$. □

2. Jordan Blocks

We know that if V has a basis of eigenvectors for a linear map $T \in \mathcal{L}(V)$, then the matrix of T for that basis is diagonal. Thus if $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a basis for V and if

$$T\mathbf{v}_i = \lambda_i \mathbf{v}_i \quad \text{for } 1 \le i \le n,$$

then

$$\mathcal{M}\big(T, \{\mathbf{v}_i\}\big) = \begin{pmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

Unfortunately, not every linear transformation has a basis of eigenvectors. Our goal in the next two sections is to find bases that are almost as good.

EXAMPLE 4. Consider the linear map $T \in \mathcal{L}(\mathbb{F}^2)$ given by the formula

$$T(x_1, x_2) = (x_1 + x_2, x_2).$$

The matrix of T for the standard basis of $\mathbb{F}^2$ is $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$, so the characteristic polynomial of T is

$$\det(zI - T) = \det \begin{pmatrix} z-1 & -1 \\ 0 & z-1 \end{pmatrix} = (z-1)^2,$$

so Proposition 3 says that the only eigenvalue of T is $\lambda = 1$. But one easily checks that the only vectors satisfying $T\mathbf{v} = \mathbf{v}$ are multiples of $(1, 0)$, so $\mathbb{F}^2$ does not have a basis consisting of eigenvectors for T.

Generalizing Example 4, we are led to look at matrices of the following form.

DEFINITION. The *Jordan matrix* (or *Jordan block*) of dimension m and eigenvalue λ is the m-by-m matrix

$$J_m(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & 0 & \cdots & 0 \\ 0 & 0 & \lambda & 1 & \cdots & 0 \\ & \vdots & & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & 0 & 0 & \cdots & \lambda \end{pmatrix}$$

Thus $J_m(\lambda)$ is an m-by-m matrix with λ's on its main diagonal and with 1's just above the main diagonal, and all other entries are 0.

PROPOSITION 5. *Let $J_m(\lambda)$ be a Jordan matrix. The only eigenvalue of $J_m(\lambda)$ is λ, and the only eigenvectors of $J_m(\lambda)$ are multiples of the standard basis vector $\mathbf{e}_1$.*

PROOF. Let $J = J_m(\lambda)$. The matrix J is upper triangular, so its characteristic polynomial is

$$P_J(z) = \det(zI - J) = (z - \lambda)^m.$$

Proposition 3 tells us that λ is the only eigenvalue of J. Next suppose that $J\mathbf{v} = \lambda\mathbf{v}$, so

$$(J - \lambda I)\mathbf{v} = \mathbf{0}.$$

But looking at the matrix $J - \lambda I$, we see that

$$(J-\lambda I)\mathbf{v} = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 \\ & \vdots & & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{m-1} \\ x_m \end{pmatrix} = \begin{pmatrix} x_2 \\ x_3 \\ x_4 \\ \vdots \\ x_m \\ 0 \end{pmatrix}.$$

Hence $(J - \lambda I)\mathbf{v} = \mathbf{0}$ if and only if

$$x_2 = x_3 = \cdots = x_m = 0,$$

which is just another way of saying that $\mathbf{v}$ is a multiple of $\mathbf{e}_1$. □

We note that the application of $J = J_m(\lambda)$ to the standard basis vectors of $\mathbb{F}^m$ exhibits a sort of shift effect,

$$\begin{array}{lclllllll} J\mathbf{e}_1 & = & \lambda\mathbf{e}_1 & & & & & & \\ J\mathbf{e}_2 & = & \mathbf{e}_1 & + & \lambda\mathbf{e}_2 & & & & \\ J\mathbf{e}_3 & = & & & \mathbf{e}_2 & + & \lambda\mathbf{e}_3 & & \\ \vdots & & & & & & \ddots & & \\ J\mathbf{e}_m & = & & & & & \mathbf{e}_{m-1} & + & \lambda\mathbf{e}_m \end{array}$$

It is often convenient to write a Jordan block as

$$J = \lambda I + N \quad \text{with} \quad N = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 \\ & \vdots & & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 \end{pmatrix}. \tag{1}$$

In other words, J is the sum of a multiple of the identity matrix and a matrix N that has 1's just above the diagonal, and 0's everywhere else. When we compute the powers of the matrix N, we find something surprising.

PROPOSITION 6. *Let N be the m-by-m matrix described in* (1). *Then*

$$N^m = 0.$$

PROOF. The effect of N on the list $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_m$ of standard basis vectors for $\mathbb{F}^m$ is to shift the list to the right, deleting $\mathbf{e}_m$ and putting a zero vector in front. In other words,

$$N\mathbf{e}_1 = \mathbf{0}, \;\; N\mathbf{e}_2 = \mathbf{e}_1, \;\; N\mathbf{e}_3 = \mathbf{e}_2, \;\; N\mathbf{e}_4 = \mathbf{e}_3, \;\; \ldots \;\; N\mathbf{e}_m = \mathbf{e}_{m-1}.$$

What happens if we apply N again? The list shifts again, yielding

$$N^2\mathbf{e}_1 = \mathbf{0},\ \ N^2\mathbf{e}_2 = \mathbf{0},\ \ N^2\mathbf{e}_3 = \mathbf{e}_1,\ \ N^2\mathbf{e}_4 = \mathbf{e}_2,\ \ \ldots\ \ N^2\mathbf{e}_m = \mathbf{e}_{m-2}.$$

Applying N again, the list shifts yet one step further. So if we apply N a total of m times, all of the $\mathbf{e}_i$ are shifted out of the list, and we're just left with zero vectors,

$$N^m\mathbf{e}_1 = \mathbf{0},\ \ N^m\mathbf{e}_2 = \mathbf{0},\ \ N^m\mathbf{e}_3 = \mathbf{0},\ \ N^m\mathbf{e}_4 = \mathbf{0},\ \ \ldots\ \ N^m\mathbf{e}_m = \mathbf{0}.$$

This shows that N^m sends all of the basis vectors $\mathbf{e}_1, \ldots, \mathbf{e}_m$ to $\mathbf{0}$, so N^m is the zero matrix. □

There is a name for matrices (or linear maps) that have the property described in Proposition 6.

