Lecture 3

Difference equations and some examples (cont'd)

We continue with some additional examples of difference equations and where they are encountered in practical applications.

Difference equations associated with interest on investments/loans

Reader Alert! We qualify that the models used in this section are extremely simple in form and not necessarily those employed in the modern-day real world.

Simple Interest vs. Compound Interest

Suppose that x_0 dollars are initially invested at an annual interest rate r. (The annual percentage rate is 100 r%. For example 6% corresponds to $r = 0.06$.)

Simple Interest Scheme: Interest is earned only on the **initial** amount invested. If x_k denotes the value of the investment after k years, then

$$\begin{aligned} x_1 &= x_0 + rx_0 \\ x_2 &= x_1 + rx_0 \\ x_3 &= x_2 + rx_0 \\ &\vdots \end{aligned}$$

to give the difference equation

$$\boxed{x_n = x_{n-1} + rx_0\,, \quad n \geq 1\,.} \tag{1}$$

It's not too difficult to determine the solution of this difference equation:

$$\begin{aligned} x_2 &= x_1 + rx_0 = x_0 + rx_0 + rx_0 = x_0 + 2rx_0 \\ x_3 &= x_2 + rx_0 = x_2 + x_0 + 2rx_0 + rx_0 = x_0 + 3rx_0\,. \end{aligned} \tag{2}$$

The value x_n can be written in terms of x_0 as

$$\boxed{x_n = (1 + nr)x_0 \quad n \geq 0\,.} \tag{3}$$

This is the solution of the d.e. in (1). We can prove it by substitution as follows. Substitution of the above expression into the RHS of (1) yields,

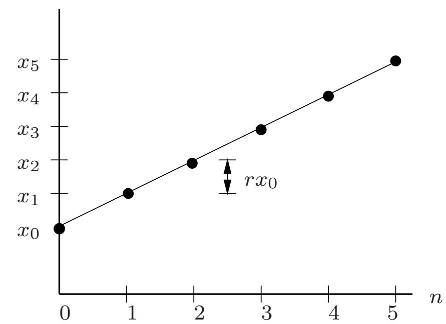
$$x_{n-1} + rx_0 = (1 + (n-1)r)x_0 + rx_0 = (1 + nr)x_0 = \text{RHS}\,. \tag{4}$$

That was quite simple. We could also prove this result by induction but shall not do so here.

From Eq. (3), the investment grows linearly in time:

$$x_n - x_{n-1} = rx_0 \quad \text{(constant)}.$$

This behaviour is shown schematically on the right.



Example: If $x_0 = 1000$ and $r = 0.06$, then:

$$x_1 = 1060$$
$$x_2 = 1120$$
$$x_3 = 1180$$
$$x_{10} = 1600$$

Of course, this is not a very good investment scheme. On the other hand, it is a nice borrowing scheme!

Compound Interest Scheme: In this scheme, the interest earned is based on the current value of the investment. Let's assume that interest on the investment is compounded (added) at regular time intervals $t_n = n\Delta t$. For example, if the interest is compounded annually, we can set $\Delta t = 1$ and t_n represents the number of years elapsed. If x_n is the value of the investment at time t_n, then the interest gained on the investment over the time interval $[t_n, t_{n+1}]$ is rx_n. As such, the value of the investment at time t_{n+1} is

$$x_{n+1} = x_n + rx_n = (1+r)x_n \,. \tag{5}$$

This is the difference equation associated with this simple compound interest scheme. From a look at the first two investment values,

$$x_1 = x_0 + rx_0 = (1+r)x_0$$
$$x_2 = x_1 + rx_1 = (1+r)x_1 = (1+r)^2 x_0 \,,$$

we can see the pattern and conjecture that the value of the investment at time t_n is

$$x_n = (1+r)^n x_0 \,. \tag{6}$$

But let's recall that we encountered the d.e. in Eq. (5) earlier - it has the form,

$$x_n = cx_{n-1} \quad \text{with} \quad c = 1 + r \,. \tag{7}$$

The solution of this difference equation was shown to be $x_n = c^n x_0$ which, in this case, becomes

$$x_n = (1+r)^n x_0 \,, \quad n = 0, 1, 2, \cdots \,, \tag{8}$$

in agreement with our earlier conjecture. Since

$$(1+r)^n = 1 + nr + \frac{n(n-1)}{2}r^2 + \cdots + r^n, \tag{9}$$

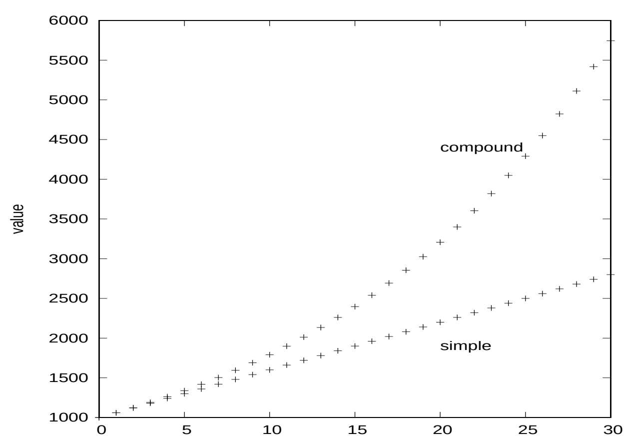
then the difference between investments with compound and simple interest at stage $n > 0$ is given by

$$(1+r)^n x_0 - (1+nr)x_0 = \left[\frac{n(n-1)}{2}r^2 + \cdots + r^n\right] x_0 > 0\,. \tag{10}$$

Example: As before $x_0 = 1000$ and $r = 0.06$. Then

$$\begin{aligned} x_1 &= 1060 \\ x_2 &= 1123.60 \\ x_3 &= 1191.02 \\ x_{10} &= 1790.85 \end{aligned}$$

As expected, the compound interest scheme will produce higher-valued investments for $n > 1$. For small n, i.e., the first few iterates, the differences between the two investment schemes are small but it doesn't take long for the compound scheme to "take off" and increase in what seems to be an exponential fashion. This is shown in the plot below, in which the values of the investments according to the two schemes, with $x_0 = 1000$ and $r = 0.06$ are plotted for $0 \leq n \leq 30$.



Interest is commonly compounded more frequently than annually, say n times a year. A very simple scheme is to divide the interest rate equally over each smaller compounding period, i.e. over the period of $\frac{1}{n}$ year $\left(= \frac{12}{n} \text{ months} = \frac{365}{n} \text{ days etc.}\right)$, the interest rate is $\frac{r}{n}$. Then after k such periods, the investment is worth

$$\left(1 + \frac{r}{n}\right)^k x_0. \tag{11}$$

After n periods $= 1$ year, the investment is worth

$$x_n = \left(1 + \frac{r}{n}\right)^n x_0. \tag{12}$$

After t years $= nt$ periods, the investment is worth

$$x_{nt} = \left(1 + \frac{r}{n}\right)^{nt} x_0. \tag{13}$$

Let us return to the prevous example with $x_0 = 1000$, $r = 0.06$ and investigate the effects of different compounding schemes. After one year, the investment is worth:

$$\begin{aligned}
1000(1.06) &= 1060 & \text{annual compounding,} \quad & n = 1 \quad (1 \text{ yr. period}) \\
1000(1.03)^2 &= 1060.90 & \text{semi-annual compounding,} \quad & n = 2 \quad \left(\tfrac{1}{2} \text{ yr. periods}\right) \\
1000(1.015)^4 &= 1061.36 & \text{quarterly compounding,} \quad & n = 4 \quad \left(\tfrac{1}{4} \text{ yr. periods}\right) \\
1000\left(1 + \tfrac{0.06}{365}\right)^{365} &= 1061.83 & \text{daily compounding,} \quad & n = 365
\end{aligned}$$

It appears as though the value of our investment after one year is approaching a limiting value as n increases, i.e. as the compounding period, $\frac{1}{n}$ year, approaches zero. A natural mathematical question is: "What is the value of the investment after one year if interest is compounded **continually**, i.e. over infinitesimal time periods dt?" The answer is that this value is the limit

$$\lim_{n\to\infty} \left(1 + \frac{r}{n}\right)^n x_0 = e^r x_0. \tag{14}$$

For the above example, $r = 0.06$, so that the resulting investment is worth

$$e^{0.06} \cdot 1000 = 1061.84\,. \tag{15}$$

You see that we have gained only **one cent** in going from daily compounding to continuous compounding.

After t years, the value of the investment is, from the expression for x_{nt},

$$\begin{aligned}
x(t) &= x_0 \lim_{n\to\infty} \left(1 + \frac{r}{n}\right)^{nt} \\
&= x_0 \left[\lim_{n\to\infty} \left(1 + \frac{r}{n}\right)^n\right]^t \\
&= x_0 e^{rt}.
\end{aligned} \tag{16}$$

In other words, our investment grows exponentially under the "Malthusian growth" model

$$\frac{dx}{dt} = rx \qquad x(0) = x_0\,, \tag{17}$$

where r is the interest rate.

Difference schemes for the numerical approximation of solutions of ODEs

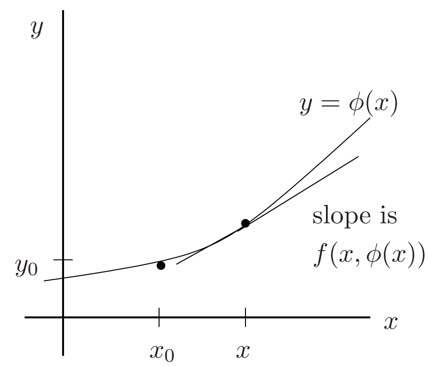
Consider the initial value problem (IVP) associated with the general first order DE

$$\frac{dy}{dx} = f(x,y) \qquad y(x_0) = y_0. \tag{18}$$

In other courses, you may have seen the fundamental result that if f is "nice enough" with respect to x and y, in particular y, at (x_0, y_0), then there exists a unique solution $y = \phi(x)$ satisfying the initial condition, i.e. $\phi(x_0) = y_0$.

Geometric interpretation: At each point (x, y) on the curve $y = \phi(x)$, the slope of the curve $y' = \phi'(x)$ is given by $f(x, y)$, i.e.

$$\phi'(x) = f(x, \phi(x)). \tag{19}$$

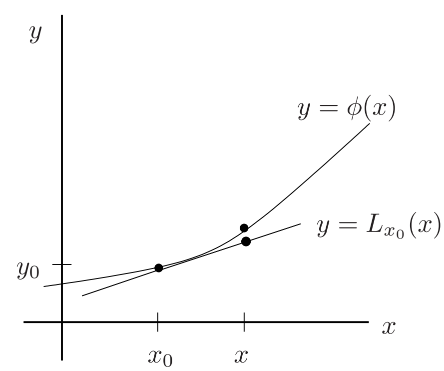


The Euler method of producing approximations $u(x)$ to the exact solution $\phi(x)$ is based on the linear or tangent-line approximation to $\phi(x)$. Recall that the linearization of $\phi(x)$ at $x = x_0$ is given by

$$L_{x_0}(x) = \phi(x_0) + \phi'(x_0)(x - x_0). \tag{20}$$

We then use the linearization $L_{x_0}(x)$ to approximate $\phi(x)$:

$$\phi(x) \cong L_{x_0} \quad \text{for} \quad x \quad \text{near} \quad x_0. \tag{21}$$



We expect that the closer x is to x_0, the better the approximation. In this example, we assume that we do not know the solution $\phi(x)$ in closed form. However, we do know that $\phi(x_0) = y_0$. In addition, since $\phi(x)$ is a solution to the DE, it follows that $\phi'(x_0) = f(x_0, y_0)$. Therefore the linear approximation becomes

