Discrete Mathematics	04/02/2019

Lecture 6: Recurrence Relations

Instructor: Goutam Paul *Scribe: Aditya Ghosh*

6.1 Introduction

A recurrence relation for a sequence $\{a_n\}_{n\geq 0}$ is an equation that expresses a_n in terms of one or more of the previous terms $a_0, a_1, \cdots, a_{n-1}$. For example, $a_n = 2a_{n-1} - a_{n-2}$, for all $n \geq 2$, is a recurrence relation. The formal definition is given below.

Definition 6.1. A recurrence relation is an equation of the form

$$a_n = f(a_{n-1}, a_{n-2}, \cdots, a_{n-k}) \text{ for all } n \geq k \tag{1}$$

with k many initial conditions which should completely determine the sequence.

Why do we need the initial conditions? Let us take the previous example:

$$a_n = 2a_{n-1} - a_{n-2}, \text{ for } n \geq 2. \tag{2}$$

Without any further assumptions on the sequence, this recurrence relation can have more than one solution. For instance, (i) $a_n = 3n$ for all $n \geq 0$ and (ii) $a_n = 5$ for all $n \geq 0$ both satisfy the above recurrence relation (which can be checked directly). In fact, if we specify $a_0 = 0, a_1 = 3$ then we get the solution (i) and if we specify $a_0 = 5, a_1 = 5$ then we get the solution (ii) (these can be easily verified by strong induction, here we omit the details). What we learn from this example is that, we need to specify some *initial* conditions (values of some terms at the beginning of the sequence) to completely specify the sequence.

It should be intuitively clear why we should specify exactly k many initial values: because the recurrence relation is silent about the relationship of $a_0, a_1, \cdots, a_{k-1}$, so that we can fix these arbitrarily and note that once these values get fixed, the whole sequence is completely determined, by virtue of the strong induction principle.

6.2 Solving Problems using Recurrence Relations

When we want to apply the idea of recurrence relations to solve combinatorial problems, there are two steps: (i) we have to formulate the recurrence relation and specify the initial values, and (ii) we need to solve the recurrence relation. Let us see some examples of solving combinatorial problems using recurrence relations.

Example 6.2. Find the number of binary sequences of length n with no consecutive zeros.

Solution. Let a_n be the number of binary sequences of length n with no consecutive zeros. Take any such sequence of length n. If the first term of the sequence is 1 then the remaining string is a binary sequence of length $(n-1)$ containing no consecutive zeros, which can be chosen in a_{n-1} ways. Else, the first digit is 0 and then the second digit must be 1 and the remaining $(n-2)$ digits make a binary sequence of length $(n-2)$ containing no consecutive zeros, which can be chosen in a_{n-2} ways. Thus, we obtain

$$a_n = a_{n-1} + a_{n-2} \text{ for all } n > 2. \tag{3}$$

Let us now find the initial values a_1 and a_2. Clearly, $a_1 = 2$ (the two binary sequences of length one are 0 and 1) and $a_2 = 3$ (excluding 00, we have 3 such sequences of length two: 01,10,11).

How to solve this recurrence relation? It is notable that the recurrence relation is exactly same as that for the well-known Fibonacci sequence and the initial values are also equal to two consecutive terms of the Fibonacci sequence. So of course a_n can be expressed as a term of the Fibonacci sequence (in fact, $a_n = F_{n+2}$); but finding an explicit formula for the Fibonacci sequence is equally hard as finding an explicit formula for a_n.

Then, how to solve it? We shall address this problem in more generality, little later. □

Example 6.3. Find the number of sequences of length n from the alphabet $A = \{0, 1, 2, 3\}$ which contains an even number of zeros.

Solution. Let x_n be the number of sequences of length n from the alphabet $A = \{0, 1, 2, 3\}$ containing an even number of zeros. Take any such sequence of length n. If the first digit is 1, then the remaining part is a sequence of the same type, but with length $(n-1)$. So the rest of the terms can be chosen in x_{n-1} ways. Same holds if the first digit is 2 or 3.

What if the first digit is 0? If the first digit is 0 then the remaining part is a sequence of length $(n-1)$ containing an odd number of zeros. This equals the total number of sequences of length $(n-1)$ minus x_{n-1}. Since the alphabet $A = \{0, 1, 2, 3\}$ has size 4, the number of sequences of length $(n-1)$ is 4^{n-1}. Thus, we obtain the recurrence relation

$$\begin{aligned} & x_n = x_{n-1} + x_{n-1} + x_{n-1} + (4^{n-1} - x_{n-1}) \\ \Longrightarrow\ & x_n = 2x_{n-1} + 4^{n-1},\ n \geq 2. \end{aligned} \tag{4}$$

Here we have specify just one initial value, namely of x_1. Clearly, there are 3 sequences of length one which contains an even number of zeros, namely $\{1\}, \{2\}, \{3\}$, so $x_1 = 3$. (Of course, zero is an even number.)