Definition. A linear map $T \in \mathcal{L}(V)$ is said to be *nilpotent* if there is some integer $j \geq 1$ such that $T^j = 0$. Similarly, a square matrix A is said to be *nilpotent* if there is some integer $j \geq 1$ such that $A^j = 0$.

3. Jordan Normal Form

We can now define matrices that are "almost, but not necessarily entirely, diagonal."

Definition. A matrix is in *Jordan Normal Form* if it looks like

$$A = \begin{pmatrix} J_1 & 0 & 0 & \cdots & 0 \\ 0 & J_2 & 0 & \cdots & 0 \\ 0 & 0 & J_3 & \cdots & 0 \\ & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & J_r \end{pmatrix},$$

where each J_i is a Jordan block. Thus J_i is an m_i-by-m_i Jordan block with eigenvalue λ_i,

$$J_i = \begin{pmatrix} \lambda_i & 1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_i & 1 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_i & 1 & \cdots & 0 \\ & \vdots & & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_i & 1 \\ 0 & 0 & 0 & 0 & \cdots & \lambda_i \end{pmatrix}.$$

EXAMPLE 7. The 6-by-6 matrix

$$A = \begin{pmatrix} 3 & 1 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 5 & 1 & 0 \\ 0 & 0 & 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 0 & 0 & 5 \end{pmatrix},$$

consists of three Jordan blocks. The first is a 2-by-2 Jordan block with eigenvalue 3, the second is a 1-by-1 Jordan block with eigenvalue 5, and the third is a 3-by-3 Jordan block with eigenvalue 5. Thus

$$A = \begin{pmatrix} J_1 & 0 & 0 \\ 0 & J_2 & 0 \\ 0 & 0 & J_3 \end{pmatrix} \quad \text{with } J_1 = \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}, \quad J_2 = \big(5\big), \quad J_3 = \begin{pmatrix} 5 & 1 & 0 \\ 0 & 5 & 1 \\ 0 & 0 & 5 \end{pmatrix}.$$

Note that different blocks need not have different eigenvalues.

As we have seen, not every linear transformation T can be diagonalized. But we can come close in the sense that we can always find a basis that puts T into Jordan normal form, as described in the following important result.

THEOREM 8. (Jordan Normal Form Theorem) *Let V be a vector space over $\mathbb{C}$, and let $T \in \mathcal{L}(V)$ be a linear transformation. Then there is a basis $\{\mathbf{v}_i\}$ for V so that the matrix $\mathcal{M}\big(T, \{\mathbf{v}_i\}\big)$ is in Jordan normal form.*

PROOF. We won't have time in class to prove this theorem, but you can find a proof in the book if you're interested. You may also someday see it proven in a more advanced mathematics course as a special case of a general theorem on finitely generated modules over principal ideal domains. □

4. The Cayley-Hamilton Theorem

We are now ready to state and prove a useful theorem about linear maps. The proof uses the Jordan Normal Form Theorem (Theorem 8).

THEOREM 9. (Cayley-Hamilton Theorem) *Let $T \in \mathcal{L}(V)$ and let $P_T(z)$ be its characteristic polynomial. Then $P_T(T) = 0$.*

EXAMPLE 10. Before proving the Cayley-Hamilton Theorem, we illustrate it using the linear map T in Example 2. That map had associated matrix $A = \left(\begin{smallmatrix}1 & 2\\ 3 & 4\end{smallmatrix}\right)$ and characteristic polynomial $P_T(z) = z^2 - 5z - 2$. We compute

$$\begin{aligned} P_T(A) &= A^2 - 5A - 2I \\ &= \begin{pmatrix}1 & 2\\ 3 & 4\end{pmatrix}^2 - 5\begin{pmatrix}1 & 2\\ 3 & 4\end{pmatrix} - 2\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} \\ &= \begin{pmatrix}7 & 10\\ 15 & 22\end{pmatrix} - \begin{pmatrix}5 & 10\\ 15 & 20\end{pmatrix} - \begin{pmatrix}2 & 0\\ 0 & 2\end{pmatrix} \\ &= \begin{pmatrix}0 & 0\\ 0 & 0\end{pmatrix}. \end{aligned}$$

Since the matrix of $P_T(T)$ is $P_T(A)$, this shows that $P_T(T) = 0$.

Proof of the Cayley-Hamilton Theorem (Theorem 9). We give the proof in the case that $\mathbb{F} = \mathbb{C}$.[1] The Jordan Normal Form Theorem (Theorem 8) says that there is a basis $\{\mathbf{v}_i\}$ for V such that the matrix $A = \mathcal{M}(T, \{\mathbf{v}_i\})$ of T is in Jordan normal form,

$$A = \begin{pmatrix} J_1 & 0 & 0 & \cdots & 0 \\ 0 & J_2 & 0 & \cdots & 0 \\ 0 & 0 & J_3 & \cdots & 0 \\ & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & J_r \end{pmatrix}.$$

Here J_i is an m_i-by-m_i Jordan block with eigenvalue λ_i,

$$J_i = \begin{pmatrix} \lambda_i & 1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_i & 1 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_i & 1 & \cdots & 0 \\ & \vdots & & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_i & 1 \\ 0 & 0 & 0 & 0 & \cdots & \lambda_i \end{pmatrix}.$$

In particular, the matrix A is upper triangular, so the matrix $zI - A$ is also upper triangular. This allows us to easily compute the characteristic polynomial

$$P_T(z) = \det(zI - A) = (z - \lambda_1)^{m_1}(z - \lambda_2)^{m_2} \cdots (z - \lambda_r)^{m_r}.$$

What happens when we substitute $z = A$ into the polynomial

$$(z - \lambda_1)^{m_1}(z - \lambda_2)^{m_2} \cdots (z - \lambda_r)^{m_r}?$$

[1] If $\mathbb{F} = \mathbb{R}$, one may "extend scalars" and treat V as if it were a complex vector space, but we haven't discussed how this process works, so we'll be content to just work with $\mathbb{F} = \mathbb{C}$.