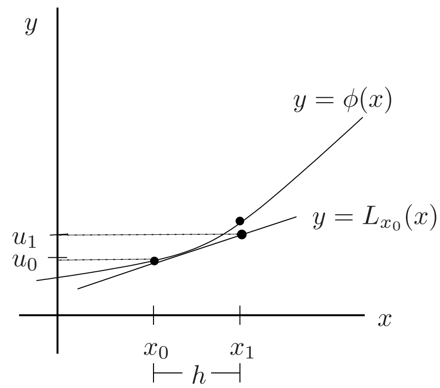
$$\phi(x) \cong y_0 + f(x_0, y_0)(x - x_0). \tag{22}$$

Now define a sequence of discrete x-values $x_0, x_1 = x_0 + h, x_2 = x_1 + h, \ldots$, so that

$$x_k = x_0 + kh, \quad k = 0, 1, 2, \ldots,$$

where $h > 0$ is a "step-size", as of yet undetermined. Now use $L_{x_0}(x)$ to produce an approximation to $\phi(x)$ at $= x_1$. We shall call this value u_1:

$$\begin{aligned} u_1 &= L_{x_0}(x_1) \\ &= y_0 + f(x_0, y_0)(x_1 - x_0) \\ &= y_0 + f(x_0, y_0)h \end{aligned}$$



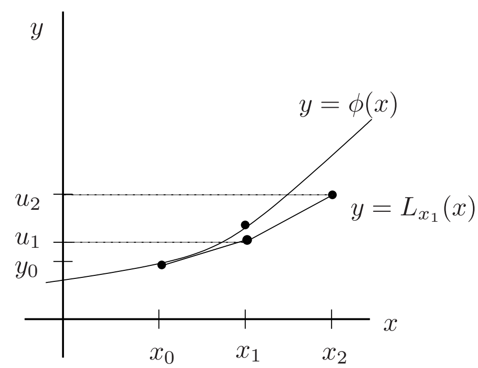
It is convenient to define $u_0 = y_0$, so that the above equation becomes

$$u_1 = u_0 + f(x_0, y_0)h. \qquad (23)$$

Thus, we claim that $\phi(x_1) \cong u_1$, as shown on the right.
We then use the point (x_1, u_1) as a starting point for a new linear approximation to $\phi(x)$. The problem is that we don't know the true value of $\phi(x_1)$ or its derivative $\phi'(x_1)$.

In the most naive form of this approximation method, we simply focus on the solution to the DE $y' = f(x, y)$ that passes through (x_1, u_1). The associated linearization to this solution is

$$L_{x_1}(x) = u_1 + f(x_1, u_1)(x - x_1). \qquad (24)$$



The approximation to $\phi(x_2)$ is then

$$\begin{aligned} u_2 &= u_1 + f(x_1, u_1)(x_2 - x_1) \\ &= u_1 + f(x_1, u_1)h. \end{aligned} \qquad (25)$$

This procedure is continued in an obvious manner to produce the difference scheme,

$$u_{k+1} = u_k + f(x_k, u_k)h, \qquad k = 0, 1, 2, \ldots \qquad (26)$$

where $x_k = x_0 + kh$.

Example: Consider the initial value problem

$$\frac{dy}{dx} = xy, \quad y(0) = 1.$$

This is a first order DE which can be solved using two methods that you probably saw in first-year

Calculus: (i) it is a linear DE, (ii) it is a separable DE. The unique solution of this IVP is (Exercise):

$$\phi(x) = e^{x^2/2}. \tag{27}$$

Here, $f(x, y) = xy$ so, for a given $h > 0$, the Euler difference scheme of Eq. (26) becomes

$$u_{k+1} = u_k + x_k u_k h, \qquad k = 0, 1, 2, \dots , \tag{28}$$

with $u_0 = y(0) = 1$. Since, in this case, $x_k = kh$, the above difference equation becomes

$$u_{k+1} = u_k + k u_k h^2 = (1 + kh^2) u_k, \qquad k = 0, 1, 2 \dots . \tag{29}$$

We expect the Euler method to yield better approximations for smaller values of the step size h. However, a smaller step size h implies a greater number of computations necessary to compute the approximation $u(x)$ over a prescribed interval of x-values. Also note that a basic limitation of this method is that errors propagate as we move away from x_0.

To illustrate, we show some computations associated with the above example for two choices of step size, $h = 0.1$ and 0.01, over the interval $0 \leq x \leq 1$. (Only a few of the iterates u_k are shown in the case $h = 0.01$.) In the tables below are listed the approximations u_k and the exact values of the solution $y_k = \phi(x_k)$ at the mesh points x_k along with the errors $e_k = u_k - y_k$. (We choose to plot the errors in this way, as opposed to the usual absolute values of the errors, to emphasize that all approximations lie *below* the exact solution curve $\phi(x)$ in this example. Question: Why do the approximations lie below the exact solution curve?)

Note the improvement in the approximations u_k yielded by decreasing the step size, for example, an error of -0.01 vs. an error of -0.1 at $x = 1$. There is, of course, a computational price to be paid: Using the smaller step size ($h = 0.01$ vs. $h = 0.1$) implies a tenfold increase in the number of computations needed to arrive at $x = 1$ from $x_0 = 0$.

Stepsize $h = 0.1$

k	x_k	u_k	$y_k = \phi(x_k)$	$e_k = u_k - y_k$
0	.00	1.00000	1.00000	
1	.10	1.00000	1.00501	-.00501
2	.20	1.01000	1.02020	-.01020
3	.30	1.03020	1.04603	-.01583
4	.40	1.06111	1.08329	-.02218
5	.50	1.10355	1.13315	-.02960
6	.60	1.15873	1.19722	-.03849
7	.70	1.22825	1.27762	-.04937
8	.80	1.31423	1.37713	-.06290
9	.90	1.41937	1.49930	-.07994
10	1.00	1.54711	1.64872	-.10161

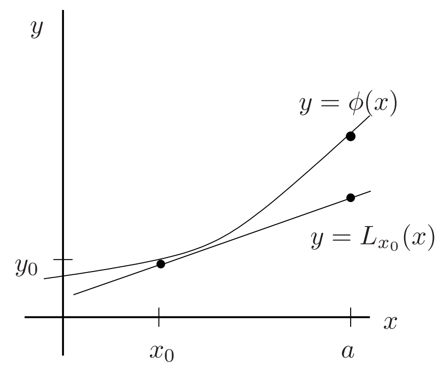
Stepsize $h = 0.01$

k	x_k	u_k	$y_k = \phi(x_k)$	$e_k = u_k - y_k$
0	.00	1.00000	1.00000	
1	.01	1.00000	1.00005	-.00005
2	.02	1.00010	1.00020	-.00010
3	.03	1.00030	1.00045	-.00015
4	.04	1.00060	1.00080	-.00020
5	.05	1.00100	1.00125	-.00025
6	.06	1.00150	1.00180	-.00030
7	.07	1.00210	1.00245	-.00035
8	.08	1.00280	1.00321	-.00040
9	.09	1.00361	1.00406	-.00045
10	.10	1.00451	1.00501	-.00050
20	.20	1.01917	1.02020	-.00103
30	.30	1.04442	1.04603	-.00161
40	.40	1.08101	1.08329	-.00228
50	.50	1.13009	1.13315	-.00306
60	.60	1.19321	1.19722	-.00400
70	.70	1.27245	1.27762	-.00517
80	.80	1.37049	1.37713	-.00664
90	.90	1.49080	1.49930	-.00851
100	1.00	1.63782	1.64872	-.01090

Estimates of the solution $\phi(x) = e^{x^2/2}$ of the initial value problem $y' = xy, \quad y(0) = 1$ on $[0, 1]$ yielded by Euler's method for two step sizes: $h = 0.1$ and $h = 0.01$.

Let us now outline a slight modification of the above scheme in order to systematically produce better estimates of the solution $\phi(x)$ at a particular value of x, say $x = a$. (Without loss of generality, we assume that $a > x$, to simplify the picture.) As before, the simplest approximation to $\phi(a)$ is afforded by the linearization to $\phi(x)$ at $x = x_0$:

$$\begin{aligned}\phi(a) &\cong L_{x_0}(a) \\ &= y_0 + f(x_0, y_0)(a - x).\end{aligned}$$



In an attempt to produce better estimates of $\phi(a)$, we partition the interval $[x_0, a]$ into n subintervals of length $\Delta x = h = \frac{a-x_0}{n}$ by introducing the partition points $x_k = x_0 + kh$, $k = 0, 1, 2, \ldots, n$.

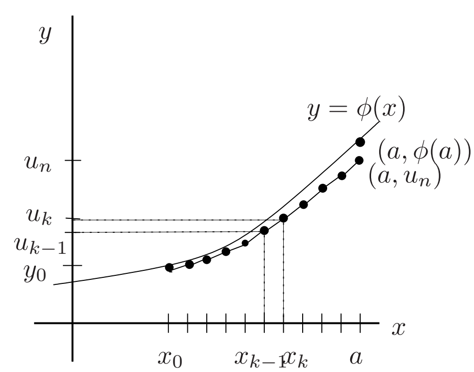
We now apply our earlier Euler method on these points x_k in order to generate the sequence of approximations $\phi(x_k) \cong u_k$, where $u_0 = y_0$ and

$$u_{k+1} = u_k + f(x_k, u_k)h \qquad k = 0, 1, 2, \ldots, n-1. \tag{30}$$

The final element in this sequence u_n, is our desired approximation to $\phi(a)$:

$$\phi(a) \cong u_n.$$

The procedure is sketched below.



The following theoretical result can be proved: Given a fixed a, and the algorithm outlined above, with $h = \frac{a-x_0}{n}$, $x_k = x_0 + kh$, $k = 0, 1, 2, \ldots, n$, define the sequence

$$\begin{aligned}u_0 &= y_0, \\ u_{k+1} &= u_k + f(x_k, u_k)h, \qquad 1 \le k \le n-1.\end{aligned}$$

Then the absolute error $E_n = |\phi(a) - u_n| \to 0$ as $n \to \infty$.

The proof is beyond the scope of this course. An intuitive picture of the process may be obtained by rearranging the above difference scheme:

$$\frac{u_{k+1} - u_k}{h} = \frac{\Delta u}{\Delta x} = f(x_k, u_k), \qquad 1 \le k \le n-1.$$

As n increases, the step size h decreases, approaching zero. The LHS of the above equation gets "closer" to a derivative $\frac{du}{dx}$ so that the equation, in some sense, gets "closer" to the DE

$$\frac{du}{dx} = f(x, u).$$

Example: Return to the initial value problem studied earlier,

$$\frac{dy}{dx} = xy, \quad y(0) = 1,$$

with exact solution $\phi(x) = e^{x^2/2}$. We shall be concerned with the approximation to $\phi(1) = e^{1/2} \cong 1.64872$, i.e. $a = 1$.

For a given value of $n \ge 1$, the step size becomes $h = \frac{1}{n}$ with corresponding mesh points $x_k = kh = \frac{k}{n}$, $k = 0, 1, 2, \ldots, n$. The Euler difference scheme becomes: $u_0 = 1$,

$$u_{k+1} = u_k + x_k u_k h, \qquad k = 1, 2, \ldots, n-1.$$

Below are listed the approximations u_n yielded by some values of n:

n	u_n
1	0
2	1.25
10	1.54711
50	1.62710
100	1.63782
500	1.64653
1000	1.64762
10000	1.64861

(The special cases $n = 10$ and 100 were computed in the table on Page 17.) As n increases, the values u_n are seen to approach the exact value $1.64872\ldots$.

Special Example: The Difference Equation $x_n = ax_{n-1} + b$

Recall that we have already studied the difference equation,

$$x_n = ax_{n-1}. \tag{31}$$

(OK, we used "c" before. Now we're going to use "a".) For example, if $|a| < 1$, then $x_n \to 0$ as $n \to \infty$. If $|a| > 1$, then $|x_n| \to \infty$ as $n \to \infty$.