How to solve this recurrence relation? Since the recurrence relation (4) expresses x_n only in terms of x_{n-1} and n, we can solve it by using (4) repeatedly:

$$\begin{aligned} x_n &= 2x_{n-1} + 4^{n-1} = 2(2x_{n-2} + 4^{n-2}) + 4^{n-1} \\ &= 2^2 x_{n-2} + 2 \cdot 4^{n-2} + 4^{n-1} \\ &= 2^2(2x_{n-3} + 4^{n-3}) + 2 \cdot 4^{n-2} + 4^{n-1} \\ &= 2^3 x_{n-3} + 2^2 \cdot 4^{n-3} + 2 \cdot 4^{n-2} + 4^{n-1} \\ &\quad \vdots \\ &= 2^{n-1} x_1 + 2^{n-2} \cdot 4 + \cdots + 2^2 \cdot 4^{n-3} + 2 \cdot 4^{n-2} + 4^{n-1} \\ &= 2 \cdot 2^{n-1} + (2^{n-1} + 2^{n-2} \cdot 4 + \cdots + 2^2 \cdot 4^{n-3} + 2 \cdot 4^{n-2} + 4^{n-1}) \\ &= 2^n + (4^n - 2^n)/(4-2) = (4^n + 2^n)/2. \end{aligned}$$

In the second last step, we used the identity $\dfrac{a^n - b^n}{a - b} = a^{n-1} + a^{n-2}b + \cdots + b^{n-1}$. □

Example 6.4 (Derangements)**.** Suppose that there are n people numbered $1, 2, \cdots, n$. Let there be n hats also numbered $1, 2, \cdots, n$. Find the number of ways to distribute the hats (one hat per person) such that no one gets the hat having the same number as their number.

Solution. Let D_n be the required number of ways to distribute the hats such that no one gets the hat having the same number as their number.

Let us assume that the first person takes hat i where $2 \le i \le n$ (so i can be chosen in $n-1$ ways). Now there are two possibilities, depending on whether or not person i takes hat 1 in return:

1. Person i takes the hat 1. Now the problem reduces to the remaining $n-2$ hats to be distributed among the remaining $n-2$ persons, which can be done in D_{n-2} ways.

2. Person i does not take the hat 1. This case is equivalent to solving the problem with $n-1$ persons and $n-1$ hats: each of the remaining $n-1$ people has precisely 1 forbidden choice from among the remaining $n-1$ hats (person i's forbidden choice is hat 1). So this can be done in D_{n-1} ways.

From this, we get the following recursion relation:

$$D_n = (n-1)(D_{n-1} + D_{n-2}) \text{ for } n > 2. \tag{5}$$

And the intial conditions are $D_1 = 0$ and $D_2 = 1$. (When there is only one person and only one hat, there can't be any derangement. When there are two persons and two hat, then derangement can occur in only one way: each person gets the hat of the other person.)

How to solve this recurrence relation? Note that the recurrence relation can be written as

$$(D_n - nD_{n-1}) = -(D_{n-1} - (n-1)D_{n-2}) \ , \ n > 2.$$

So, if we denote $(D_n - nD_{n-1}) = E_n$ then we easily get

$$E_n = -E_{n-1} = E_{n-2} = \cdots = (-1)^{(n-2)}E_2 = (-1)^n(D_2 - 2D_1) = (-1)^n.$$

Thus, we arrive at

$$D_n = nD_{n-1} + (-1)^n \ , \ n \ge 2. \tag{6}$$

A clever manipulation makes it easy to be solved:

$$\frac{D_n}{n!} - \frac{D_{n-1}}{(n-1)!} = \frac{(-1)^n}{n!}. \tag{7}$$

The advantage of writing in this form is that, if we sum up this equation for succesive values of n, the LHS telescopes. Thus, we get

$$\frac{1}{2!} - \frac{1}{3!} + \cdots + \frac{(-1)^n}{n!} = \sum_{k=2}^{n} \frac{D_k}{k!} - \frac{D_{k-1}}{(k-1)!} = \frac{D_n}{n!} - \frac{D_1}{1!} = \frac{D_n}{n!}.$$

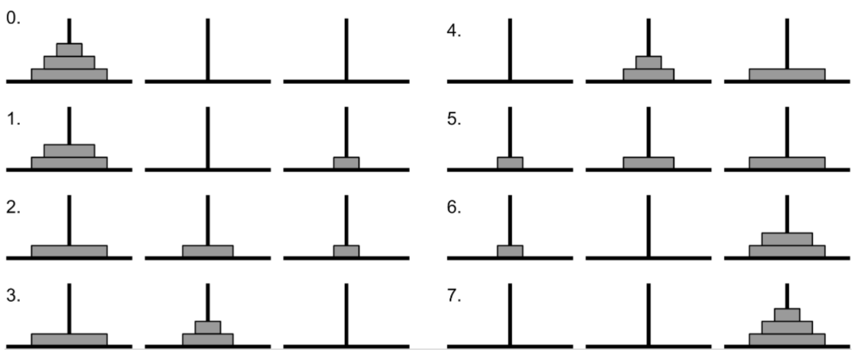
Which can also be written as $D_n = n!\Big(1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \cdots + \frac{(-1)^n}{n!}\Big)$. □

Example 6.5 (Tower of Hanoi puzzle)**.** Given a stack of n disks arranged from largest on the bottom to smallest on top placed on a rod A, together with two empty rods (B and C), you have to move the entire stack to another rod (say C), obeying the following simple rules:

1. Only one disk can be moved at a time.
2. Each move consists of taking the upper disk from one of the stacks and placing it on top of another stack or on an empty rod.
3. No larger disk may be placed on top of a smaller disk.

