SOLUTION SETS OF RECURRENCE RELATIONS

SEBASTIAN BOZLEE
UNIVERSITY OF COLORADO AT BOULDER

The first section of these notes describes general solutions to linear, constant-coefficient, homogeneous recurrence relations of arbitrary order, including the case that eigenvalues are repeated. Our proofs will use linear algebra and a touch of abstract algebra instead of generating functions. Accordingly, I hope that these notes will be accessible to a student with a course or two of undergraduate linear algebra and some experience with arguments involving divisibility.

We will then apply these results to sketch a derivation of the general solutions to ordinary differential equations.

1. General solutions to recurrence relations

For the purpose of these notes, a sequence is a sequence of complex numbers (although our results should hold if we replace $\mathbb{C}$ by any algebraically closed field). We will either write a sequence as a list of numbers enclosed in parentheses or by a single expression enclosed in parentheses. It will be convenient to index all sequences by the variable j, starting with $j = 0$. To illustrate these conventions, (j) denotes the sequence of numbers $(0, 1, 2, 3, 4, \ldots)$ and (2^j) is the sequence $(1, 2, 4, 8, 16, \ldots)$.

The set of all sequences of complex numbers we denote by $\mathbb{C}^{\mathbb{N}}$. This has the structure of an infinite-dimensional complex vector space where the addition of two sequences (a_j) and (b_j) is done coordinatewise (i.e. $(a_j) + (b_j) = (a_j + b_j)$) and scalar multiplication is defined by $c(a_j) = (ca_j)$. We will see that solutions to recurrence relations form finite-dimensional subspaces of $\mathbb{C}^{\mathbb{N}}$.

A **linear homogeneous recurrence relation** of order k (hereafter, a **recurrence relation of order** k) is an equation of the form

$$a_j = c_1 a_{j-1} + c_2 a_{j-2} + \cdots + c_k a_{j-k}. \tag{1}$$

where $c_1, c_2, \ldots, c_k$ are complex numbers. A **solution** to a recurrence relation is a complex-valued sequence $(a_1, a_2, \ldots)$ such that (1) holds for each $j \geq k$. Such a sequence is called a **recurrence sequence of order** k. Familiar examples of recurrence sequences include geometric sequences and the Fibonacci numbers.

Date: January 23, 2016.

To each recurrence relation of order k,

$$a_j = c_1 a_{j-1} + c_2 a_{j-2} + \cdots + c_k a_{j-k}, \tag{2}$$

there is associated a **characteristic polynomial** of degree k,

$$f(\lambda) = \lambda^k - c_1\lambda^{k-1} - c_2\lambda^{k-2} - \cdots - c_{k-1}\lambda - c_k.$$

The roots of the characteristic polynomial are called the **eigenvalues** of the recurrence relation. We shall see that the characteristic polynomial of a recurrence relation is closely related to its solutions.

Let Λ be the linear operator on sequences defined by $\Lambda(a_j) = (a_{j+1})$. The effect of Λ is to shift the elements of the sequence by one space. For example,

$$\Lambda(3, 1, 4, 1, 5, \ldots) = (1, 4, 1, 5, \ldots).$$

Given a polynomial $f(\lambda) = \lambda^k - c_1\lambda^{k-1} - c_2\lambda^{k-2} - \cdots - c_{k-1}\lambda - c_k$, we define a linear operator $f(\Lambda)$ by

$$f(\Lambda) = \Lambda^k - c_1\Lambda^{k-1} - c_2\Lambda^{k-2} - \cdots - c_{k-1}\Lambda - c_k I.$$

where Λ^k is defined to be the composition of Λ with itself k times, i.e. a shift by k spaces, and I is the identity operator. Given these definitions, for an arbitrary sequence (a_j),

$$\begin{aligned} f(\Lambda)(a_j) &= (\Lambda^k - c_1\Lambda^{k-1} - c_2\Lambda^{k-2} - \cdots - c_{k-1}\Lambda - c_k I)(a_j) \\ &= (a_{j+k} - c_1 a_{j+k-1} - c_2 a_{j+k-2} - \cdots - c_{k-1}a_{j+1} - c_k a_j). \end{aligned}$$

So, $f(\Lambda)(a_j) = (0)$ if and only if (a_j) satisfies the recurrence relation (2).

In other words, $\ker f(\Lambda)$ is the solution set of (2). Since the kernel of a linear map is a vector space, the solution set is a vector space. Therefore all we have to do to describe the solution set of a recurrence relation is to find a basis for $\ker f(\Lambda)$. We will spend the rest of these notes deriving a "nice" basis for $\ker f(\Lambda)$ related to the factorization of $f(\lambda)$. Its basis vectors will be called **fundamental solutions**.