Let's consider the first factor $(A-\lambda_1 I)^{m_1}$. It looks like

$$
\begin{aligned}
&(A-\lambda_1 I)^{m_1}\\
&\quad=\begin{pmatrix}
J_1-\lambda_1 I & 0 & 0 & \cdots & 0\\
0 & J_2-\lambda_1 I & 0 & \cdots & 0\\
0 & 0 & J_3-\lambda_1 I & \cdots & 0\\
 & \vdots & & \ddots & \vdots\\
0 & 0 & 0 & \cdots & J_r-\lambda_1 I
\end{pmatrix}^{m_1}\\
&\quad=\begin{pmatrix}
(J_1-\lambda_1 I)^{m_1} & 0 & 0 & \cdots & 0\\
0 & (J_2-\lambda_1 I)^{m_1} & 0 & \cdots & 0\\
0 & 0 & (J_3-\lambda_1 I)^{m_1} & \cdots & 0\\
 & \vdots & & \ddots & \vdots\\
0 & 0 & 0 & \cdots & (J_r-\lambda_1 I)^{m_1}
\end{pmatrix}\\
&\quad=\begin{pmatrix}
0 & 0 & 0 & \cdots & 0\\
0 & (J_2-\lambda_1 I)^{m_1} & 0 & \cdots & 0\\
0 & 0 & (J_3-\lambda_1 I)^{m_1} & \cdots & 0\\
 & \vdots & & \ddots & \vdots\\
0 & 0 & 0 & \cdots & (J_r-\lambda_1 I)^{m_1}
\end{pmatrix},
\end{aligned}
$$

where the 0 in the upper left corner is there because Proposition 6 tells us that $(J_1-\lambda_1 I)^{m_1}=0$.

Similarly,

$$
(A-\lambda_2 I)^{m_2}=\begin{pmatrix}
(J_1-\lambda_2 I)^{m_2} & 0 & 0 & \cdots & 0\\
0 & 0 & 0 & \cdots & 0\\
0 & 0 & (J_3-\lambda_2 I)^{m_2} & \cdots & 0\\
 & \vdots & & \ddots & \vdots\\
0 & 0 & 0 & \cdots & (J_r-\lambda_2 I)^{m_2}
\end{pmatrix},
$$

(notice the 0 in place of $(J_2-\lambda_2 I)^{m_2}$), and so on until we get to

$$
(A-\lambda_r I)^{m_r}=\begin{pmatrix}
(J_1-\lambda_r I)^{m_r} & 0 & 0 & \cdots & 0\\
0 & (J_2-\lambda_r I)^{m_r} & 0 & \cdots & 0\\
0 & 0 & (J_3-\lambda_r I)^{m_r} & \cdots & 0\\
 & \vdots & & \ddots & \vdots\\
0 & 0 & 0 & \cdots & 0
\end{pmatrix}.
$$

If we multiply them together, we get the following product, where we write a $*$ to indicate a possibly non-zero block. (If you want to be precise, you can say that each $*$ looks like $(J_i - \lambda_k)^{m_k}$ for some i and k.)

$$
\begin{aligned}
P_T(A) &= (A-\lambda_1)^{m_1}(A-\lambda_2)^{m_2}\cdots(A-\lambda_r)^{m_r} \\
&= \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ 0 & * & 0 & \cdots & 0 \\ 0 & 0 & * & \cdots & 0 \\ & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & * \end{pmatrix} \times \begin{pmatrix} * & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & * & \cdots & 0 \\ & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & * \end{pmatrix} \\
&\qquad \times \begin{pmatrix} * & 0 & 0 & \cdots & 0 \\ 0 & * & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & * \end{pmatrix} \times \cdots \times \begin{pmatrix} * & 0 & 0 & \cdots & 0 \\ 0 & * & 0 & \cdots & 0 \\ 0 & 0 & * & \cdots & 0 \\ & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix} \\
&= \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}.
\end{aligned}
$$

This proves that $P_T(A)$ is the zero matrix. Hence

$$\mathcal{M}\bigl(P_T(T), \{\mathbf{v}_i\}\bigr) = P_T\Bigl(\mathcal{M}\bigl(T, \{\mathbf{v}_i\}\bigr)\Bigr) = P_T(A) = 0.$$

Thus the matrix associated to $P_T(T)$ is the zero matrix, so $P_T(T)$ is the zero linear map. This concludes the proof of the Cayley-Hamilton Theorem. □

Exercises for Chapter 1

Problem # C.1. Let $T \in \mathcal{L}(V)$, and suppose that $\mathbf{v}_1, \ldots, \mathbf{v}_n$ is a basis of eigenvectors for T, say

$$T\mathbf{v}_1 = \lambda_1\mathbf{v}_1, \quad T\mathbf{v}_2 = \lambda_2\mathbf{v}_2, \ldots \quad T\mathbf{v}_n = \lambda_n\mathbf{v}_n.$$

Some of the eigenvalues might be the same, so let

$$\mu_1, \mu_2, \ldots, \mu_k$$

be a list of the distinct eigenvalues in the set $\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$. Let $F(z)$ be the polynomial

$$F(z) = (z-\mu_1)(z-\mu_2)\ldots(z-\mu_k).$$

Prove that $F(T) = 0$.