Find the minimal number of moves to execute this task.

Solution. Let x_n be the required number of moves, when there are n disks in rod A initially. It is easy to see that $x_1 = 1$ and $x_2 = 2$. For 3 disks, the diagram below shows that $x_3 = 7$.



Now, in order to find a recursion for x_n, note that for moving n disks from rod A to rod C (using minimal number of moves):

1. First we have to transfer $n-1$ disks from rod A to rod B (otherwise we can't move the largest disk to rod C). The minimum number of moves to do this is same as those needed to transfer $n-1$ disks from rod A to rod C, which is x_{n-1}. (As we can see above, with three disks it takes 3 moves to transfer two disks from rod A to rod C.)
2. Next, transfer the largest disk from rod A to rod C (in 1 move).
3. Finally, transfer the remaining $(n-1)$ disks from rod B to rod C. Again, the minimum number of moves to do this is same as those needed to transfer $(n-1)$ disks from rod A to rod C, which is x_{n-1}.

Therefore, we obtain the recurrence relation

$$x_n = 2x_{n-1} + 1, \quad \text{for } n \geq 2. \tag{8}$$

We can solve this recurrence relation by using it iteratively, as we did for equation (4).

$$x_n = 2x_{n-1} + 1 = 2(2x_{n-2} + 1) + 1 = 2^2 x_{n-2} + 2 + 1 = \cdots = 2^{n-1}x_1 + (2^{n-2} + \cdots + 2 + 1),$$

and $x_1 = 1$, so $x_n = 2^{n-1} + \cdots + 2 + 1 = 2^n - 1$. □

Remark. There is a story about an Indian temple in Kashi Vishwanath which contains a large room with three time-worn posts in it, surrounded by 64 golden disks. Brahmin priests, acting out the command of an ancient prophecy, have been moving these disks in accordance with the immutable rules of Brahma since that time. The puzzle is therefore also known as the Tower of Brahma puzzle. According to the story, when the last move of the puzzle is completed, the world will end.

If the story was true, and if the priests were able to move disks at a rate of one per second, using the smallest number of moves it would take them $2^{64} - 1$ seconds or roughly 585 billion years to finish, which is about 42 times the current age of the Universe. So even if the story was true, we were safe!

Example 6.6. Find the number of functions $f : \{1, 2, 3, 4, 5\} \to \{1, 2, 3, 4, 5\}$ which satisfy $f(f(x)) = x$ for each $x \in \{1, 2, 3, 4, 5\}$.

Solution. Suppose T_n is the number of functions f from $A_n = \{1, 2, \ldots, n\}$ to itself with the property $f(f(x)) = x$. Now, for any function $f : A_n \to A_n$ satisfying $f(f(x)) = x$, note that either $f(n) = n$ which is possible in T_{n-1} ways, because remaining $n - 1$ values of f can be fixed in T_{n-1} ways; or $f(n) = k$ (for some $k = 1, 2, \ldots, n - 1$) then $f(k) = f(f(n)) = n$, so we need to fix only the remaining $(n - 2)$ f-values which can be done in T_{n-2} ways. Since there are $n - 1$ choices for k, namely $1, 2, \ldots, n - 1$, so we get that,

$$T_n = T_{n-1} + (n - 1)T_{n-2} \text{ for } n > 2. \tag{9}$$

And note that, $T_1 = 1, T_2 = 2$. So using the above recursion relation, we get $T_3 = 4, T_4 = 10$ and finally $T_5 = 26$. □

Note, in the last example, we were not asked to find T_n explicitly and actually it turns out that there is no simple formula for T_n.

6.3 Solving Linear Recurrence Relations

Let us once again go back to the definition of recurrence relations. If the function f that relates a_n with $a_{n-1}, \cdots, a_{n-k}$ via equation (1) is a linear function of the inputs $a_{n-1}, \cdots, a_{n-k}$, then we say that the recurrence relation is linear.

Observe that all the examples described in last subsection have recurrence relations which are linear in nature. A good example of a non-linear recurrence relation is the one for the *Catalan numbers*. A fairly detailed discussion on it can be found in the Wikipedia page on it.

Linear recurrence relations can be further divided into several classes. Let us begin with the simplest of them.

Definition 6.7. A *linear homogeneous recurrence relation of degree k with constant coefficients* is a recurrence relation of the form

$$a_n = c_1 a_{n-1} + c_2 a_{n-2} + \cdots + c_k a_{n-k}$$

where $c_1, c_2, \cdots, c_k \in \mathbb{R}$ and $c_k \neq 0$.

Note, in this section we would be interested in finding general solutions to this recurrence relation. Initial conditions determine particular solutions out of the family of general solutions. So while finding general solutions, we would not pay attention to the initial conditions.