Consider the correspondence between characteristic polynomials and solution sets

$$f(\lambda) \mapsto \ker f(\Lambda).$$

In the Lemma 2, we show that some of the "divisibility structure" of polynomials is carried over by this correspondence to the "containment structure" of subspaces of $\mathbb{C}^{\mathbb{N}}$. (For readers familiar with lattices, this is the first part of a proof that the correspondence above is a lattice homomorphism.)

First, however, we will prove a lemma about polynomials. The proof can probably be skipped on a first read, and can be found with more context in any book on abstract algebra (we are essentially proving that $\mathbb{C}[\lambda]$ is a PID).

Lemma 1. *If $f(\lambda)$ and $g(\lambda)$ are polynomials and $h(\lambda) = \gcd(f(\lambda), g(\lambda))$, then there exist polynomials $p(\lambda)$ and $q(\lambda)$ so that*

$$p(\lambda)f(\lambda) + q(\lambda)g(\lambda) = h(\lambda).$$

Proof. Let $I = \{a(\lambda)f(\lambda) + b(\lambda)g(\lambda) \mid a(\lambda), b(\lambda) \text{ are polynomials }\}$. Note that the sum of any two elements of I lies in I and a polynomial multiple of any element of I lies in I. We will now show that I is equal to the set

$$J = \{c(\lambda)k(\lambda) \mid c(\lambda) \text{ is a polynomial }\}$$

for some polynomial $k(\lambda)$ in I.

Tentatively, we select $k(\lambda)$ to be a polynomial of I of least degree. Let $d(\lambda)$ be some polynomial in I. By the division algorithm for polynomials, we can write

$$d(\lambda) = c(\lambda)k(\lambda) + r(\lambda)$$

where the remainder $r(\lambda)$ is either 0 or a polynomial of degree strictly less than the degree of $k(\lambda)$. Rearranging the equation above,

$$d(\lambda) - c(\lambda)k(\lambda) = r(\lambda).$$

Since $d(\lambda) \in I$ and $k(\lambda) \in I$, and I is closed under addition and polynomial multiples, $d(\lambda) - c(\lambda)k(\lambda) \in I$, so $r(\lambda) \in I$. Recall $k(\lambda)$ which was chosen as the polynomial of least degree in I, and $r(\lambda)$ was chosen to be zero or to have degree less than $k(\lambda)$. Therefore $r(\lambda) = 0$, so $k(\lambda) \mid d(\lambda)$. Since $d(\lambda)$ was arbitrary, $I \subseteq J$.

On the other hand, since J consists of multiples of an element of I and I contains all multiples of its elements, $J \subseteq I$.

Since $k(\lambda) \in I$, we have that $k(\lambda) = p_1(\lambda)f(\lambda) - q_1(\lambda)g(\lambda)$. Now, since $I = J$ and $f(\lambda) \in I$, $k(\lambda) \mid f(\lambda)$. Similarly, $k(\lambda) \mid g(\lambda)$. That is, $k(\lambda)$ is a common divisor of $f(\lambda)$ and $g(\lambda)$. Therefore, by definition of gcd, $k(\lambda) \mid h(\lambda)$. On the other hand, we note that since $h(\lambda)$ divides $f(\lambda)$ and $g(\lambda)$, $h(\lambda)$ divides $p_1(\lambda)f(\lambda)$ and $q_1(\lambda)g(\lambda)$. It follows that $h(\lambda)$ divides their sum, $p_1(\lambda)f(\lambda) + q_1(\lambda)g(\lambda) = k(\lambda)$. Since $k(\lambda)$ divides $h(\lambda)$ and vice versa, $h(\lambda) = Ck(\lambda)$ for some constant C. Then

$$h(\lambda) = Cp_1(\lambda)f(\lambda) + Cq_1(\lambda)g(\lambda).$$

Taking $p(\lambda) = Cp_1(\lambda)$ and $q(\lambda) = Cp_2(\lambda)$, we have the result. ∎

Lemma 2. *If $f(\lambda)$ and $g(\lambda)$ are polynomials and $h(\lambda) = \gcd(f(\lambda), g(\lambda))$, then*

$$\ker h(\Lambda) = \ker f(\Lambda) \cap \ker g(\Lambda).$$

In particular, if $g(\lambda) \mid f(\lambda)$, then $\ker g(\Lambda)$ is a subspace of $\ker f(\Lambda)$.