This shows that for some linear maps T, there may be a non-zero polynomial $F(z)$ with $F(T) = 0$ whose degree is smaller than the degree of $P_T(z)$.

Problem # C.2. Let B be the matrix

$$B = \begin{pmatrix} 0 & 0 & \dots & 0 & -c_0 \\ 1 & 0 & \dots & 0 & -c_1 \\ 0 & 1 & \dots & 0 & -c_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & -c_{n-1} \end{pmatrix}.$$

Prove that

$$\det(zI - B) = z^n + c_{n-1}z^{n-1} + \cdots + c_1 z + c_0.$$

This exercise shows that every monic polynomial is the characteristic polynomial of some matrix. The matrix B is often called the *companion matrix* to the polynomial.

Problem # C.3. We have seen that the set of linear maps $\mathcal{L}(V)$ is a vector space. Let $T \in \mathcal{L}(V)$, and for each $k = 1, 2, 3, \dots$, let

$$U_k = \operatorname{Span}(I, T, T^2, \dots, T^k) \subset \mathcal{L}(V).$$

So U_k is a subspace of $\mathcal{L}(V)$, and clearly $\dim(U_k) \le k + 1$, since it has a spanning set consisting of $k + 1$ vectors. Prove that

$$\dim(U_k) \le \dim(V) \quad \text{for all } k.$$

Problem # C.4. If $T \in \mathcal{L}(V)$ is a nilpotent, prove that 0 is the only eigenvalue of T.

Problem # C.5. Let $n = \dim(V)$ and let $T \in \mathcal{L}(V)$ be a nilpotent linear map. Prove that $T^n = 0$. (By definition, the fact that T is nilpotent means that $T^j = 0$ for some j. This exercise asks you to prove that it suffices to take $j = n$.)

Problem # C.6. Let $J = J_m(\lambda)$ be a Jordan block matrix, and let $\mathbf{e}_m = (0, 0, \dots, 0, 1)$ be the last vector in the standard basis for $\mathbb{F}^m$. Prove that

$$\{\mathbf{e}_m, J\mathbf{e}_m, J^2\mathbf{e}_m, \dots, J^{m-1}\mathbf{e}_m\}$$

is a basis for $\mathbb{F}^m$.

CHAPTER 2

Linear Recursions

A linear recursion is a sequence of numbers

$$a_1, a_2, a_3, a_4, \ldots$$

in which each successive entry in the sequence is created by taking a linear combination of the previous few entries. A good intuition is to view a_n as representing the value of some quantity at time n. For example, a_n could be the population of a city in year n, or the amount of money in circulation in year n, or the amount of CO_2 in the atmosphere in year n. In many situations, one can create a model in which a_n is determined (at least approximately) by a linear function of the values in the previous two or three or four years.

1. What is a Linear Recursion?

We start with a formal definition.

DEFINITION. Let

$$L(x_1, \ldots, x_d) = c_1x_1 + c_2x_2 + \cdots + c_dx_d$$

be a linear function with coefficients $c_1, \ldots, c_d \in \mathbb{F}$, and let $a_1, \ldots, a_d \in \mathbb{F}$ be some initial values. The *linear recursion* generated by the linear function L and the chosen initial values is the list of numbers

$$a_1, a_2, a_3, a_4, \ldots$$

created by setting each successive a_n equal to L evaluated at the previous d entries in the sequence. Thus in terms of formulas,

$$\begin{aligned} a_{d+1} &= L(a_d, a_{d-1}, \ldots, a_2, a_1), \\ a_{d+2} &= L(a_{d+1}, a_d, \ldots, a_3, a_2), \\ a_{d+3} &= L(a_{d+2}, a_{d+1}, \ldots, a_4, a_3), \end{aligned}$$

and so on. So in general,

$$\begin{aligned} a_n &= L(a_{n-1}, a_{n-2}, \ldots, a_{n-d+1}, a_{n-d}) \\ &= c_1a_{n-1} + c_2a_{n-2} + \cdots + c_da_{n-d}. \end{aligned}$$

Example 11. The most famous linear recursion is undoubtedly the Fibonacci sequence

$$1,\ 1,\ 2,\ 3,\ 5,\ 8,\ 13,\ 21,\ 34,\ 55,\ \ldots\ . \tag{2}$$

It is generated by the following initial values and linear function:

$$F_1 = 1, \qquad F_2 = 1, \qquad L(x_1, x_2) = x_1 + x_2.$$

Thus after the first two terms, the subsequent terms are generated using the familiar formula

$$F_n = F_{n-1} + F_{n-2} \quad \text{for } n = 3, 4, 5, \ldots.$$

Fibonacci described the sequence that now bears his name when posed a question about rabbit population growth!

One of our goals is to find explicit formulas for the n^{th} term of a linear recursion, and to use these explicit formulas to estimate how fast a_n grows.

2. An Example: The Fibonacci Sequence

In this section we work out in detail an explicit formula and estimated growth rate for the Fibonacci sequence

$$F_1 = 1, \qquad F_2 = 1, \qquad F_n = F_{n-1} + F_{n-2} \quad \text{for } n = 3, 4, 5, \ldots.$$

The first ten terms of the Fibonacci sequence (2) don't look that large, but don't be fooled. The value of F_n grows very rapidly as n increases. For example,

$$F_{100} = 354224848179261915075 \approx 3.54 \cdot 10^{20}.$$

The key to analyzing linear recursions is to reformulate them in terms of linear transformations. For the Fibonacci sequence, we look at the matrix

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

If we apply A to a vector $\mathbf{v} = (x_1, x_2) \in \mathbb{F}^2$, we get

$$A\mathbf{v} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 \\ x_1 \end{pmatrix}.$$