We now consider the behaviour of solutions to the following difference equation,

$$x_n = ax_{n-1} + b\,, \tag{32}$$

where a and b are constants.

With perhaps some stretch of the imagination, we can consider this d.e. as a discrete model for population growth, in the case $a > 0$. The physical interpretation of the constant term "b" could be as follows: In addition to the birth/death processes of this bacterium assumed earlier, i.e. at each discrete time interval t_k, replacing the previous population x_{k-1} by the new population ax_{k-1}, we somehow add ($b > 0$) or subtract ($b < 0$) a constant amount of bacteria to/from the Petri dish.

Perhaps a more tractable application involves investments with compound interest (see earlier section). Suppose that x_0 dollars are invested at an annual interest rate r. Also suppose that after each period of compounding, a fixed amount b is either added ($b > 0$) or withdrawn ($b < 0$) from the investment. Then Eq. (32) will model the value of the investment, with $a = 1 + r$.

Based on our earlier discussions in this course, it would be a natural course of action to try to find the solution to Eq. (32) by examining the first few iterates x_n, given an initial value x_0:

$$\begin{aligned} x_1 &= ax_0 + b && (33)\\ x_2 &= ax_1 + b = a(ax_0 + b) + b && (34)\\ &= a^2x_0 + ab + b && (35)\\ x_3 &= ax_2 + b = a^3x_0 + a^2b + ab + b. && (36) \end{aligned}$$

There seems to be a general pattern evident in these first few iterates. For a given n, the form of x_n involves a term of the form $a^n x_0$. The remainder of the expression for x_n appears to have a common factor b multiplying a polynomial in a. The general expression for x_n appears to have the form,

$$x_n = a^n x_0 + bS_n\,, \quad n \geq 1\,, \tag{37}$$

where

$$S_n = 1 + a + a^2 + \cdots + a^{n-1}\,, \quad n \geq 1\,. \tag{38}$$

We can check whether or not expression (37) for x_n satisfies the difference equation in (32) by substituting it into both sides separately.

$$\begin{aligned}
\text{RHS} &= ax_{n-1} + b \\
&= a(a^{n-1}x_0 + bS_{n-1}) + b \\
&= a^n x_0 + ab(1 + a + \cdots a^{n-2}) + b \\
&= a^n x_0 + b(a + a^2 + \cdots a^{n-1}) + b \\
&= a^n x_0 + b(1 + a + a^2 + \cdots a^{n-1}) \\
&= a^n x_0 + bS_n\,. \\
\text{LHS} &= x_n \\
&= a^n x_0 + bS_n\,.
\end{aligned} \tag{39}$$

Since RHS=LHS, the expression for x_n in (37) satisfies the difference equation in (32).

Just to go a little further, S_n is the partial sum of n terms of a geometric series with first term 1 and ratio a. If $a = 1$, then $S_n = n$. If $a \neq 1$, then

$$S_n = \frac{1 - a^n}{1 - a} \quad n \geq 1\,. \tag{40}$$

Thus, the solution to the difference equation $x_n = ax_{n-1} + b$, expressed in terms of x_0, is

$$\boxed{x_n = \begin{cases} a^n x_0 + b\frac{1-a^n}{1-a}, & a \neq 1, \\ x_0 + bn, & a = 1, \end{cases} \qquad n = 1, 2, \ldots\,.} \tag{41}$$

The next goal is to determine the possible asymptotic behaviours of the sequences $\{x_n\}$ as $n \to \infty$ in terms of the parameters a and b. In principle, this can be done from a knowledge of the solution in Eq. (41) but a complete analysis is rather complicated. Indeed, this is the way in which the d.e. in (32) was analyzed in many textbooks in the past, for example, the book by Goldberg (1958), a secondary reference for this course. A summary of such an analysis is presented as an Appendix to this lecture. There is a much easier method, as we now show.

The idea is to look for "fixed point" solutions to the d.e. in Eq. (32) – points which remain constant under the iteration procedure. Recall that 0 is a fixed point – the only fixed point – of the d.e.,

$$x_n = cx_{n-1}\,. \tag{42}$$

We now look for fixed points to the d.e. in (32), which we copy below for convenience,

$$x_n = ax_{n-1} + b\,, \tag{43}$$

i.e., values $\bar{x} \in \mathbb{R}$ such that

$$\bar{x} = a\bar{x} + b\,. \tag{44}$$

(We'll be using bars, i.e., $\bar{x}$, to denote fixed points throughout this course.)
We could simply solve for $\bar{x}$ in (44), i.e.,

$$\bar{x} = \frac{b}{1-a}\,, \quad a \neq 1\,, \tag{45}$$

noting, as we wrote above, that the expression is valid only when $a \neq 1$, which means "almost always". What about the special case $a = 1$? We have already solved this d.e. – the solution is given in 41), In this case, the d.e. in)43) becomes

$$x_n = x_0 + bn\,. \tag{46}$$

The x_n are simply **translates** of x_0. In the very special case $b = 0$, then $x_n = x_0$ for all $n \geq 1$, i.e., every point x_0 is a fixed point. For $b \neq 0$, we simply keep stepping to the right (if $b > 0$) or to the left (if $b < 0$) on the real line $\mathbb{R}$. In other words, there are **no** fixed points.

For the case $a \neq 1$, we're going to do something quite "radical." We're going to make a "change of variables". For a given sequence $\{x_n\}$ which corresponds to an initial value x_0, we're going to define the following sequence,

$$y_n = x_n - \bar{x}\,, \tag{47}$$

where $\bar{x}$ is the fixed point of the d.e. as given in Eq. (45). In other words, the y_n are the **relative positions of the iterates x_n with respect to the fixed point $\bar{x}$ of the d.e.**. This is important and we shall return to this idea a little later.

Let us now substitute (47) into the d.e. in (43). We'll have to rewrite it as follows,

$$x_n = y_n + \bar{x}\,. \tag{48}$$

Now substitute into (43),

$$y_n + \bar{x} = a(y_{n-1} + \bar{x}) + b\,, \tag{49}$$

which we can rewrite as

$$y_n = ay_{n-1} + a\bar{x} + b - \bar{x}\,. \tag{50}$$

But $\bar{x}$ satisfies the fixed point equation in (44), implying that the above d.e. becomes

$$\boxed{y_n = ay_{n-1}\,, \quad n \geq 1 \quad (a \neq 1)\,.} \tag{51}$$

Wow! By the simple change of variable in (47) we have transformed our original d.e. in (43) to the simpler d.e. in (51). And we know all of the properties of this d.e. from our analysis in the previous lecture! We can quickly deduce the following results.

- Firstly, we know that $y = 0$ is the fixed point of this d.e.. But from the change of variable in (47), we know that this corresponds to the point $\bar{x}$ associated with the x_n. And we know that this is the fixed point of the original d.e. (43). So far, so good.

- In the case $|a| < 1$, we know that all $y_n \to 0$ as $n \to \infty$. From our change of variable in (47), this implies that all x_n in our original d.e. in (43) behave as $x_n \to \bar{x}$.

  In the subcase $0 < a < 1$, the y_n approach 0 monotonically as $n \to \infty$ which implies that the x_n approach $\bar{x}$ monotonically as $n \to \infty$.

  In the subcase $-1 < a < 0$, the y_n approach 0 in an alternating fashion which implies that the x_n approach $\bar{x}$ in an alternating fashion as $n \to \infty$.

  We conclude that in all of the above cases, the fixed point $\bar{x}$ is **attractive** under the iteration process defined by (43).

- In the case $|a| > 1$, we know that the magnitudes $|y_n| \to \infty$ as $n \to \infty$. This implies that the fixed point $y = 0$ under the iteration process defined by (51) is **repulsive**. From our change of variable, this implies that the fixed point $x = \bar{x}$ is **repulsive** under the iteration process defined by (43).

  In the subcase $a > 1$, the iterates y_n will move away from the fixed point 0 monotonically which implies that the iterates x_n will move away from $\bar{x}$ monotonically: If $x_0 > 0$, then $x_n \to +\infty$ monotonically and if $x_0 < 0$, then $x_n \to -\infty$ monotonically.

  In the subcase $a < -1$, the iterates y_n will move away from the fixed point 0 in an alternating fashion which applies that the iterates x_n will move away from $\bar{x}$ in an alternating fashion.

This is a very important result worthy of further discussion. The change of variable,

$$y_n = x_n - \bar{x}\,, \tag{52}$$

transforms the difference equation/dynamical system,

$$x_n = ax_{n-1} + b\,, \tag{53}$$

into the difference equation/dynamical system,

$$y_n = ay_{n-1}\,, \tag{54}$$

which is much easier to study – we performed an analysis of this d.e. in Lecture 1 with little problem. But just as important: **By examining the iterates y_n, we are actually viewing the iterates x_n while "standing" at the fixed point $\bar{x}$ of the original d.e..** In other words, we are viewing the behaviour of the iterates x_n **relative to the fixed point** $\bar{x}$. In this way, we can obtain a complete picture of the behaviour of iterates x_n without (i) having to find the general solution in Eq. (41) and then (ii) having to perform the tedious analysis presented in the Appendix.

Finally, let us recall one additional idea from our earlier discussion. The iteration procedure,

$$x_n = ax_n + b \tag{55}$$

may be written in the form,

$$x_n = f(x_{n-1}), \tag{56}$$

as before but this time the function $f(x)$ is given by

$$f(x) = ax + b. \tag{57}$$

Recall that the **fixed point** of $f(x)$, i.e., a point $\bar{x}$ such that $f(\bar{x}) = \bar{x}$, is the fixed point of the dynamical system in (56) since

$$\bar{x} = f(\bar{x}) = \bar{x}. \tag{58}$$

In other words, if $x_0 = \bar{x}$ then the iterates in Eq. (56) are $x_n = \bar{x}$. The fixed point of the function $f(x) = ax + b$ is the fixed point $\bar{x}$ determined earlier:

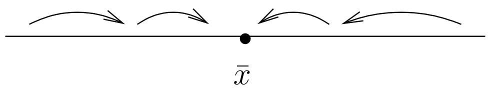
$$f(\bar{x}) = \bar{x} \quad \Longrightarrow \quad a\bar{x} + b = \bar{x} \quad \Longrightarrow \quad \bar{x} = \frac{b}{1-a}, \qquad a \neq 1. \tag{59}$$

The above analysis implies that we can take the phase portraits that were derived in the first lecture for the dynamical system $x_n = cx_{n-1}$ and use them here for the dynamical system $x_n = ax_{n-1} + b$ with a simple change – we change the fixed point 0 used before to the fixed point $\bar{x}$. Of course, in the special case $b = 0$, the fixed point is $\bar{x} = 0$. The results are presented below.

Phase portraits of the dynamical system $x_n = ax_{n-1} + b$

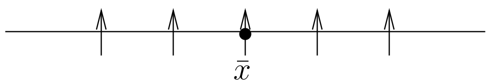
Case A: $a > 0$

$0 < c < 1$



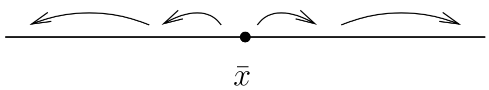
$\bar{x}$ is an *attractive* fixed point

$c = 1$



each point x is a fixed point

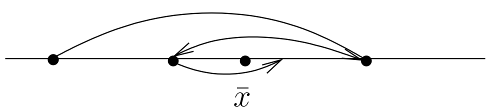
$c > 1$



$\bar{x}$ is a *repulsive* fixed point

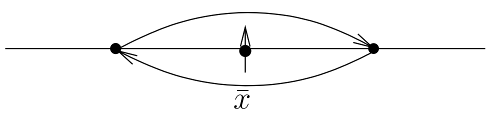
Case B: $a < 0$

$-1 < c < 0$



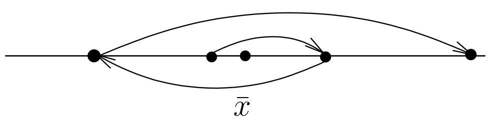
$\bar{x}$ is an *attractive* fixed point

$c = -1$



$\bar{x}$ is a *neutral* fixed point

$c < -1$



$\bar{x}$ is a *repulsive* fixed point

Appendix: A more "traditional" analysis of the dynamical system $x_n = ax_{n-1} + b$

Note: This section is presented for purposes of extra information only. It was not covered in the lecture.