Proof. Let $h(\lambda) = \gcd(f(\lambda), g(\lambda))$. By the preceding lemma, there are polynomials $p(\lambda), q(\lambda)$ so that $p(\lambda)f(\lambda) + q(\lambda)g(\lambda) = h(\lambda)$. Let $(a_j) \in \ker f(\Lambda) \cap \ker g(\Lambda)$. Then

$$h(\Lambda)(a_j) = (p(\Lambda)f(\Lambda) + q(\Lambda)g(\Lambda))(a_j) = p(\Lambda)(0) + q(\Lambda)(0) = (0).$$

So $\ker f(\Lambda) \cap \ker g(\Lambda) \subseteq \ker h(\Lambda)$.

Conversely, let $(a_j) \in \ker h(\Lambda)$. Since $h(\lambda) \mid f(\lambda)$, there is a polynomial $p(\lambda)$ so that $f(\lambda) = p(\lambda)h(\lambda)$. Then

$$f(\Lambda)(a_j) = p(\Lambda)h(\Lambda)(a_j) = p(\Lambda)(0) = (0).$$

So, $(a_j) \in \ker f(\Lambda)$. Similarly, $(a_j) \in \ker g(\Lambda)$. Therefore, $(a_j) \in \ker f(\Lambda) \cap \ker g(\Lambda)$, so $\ker f(\Lambda) \cap \ker g(\Lambda) \supseteq \ker h(\Lambda)$. The result follows. □

We will also require the following fact.

Lemma 3. *The solution set of an order k recurrence has dimension k. Equivalently, $\dim \ker f(\Lambda) = \deg f(\lambda)$.*

Proof. Let $f(\lambda)$ be the characteristic polynomial of the recurrence. The solution set is $\ker f(\Lambda)$. Let $\varphi : \ker f(\Lambda) \to \mathbb{C}^k$ be the map that takes a solution (a_j) to its first k values, $(a_0, a_1, \ldots, a_{k-1})$. This is clearly linear. It is bijective since any choice of initial values $a_0, a_1, \ldots, a_{k-1}$ uniquely determines a solution sequence via the recurrence relation: i.e.

$$\begin{aligned} a_k &= c_1 a_{k-1} + c_2 a_{k-2} + \cdots + c_k a_0, \\ a_{k+1} &= c_1 a_k + c_2 a_{k-1} + \cdots + c_k a_1, \\ &\vdots \end{aligned}$$

Therefore φ is an isomorphism of vector spaces. Since $\mathbb{C}^k$ has dimension k, we conclude that $\ker f(\Lambda)$ has dimension k. □

Suppose $f(\Lambda) = \prod_{l=1}^{q} (\Lambda - \lambda_l)^{k_l}$ is a characteristic polynomial, where the λ_l are distinct. By Lemma 2, $\ker((\Lambda - \lambda_l)^{k_l})$ is a subspace of $\ker f(\Lambda)$ for each l. As a first step toward finding a fundamental solution set for the corresponding recurrence relation, we start by finding a basis of $\ker((\Lambda - \lambda_l)^{k_l})$ for each l.

The case where $\lambda_l = 0$ is handled separately.

Lemma 4. $\ker \Lambda^n$ *has the basis* $\mathcal{B}_0^n = \{(\delta_{j0}), (\delta_{j1}), \dots, (\delta_{j,n-1})\}$, *where*

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

Proof. $\mathcal{B}_0^n$ is clearly linearly independent (its elements have nonzero entries in different coordinates) and clearly belongs to $\ker \Lambda^n$ (since shifting by n gets rid of all nonzero entries). Since $|\mathcal{B}_0^n| = n$ and $\ker \Lambda^n$ has dimension n, $\mathcal{B}_0^n$ is a basis. $\square$

Lemma 5. *Let* λ *be a nonzero complex number and let* n *be a positive integer. Then* $\ker((\Lambda - \lambda)^n)$ *has the basis* $\mathcal{B}_\lambda^n = \{(\lambda^j), (j\lambda^j), \dots, (j^{n-1}\lambda^j)\}$.

Proof. We will first prove that for each n, $\{(\lambda^j), (j\lambda^j), \dots, (j^{n-1}\lambda^j)\} \subseteq \ker((\Lambda - \lambda)^n)$. The proof is by induction.

Let $n = 1$. Then $(\Lambda - \lambda)(\lambda^j) = (\lambda^{j+1} - \lambda\lambda^j) = (0)$, so $(\lambda_j) \in \ker(\Lambda - \lambda)$.