Notice that useful $x_1 + x_2$ in the lower left corner of $A\mathbf{v}$. If we take the coordinates of $\mathbf{v}$ to be consecutive Fibonacci numbers, then we find that

$$\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} F_{n-1} \\ F_{n-2} \end{pmatrix} = \begin{pmatrix} F_{n-1} + F_{n-2} \\ F_{n-1} \end{pmatrix} = \begin{pmatrix} F_n \\ F_{n-1} \end{pmatrix}.$$

And if we apply the matrix A again, then we get

$$\begin{pmatrix}1&1\\1&0\end{pmatrix}^2\begin{pmatrix}F_{n-1}\\F_{n-2}\end{pmatrix}=\begin{pmatrix}1&1\\1&0\end{pmatrix}\begin{pmatrix}F_n\\F_{n-1}\end{pmatrix}=\begin{pmatrix}F_{n+1}\\F_n\end{pmatrix}.$$

And so on.

In general, if we start with $\mathbf{v}=(1,0)$, we obtain the formula

$$\begin{pmatrix}F_{n+1}\\F_n\end{pmatrix}=\begin{pmatrix}1&1\\1&0\end{pmatrix}^n\begin{pmatrix}1\\0\end{pmatrix}=A^n\begin{pmatrix}1\\0\end{pmatrix}. \tag{3}$$

In some sense, this gives a formula for F_n, but it's only useful if we have some convenient way of computing $A^n=\left(\begin{smallmatrix}1&1\\1&0\end{smallmatrix}\right)^n$ when n is large.

If A were a diagonal matrix, then it would be easy to compute A^n. In order to turn A into a diagonal matrix, we look for a basis of $\mathbb{F}^2$ that consists of eigenvectors of A. This is easy to do. The eigenvalues of A are the roots of

$$\det(zI-A)=\det\begin{pmatrix}z-1&-1\\-1&z\end{pmatrix}=z^2-z-1,$$

which are the numbers

$$\alpha=\frac{1+\sqrt{5}}{2}\quad\text{and}\quad\beta=\frac{1-\sqrt{5}}{2}.$$

A little bit of algebra yields eigenvectors

$$\mathbf{v}_1=\begin{pmatrix}\alpha\\1\end{pmatrix}\ \text{ and }\ \mathbf{v}_2=\begin{pmatrix}\beta\\1\end{pmatrix}\ \text{ satisfying }\ A\mathbf{v}_1=\alpha\mathbf{v}_1\ \text{ and }\ A\mathbf{v}_2=\beta\mathbf{v}_2.$$

So if we form the matrix B whose columns are $\mathbf{v}_1$ and $\mathbf{v}_2$, then

$$B=\begin{pmatrix}\alpha&\beta\\1&1\end{pmatrix}\quad\text{satisfies}\quad AB=B\begin{pmatrix}\alpha&0\\0&\beta\end{pmatrix}.$$

In other words, if we let Δ be the diagonal matrix

$$\Delta=\begin{pmatrix}\alpha&0\\0&\beta\end{pmatrix},\quad\text{then we get}\quad A=B\Delta B^{-1}.$$

It is now easy to compute A^n, since powers of a diagonal matrix are easy to compute. Here's what we get:

$$\begin{aligned}A^n&=(B\Delta B^{-1})^n\\&=\underbrace{(B\Delta B^{-1})(B\Delta B^{-1})(B\Delta B^{-1})\cdots(B\Delta B^{-1})}_{n\text{ copies of }B^{-1}\Delta B}\\&=B\Delta^nB^{-1}\quad\text{since the }B^{-1}B\text{ products cancel.}\end{aligned} \tag{4}$$

We also have

$$\Delta^n = \begin{pmatrix} \alpha^n & 0 \\ 0 & \beta^n \end{pmatrix} \quad \text{and} \quad B^{-1} = \begin{pmatrix} \frac{1}{\alpha-\beta} & \frac{-\beta}{\alpha-\beta} \\ \frac{-1}{\alpha-\beta} & \frac{\alpha}{\alpha-\beta} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{1-\sqrt{5}}{2\sqrt{5}} \\ \frac{-1}{\sqrt{5}} & \frac{1+\sqrt{5}}{2\sqrt{5}} \end{pmatrix}. \tag{5}$$

We use these formula to derive a beautiful closed formula for the n^{th} Fibonacci number.

$$\begin{aligned}
\begin{pmatrix} F_{n+1} \\ F_n \end{pmatrix} &= A^n \begin{pmatrix} 1 \\ 0 \end{pmatrix} && \text{from (3),} \\
&= B\Delta^n B^{-1} \begin{pmatrix} 1 \\ 0 \end{pmatrix} && \text{from (4),} \\
&= \begin{pmatrix} \alpha & \beta \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \alpha^n & 0 \\ 0 & \beta^n \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{1-\sqrt{5}}{2\sqrt{5}} \\ \frac{-1}{\sqrt{5}} & \frac{1+\sqrt{5}}{2\sqrt{5}} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} && \text{from (5),} \\
&= \begin{pmatrix} \alpha & \beta \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \alpha^n & 0 \\ 0 & \beta^n \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{5}} \\ \frac{-1}{\sqrt{5}} \end{pmatrix} && \text{matrix multiplication,} \\
&= \begin{pmatrix} \alpha & \beta \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{5}}\alpha^n \\ -\frac{1}{\sqrt{5}}\beta^n \end{pmatrix} && \text{matrix multiplication,} \\
&= \begin{pmatrix} \frac{1}{\sqrt{5}}(\alpha^{n+1} - \beta^{n+1}) \\ \frac{1}{\sqrt{5}}(\alpha^n - \beta^n) \end{pmatrix} && \text{matrix multiplication.}
\end{aligned}$$

Equating the bottom entries of these vectors gives an explicit closed expression for the n^{th} Fibonacci number,

$$F_n = \frac{\alpha^n - \beta^n}{\sqrt{5}} = \frac{1}{\sqrt{5}} \left[\left(\frac{1+\sqrt{5}}{2} \right)^n + \left(\frac{1-\sqrt{5}}{2} \right)^n \right]. \tag{6}$$

This famous formula is known as *Binet's Formula*, although it had been discovered by others long before Binet found it.