Here we present a more traditional analysis, as often presented in older textbooks of the behaviour of solutions to the difference equation,

$$x_n = ax_{n-1} + b\,, \tag{60}$$

from a knowledge of the formal solution as determined earlier and presented in Eq. (41) and copied below for convenience,

$$\boxed{x_n = \begin{cases} a^n x_0 + b\frac{1-a^n}{1-a}, & a \neq 1, \\ x_0 + bn, & a = 1, \end{cases} \qquad n = 1, 2, \ldots\,.} \tag{61}$$

A complete analysis is rather complicated, so we outline only some major features.

1. $a = 1$. This case is particularly simple, so we deal with it first. The difference equation becomes

   $$x_n = x_{n-1} + b$$

   so that (see above)

   $$x_n = x_0 + nb.$$

   i. If $b = 0$, then $x_n = x_0$ for all n.

   ii. If $b > 0$, then $x_n - x_{n-1} = b > 0$ so that $\{x_n\}$ is an increasing sequence and $x_n \to +\infty$ as $n \to \infty$.

   iii. If $b < 0$, then $x_n - x_{n-1} = b < 0$ so that $\{x_n\}$ is a decreasing sequence, with $x_n \to -\infty$ as $n \to \infty$.

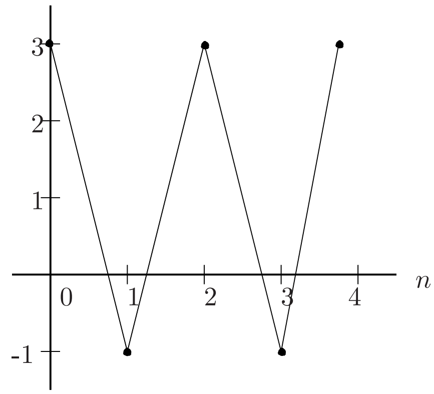
2. $a = -1$. In this case,

   $$\begin{aligned} S_n &= 1 + a + \cdots + a^{n-1} \\ &= \begin{cases} 1, & n \quad \text{odd} \\ 0, & n \quad \text{even.} \end{cases} \end{aligned}$$

   Substitution into (61) yields

   $$x_n = \begin{cases} -x_0 + b, & n \quad \text{odd} \\ x_0, & n \quad \text{even.} \end{cases}$$

   Thus, $\{x_n\}$ is an oscillating sequence. In the special case $b = 0$, $x_n = (-1)^n x_0$.



Plot of x_n vs. n for $x_n = -x_{n-1} + 2$, $x_0 = 3$ ($a = -1$, $b = 2$).

3. $|a| < 1$. From (61),

$$x_n = a^n x_0 + b\frac{1-a^n}{1-a}$$

and the fact that $a^n \to 0$ as $n \to \infty$, we have

$$\begin{aligned} \lim_{n\to\infty} x_n &= \lim_{n\to\infty} \left[a^n x_0 + b\frac{1-a^n}{1-a}\right] \\ &= \frac{b}{1-a}. \end{aligned} \tag{62}$$

Thus the sequence $\{x_n\}$ converges to the value $\bar{x} = \frac{b}{1-a}$. Note that $\bar{x}$ is the *fixed point* of the function $f(x) = ax+b$ that defines the difference equation/dynamical system being studied here:

$$\begin{aligned} x_n &= f(x_{n-1}) \\ &= ax_{n-1} + b. \end{aligned}$$

That $\bar{x}$ is a fixed point of f means that $\bar{x} = f(\bar{x})$. Let us check this:

$$\begin{aligned} f(\bar{x}) &= a\bar{x} + b \\ &= a\frac{b}{1-a} + b \\ &= \frac{ab + b(1-a)}{1-a} \\ &= \frac{b}{1-a} \\ &= \bar{x}. \end{aligned}$$

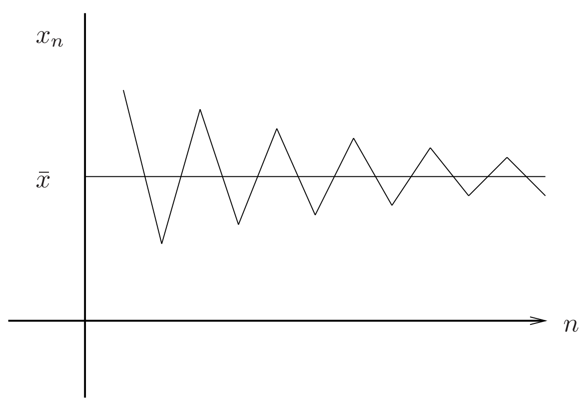
Thus, if $x_0 = \bar{x}$, then $x_n = \bar{x}$ for $n = 1, 2, 3, \ldots$. If we start at the fixed point $\bar{x}$ of f, we stay there forever. On the other hand, for all other starting points $x_0 \neq \bar{x}$, we approach $\bar{x}$ arbitrarily closely, i.e. $\bar{x}$ is the limit point of the sequence. In this case, $|a| < 1$, $\bar{x}$ is said to be an *attractive* fixed point. (It's also the only one.)

A more detailed analysis of the behaviour of the iterates $\{x_n\}$ is possible, but beyond the scope of this section. We shall see later that a graphical analysis of the iteration procedure $x_k = f(x_{k-1})$ can reveal much of this qualitative behaviour. Let us simply summarize the behaviour of the $\{x_n\}$ sequence, dependent upon a, as follows:

i) If $0 < a < 1$, then a^n decreases monotonically to zero as n increases. The result is that the x_n approach $\bar{x}$ monotonically:

  - from above if $x_0 > \bar{x}$, i.e. $x_0 > x_1 > x_2 \dots x_n \to \bar{x}$
  - from below if $x_0 < \bar{x}$, i.e. $x_0 < x_1 < x_2 \dots x_n \to \bar{x}$
  - of course, if $x_0 = \bar{x}$, then $x_n = \bar{x}$.

ii) If $-1 < a < 0$, then the a^n oscillate in sign but decrease in magnitude to zero as n increases. The result is a "damped oscillatory" motion of the x_n toward $\bar{x}$.



"Damped oscillatory" convergence to $\bar{x}$.

4. $|a| > 1$. Since $|a^n|$ increases with n, we expect the iterates x_n to exhibit divergence, with $|x_n| \to \infty$ as $n \to \infty$. From Eq. (37) and (38),

$$\begin{aligned} x_n - x_{n-1} &= a^n x_0 - a^{n-1} x_0 + S_n - S_{n-1} \\ &= a^{n-1}\big[(a-1)x_0 + b\big] \end{aligned} \tag{63}$$

i. $a > 1$. In this case

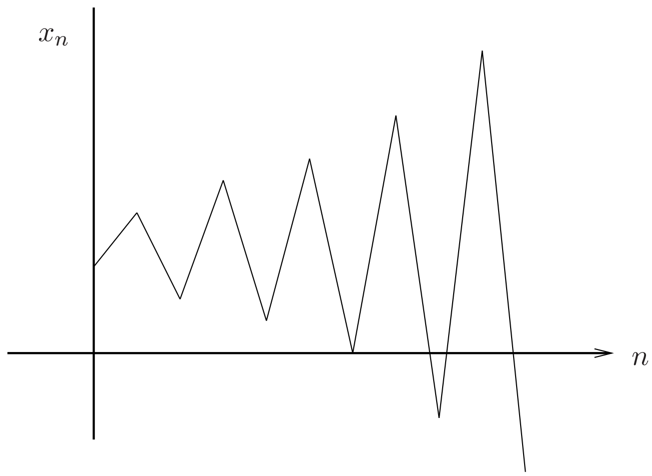
$$x_n - x_{n-1} > 0$$

i.e. $\{x_n\}$ is a monotonically increasing sequence. (Note that the differences themselves, $x_n - x_{n-1}$, increase geometrically as a^{n-1}.) Thus $x_n \to +\infty$ as $n \to \infty$, i.e. the series diverges.

ii. $a < -1$. From (63), $x_n - x_{n-1}$ alternates in sign, implying that the sequence $\{x_n\}$ is oscillatory. However

$$|x_n - x_{n-1}| = K|a|^{n-1},$$

where $K = |(a-1)x_0 + b|$ so that $|x_n - x_{n-1}| \to \infty$ as $n \to \infty$. Thus, the series diverges, oscillating with greater and greater amplitude as $n \to \infty$. This behaviour is illustrated in the following figure.



Oscillatory divergence of x_n, $|x_n| \to \infty$ as $n \to \infty$.

We summarize the various behaviours of the iteration sequence in the table below. The reader is encouraged to draw sample plots of these sequences.

BEHAVIOUR OF THE SOLUTION SEQUENCE $\{x_n\}$

$x_{n+1} = ax_k + b \quad n = 0, 1, 2, \ldots$ with fixed point $\bar{x} = \frac{b}{1-a}$

Parameters/Initial Conditions				Behaviour of Solution	
Row	a	b	x_0	for $n = 1, 2, 3, \ldots$	the sequence $\{x_n\}$ is
(A)	$a \neq 1$		$x_0 = \bar{x}$	$x_n = \bar{x}$	constant $(= \bar{x})$
(B)	$a > 1$		$x_0 > \bar{x}$	$x_k > \bar{x}$	monotone increasing, diverges to $+\infty$
(C)	$a > 1$		$x_0 < \bar{x}$	$x_k < \bar{y}$	monotone decreasing, diverges to $-\infty$
(D)	$0 < a < 1$		$x_0 > \bar{x}$	$x_k > \bar{x}$	monotone decreasing, converges to limit $\bar{x}$
(E)	$0 < a < 1$		$x_0 < \bar{x}$	$x_k < \bar{x}$	monotone increasing, converges to limit $\bar{x}$
(F)	$-1 < a < 0$		$x_0 \neq \bar{x}$		damped oscillatory, converges to limit $\bar{x}$
(G)	$a = -1$		$x_0 \neq \bar{x}$		divergent, oscillates finitely
(H)	$a < -1$		$x_0 \neq \bar{x}$		divergent, oscillates infinitely
(I)	$a = 1$	$b = 0$		$x_k = x_0$	constant $(= x_0)$
(J)	$a = 1$	$b > 0$		$x_k > x_0$	monotone increasing, diverges to $+\infty$
(K)	$a = 1$	$b < 0$		$x_k < x_0$	monotone decreasing, diverges to $-\infty$

Lecture 4

A brief review of sequences and their properties

As in the example of the previous section, we shall be interested in the qualitative behaviour of sequences $\{x_k\}$ that are solutions to difference equations. Of primary concern is whether such sequences converge to a limit, oscillate or diverge to infinity. "Desirable" behaviour of such solutions depends, of course, on the particular application. For example, if the d.e. models the value of an investment, a growth of the $\{x_k\}$ is desirable. If it models the temperature of a room in response to a thermostat feedback control, a convergence to a limit point is desirable.

For this purpose, it is important to recall some fundamental concepts involving sequences. Below are listed the major properties of sequences. For more details, it is recommended that you consult a calculus textbook.