Suppose $\{(\lambda^j), (j\lambda^j), \dots, (j^{n-1}\lambda^j)\} \subseteq \ker((\Lambda - \lambda)^n)$ for $n = k$, for some integer $k \geq 1$. Then by Lemma 2, $\{(\lambda^j), (j\lambda^j), \dots, (j^{n-1}\lambda^j)\} \subseteq \ker((\Lambda - \lambda)^{n+1})$. It remains to show that $(j^n\lambda^j) \in \ker((\Lambda - \lambda)^{n+1})$. Now,

$$\begin{aligned} (\Lambda - \lambda)^{n+1}(j^n\lambda^j) &= (\Lambda - \lambda)^n((j+1)^n\lambda^{j+1} - j^n\lambda^{j+1}) \\ &= (\Lambda - \lambda)^n \left(\lambda\binom{n}{1}j^{n-1}\lambda^j + \lambda\binom{n}{2}j^{n-2}\lambda^j + \cdots + \lambda\binom{n}{n}\lambda^j \right) \\ &= (0) \end{aligned}$$

where we have used the binomial theorem on the second line and the induction hypothesis on the third. Therefore, $\{(\lambda^j), (j\lambda^j), \dots, (j^n\lambda^j)\} \subseteq \ker((\Lambda - \lambda)^{n+1})$. By induction, $\{(\lambda^j), (j\lambda^j), \dots, (j^{n-1}\lambda^j)\} \subseteq \ker((\Lambda - \lambda)^n)$ for all $n \in \mathbb{N}$.

Next we will show that $(j^n\lambda^j) \notin \operatorname{Span}\{(\lambda^j), (j\lambda^j), \dots, (j^{n-1}\lambda^j)\}$ for all n. The result is trivial for $n = 1$. For $n > 1$, suppose that $K_1, K_2, \dots, K_n$ are complex numbers so that

$$(j^n\lambda^j) = K_1(\lambda^j) + K_2(j\lambda^j) + \cdots + K_n(j^{n-1}\lambda^j).$$

But then

$$j^n\lambda^j = K_1\lambda^j + K_2 j\lambda^j + \cdots + K_n j^{n-1}\lambda^j,$$

which implies (this is where we use that λ is nonzero)

$$j^n = K_1 + K_2 j + \cdots + K_n j^{n-1}$$

for all j. This is impossible, since the left hand side is an nth degree polynomial and the right hand side is an $(n-1)$st degree polynomial. So $(j^n\lambda^j) \notin \operatorname{Span}\{(\lambda^j), (j\lambda^j), \dots, (j^{n-1}\lambda^j\}$ for all n.

Therefore, for all n, $\{(\lambda_j), (j\lambda^j), \ldots, (j^{n-1}\lambda^j)\}$ is a linearly independent subset of $\ker((\Lambda - \lambda)^n)$ containing n vectors. Since $\ker((\Lambda-\lambda)^n)$ has dimension n, $\{(\lambda_j), (j\lambda^j), \ldots, (j^{n-1}\lambda^j)\}$ is a basis of $\ker((\Lambda-\lambda)^n)$. □

So far we have characterized the solution sets of recurrence relations with a single eigenvalue, possibly repeated. For the remaining recurrences it suffices to piece together the solutions corresponding to each eigenvalue.

Theorem 6. *Let $f(\lambda) = \prod_{l=1}^{q}(\lambda - \lambda_l)^{k_l}$ be the characteristic polynomial of a kth order recurrence relation with q distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_q$ with respective multiplicities $k_1, k_2, \ldots, k_q$. Then the solution set of the recurrence has the basis of fundamental solutions*

$$\bigcup_{l=1}^{q} \mathcal{B}_\lambda^{k_l}.$$

Proof. The proof is by induction on q. The $q = 1$ case was proven in Lemmas 4 and 5.

Suppose the result holds for some integer $q \geq 1$. Let $f(\lambda) = \prod_{l=1}^{q+1}(\lambda - \lambda_l)^{k_l}$ be the characteristic polynomial of a kth order recurrence relation with $q+1$ distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_{q+1}$ with respective multiplicities $k_1, k_2, \ldots, k_{q+1}$. We split $\ker f(\Lambda)$ into two subspaces $U = \ker\left(\prod_{l=1}^{q}(\Lambda - \lambda_l)^{k_l}\right)$ and $V = \ker((\Lambda - \lambda_{q+1})^{k_{q+1}})$, with U corresponding to the first q eigenvalues, and V corresponding to the $(q+1)$st eigenvalue. By the induction hypothesis, the first subspace, U has the basis $\bigcup_{l=1}^{q} \mathcal{B}_\lambda^{k_l}$. By Lemmas 4 and 5, the second subspace V has the basis $\mathcal{B}_\lambda^{k_{q+1}}$.