We can use Binet's formula to measure how fast the Fibonacci sequence grows. Notice that

$$\frac{1+\sqrt{5}}{2} = 1.618\ldots \quad \text{and} \quad \frac{1-\sqrt{5}}{2} = -0.618\ldots.$$

Since $\left|\frac{1-\sqrt{5}}{2}\right| < 1$, we see that $\left(\frac{1-\sqrt{5}}{2}\right)^n \to 0$ very rapidly as $n \to \infty$. So for large values of n we have

$$F_n = \left(\frac{1+\sqrt{5}}{2} \right)^n + (\text{something very tiny}).$$

Hence

$$F_n \approx \left(\frac{1+\sqrt{5}}{2}\right)^n \approx 1.61803398875^n \approx 10^{0.209n}.$$

So F_n has roughly $0.209n$ digits. Let's check. We saw earlier that $F_{100} \approx 3.54 \cdot 10^{20}$ has 21 digits, while $0.209 \times 100 = 20.9$, which is a pretty good estimate. And F_{100000} will have more than $20,000$ digits, so we probably don't want to try to write it down exactly!

3. The Matrix Associated to a Linear Recursion

In Section 2 we described the Fibonacci sequence in terms of the powers of matrix. We now generalize this construction to arbitrary linear recursions. We let

$$a_1, a_2, a_3, \ldots$$

be a linear recursion generated by the linear function

$$L(x_1, \ldots, x_d) = c_1x_1 + c_2x_2 + \cdots + c_dx_d.$$

We associate to L the d-by-d matrix

$$A_L = \begin{pmatrix} c_1 & c_2 & \cdots & c_{d-1} & c_d \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & 0 \end{pmatrix}. \tag{7}$$

Example 12. The recursion

$$a_n = 4a_{n-1} - 6a_{n-2} + 4a_{n-3}$$

is generated by the linear form

$$L(x_1, x_2, x_3) = 4x_1 - 6x_2 + 4x_3.$$

The associated matrix is

$$A_L = \begin{pmatrix} 4 & -6 & 4 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}. \tag{8}$$

With this notation, the recursion satisfied by the a_n gives a matrix equation

$$\begin{pmatrix} a_{n+d} \\ a_{n+d-1} \\ \vdots \\ \\ a_{n+1} \end{pmatrix} = \begin{pmatrix} c_1 & c_2 & \cdots & c_{d-1} & c_d \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} a_{n+d-1} \\ a_{n+d-2} \\ \vdots \\ \\ a_n \end{pmatrix}. \tag{9}$$

To ease notation, we define vectors

$$\mathbf{v}_n = \begin{pmatrix} a_{n+d-1} \\ a_{n+d-2} \\ \vdots \\ \vdots \\ a_n \end{pmatrix},$$

which allows us to write the matrix formula (9) in the succint form

$$\mathbf{v}_{n+1} = A_L \mathbf{v}_n. \tag{10}$$

Using (10) repeatedly, we find that

$$\mathbf{v}_n = A_L \mathbf{v}_{n-1} = A_L^2 \mathbf{v}_{n-2} = A_L^3 \mathbf{v}_{n-3} = \cdots = A_L^{n-1} \mathbf{v}_1.$$

The coordinates of the vector $\mathbf{v}_1$ are the initial values of the recursion, while the coordinates of $\mathbf{v}_n$ are later terms in the recursion. Thus in order to compute the recursion, we need to compute the powers A_L^n of the matrix A_L.

The first step is to compute the characteristic polynomial.

Proposition 13. *The characteristic polynomial of the matrix L is the polynomial*

$$f_L(z) = z^d - c_1 z^{d-1} - \cdots - c_{d-1} z - c_d.$$

In other words,

$$f_L(z) = \det(zI - A_L) = \det \begin{pmatrix} z - c_1 & -c_2 & \cdots & -c_{d-1} & -c_d \\ -1 & z & 0 & \cdots & 0 \\ 0 & -1 & z & \cdots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & -1 & z \end{pmatrix}.$$

Proof. Expanding $\det(zI - A_L)$ down the first column, there are only two terms, so we find that

$$\begin{aligned} &\det(zI - A_L) \\ &= (z - c_1) \det \begin{pmatrix} z & 0 & \cdots & 0 \\ -1 & z & \cdots & 0 \\ & \ddots & \ddots & \vdots \\ 0 & 0 & -1 & z \end{pmatrix} + \det \begin{pmatrix} -c_2 & \cdots & -c_{d-1} & -c_d \\ -1 & z & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & -1 & z \end{pmatrix}. \end{aligned}$$

The first matrix is lower triangular, so its determinant is just z^{d-1}. The second matrix looks very much like our original matrix, with a non-zero top row, z's on the digaonal, and -1's just below the diagonal. So expanding the second matrix along it's first column, we get

$$\det(zI - A_L) = (z - c_1)z^{d-1} - c_2 z^{d-2} + \det \begin{pmatrix} * & \cdots & * \\ \vdots & \ddots & \vdots \\ * & \cdots & * \end{pmatrix},$$

where the new remaining matrix again has non-zero top row, z's on the digaonal, and -1's just below the diagonal. More precisely, the top row now consists of $-c_3, -c_4, \ldots, -c_d$. Continuing in this fashion, we find that

$$\det(zI - A_L) = (z - c_1)z^{d-1} - c_2 z^{d-2} - \cdots - c_{d-1}z - c_d = f_L(z).$$

This concludes the proof of Proposition 13. □

4. Diagonalizable Linear Recursions

The procedure that we used to find a closed formula for the Fibonacci sequence in Section 2 was helped by the fact that the associated matrix $\left(\begin{smallmatrix} 1 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ is diagonalizable. In this section we consider general linear recursions whose matrices are diagonalizable. Later in Section 6 we use Jordan normal form to handle the non-diagonalizable case.