Given an infinite sequence of real numbers $\{a_n\}_{n=0}^{\infty}$, the sequence is said to be:

1) **increasing** if $a_n \le a_{n+1}$ for all n, **strictly increasing** if $a_n < a_{n+1}$, for all n,

2) **decreasing** if $a_n \ge a_{n+1}$ for all n, **strictly decreasing** if $a_n > a_{n+1}$ for all n,

3) **monotone** if either (strictly) increasing or (strictly) decreasing,

4) **alternating** if a_n and a_{n+1} have opposite signs for all n, i.e. $a_n a_{n+1} < 0$,
   oscillating if $b_n = a_{n+1} - a_n$ and $b_{n+1} = a_{n+2} - a_{n+1}$ have opposite signs for all n,

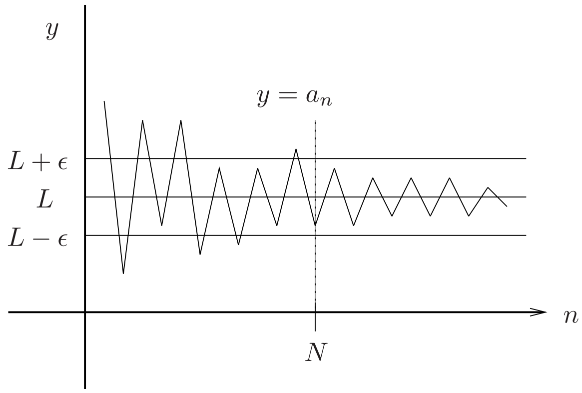
5) **bounded above** if there exists a constant $A \in \mathbb{R}$ such that $a_n \le A$ for all n,
   bounded below if there exists a constant $B \in \mathbb{R}$ such that $a_n \ge B$ for all n,
   bounded if it is bounded both above and below,

6) **convergent** if $\lim_{n\to\infty} a_n = L$ exists, i.e. is finite.

It is also very important to review the meaning of the limit of a sequence:

> $\lim_{n\to\infty} a_n = L$ means: Given any $\varepsilon > 0$, there exists an $N > 0$ (or more explicitly $N(\varepsilon)$) such that $|a_n - L| < \varepsilon$ for all $n \ge N$.

Note that this is a very strong requirement on the sequence elements as illustrated below:



Not even one element a_n can be outside the "ribbon" $L - \varepsilon < y < L + \varepsilon$ for $n \geq N$.

Practice Problems

1. A difference equation over the set $0, 1, 2, \ldots$ and a function are given. In each case, show that the function is a solution of the difference equation. (C, C_1, and C_2 denote arbitrary constant.)

(a)	$y_{k+1} - y_k = 0$	$y_k = 5$
(b)	$y_{k+1} - y_k = 0$	$y_k = C$
(c)	$y_{k+1} - y_k = 1$	$y_k = k$
(d)	$y_{k+1} - y_k = 1$	$y_k = k + C$
(e)	$y_{k+1} - y_k = k$	$y_k = \frac{k(k-1)}{2}$
(f)	$y_{k+1} - y_k = k$	$y_k = \frac{k(k-1)}{2} + C$
(g)	$y_{k+2} - 3y_{k+1} + 2y_k = 0$	$y_k = C_1 + C_2 \cdot 2^k$
(h)	$y_{k+2} - 3y_{k+1} + 2y_k = 1$	$y_k = C_1 + C_2 \cdot 2^k - k$
(i)	$y_{k+2} - y_k = 0$	$y_k = C_1 + C_2(-1)^k$
(j)	$y_{k+1} = \frac{y_k}{1+y_k}$	$y_k = \frac{C}{1+Ck}$

(a) For the difference equations in Problem 1(d), (f), and (j), find the particular solutions satisfying the initial condition $y_0 = 1$.

(b) For the difference equations in Problem 1(g), (h), and (i), find the particular solutions satisfying the two initial conditions $y_0 = 1$ and $y_1 = 2$.

Answers:

(a) (d) $y_k = k + 1$. (f) $y_k = \dfrac{k(k-1)}{2} + 1$. (j) $y_k = \dfrac{1}{1+k}$.

(b) (g) $y_k = 2^k$. (h) $y_k = -1 + 2^{k+1} - k$. (i) $y_k = \frac{3}{2} - \frac{1}{2}(-1)^k$.

2. Each of the following difference equations is assumed to be defined over the set of k-values $0, 1, 2, \ldots$. In addition, suppose y_0 is prescribed in each case and is equal to 2. Find the solution of the difference equation, write out the first six values of y in sequence form, and describe the apparent behaviour of y in the sequence.

(a) $y_{k+1} = 3y_k - 1$

(b) $y_{k+1} = y_k + 2$

(c) $y_{k+1} + y_k - 2 = 0$

(d) $3y_{k+1} = 2y_k + 3$

(e) $2y_{k+1} + y_k - 3 = 0$

(f) $y_{k+1} + 3y_k = 0$

Answers:

(a) $y_k = \frac{1}{2}(3^{k+1} + 1)$; 2, 5, 14, 41, 122, 365; y increases without bound.

(c) $y_k = 1 + (-1)^k$; 2, 0, 2, 0, 2, 0; y oscillates between the values 2 and 0.

(e) $y_k = 1 + \left(-\frac{1}{2}\right)^k$; 2, $\frac{1}{2}, \frac{5}{4}, \frac{7}{8}, \frac{17}{16}, \frac{31}{32}$; y approaches the value 1 but is alternately above and below 1.

3. Each of the following difference equations is assumed to be defined over the set of k-values $0, 1, 2, \ldots$. In each case (i) find the solution of the equation with the indicated value of y_0, (ii) characterize the behaviour of the (solution) sequence $\{y_k\}$ according to the table in the previous section. As well, (iii) draw a graph of this sequence $\{y_k\}$, labeling both axes and indicating the scale used on each.

(a)	$y_{k+1} = 3y_k - 1$	$y_0 = \frac{1}{2}$
(b)	$y_{k+1} = 3y_k - 1$	$y_0 = 1$
(c)	$y_{k+1} + 3y_k + 1 = 0$	$y_0 = 1$
(d)	$2y_{k+1} - y_k = 2$	$y_0 = 4$
(e)	$3y_{k+1} + 2y_k = 1$	$y_0 = 1$
(f)	$y_{k+1} = y_k - 1$	$y_0 = 5$
(g)	$y_{k+1} + y_k = 0$	$y_0 = 1$

Answers:

(a) $y_k = \frac{1}{2}$, row (A) in table.

(b) $y_k = \frac{1}{2}(3^k + 1)$, row (B) in table.

(c) $y_k = \frac{1}{4}[5(-3)^k - 1]$, row (H) in table.

(d) $y_k = 2\left[\left(\frac{1}{2}\right)^k + 1\right]$, row (D) in table.

(e) $y_k = \frac{1}{5}\left[4\left(-\frac{2}{3}\right)^k + 1\right]$, row (F) in table.

(f) $y_k = 5 - k$, row (K) in table.

(g) $y_k = (-1)^k$, row (G) in table.

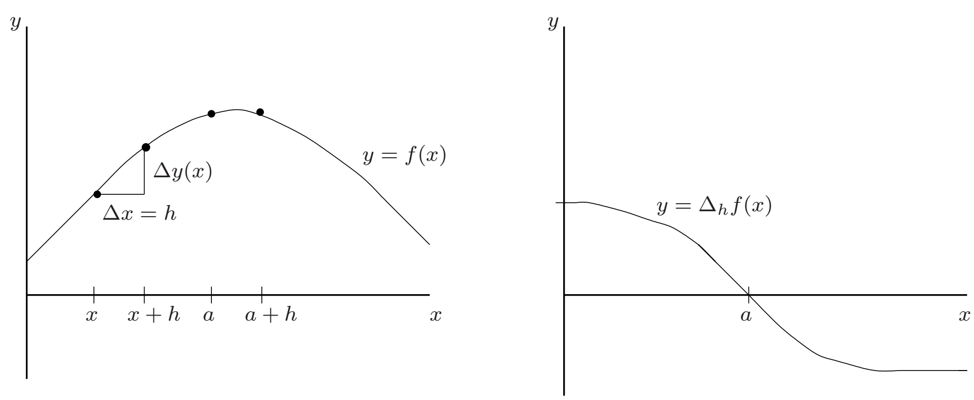
Some theory of difference equations

Finite differences

In this section, we briefly examine the idea of **finite differences** as applied to functions. Historically, this idea was important not only in the analysis of functions (a kind of discrete version of calculus) and the development of numerical methods but also of time series, which are particular examples of sequences. Over the years, and especially after the idea of "chaos" and "chaotic sequences" became a well-established principle, the analysis of time series has become a very rich subject. The purpose of this section is simply to give the reader an idea of the connection of finite differences and difference equations.

Let $f : \mathbb{R} \to \mathbb{R}$ be given function and $h > 0$ a "difference interval". For simplicity, we assume that f is defined for all values $x \in \mathbb{R}$. Now define a **(first) difference operator** Δ_h that acts on f to produce a new function g as follows:

$$g(x) = \Delta_h f(x) = f(x + h) - f(x). \tag{64}$$



The construction is illustrated above. We present a few examples below.

1) $f(x) = ax + b$, $a, b \in \mathbb{R}$ (linear function)

$$\begin{aligned} \Delta_h f(x) &= f(x + h) - f(x) \\ &= a(x + h) + b - [ax + b] \\ &= ah \end{aligned}$$

As expected, $g(x) = \Delta_h f(x)$ is a constant function.

2) $f(x) = x^2$

$$\begin{aligned} \Delta_h f(x) &= (x + h)^2 - x^2 \\ &= 2hx + h^2 \end{aligned}$$

The function $g(x) = \Delta_h f(x)$ is a straight line with slope $2h$ and y-intercept h^2.

We can apply the Δ_h difference operator more than once on a function, for example,

$$\begin{aligned}\Delta_h^2 f(x) &= \Delta_h(\Delta_h f(x)) \\ &= \Delta_h(f(x+h) - f(x)) \\ &= (f(x+h+h) - f(x+h)) - (f(x+h) - f(x)) \\ &= f(x+2h) - 2f(x+h) + f(x). \end{aligned} \tag{65}$$

Returning to the above examples:

1) $f(x) = ax + b$

$$\Delta_h^2 f(x) = \Delta_h(\Delta_h f(x)) = ah - ah = 0\,,$$

i.e., the second difference of a linear function is identically zero.

2) $f(x) = x^2$

$$\begin{aligned}\Delta_h^2 f(x) = \Delta_h(\Delta_h f(x)) &= 2h(x+h) + h^2 - (2hx + h^2) \\ &= 2h^2\,,\end{aligned}$$

which is constant, i.e., independent of x.

The **rate of change** of $f(x)$ over the interval $[x, x+h]$ is defined as

$$\frac{f(x+h) - f(x)}{h} = \frac{\Delta_h f(x)}{h}. \tag{66}$$

Returning again to the above examples

1) $f(x) = ax + b$

$$\frac{\Delta_h f(x)}{h} = \frac{ah}{h} = a$$

2) $f(x) = x^2$

$$\frac{\Delta_h f(x)}{h} = \frac{2hx + h^2}{h} = 2x + h.$$

You will recall that Calculus is concerned with the behaviour of the function f as well as its rate of change in the limit $h \to 0$:

a) If $f(x)$ is continuous at x then

$$\lim_{h\to 0} \Delta_h f(x) = \lim_{h\to 0} \big[f(x+h) - f(x)\big] = 0.$$

b) If $f(x)$ is differentiable at x then the limit

$$f'(x) = \lim_{h \to 0} \frac{\Delta_h f(x)}{h} = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

exists.