By Lemma 2, since the polynomials $\prod_{l=1}^{q}(\lambda - \lambda_l)^{k_l}$ and $(\lambda - \lambda_{q+1})^{k_{q+1}}$ that determine each subspace are relatively prime, $U \cap V = \ker I = \{(0)\}$. It follows

$$\begin{aligned}
\dim(U+V) &= \dim U + \dim V - \dim U \cap V \\
&= \dim U + \dim V \\
&= \sum_{l=1}^{q} k_l + k_{q+1} \\
&= \deg f(\lambda) \\
&= \dim \ker f(\Lambda).
\end{aligned}$$

Since $U + V$ is a subspace of $\ker f(\Lambda)$ with dimension equal to that of $\ker f(\Lambda)$, $U + V = \ker f(\Lambda)$. Furthermore, since $U \cap V = \{(0)\}$, $U + V$ is a direct sum. Therefore, the union of the bases, $\bigcup_{l=1}^{q+1} \mathcal{B}_\lambda^{k_l}$, is a basis of $\ker f(\Lambda)$. The result follows by induction. □

By comparing fundamental solution sets, we obtain the following corollary. (Readers familiar with lattices may recognize that this completes the proof that $f(\lambda) \mapsto \ker f(\Lambda)$ is a lattice homomorphism.)

Corollary 7. *If $f(\lambda)$ and $g(\lambda)$ are polynomials and $h(\lambda) = \operatorname{lcm}(f(\lambda), g(\lambda))$, then*

$$\ker h(\Lambda) = \ker f(\Lambda) + \ker g(\Lambda).$$

In particular, $\ker f(\Lambda) \subseteq \ker g(\Lambda)$ *only if* $f(\lambda) \mid g(\lambda)$.

2. General solutions to differential equations

A **linear, constant-coefficient, homogeneous differential equation of order** k (hereafter, a **differential equation of order** k) is an equation of the form

$$y^{(k)} = c_1 y^{(k-1)} + \cdots + c_{k-1} y' + c_k y, \tag{3}$$

where y is understood to be a complex-valued function on the complex plane, and $y^{(n)}$ denotes the nth derivative of y. A **solution** is a differentiable function $y : \mathbb{C} \to \mathbb{C}$ so that the above equation holds. (Readers who are uncomfortable with derivatives of complex functions may assume $y : \mathbb{R} \to \mathbb{C}$ instead; in this case our argument will only give **analytic** solutions - that is, solutions $y : \mathbb{R} \to \mathbb{C}$ so that y is equal to its Taylor series. These are actually the only solutions.)

Complex differentiable functions are analytic, so we may rewrite the equation above using the Taylor series of y. If $y : \mathbb{C} \to \mathbb{C}$ is an analytic function with Taylor series $y(x) = \sum_{i=0}^{\infty} (a_i/i!) x^i$, let

$$S(y) = (a_j)$$

denote the "sequence of derivatives" of y. S is an injective linear function. Notice $S(y') = \Lambda S(y)$. Applying S to both sides of differential equation (3), we obtain an equation on sequences:

$$\Lambda^k S(y) = c_1 \Lambda^{k-1} S(y) + \cdots + c_{k-1} \Lambda S(y) + c_k S(y).$$

In other words, if y satisfies (3), then $S(y) = (a_j)$ is a solution to the recurrence relation

$$a_j = c_1 a_1 + \cdots + c_{k-1} a_{j-k+1} + c_k a_{j-k}.$$

Conversely, since S is injective, y solves differential equation (3) only if $S(y)$ solves the above recurrence relation. Therefore to find the solutions of (3), all we have to do is find the solutions of the corresponding recurrence relation and then see what the solutions are a Taylor series of.

Theorem 8. *Consider the differential equation*

$$y^{(k)} = c_1 y^{(k-1)} + \cdots + c_{k-1} y' + c_k y. \tag{4}$$

Let $f(\lambda)$ be the polynomial $\lambda^k - c_1\lambda^{k-1} - \cdots - c_{k-1}\lambda - c_k$, and let $\lambda_1, \ldots, \lambda_q$ be its distinct roots, with respective multiplicities $k_1, \ldots, k_q$. Then the set of solutions to (4) is

$$\left\{ \sum_{i=1}^{q} \sum_{l=1}^{k_i} K_{i,l} x^l e^{\lambda_i x} \mid K_{i,l} \in \mathbb{C} \text{ for all } i, l \right\}$$