Let

$$a_1, a_2, a_3, \ldots$$

be a linear recursion generated by the linear function

$$L(x_1, \ldots, x_d) = c_1 x_1 + c_2 x_2 + \cdots + c_d x_d$$

and initial values $a_1, a_2, \ldots, a_d$, and let A_L be the associated matrix (7). For the rest of this section we make the following assumption:

The matrix A_L has a basis of eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_d \in \mathbb{C}^d$ satisfying $A_L \mathbf{v}_i = \lambda_i \mathbf{v}_i$.

We let B_L be the matrix whose columns are the eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_d$ and Δ_L the diagonal matrix with entries $\lambda_1, \ldots, \lambda_d$,

$$B_L = (\mathbf{v}_1\ \mathbf{v}_2\ \cdots\ \mathbf{v}_d) \quad \text{and} \quad \Delta_L = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_d \end{pmatrix}.$$

Then

$$A_L = B_L \Delta_L B_L^{-1},$$

so we can easily compute power of A_L via the formula

$$A_L^n = B_L \Delta_L^n B_L^{-1} \quad \text{with} \quad \Delta_L = \begin{pmatrix} \lambda_1^n & 0 & \cdots & 0 \\ 0 & \lambda_2^n & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_d^n \end{pmatrix}.$$

Finally, we obtain a formula for a_n, similar to Binet's formula (6), by observing that

$$\boxed{\begin{aligned} a_n &= \text{last coordinate of } \mathbf{v}_n \\ &= \text{last coordinate of } A_L^{n-1}\mathbf{v}_1 \\ &= \text{last coordinate of } B_L \Delta_L^{n-1} B_L^{-1} \mathbf{v}_1 \end{aligned}} \tag{11}$$

Example 14. We continue with the recursion from Example 12 given by

$$a_n = 4a_{n-1} - 6a_{n-2} + 4a_{n-3}. \tag{12}$$

We also take initial values

$$a_1 = a_2 = a_3 = 1.$$

We compute the first few terms using (12),

$$1, 1, 1, 2, 6, 16, 36, 72, 136, 256, 496, 992, 2016, \ldots .$$

We can compute the characteristic polynomial of the associated matrix A_L directly, or we can use Proposition 13. In any case, we find that

$$f_L(z) = z^3 - 4z^2 + 6z - 4 = (z-2)(z-1-i)(z-1+i),$$

so the three eigenvalues

$$\lambda_1 = 2, \quad \lambda_2 = 1+i, \quad \lambda_3 = 1-i$$

are distinct. We know that this means that the associated eigenvectors form a basis for $\mathbb{C}^3$. After a little bit of work, we find eigenvectors

$$\mathbf{v}_1 = \begin{pmatrix} 4 \\ 2 \\ 1 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 2i \\ 1+i \\ 1 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} -2i \\ 1-i \\ 1 \end{pmatrix}.$$

We next form the matrices

$$B_L = \begin{pmatrix} 4 & 2i & -2i \\ 2 & 1+i & 1-i \\ 1 & 1 & 1 \end{pmatrix} \quad \text{and} \quad \Delta_L = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1+i & 0 \\ 0 & 0 & 1-i \end{pmatrix}.$$

We also need to compute the inverse matrix

$$B_L^{-1} = \begin{pmatrix} \frac{1}{2} & -1 & 1 \\ -\frac{1}{4} + \frac{1}{4}i & \frac{1}{2} - i & i \\ -\frac{1}{4} - \frac{1}{4}i & \frac{1}{2} + i & -i \end{pmatrix}.$$

Finally, we are ready to use (11) to compute

$$a_n = \text{last coordinate of } B_L \Delta_L^{n-1} B_L^{-1} \mathbf{v}_1.$$

So we need to compute the last coordinate of the product

$$\begin{pmatrix} 4 & 2i & -2i \\ 2 & 1+i & 1-i \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 2^{n-1} & 0 & 0 \\ 0 & (1+i)^{n-1} & 0 \\ 0 & 0 & (1-i)^{n-1} \end{pmatrix} \begin{pmatrix} \frac{1}{2} & -1 & 1 \\ -\frac{1}{4} + \frac{1}{4}i & \frac{1}{2} - i & i \\ -\frac{1}{4} - \frac{1}{4}i & \frac{1}{2} + i & -i \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

After doing some algebra, we find that

$$a_n = \frac{1}{4} \cdot 2^n + \frac{1}{4} \cdot (1+i)^n + \frac{1}{4} \cdot (1-i)^n.$$

Further, since $|1 \pm i| = \sqrt{2} < 2$, we see that the 2^n term dominates when n gets big, so in particular,

$$\lim_{n \to \infty} \frac{a_n}{2^n} = \frac{1}{4}.$$

5. Powers of Jordan Block Matrices

In order to deal with general linear recursions whose associated matrix is not diagonalizable, we need to figure out how to compute powers of matrices that are in Jordan block form. We start by using Proposition 6 and the binomial theorem to write down a simple expression for the powers of a single Jordan block.

Proposition 15. *Let $J = \lambda I + N$ be an m-by-m Jordan block with eigenvalue λ. Then for all $k \geq 0$ we have*

$$J^k = (\lambda I + N)^k = \sum_{i=0}^{m-1} \binom{k}{i} \lambda^{k-i} N^i.$$

Note that the sum only has m terms, no matter how large k becomes. You are probably familiar with the binomial coefficient

$$\binom{k}{i} = \frac{k!}{i!(k-i)!}.$$

(If $i = 0$, by convention we set $\binom{k}{0} = 1$.)