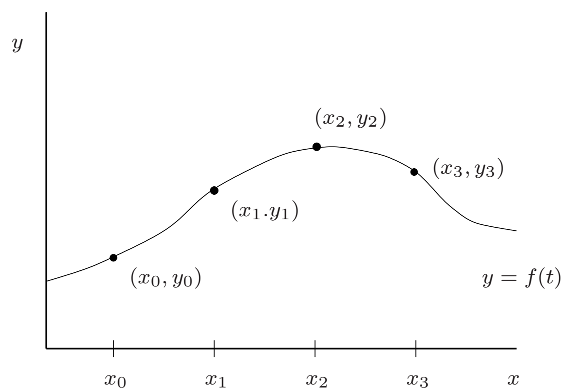
In principle, we can examine the differences $\Delta_h f(x)$, $\Delta_h^2 f(x)$, etc. at arbitrary values of $x \in \mathbb{R}$. In practice, however, we shall focus on a particular discrete set of x-values defined by a starting point $x_0 \in \mathbb{R}$ and

$$x_k = x_0 + kh, \qquad k = 0, 1, 2, \ldots . \tag{67}$$

We then define

$$y_k = f(x_k), \qquad k = 0, 1, 2, 3, \ldots . \tag{68}$$

Our study of the function $f(x)$ is thus restricted to the values y_k it assumes on the discrete set of points x_k, $k = 0, 1, 2, \ldots$.



In light of the discussion in the previous chapter, the set of points (x_k, y_k), $k = 0, 1, 2, \ldots$ may have been obtained by the sampling of a physical or biological system at equal time intervals, e.g. the degree of radioactivity of a rock sample or the number of bacteria in a Petri dish. Here, however we dispense with the notion of these points as samples of a function $f(x)$ and focus only on the sequence $y = \{y_k\}$ itself.

Now define the first differences of the sequence $\{y_k\}$ as

$$\Delta y_k = y_{k+1} - y_k \tag{69}$$

as well as the second differences

$$\begin{aligned} \Delta^2 y_k = \Delta(\Delta y_k) &= \Delta(y_{k+1} - y_k) \\ &= (y_{k+2} - y_{k+1}) - (y_{k+1} - y_k) \\ &= y_{k+2} - 2y_{k+1} + y_k. \end{aligned} \tag{70}$$

Note that we have dropped the subscript "h" from our operator symbol "Δ". We have conveniently forgotten about the origin of these sequence values: y is now considered a function of non-negative integer variables, i.e. $y_k = y(k)$. In some books, you will find the sequence values written as $y(k)$, $y(k+1), \ldots$.

Example: Let $f(x) = x^2$, $h = 1$, $x_0 = 0$ so that $x_k = k$. Then $y_k = f(x_k) = k^2$. Now construct a table of values of the first few differences:

k	y_k	Δy_k	$\Delta^2 y_k$	$\Delta^3 y_k$
0	0	1	2	0
1	1	3	2	0
2	4	5	2	0
3	9	7	2	
4	16	9		
5	25			

In this example, we note that the sequence elements y_k obey the equation $\Delta^2 y_k = 2$ or

$$y_{k+2} - 2y_{k+1} + y_k = 2 \qquad k = 0, 1, 2, \ldots . \tag{71}$$

This is a **second order linear difference equation** (with constant coefficients) since it involves the second order difference operator Δ^2. Once again, a solution of this difference equation is the sequence $y = (y_0, y_1, y_2, \ldots) = (0, 1, 4, 9, \ldots, k^2, \ldots)$. If we rewrote the difference equation as the recursion relation

$$y_{k+2} = 2y_{k+1} - y_k + 2, \qquad k \geq 0, \tag{72}$$

then the above sequence $y_k = k^2$ can be generated if we provide the two initial conditions $y_0 = 0$, $y_1 = 1$.

Note that the sequence $y_k = k^2$ is also observed to obey the equation $\Delta y_k = 2k + 1$ or

$$y_{k+1} - y_k = 2k + 1, \qquad k \geq 0. \tag{73}$$

This is a **linear first order difference equation** (with nonconstant coefficients). If it were rewritten as

$$y_{k+1} = y_k + 2k + 1, \qquad k \geq 0, \tag{74}$$

then the sequence $y_k = k^2$ can be generated if we provide the single initial condition $y_0 = 0$.

It is helpful to discuss the differential equation analogues of these difference equations. The function $y = f(x) = x^2$ is the unique solution of:

1) the second order differential equation $\dfrac{d^2y}{dx^2} = 2$ satisfying the initial conditions $y(0) = 0$, $y'(0) = 0$.

2) the first order differential equation $\dfrac{dy}{dx} = 2x$ satisfying the initial condition $y(0) = 0$.

Difference Equations

As stated earlier, the general form of an nth order difference equation is

$$y_k = f(k, y_{k-1}, y_{k-2}, \dots, y_{k-n}) \qquad k \geq n. \tag{75}$$

A solution to this equation is a sequence $y = (y_0, y_1, y_2, \dots)$ whose elements satisfy Eq. (75) for all permissible values of k.

For the remainder of these course notes, we shall adopt the following convenient abbreviations:

DE refers to "differential equation"
d.e. refers to "difference equation"

As in the case of DEs, there are d.e.'s that have no solutions, one solution or an infinity of solutions.

Examples:

1) The d.e. $(y_{k+1} - y_k)^2 + y_k^2 = -1$ has no real solutions.

2) The d.e. $y_{k+1} - y_k = 0$ has an infinity of solutions, e.g. $y_k = 0$, $y_k = 5$, $y_k = e^{\pi}$. In general, $y_k = C$ where C is an arbitrary constant since, trivially,

$$y_{k+1} - y_k = C - C = 0 \quad \text{for all } k \geq 0\,.$$

3) The d.e. $y_{k+1} - y_k = 1$ has an infinity of solutions: $y_k = k + C$ where C is an arbitrary constant. This is easily verified by substitution:

$$\begin{aligned} \text{LHS} &= y_{k+1} - y_k \\ &= [(k+1) + C] - [k + C] \\ &= 1 \\ &= \text{RHS}\,. \end{aligned}$$

4) The d.e. $y_{k+2} - 3y_{k+1} + 2y_k = 0$ has the general solution $y_k = C_1 + C_2 \cdot 2^k$, where C_1 and C_2 are independent arbitrary constants. We verify this by substituting the solution into both sides of the d.e.:

$$\begin{aligned} y_{k+2} - 3y_{k+1} + 2y_k &= C_1 + C_2 \cdot 2^{k+2} - 3[C_1 + C_2 \cdot 2^{k+1}] + 2[C_1 + C_2 \cdot 2^k] \\ &= C_1[1 - 3 + 2] + 2^k C_2[2^2 - 3 \cdot 2 + 2] \\ &= 0. \end{aligned}$$

Since the RHS $= 0$ for all $k \geq 0$, the d.e. is satisfied.

Linear Difference Equations

The general form for an nth order linear difference equation in the sequence $y = (y_0, y_1, y_2, \ldots)$ is

$$f_0(k)y_{k+n} + f_1(k)y_{k+n-1} + \cdots + f_n(k)y_k = g(k) \quad k \geq 0. \tag{76}$$

This d.e. is linear because all terms are linear in the sequence values y_i: There are no products of sequence values, e.g. $y_i y_j$, nor are there any nonlinear functions of these values, e.g. $\sqrt{y_k}$. It is also assumed that $f_0(k) \neq 0$ and $f_n(k) \neq 0$, $k \geq 0$. In this way, y_{k+n} is related to y_k for all k, in order to preserve the "nth order" nature of the d.e..

Note the similarity between Eq. (76) and the general form of an nth order differential equation in the function $y(x)$:

$$a_0(x)y^{(n)} + a_1(x)y^{(n-1)} + \cdots + a_n(x)y = b(x), \tag{77}$$

where $y^{(k)} = \dfrac{d^k y}{dx^k}$. In (77), y is a function of the continuous variable $x \in \mathbb{R}$. In (76), y is a function of the discrete integer variable $k \geq 0$.

Example: We return to the d.e.

$$y_{k+2} - 3y_{k+1} + 2y_k = 0, \tag{78}$$

with general solution

$$y_k = C_1 + C_2 \cdot 2^k. \tag{79}$$

(How we obtain this "general solution" will be the subject of later sections.) If this expression for y_k represents the general solution, then we should be able to generate all possible solutions to the d.e. (2.15) from it. As we discuss in more detail below, for any pair of real numbers A, B, we should be able to find particular values of the constants C_1 and C_2 so that the sequence y_k assumes the initial values $y_0 = A$ and $y_1 = B$. By setting $k = 0$ and $k = 1$, respectively, we have the conditions

$$C_1 + C_2 = A$$
$$C_1 + 2C_2 = B.$$

The unique solution to this linear system of equations is $C_1 = 2A - B$ and $C_2 = B - A$. Therefore, the solution to the above d.e. with initial values $y_0 = A$ and $y_1 = B$ is given by

$$y_k = 2A - B + (B - A)2^k, \quad k \geq 0. \tag{80}$$

At this point you may well ask, "Why go through all the trouble of finding a general solution with arbitrary constants C_1 and C_2 whose particular values may then have to be determined? Why can't we just start the sequence with our initial values $y_0 = A$ and $y_1 = B$ and use recursion, as we did with first order d.e.'s in the previous chapter, to obtain the sequence elements y_k for $k \geq 2$ in terms of A and B? The answer is that the general form for the y_k as functions of k may not be so easy to determine.

Suppose we rewrite the above d.e. as a recursion relation:

$$y_{k+2} = 3y_{k+1} - 2y_k, \qquad k \geq 0. \tag{81}$$

Then

$$\begin{aligned} y_2 &= 3B - 2A, \\ y_3 &= 3y_2 - 2y_1 \\ &= 7B - 6A, \\ &\vdots \end{aligned}$$

It would probably be quite difficult to come up with the general closed form expression for the y_k in (80) from this procedure.

As we shall also see later, the general form (80) immediately reveals some qualitative properties of the solution $\{y_k\}$ that may not be obvious from a look at the first few, or even many, iterates. For example, from (80), if $A \neq B$, then the magnitudes $|y_k|$ will grow exponentially, essentially doubling with each increment in k, due to the factor 2^k. ("That's really how I would like to see my stock option grow!" you say.)

The idea of recursive computation of the solution elements y_k to the nth order linear d.e. in Eq. (76) is, however, important from another perspective. The recursion scheme associated with (76) is

$$y_{k+n} = \frac{1}{f_0(k)}\left[g(k) - f_1(k)y_{k+n-1} - f_2(k)y_{k+n-2} - \cdots - f_n(k)y_k\right], \quad k \geq 0. \tag{82}$$

Recall that we are assuming that $f_0(k) \neq 0$ and $f_n(k) \neq 0$. Given n "starting values" of the sequence $y_0 = A_0$, $y_1 = A_1, \ldots y_{n-1} = A_{n-1}$ we may compute $y_n(A_0, \ldots, A_{n-1})$, with $k = 0$ in (82). Setting $k = 1$ in (82), we use $y_1, \ldots, y_n$ to compute y_{n+1}, etc.. In this way, it appears that a unique solution to Eq. (76) may be constructed satisfying the n initial conditions $A_0, \ldots, A_{n-1}$.

The recursion procedure provides the basis for the proof of the following important theorem on existence and uniqueness of solutions to linear nth order difference equations.

Theorem 1: (Existence-uniqueness) Given the nth order linear d.e.,

$$f_0(k)y_{k+n} + f_1(k)y_{k+n-1} + \cdots + f_n(k)y_k = g(k) \quad k \geq 0\,, \tag{83}$$

with $f_0(k) \neq 0$, $f_n(k) \neq 0$, $k \geq 0$, then for any n-tuple $A = (A_0, A_1, \ldots, A_{n-1})$, there exists a unique solution $y = (y_0, y_1, y_2, \ldots)$ to (76) with initial values $y_k = A_k$, $k = 0, 1, \ldots, n-1$.