Let us unravel the terms in the definition. We already know that *linear* refers to the fact that $a_{n-1}, a_{n-2}, \cdots, a_{n-k}$ appear in separate terms and to the 1-st power. Next, *homogeneous* refers to the fact that the total degree of each term is the same and that there is no constant term. If there is a constant term (which is allowed to be a function of n), we call it to be a *linear non-homogeneous* recurrence relation. *Constant coefficients* refers to the fact that $c_1, c_2, \cdots, c_k$ are fixed real numbers that do not depend on n. Finally, *degree k* refers to the fact that the expression for a_n contains the previous terms upto a_{n-k}. (Thats why we wrote $c_k \neq 0$.)

If we look at the recurrence relations discussed in last subsection (equations (2) to (8)), we find that

Equation no.	(2)	(3)	(4)	(5)	(6)	(8)	(9)
Homogeneous ?	yes	yes	no	yes	no	no	yes
Const. coeff. ?	yes	yes	yes	no	no	yes	no

What about equation (7)? For the sequence D_n , equation (7) is a linear non-homogeneous equation with *non-constant* coefficients. However, for the sequence $D_n/n!$, it is a linear non-homogeneous equation with *constant* coefficients! It is obvious that linear recurrence relations with non-constant coefficients are usually harder to solve than those with constant coefficients. However, as the case of equation (7) suggests, a clever manipulation might turn it into a linear recurrence relation with constant coefficients for some new sequence. We shall now see a method for solving linear recurrence relations with constant coefficients. First We shall focus on homogeneous ones and then we shall use it for non-homogeneous ones. We begin with the observation:

Lemma 6.8. (a) If a_n is a solution of (10) then so is ca_n for any constant c.

(b) If a_n and b_n are two solutions of (10) then $a_n + b_n$ is another solution.

Note, (a) follows from homogeneity and (b) follows from linearity of the recurrence relation.

Method of Characteristic Equations: Suppose we have a linear homogeneous recurrence relation of degree k with constant coefficients of the form

$$a_n = c_1 a_{n-1} + c_2 a_{n-2} + \cdots + c_k a_{n-k} \tag{10}$$

where $c_1, c_2, \cdots, c_k \in \mathbb{R}$ and $c_k \neq 0$.

Observe that $a_n = \lambda^n$ ($\lambda \neq 0$) is a solution of (10) if and only if

$$\lambda^n = c_1\lambda^{n-1} + c_2\lambda^{n-2} + \cdots + c_k\lambda^{n-k} \iff \lambda^k = c_1\lambda^{k-1} + c_2\lambda^{k-2} + \cdots + c_k. \tag{11}$$

Thus $a_n = \lambda^n$ is a solution of (10) if and only if λ is a solution of (11). The last equation is known as the characteristic equation of the linear recurrence relation (10). From roots of this equation, we can find the solution of the linear recurrence relation, as we shall discuss below. First we shall consider the case when the roots of the characteristic equation are all distinct.

6.3.1 Characteristic equation has non-repeated roots

Theorem 6.9. Let $c_1, c_2, \cdots, c_k$ be real numbers. Suppose that the characteristic equation

$$\lambda^k = c_1\lambda^{k-1} + c_2\lambda^{k-2} + \cdots + c_k$$

has k distinct roots $\lambda_1, \lambda_2, \cdots, \lambda_k$. Then a sequence $\{a_n\}$ is a solution of the recurrence relation

$$a_n = c_1a_{n-1} + c_2a_{n-2} + \cdots + c_ka_{n-k} \ , n \geq k$$

if and only if $a_n = d_1\lambda_1^n + d_2\lambda_2^n + \cdots + d_k\lambda_k^n$ for all $n \geq 0$ where $d_1, d_2, \cdots, d_n$ are constants.

Proof. Let us prove the if part first. Since $\lambda_1, \lambda_2, \cdots, \lambda_k$ are the roots of the characteristic equation, so for each $1 \leq i \leq k$, we know that $\{\lambda_i^n\}_{n\geq 0}$ is a solution of the recurrence relation. Hence the last lemma implies that $a_n = c_1\lambda_1^n + \cdots + c_k\lambda_k^n$ is a solution of the recurrence relation.

Next, for the only if part, suppose that $\{a_n\}_{n\geq 0}$ is a solution. We need to find constants $d_1, \cdots, d_k$ such that $a_n = d_1\lambda_1^n + d_2\lambda_2^n + \cdots + d_k\lambda_k^n$ holds for every $n \geq 0$.

For any fixed $d_1, \cdots, d_k$, lets call $d_1\lambda_1^n + d_2\lambda_2^n + \cdots + d_k\lambda_k^n = b_n$. We shall find these constants such that b_n matches with a_n in the initial values. That is, we need to find the constants such that

$$\begin{pmatrix} b_0 \\ b_1 \\ b_2 \\ \vdots \\ b_{k-1} \end{pmatrix} = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & \cdots & \lambda_k \\ \lambda_1^2 & \lambda_2^2 & \cdots & \lambda_k^2 \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_1^{k-1} & \lambda_2^{k-1} & \cdots & \lambda_k^{k-1} \end{pmatrix} \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_k \end{pmatrix} = \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_{k-1} \end{pmatrix} \tag{12}$$

The matrix in the LHS is the transpose of a special matrix called Vandermont matrix, so its determinant is given by $= \prod_{1\leq i<j\leq k}(\lambda_j - \lambda_i)$, which is non-zero, since $\lambda_1, \cdots, \lambda_k$ are all distinct! Thus, the matrix is invertible and therefore the above system of equation always has a solution.