PROOF. The binomial theorem says that

$$(\lambda I + N)^k = \sum_{i=0}^{k} \binom{k}{i} \lambda^{k-i} N^i.$$

(We are using the fact that the matrices λI and N commute with each other, which is obvious since the identity matrix commutes with every other matrix.) The sum that we get from the binomial theorem runs from $i = 0$ to $i = k$, but Proposition 6 says that $N^i = 0$ when $i \geq m$, so only the terms with $i < m$ may be nonzero. $\square$

EXAMPLE 16. The matrix J^k is upper triangular. We illustrate with $m = 4$,

$$\begin{aligned} \begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{pmatrix}^k &= \lambda^k I + k\lambda^{k-1} N + \binom{k}{2} \lambda^{k-2} N^2 + \binom{k}{3} \lambda^{k-3} N^3 \\ &= \begin{pmatrix} \lambda^k & k\lambda^{k-1} & \binom{k}{2}\lambda^{k-2} & \binom{k}{3}\lambda^{k-3} \\ 0 & \lambda^k & k\lambda^{k-1} & \binom{k}{2}\lambda^{k-2} \\ 0 & 0 & \lambda^k & k\lambda^{k-1} \\ 0 & 0 & 0 & \lambda^k \end{pmatrix}. \end{aligned}$$

We leave it to the reader to check these statements.

6. Closed Formulas for General Linear Recursions

*** To Be Completed ***

Exercises for Chapter 2

Problem # C.7. In 1202 Leonardo of Pisa (also known as Leonardo Fibonacci) published his *Liber Abbaci*, a highly influential book of practical mathematics. In this book Leonardo posed the following Rabbit Problem.

> In the first month, start with a pair of baby rabbits. One month later they have grown up. The following month the pair of grown rabbits produce a pair of babies, so now we have one pair of grown rabbits and one pair of baby rabbits. Each month thereafter, each pair of grown rabbits produces a new pair of babies, and every pair of baby rabbits grows up. How many pairs of rabbits will there be at the end of one year?

Show that the number of rabbits after n months is given by the n^{th} Fibonacci number, and compute the answer to Fibonacci's problem. Notice that even this simple model for population growth yields a population that grows exponentially.

Problem # C.8. Use Binet's formula (6) to prove that the Fibonacci sequence satisfies the following formulas.
(a) $F_{n+1}F_{n-1} - F_n^2 = (-1)^n$.

(b) $F_n^2 + F_{n-1}^2 = F_{2n-1}$.

(c) $F_nF_{n-1} + F_nF_{n+1} = F_{2n}$.

(d) $\displaystyle\sum_{n=1}^{k} F_n = F_{k+2} - 1$.

(e) $\displaystyle\sum_{n=1}^{k} F_n^2 = F_kF_{k+1}$.

(f) Prove that at least one of the numbers $5F_n^2 + 4$ and $5F_n^2 - 4$ is a perfect square.

Problem # C.9. Find a closed form solution for each of the following linear recursions.
(a) $a_1 = 1, \quad a_2 = 3, \quad a_n = a_{n-1} + a_{n-2}$. This is called the *Lucas sequence.*
(b) $a_1 = 1, \quad a_2 = 5, \quad a_n = 4a_{n-1} + 5a_{n-2}$. What happens if instead we start with $a_1 = 1$ and $a_2 = 1$?
(c) $a_1 = 6, \quad a_2 = -14, \quad a_n = 6a_{n-1} - 25a_{n-2}$.

Problem # C.10. Suppose that $a_1, a_2, a_3, \dots$ is a linear recursion associated to a non-zero linear function

$$L(x_1, \dots, x_d) = c_1x_1 + c_2x_2 + \cdots + c_dx_d,$$

let

$$F(z) = z^d - c_1z^{d-1} - \cdots - c_{d-1}z - c_d$$

be the associated polynomial, and suppose that $F(z)$ factors over $\mathbb{C}$ as

$$F(z) = (z - \lambda_1)(z - \lambda_2)\cdots(z - \lambda_d).$$

Suppose further that

$$|\lambda_1| > |\lambda_2| > |\lambda_3| > \cdots > |\lambda_d|.$$

(a) Prove that the limit

$$\lim_{n\to\infty} |a_n|^{1/n} \tag{13}$$

exists and is equal to one of the numbers $|\lambda_1|, |\lambda_2|, \dots, |\lambda_d|$.
(b) Why doesn't the limit (13) have to equal $|\lambda_1|$. Give an example with $d = 2$ where the limit (13) is equal to $|\lambda_2|$.

(c) What might go wrong if $|\lambda_1| = |\lambda_2|$? For example, consider the linear recursion

$$a_1 = 6, \quad a_2 = -14, \quad a_n = 6a_{n-1} - 25a_{n-2}.$$

Does the limit $\lim_{n\to\infty} |a_n|^{1/n}$ even exist? (This is hard.)

Problem # C.11. (a) Compute the following binomial coefficients:

$$\text{(i)} \binom{5}{2}. \quad \text{(ii)} \binom{8}{3}. \quad \text{(iii)} \binom{100}{98}.$$

(Hint: You should be able to do (iii) with paper and pencil in less than 10 seconds.)

(b) Using the definition $\binom{n}{k} = \frac{n!}{k!(n-k)!}$, prove the identity

$$\binom{n+1}{k+1} = \binom{n}{k} + \binom{n}{k+1}.$$

(c) Use (b) and induction to prove that

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}.$$

This is called the *binomial theorem*.

(d) Prove the following identity (this one is harder):

$$\sum_{k=0}^{n} \binom{n}{k}^2 = \binom{2n}{n}.$$