Note: A more general existence-uniqueness result can be proved, involving an extension of the proof to the above theorem, for which the solution $y = (y_0, y_1, \ldots)$ assumes n consecutive prescribed values, but not necessarily at the beginning of the sequence, i.e.

$$y_l = A_0, \quad y_{l+1} = A_1, \ldots \quad y_{l+n-1} = A_{n-1},$$

for an $l \geq 0$. The extension involves a backward computation of unique elements $y_{l-1}, y_{l-1}, \ldots y_0$.

Before going on with some more theory, let us address a concern that you may be having at this point: "Why study higher order equations?" As in the case of DEs, higher order d.e.'s have applications in science and engineering. Numerical schemes to solve second or higher order ordinary or partial differential equations necessarily involve second or higher order differences of function values on the discrete lattice points used in the calculation.

An example from biology involves the modelling of populations of annual plants. In a given year, $k > 0$, there may be S_k seeds available for germination into plants. However, these seeds may be several years old. Thus S_k is a combination of seeds produced in year $k-1$ that survived one winter; seeds produced in year $k-2$ that survived a winter, did not germinate, survived another winter; ... year $k-n$ .... There is a fair amount of book-keeping to be done to eventually arrive at an nth order linear d.e. of the form

$$P_k = \sum_{l=1}^{n} a_l P_{k-l}.$$

We shall return to this model in a later section.

Associated with the nth order linear d.e. in Eq. (76) is the **homogeneous** d.e.

$$f_0(k)y_{k+n} + f_1(k)y_{k+n-1} + \cdots + f_n(k)y_k = 0, \quad k \geq 0. \tag{84}$$

The sequence $y_k = 0$, $k \geq 0$ is always a solution to Eq. (84). It is known as the **trivial solution**.

Theorem 2: $y^{(1)} = (y_0^{(1)}, y_1^{(1)}, y_2^{(1)}, \ldots)$ and $y^{(2)} = (y_0^{(2)}, y_1^{(2)}, y_2^{(2)}, \ldots)$ denote two solutions to Eq. (84). Then the sequence

$$\begin{aligned} y &= C_1 y^{(1)} + C_2 y^{(2)} \\ &= (C_1 y_0^{(1)} + C_2 y_0^{(1)}, C_1 y_1^{(1)} + C_2 y_1^{(1)}, \ldots), \end{aligned} \tag{85}$$

where C_1 and C_2 are arbitrary constants, is also a solution.

Comment: Before we prove this result, we should perhaps address a possible concern: How do we know that we can obtain two possibly different solutions to (84)? Answer: By the Existence-Uniqueness Theorem 1, which also applies to homogeneous d.e.'s in which $g(k) = 0$, $k \geq 0$. We may obtain two different solutions by changing the initial conditions.

Proof: By substitution of y in (85) into (84):

$$\begin{aligned}
\text{L.H.S.} &= f_0(k)y_{k+n} + \cdots + f_n(k)y_k \\
&= f_0(k)[C_1 y^{(1)}_{k+n} + C_2 y^{(2)}_{k+n}] + \cdots + f_n(k)[C_1 y^{(1)}_k + C_2 y^{(2)}_k] \\
&= C_1[f_0(k)y^{(1)}_{k+n} + \cdots + f_n(k)y^{(1)}_k] + C_2[f_0(k)y^{(2)}_{k+n} + \cdots + f_n(k)y^{(1)}_k] \\
&= C_1 \cdot 0 + C_2 \cdot 0 \\
&= 0.
\end{aligned}$$

■

Note the importance of the linearity of the homogeneous d.e. in deriving this result. The following result is an important first step toward the construction of "complete" or general solutions to linear nth order d.e.'s of the form (76).

Theorem 3: Let $y^{(p)} = \{y^{(p)}_k\}$ be a solution to Eq. (76) and $y^{(h)} = \{y^{(h)}_k\}$ be a solution to its associated homogeneous d.e. in Eq. (84). Then

$$y = y^{(p)} + C_1 y^{(h)} \tag{86}$$

is also a solution to (76), where C_1 is an arbitrary constant.

Proof: Substitution (Exercise). ■

Of course, there is nothing to stop us from adding other solutions of the homogeneous linear d.e. in (84) to produce a sequence that satisfies the inhomogeneous d.e. in (76). The question is, "How many should we really add?" The answer is that there always exists a set of n "linearly independent" solutions to (84) – call them $y^{(h,1)}, y^{(h,2)}, \dots y^{(h,n)}$ – from which we can construct *all* solutions to (84). What we mean by linear independence will be discussed in the next section.

The **general solution** to the inhomogeneous nth order linear d.e. can then be written as

$$y = y^{(p)} + C_1 y^{(h,1)} + C_2 y^{(h,2)} + \cdots + C_n y^{(h,n)}, \tag{87}$$

where $C_1, \dots, C_N$ are arbitrary constants. Note that By "general solution", we mean that **any** solution to (76) can be obtained from Eq. (87) by means of a particular set of values of the constants $C_1, \dots, C_N$. More on this later.

Lecture 5

Linear difference equations (cont'd)

Fundamental sets of solutions to linear difference equations

We now establish the conditions for which the expression in Eq. (87) is a general solution to the nth order linear d.e. in Eq. (76), which we reproduce here,

$$f_0(k)y_{k+n} + f_1(k)y_{k+n-1} + \cdots + f_n(k)y_k = g(k) \quad k \geq 0. \tag{88}$$

Note: It is important to point out that in Eq. (88) we assume that the coefficients $f_0(k)$ and $f_n(k)$ are nonzero. (All of the other coefficients could be zero.) In this way, as was mentioned previously, we are preserving the "nth order" nature of the d.e.. We are assuming that not all $g(k)$ are zero – otherwise, the above d.e., would be homogeneous, in which case we would only have to be concerned about finding solutions to homogeneous equations, a subset of the discussion which follows.

For simplicity of notation, it is convenient to rewrite this equation in the following "normalized" form:

$$y_{k+n} + a_1(k)y_{k+n-1} + \cdots + a_n(k)y_k = r(k), \quad k \geq 0, \tag{89}$$

where $a_i(k) = f_i(k)/f_0(k)$, $1 \leq i \leq n$, and $r(k) = g(k)/f_0(k)$, $k \geq 0$. Recall that division is possible due to the assumption that $f_0(k) \neq 0$, $k \geq 0$. Also recall that $a_n(k) = f_n(k)/f_0(k) \neq 0$ by assumption.

It will be necessary to focus on the homogeneous d.e. associated with (89):

$$y_{k+n} + a_1(k)y_{k+n-1} + \cdots + a_n(k)y_k = 0, \quad k \geq 0. \tag{90}$$

For simplicity, our discussion will be restricted to the study of second order d.e.'s, i.e. $n = 2$ in Eqs. (89) and (90), which we write out explicitly below:

$$y_{k+2} + a_1(k)y_{k+1} + a_2(k)y_k = r(k) \tag{91}$$

and

$$y_{k+2} + a_1(k)y_{k+1} + a_2(k)y_k = 0. \tag{92}$$

The results developed below can then be extended to higher order d.e.'s in a straightforward way. The following is an important first step:

Theorem 4: Let $y^{(1)} = (y_0^{(1)}, y_1^{(1)}, \ldots)$ and $y^{(2)} = (y_0^{(2)}, y_1^{(2)}, \ldots)$ be two solutions to the homogeneous d.e. in Eq. (92) with the property that the determinant

$$D = \begin{vmatrix} y_0^{(1)} & y_0^{(2)} \\ y_1^{(1)} & y_1^{(2)} \end{vmatrix} \neq 0. \tag{93}$$

Then the **general solution** to (92) may be written as

$$Y = C_1 y^{(1)} + C_2 y^{(2)}, \tag{94}$$

i.e. $Y = \{Y_k\}$, where $Y_k = C_1 y_k^{(1)} + C_2 y_k^{(2)}$. **Note that by "general solution" we mean that any solution to** (92) **may be written in the form** (94) **for appropriate values of** C_1 **and** C_2**.**

Proof: Let y be a solution to Eq. (92) with initial conditions $y_0 = A_0$, $y_1 = A_1$, where A_0 and A_1 are arbitrary real numbers. Recall that from Theorem 1 (last lecture) on existence/uniqueness for linear nth order d.e.'s (here $n = 2$), only one solution with these initial conditions exists. We now determine whether or not it will be possible to find values of C_1 and C_2 in (94) such that $Y = y$, i.e., that we can generate that unique solution from our general solution. It is sufficient that we construct $Y_0 = A_0$ and $Y_1 = A_1$. From existence/uniqueness, this will imply that $Y_k = y_k$ for all $k \geq 0$. Thus:

$$\begin{aligned} Y_0 = A_0 \quad &\text{implies} \quad C_1 y_0^{(1)} + C_2 y_0^{(2)} = A_0, \\ Y_1 = A_1 \quad &\text{implies} \quad C_1 y_1^{(1)} + C_2 y_1^{(2)} = A_1. \end{aligned}$$

This is a linear system of equations in the unknowns C_1 and C_2. The existence of a unique solution is guaranteed since the determinant D of the system is assumed to be nonzero by (93). **Therefore a unique set of values** (C_1, C_2) **exists for any set of initial conditions** (A_0, A_1)**.** The proof is complete. ■

A set of solutions $y^{(1)}, y^{(2)}$ to the second order homogeneous d.e. in Eq. (92) for which the determinant in (93) is nonzero is known as a **fundamental set** of solutions. This property is essentially the linear independence referred to in the previous section.

It is indeed intriguing that such linear independence of two solutions $y^{(1)}$ and $y^{(2)}$ can be determined on the basis of their first sets of coefficients $\{y_0^{(1)}, y_1^{(1)}\}$ and $\{y_0^{(2)}, y_0^{(2)}\}$ as dictated by the determinant in (93). (We'll comment on this property a little later.) However, as we show below, this linear independence is actually propagated throughout the sequences.

Theorem 5: Let $y^{(1)} = \{y_k^{(1)}\}$ and $y^{(2)} = \{y_k^{(2)}\}$ be two solutions to the homogeneous d.e. in Eq. (92). Define

$$W_k = \begin{vmatrix} y_k^{(1)} & y_k^{(2)} \\ y_{k+1}^{(1)} & y_{k+1}^{(2)} \end{vmatrix}, \qquad k \geq 0. \tag{95}$$

Then

$$W_{k+1} = a_2(k) W_k, \qquad k \geq 0. \tag{96}$$

Proof: Since $y^{(1)}$ and $y^{(2)}$ are solutions to (92) we have

$$y_{k+2}^{(1)} + a_1(k) y_{k+1}^{(1)} + a_2(k) y_k^{(1)} = 0, \tag{a}$$
$$y_{k+2}^{(2)} + a_1(k) y_{k+1}^{(2)} + a_2(k) y_k^{(2)} = 0. \tag{b}$$

Now multiply (a) by $y_{k+1}^{(2)}$ and (b) by $y_{k+1}^{(1)}$ and subtract:

$$y_{k+2}^{(1)}y_{k+1}^{(2)} - y_{k+2}^{(2)}y_{k+1}^{(1)} + a_2(k)\left[y_{k+1}^{(2)}y_k^{(1)} - y_{k+1}^{(1)}y_k^{(2)}\right] = 0 \tag{97}$$

or

$$y_{k+1}^{(1)}y_{k+2}^{(2)} - y_{k+2}^{(1)}y_{k+1}^{(2)} = a_2(k)\left[y_k^{(1)}y_{k+1}^{(2)} - y_{k+1}^{(1)}y_k^{(2)}\right]. \tag{98}$$

But from (93), this equation is simply

$$W_{k+1} = a_2(k)W_k, \tag{99}$$

as desired. ∎

Recall the condition that $f_2(k) \neq 0$, implying that $a_2(k) \neq 0$, for all k, for linear second order d.e.'s, in order to preserve a connection between y_k and y_{k+2} for all $k \geq 0$. Also note that $W_0 = D$ in (93). From Eq. (96),

$$W_k = a_2(0)a_2(1)a_2(2)\ldots a_2(k-1)W_0, \qquad k \geq 1. \tag{100}$$

This implies the following important result:

> If $y^{(1)}$ and $y^{(2)}$ are solutions to Eq. (92), then either (i) $W_k \neq 0$ for all k, whereupon $y^{(1)}$ and $y^{(2)}$ form a fundamental set, or (ii) $W_k = 0$ for all k.