So, for any values of $a_0, \cdots, a_{k-1}$, there exists constants $d_1, \cdots, d_k$, such that $d_1\lambda_1^n + d_2\lambda_2^n + \cdots + d_k\lambda_k^n = b_n$ matches with a_n at the initial k values. Now, since $\lambda_1, \cdots, \lambda_k$ are the roots of the characteristic equation of a_n, so b_n satisfies same recurrence relation as a_n, with same initial conditions. So principle of strong induction implies that b_n must equal a_n at every n. In other words, we get $a_n = d_1\lambda_1^n + d_2\lambda_2^n + \cdots + d_k\lambda_k^n$ for all $n \geq 0$. □

Example 6.10. Let us now find the general solution of the recurrence relation for Fibonacci numbers $a_n = a_{n-1} + a_{n-2}, n \geq 2$. The characteristic equation for this recurrence relation is $x^2 = x + 1$ which has two distinct roots $\lambda_1 = (1+\sqrt{5})/2$ and $\lambda_2 = (1-\sqrt{5})/2$. Therefore, the last theorem implies that there exists constants d_1, d_2 such that

$$a_n = d_1\Big(\frac{1+\sqrt{5}}{2}\Big)^n + d_2\Big(\frac{1-\sqrt{5}}{2}\Big)^n, \text{ for all } n \geq 0.$$

The above is the general solution. To find the particular solution for the Fibonacci sequence, we shall use the initial conditions $F_0 = 0$ and $F_1 = 1$. So, d_1, d_2 must satisfy $d_1 + d_2 = 0$ and $d_1\lambda_1 + d_2\lambda_2 = 1$, which gives $d_1 = -d_2 = 1/\sqrt{5}$. Therefore, we get

$$a_n = F_n = \frac{1}{\sqrt{5}}\Big(\frac{1+\sqrt{5}}{2}\Big)^n - \frac{1}{\sqrt{5}}\Big(\frac{1-\sqrt{5}}{2}\Big)^n, \text{ for all } n \geq 0. \qquad \square$$

6.3.2 Characteristic equation has repeated roots

What if the roots of the characteristic equation are not distinct? Suppose λ is a root of the characteristic equation with multiplicity $m \geq 1$. It then turns out that $a_n = \lambda^n, n\lambda^n, \cdots, n^{m-1}\lambda^n$ all satisfy the recurrence relation corresponding to that characteristic equation. And in fact any general solution of the recurrence relation must be a linear combination of such solutions. These facts are made precise in the following theorem.

Theorem 6.11. Let $c_1, c_2, \cdots, c_k$ be real numbers. Suppose that the characteristic equation

$$\lambda^k = c_1\lambda^{k-1} + c_2\lambda^{k-2} + \cdots + c_k$$

has t distinct roots $\lambda_1, \lambda_2, \cdots, \lambda_t$with multiplicities $m_1, m_2, \cdots, m_t$ respectively, such that $m_i \geq 1$ for $i = 1, 2, \cdots, t$ and $m_1 + m_2 + \cdots + m_t = k$. Then a sequence $\{a_n\}_{n\geq 0}$ is a solution of the recurrence relation

$$a_n = c_1 a_{n-1} + c_2 a_{n-2} + \cdots + c_k a_{n-k},\ n \geq k$$

if and only if

$$a_n = (\alpha_{1,0} + \alpha_{1,1}n + \cdots + \alpha_{1,m_1-1}n^{m_1-1})\lambda_1^n + \cdots + (\alpha_{t,0} + \alpha_{t,1}n + \cdots + \alpha_{t,m_t-1}n^{m_t-1})\lambda_t^n$$

for every $n \geq 0$, where $\alpha_{i,j}$ are constants for $1 \leq i \leq t$ and $0 \leq j \leq m_i - 1$.

Proof. First we shall prove the if part. In view of **lemma 6.8.**, it suffices to show that if λ is a root of the characteristic equation with multiplicity $m \geq 1$, then $\lambda^n, n\lambda^n, \cdots, n^{m-1}\lambda^n$ all are solutions of the recurrence relation. Now, we know that any polynomial in n of degree less than m can be uniquely expressed as a linear combination of $\binom{n}{0}, \binom{n}{1}, \cdots, \binom{n}{m-1}$ (these is similar to Newton's forward interpolation). So, invoking **lemma 6.8.** once again, it suffices to show that $\lambda^n, \binom{n}{1}\lambda^n, \cdots, \binom{n}{m-1}\lambda^n$ all are solutions of the recurrence relation.

To prove this, fix any $0 \leq j \leq m-1$. We need to show that $a_n = \binom{n}{j}\lambda^n$ is a solution of the recurrence relation, that is,

$$\binom{n}{j}\lambda^n = c_1\binom{n-1}{j}\lambda^{n-1} + \cdots + c_k\binom{n-k}{j}\lambda^{n-k}. \qquad (*)$$

Observe that λ is a zero of multiplicity m of the polynomial $P(x) = x^n - c_1x^{n-1} - \cdots - c_kx^{n-k}$. And $j < m$, so the j-th derivative of $P(x)$ vanishes at $x = \lambda$. Hence,

$$0 = \frac{d^j}{dx^j}P(x)\Big|_{x=\lambda} = j!\left[\binom{n}{j}\lambda^{n-j} - c_1\binom{n-1}{j}\lambda^{n-j-1} - \cdots - c_k\binom{n-k}{j}\lambda^{n-j-k}\right]$$

which proves $(*)$ and thus completes the proof of the if part.

Next, we shall prove the only if part. Suppose a_n is a solution of the recurrence relation. Consider the generating function $g(x) = \sum_{n=0}^{\infty} a_n x^n$. If we multiply it by x^j, then it becomes $x^j g(x) = a_0 x^j + a_1 x^{j+1} + \cdots$. Therefore, the coefficient of x^n in $(c_1 x + c_2 x^2 + \cdots + c_k x^k) g(x)$ is $c_1 a_{n-1} + c_2 a_{n-2} \cdots + c_k a_{n-k}$ which is same as a_n for all $n \geq k$. Thus, the coefficient of x^n in $N(x) = (1 - (c_1 x + c_2 x^2 + \cdots + c_k x^k)) g(x)$ is 0 for all $n \geq k$. Hence, $N(x)$ is a polynomial of degree less than k.

Now, observe that the polynomial $D(x) = 1 - (c_1 x + c_2 x^2 + \cdots + c_k x^k)$ is similar to the characteristic polynomial. Indeed, $x^k D(1/x) = x^k - c_1 x^{k-1} - c_2 x^{k-2} - \cdots - c_k$ is precisely the characteristic polynomial. Since we assumed that the characteristic equation has t distinct roots $\lambda_1, \lambda_2, \cdots, \lambda_t$ with multiplicities $m_1, m_2, \cdots, m_t$ respectively, such that $m_i \geq 1$ for $i = 1, 2, \cdots, t$ and $m_1 + m_2 + \cdots + m_t = k$, so we can write

$$x^k D(1/x) = x^k - c_1 x^{k-1} - c_2 x^{k-2} - \cdots - c_k = (x - \lambda_1)^{m_1} \cdots (x - \lambda_t)^{m_t}$$

for all $x \neq 0$. This can be rewritten as $D(x) = (1 - \lambda_1 x)^{m_1} \cdots (1 - \lambda_t x)^{m_t}$ for all x. And since $N(x) = D(x) g(x)$, so we obtain that

$$g(x) = \frac{N(x)}{(1 - \lambda_1 x)^{m_1} \cdots (1 - \lambda_t x)^{m_t}}.$$

Now, we can use partial fraction decomposition. Since $N(x)$ is a polynomial of degree less than k, so we know that

$$\frac{N(x)}{(1 - \lambda_1 x)^{m_1} \cdots (1 - \lambda_t x)^{m_t}} = \sum_{i=1}^{t} \Big(\frac{A_{i,1}}{(1 - \lambda_i x)} + \frac{A_{i,2}}{(1 - \lambda_i x)^2} + \cdots + \frac{A_{i,m_i}}{(1 - \lambda_i x)^{m_i}} \Big)$$

for some constants $A_{i,j}$, where $1 \leq i \leq t$ and $1 \leq j \leq m_i$.

Therefore, $g(x)$ is a linear combination of terms of the form $\frac{A}{(1 - \lambda x)^m}$. We know that a_n is the coefficient of x^n in $g(x)$. And the coefficient of x^n in $A(1 - \lambda x)^{-m}$ is known to be $\binom{n+m-1}{m-1}$. In fact, using $g(x) = \sum_{i=1}^{t} \sum_{j=1}^{m_i} A_{i,j} (1 - \lambda_i x)^{-j}$ we can explicitly find the coefficient of x^n in both sides and thus obtain

$$a_n = \sum_{i=1}^{t} \sum_{j=1}^{m_i} A_{i,j} \binom{n+j-1}{j-1} \lambda_i^n.$$

Now, for each $1 \leq i \leq t$, $\sum_{j=1}^{m_i} A_{i,j} \binom{n+j-1}{j-1}$ is a polynomial in n of degree at most $m_i - 1$, therefore,

$$a_n = \sum_{i=1}^{t} \Big(\alpha_{i,0} + \alpha_{i,1} n + \cdots + \alpha_{i,m_i-1} n^{m_i-1} \Big) \lambda_i^n$$

where $\alpha_{i,j}$ are constants for $1 \leq i \leq t$ and $0 \leq j \leq m_i - 1$. This completes the proof! □

Example 6.12. Find the general solution of the recurrence relation $a_n = 5a_{n-1} - 8a_{n-2} + 4a_{n-3}, n \geq 3$.

Solution. The characteristic equation for this recurrence relation is $x^3 - 5x^2 + 8x - 4 = 0$ or, $(x-2)^2(x-1) = 0$. It has one root $x = 2$ with multiplicity 2 and the other root $x = 1$ with multiplicity 1. Therefore, in light of the last theorem, the general solution is given by $a_n = \alpha_1 \cdot 2^n + \alpha_2 \cdot n\ 2^n + \beta \cdot 1^n$ where $\alpha_1, \alpha_2, \beta$ are constants. To find particular solution, we need to find these from the initial conditions. □

6.3.3 Linear non-homogeneous equations with constant coefficients

Now we shall see how to solve linear non-homogeneous equations, of the form

$$a_n = c_1 a_{n-1} + c_2 a_{n-2} + \cdots + c_k a_{n-k} + g(n),\ n \geq k$$

where $g(n)$ is a function of n which does not depend on the terms of the sequences.

Theorem 6.13. Suppose x_n is one solution of the linear non-homogeneous recurrence relation

$$a_n = c_1 a_{n-1} + c_2 a_{n-2} + \cdots + c_k a_{n-k} + g(n),\ n \geq k \tag{13}$$

Then y_n is another solution if and only if $y_n = x_n + h_n$ for all $n \geq 0$, where h_n is a solution of the linear homogeneous recurrence relation

$$a_n = c_1 a_{n-1} + c_2 a_{n-2} + \cdots + c_k a_{n-k},\ n \geq k.$$

Proof. If part is easy to show. Let $y_n = x_n + h_n$ for all $n \geq 0$ where x_n and h_n are as above. Then,

$$y_n = x_n + h_n = \sum_{i=1}^{k} c_i x_{n-i} + g(n) + \sum_{i=1}^{k} c_i h_{n-i} = \sum_{i=1}^{k} c_i (x_{n-i} + h_{n-i}) + g(n) = \sum_{i=1}^{k} c_i y_{n-i} + g(n)$$

which shows that y_n satisfies the given recurrence relation (13).

Next, to show the only if part, let x_n and y_n both satisfy (13). Define $h_n = y_n - x_n$. We shall show that h_n satisfies the homogeneous equation corresponding to (13). Indeed, for any $n \geq k$, we have

$$\begin{aligned} h_n = y_n - x_n &= \Big(\sum_{i=1}^{k} c_i y_{n-i} + g(n)\Big) - \Big(\sum_{i=1}^{k} c_i x_{n-i} + g(n)\Big) \\ &= \sum_{i=1}^{k} c_i (y_{n-i} - x_{n-i}) = \sum_{i=1}^{k} c_i h_{n-i} \end{aligned}$$

which shows that h_n satisfies the homogeneous recurrence relation corresponding to (13). □

Example 6.14. Find the general solution of $a_n = 3a_{n-1} - 2^{n-1}, n \geq 1$.

Solution. Since $3 \cdot 2^{n-1} - 2^{n-1} = 2^n$, so $a_n = 2^n$ is one solution of the given non-homogeneous recurrence relation. Now, the homogeneous recurrence relation corresponding to it is $a_n = 3a_{n-1}$ which has characteristic equation $x = 3$, so the general solution to this (homogeneous) equation is given by $h_n = c \cdot 3^n$ where c is a constant. Therefore, in light of the above theorem, the general solution to the non-homogeneous equation $a_n = 3a_{n-1} - 2^{n-1}, n \geq 1$, is given by $a_n = c \cdot 3^n + 2^n$ where c is a constant. □

The following theorem says that in some situations there is a particular solution of a certain form, depending upon the form of the non-homogeneous part $g(n)$ in equation (13).

Theorem 6.15. Suppose the non-homogeneous term $g(n)$ in equation (13) is of the form $p(n)s^n$, where $p(n)$ is a polynomial of degree r and s is a constant.

(a) If s is not a root of the characteristic equation, then there is a particular solution of the form $q(n)s^n$, where q is a polynomial of degree at most r.

(b) If s is a root of the characteristic equation of multiplicity t, then there is a particular solution of the form $n^t q(n)s^n$, where q is a polynomial of degree at most r.

Proof. We shall use the method of undetermined coefficients. Let us write $p(n) = \sum_{j=0}^{r} \alpha_j n^j$. If the number s in is not a root of the characteristic equation, then we plug $a_n = (\beta_r n^r + \beta_{r-1}n^{r-1} + \cdots + \beta_1 n + \beta_0)s^n$ into the recurrence relation and get

$$\sum_{j=0}^{r} \beta_j n^j s^n = \sum_{i=1}^{k} c_i s^{n-i}\Big(\sum_{j=0}^{r} \beta_j (n-i)^j\Big) + \sum_{j=0}^{r} \alpha_j n^j s^n \tag{14}$$

Both sides of the equation are polynomials in n (note that s is constant), and two polynomials must be identical. So if we equate the coefficients of same powers of n on both sides, this results in a system of $r+1$ linear equations in $r+1$ unknowns (the β_is) to solve. Let us unravel the form of these equations. Equating the coefficients of n^r, we get

$$\beta_r(s^n - c_1 s^{n-1} - c_2 s^{n-2} - \cdots - c_k s^{n-k}) = \alpha_r s^n.$$

Since s is not a root of the characteristic equation, so the bracketed part in the LHS is non-zero, from which we get β_r. Next, equating the coefficients of n^{r-1}, we get

$$\beta_{r-1}(s^n - c_1 s^{n-1} - c_2 s^{n-2} - \cdots - c_k s^{n-k}) = -\binom{r}{1}\beta_r \sum_{i=1}^{k} c_i \cdot i s^{n-i} + \alpha_{r-1}s^n.$$

So, after substituting the value of β_r, we get β_{r-1}. In this way, it turns out that the system of equations for $\beta_0, \cdots, \beta_r$ can be solved by backward substitution method. Thus, the system of equations is always solvable, which completes the proof of (a).

When s is a root of multiplicity t of the characteristic equation, then $s^n, ns^n, \cdots, n^t s^n$ satisfies the homogeneous equation, as we showed last time. Hence if we want to plug any polynomial into the recurrence to compensate for the homogeneous part, we must care about only the terms $n^t, n^{t+1}, \cdots$. Keeping this in mind, we plug $a_n = n^t(\beta_r n^r + \beta_{r-1}n^{r-1} + \cdots + \beta_1 n + \beta_0)s^n$ into the recurrence and obtain a similar (but complicated) system of equations for $\beta_0, \cdots, \beta_r$ which can be solved in backward substitution method, as before. This completes the proof of (b). □

Note that non-homogeneous terms like n or 5^n are of the right form to apply the last theorem : there is a 1 which is not written down. The term n is the same as $n \cdot 1^n$, and 5^n is the same as $1 \cdot 5^n$.

In some combinatorics texts, you will find tables of other situations in which the form of a particular solution is known. We will deal only with the cases covered by the last theorem, and the next one.

Theorem 6.16. If the non-homogeneous term $g(n)$ is of the form $p_1(n)s_1^n + p_2(n)s_2^n + \cdots + p_m(n)s_m^n$, then there is a particular solution of the form $f_1(n) + f_2(n) + \cdots + f_m(n)$ where, for each i, $f_i(n)$ is a particular solution to the non-homogeneous recurence

$$a_n = c_1 a_{n-1} + c_2 a_{n-2} + \cdots + c_k a_{n-k} + p_i(n)s_i^n.$$

Proof. For each $1 \leq i \leq m$, the previous theorem guarantees the existence of a particular solution $f_i(n)$ of the recurrence

$$a_n = c_1 a_{n-1} + c_2 a_{n-2} + \cdots + c_k a_{n-k} + p_i(n)s_i^n.$$

Then, by linearity, $f_1(n)+f_2(n)+\cdots+f_m(n)$ becomes a particular solution of the recurrence relation

$$a_n = c_1 a_{n-1} + c_2 a_{n-2} + \cdots + c_k a_{n-k} + \sum_{i=1}^{m} p_i(n)s_i^n.$$

□

Example 6.17. Solve $a_n = a_{n-1} + 2a_{n-2} + 2^n + 4$, $n \geq 2$, $a_0 = 0, a_1 = 1/3$.

Solution. The associated homogeneous recurrence is $a_n = a_{n-1} + 2a_{n-2}$. Its characteristic equation is $x^2 = x + 2$, or $(x+1)(x-2) = 0$. Thus, $(-1), 2$ are the only roots and each root has multiplicity 1. The general solution to the associated homogeneous recurrence is $c_1 \cdot (-1)^n + c_2 \cdot 2^n$.

To find a particular solution to the non-homogeneous recurrence, we add together particular solutions to the two 'simpler' non-homogeneous recurrences: (i) $a_n = a_{n-1}+2a_{n-2}+2^n$ and (ii) $a_n = a_{n-1} + 2a_{n-2} + 4$.

To solve (ii), note that $4 = 4 \cdot 1^n$ and 1 is not a root of the characteristic equation, so by **Theorem 6.15**(a), it has a particular solution of the form $c \cdot 1^n = c$. Plugging this into the equation (ii), we get $c = c + 2c + 4$ which gives $c = -2$.

To solve (i), note that $2^n = 1 \cdot 2^n$ and 2 is a root with multiplicity one of the characteristic equation, so by **Theorem 6.15**(b), it has a particular solution of the form $n \cdot d \cdot 2^n$. Plugging this into the equation (i), we get $dn2^n = d(n-1)2^{n-1} + 2d(n-2)2^{n-2} + 2^n$ which gives $d = 2/3$.

Thus, in view of **Theorem 6.16**, a particular solution of the non-homogeneous recurrence relation is $-2 + \frac{2}{3}n2^n$.

Now, **Theorem 6.13** tells us that the general solution of the non-homogeneous equation is given by $a_n = c_1 \cdot (-1)^n + c_2 \cdot 2^n + \frac{2}{3}n2^n - 2$. Using the initial conditions $a_0 = 0$ and $a_1 = 1$, we find that $c_1 = c_2 = 1$, so that the required particular solution of the recurrence relation is given by $a_n = (-1)^n + 2^n + \frac{2}{3}n2^n - 2$, $n \geq 0$. □