This result implies that we can check pairs of solutions for linear independence by applying the determinant test, i.e. by examining W_k for appropriate pairs of consecutive elements of each sequence, at any $k \geq 0$:

$$\begin{array}{ll} y^{(1)}: & y_0^{(1)}, y_1^{(1)}, y_2^{(1)}, \ldots \quad y_k^{(1)}, y_{k+1}^{(1)} \cdots \\ y^{(2)}: & y_0^{(2)}, y_1^{(2)}, y_2^{(2)}, \ldots \quad y_k^{(2)}, y_{k+1}^{(2)} \cdots \\ & \quad W_0 \quad\; W_1 \qquad\quad W_k \end{array}$$

The reader familiar with DEs will note that the above results are discrete analogues of the result for the **Wronskian** between two solutions of a homogeneous DE (Abel's Theorem).

Example: Given the linear homogeneous d.e.

$$y_{k+2} - 5y_{k+1} + 6y_k = 0, \tag{101}$$

two sets of solutions are $y_k^{(1)} = 2^k$, $y_k^{(2)} = 3^k$. Let's check this by substitution:

- $y_k = 2^k$:

$$\begin{aligned} y_{k+2} - 5y_{k+1} + 6y_k &= 2^{k+2} - 5 \cdot 2^{k+1} + 6 \cdot 2^k \\ &= 2^k[4 - 10 + 6] \\ &= 0. \end{aligned} \tag{102}$$

- $y_k = 3^k$:

$$\begin{aligned} y_{k+2} - 5y_{k+1} + 6y_k &= 3^{k+2} - 5 \cdot 3^{k+1} + 6 \cdot 3^k \\ &= 3^k[9 - 15 + 6] \\ &= 0\,. \end{aligned} \tag{103}$$

In the next section, we shall show how these solutions may be obtained. The determinant $D = W_0$ in (91) is

$$W_0 = \begin{vmatrix} y_0^{(1)} & y_0^{(2)} \\ y_1^{(1)} & y_1^{(2)} \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ 2 & 3 \end{vmatrix} = 1.$$

Therefore $y^{(1)}$ and $y^{(2)}$ form a fundamental set of solutions. Furthermore,

$$W_k = \begin{vmatrix} y_k^{(1)} & y_k^{(2)} \\ y_{k+1}^{(1)} & y_{k+1}^{(2)} \end{vmatrix} = \begin{vmatrix} 2^k & 3^k \\ 2^{k+1} & 3^{k+1} \end{vmatrix} = 2^k \cdot 3^{k+1} - 2^{k+1} \cdot 3^k = 2^k 3^k = 6^k.$$

This is in agreement with Eq. (100), since $W_0 = 1$ and $a_2(i) = 6$, $i \geq 0$.

We may now address the problem of complete or general solutions to linear second order difference equations having the form of Eq. (91). The following result essentially completes the picture.

Theorem 6: Assume the following:

- $y^{(p)}$ is a particular solution of the inhomogeneous d.e.,

$$y_{k+2} + a_1(k)y_{k+1} + a_2(k)y_k = r(k)\,. \tag{104}$$

- $y^{(1)}$ and $y^{(2)}$ form a fundamental set of solutions of the associated homogeneous d.e.,

$$y_{k+2} + a_1(k)y_{k+1} + a_2(k)y_k = 0\,. \tag{105}$$

Then the general solution of the inhomogeneous d.e. in (104)is given by

$$y = y^{(p)} + C_1 y^{(1)} + C_2 y^{(2)}, \tag{106}$$

where C_1 and C_2 are constants.

Proof: The fact that y is a solution of (91) follows from Theorem 3. Now define $Y = y - y^{(p)}$. By substitution, it is easy to show that Y must be a solution to the homogeneous equation (105). Now suppose that the initial values of y are $y_0 = A_0$, $y_1 = A_1$, where A_0 and A_1 arbitrary. This implies that the initial values of Y are

$$Y_0 = A_0 - y_0^{(p)}, \quad Y_1 = A_1 - y_1^{(p)}.$$

It now remains to show that

$$Y = y - y^{(p)} = C_1 y^{(1)} + C_2 y^{(2)}$$

for a unique set of values C_1 and C_2. It is necessary to match the first two sequence values of both sides:

$$\begin{aligned} k = 0: &\quad C_1 y_0^{(1)} + C_2 y_0^{(2)} = A_0 - y_0^{(p)} \\ k = 1: &\quad C_1 y_1^{(1)} + C_2 y_1^{(2)} = A_1 - y_1^{(p)}. \end{aligned}$$

The hypothesis that $y^{(1)}$ and $y^{(2)}$ form a fundamental set of solutions to (105) ensures the existence of C_1 and C_2. ■

Example: The linear second order d.e.

$$y_{k+2} - 5y_{k+1} + 6y_k = 10. \tag{107}$$

From the previous example, we know that $y_k^{(1)} = 2^k$ and $y_k^{(2)} = 3^k$ form a fundamental set of solutions to the homogeneous d.e. (101) associated with Eq. (107). The particular solution for Eq. (107) is $y_k = 5$. Check by substitution:

$$\begin{aligned} y_{k+2} - 5y_{k+1} + 6y_k &= 5 - 5 \cdot 5 + 6 \cdot 5 \\ &= 5[1 - 5 + 6] \\ &= 10\,. \end{aligned} \tag{108}$$

(How we obtain this particular solution will be discussed later.) From Theorem 6, the general solution to Eq. (107) is

$$y_k = 5 + C_1 \cdot 2^k + C_2 3^k.$$

For given initial conditions $y_0 = A_0$ and $y_1 = A_1$, we may solve for C_1 and C_2:

$$\begin{aligned} C_1 + C_2 &= A_0 - 5 \\ 2C_1 + 3C_2 &= A_1 - 5. \end{aligned}$$

yielding $C_1 = 3A_0 - A_1 - 10$, $C_2 = A_1 - 2A_0 + 5$. Thus,

$$y_k = 5 + (3A_0 - A_1 - 10) \cdot 2^k + (A_1 - 2A_0 + 5) \cdot 3^k$$

is the solution with initial conditions A_0 and A_1.

Let's check this result – always a good thing to do:

- $k = 0$:

$$\begin{aligned} y_0 &= 5 + (3A_0 - A_1 - 10) + (A_1 - 2A_0 + 5) \\ &= A_0\,. \end{aligned} \tag{109}$$

- $k = 1$:

$$\begin{aligned} y_1 &= 5 + 2(3A_0 - A_1 - 10) + 3(A_1 - 2A_0 + 5) \\ &= A_1\,. \end{aligned} \tag{110}$$

Good.

We now have the important results to characterize general solutions of second order linear difference equations. In the next section, we consider an important family of d.e.'s for which solutions can always be found – the case of d.e.'s with constant coefficients. This family of d.e.'s has many applications in science and economics.

A final comment on general solutions: We have said that a general solution to a linear, second order difference equation is able to generate any solution y_k with prescribed initial conditions $y_0 = A_0$ and $y_1 = A_1$, where $A_1, A_2 \in \mathbb{R}$. We also mentioned repeated the "analogy" between difference equations and differential equations. But in the case of differential equations, the usual initial conditions imposed on a solution $y(x)$ at, say, $x = 0$, are $y(0) = y_0$ and $y'(0) = w_0$, where $y_0, w_0 \in \mathbb{R}$ – in other words, the values of the solution $y(0)$ and its derivative $y'(0)$ at $x = 0$. If difference equations are an "analogy", then shouldn't we be imposing the initial value $y_0 = A_0$ and initial difference $\Delta y_0 = y_1 - y_0 = D_0$ where $A_0, D_0 \in \mathbb{R}$? The answer is **Yes, indeed, we could do this.** The two sets of initial value conditions are, in fact, equivalent. The imposition of the initial value $y_0 = A_0$ is, of course, the same in both schemes. The imposition of the initial difference $\Delta y_0 = D_0$ is the same as imposing the initial condition,

$$y_1 = y_0 + (y_1 - y_0) = y_0 + D_0 = A_1\,, \tag{111}$$

Specifying D_0 on the first difference implies specifying $y_1 = y_0 + D_0$. The fact that A_0 and D_0 can be arbitrarily chosen implies that A_0 and A_1 can be arbitrarily chosen. So the two schemes are equivalent. It turns out that the scheme that we are using which, in fact, is the "standard" scheme, is more convenient to work with.

Problems

1. Show that the functions $y^{(1)}$ and $y^{(2)}$ given by

$$y_k^{(1)} = 1 \quad \text{and} \quad y_k^{(2)} = (-1)^k\,, \quad k \geq 0\,,$$

are solutions of the difference equation

$$y_{k+2} - y_k = 0$$

and that they form a fundamental set. Find the general solution of the difference equation and a particular solution for which $y_0 = 0$ and $y_1 = 2$.

2. Find the particular solution of the difference equation in Problem 1 which satisfies the initial conditions $y_1 = 2$ and $y_2 = 0$. Note that this is the same solution as that obtained in Problem 1. Explain.

3. In each of the following parts, a difference equation and three functions, $y^{(1)}, y^{(2)}$ and $y^{(p)}$, are given. Show in each case that (a) $y^{(1)}$ and $y^{(2)}$ are solutions of the corresponding homogeneous equation; (b) $y^{(1)}$ and $y^{(2)}$ form a fundamental set of solutions; (c) $y^{(p)}$ is a solution of the complete equation. Then (d) use Theorem 2.6 to write the general solution of the complete equation, and (e) find the particular solution satisfying the initial conditions $y_0 = 0$ and $y_1 = 1$.

(i)	$y_{k+2} - 7y_{k+1} + 12y_k = 2$	$y_k^{(1)} = 3^k$	$y_k^{(2)} = 4^k$	$y_k^{(p)} = \frac{1}{3}$
(ii)	$y_{k+2} - 4y_k = 9$	$y_k^{(1)} = 2^k$	$y_k^{(2)} = (-2)^k$	$y_k^{(p)} = -3$
(iii)	$2y_{k+2} + 3y_{k+1} - 2y_k = 3k + 1$	$y_k^{(1)} = \left(\frac{1}{2}\right)^k$	$y_k^{(2)} = (-2)^k$	$y_k^{(p)} = k - 2$
(iv)	$y_{k+2} - 4y_{k+1} + 4y_k = 1$	$y_k^{(1)} = 2^k$	$y_k^{(2)} = k \cdot 2^k$	$y_k^{(p)} = 1$
(v)	$8y_{k+2} - 6y_{k+1} + +y_k = 2^k$	$y_k^{(1)} = \left(\frac{1}{2}\right)^k$	$y_k^{(2)} = \left(\frac{1}{4}\right)^k$	$y_k^{(p)} = \frac{1}{21}2^k$

Linear Difference Equations with Constant Coefficients

In this important family of d.e.'s, the coefficients $a_1(k) \ldots a_n(k)$ in Eq. (89) are constants, i.e. they do not depend upon k. The inhomogeneous terms on the RHS, however, are allowed to vary with k. Eq. (89) then has the form

$$y_{k+n} + a_1 y_{n+n-1} + \cdots + a_n y_k = r_k, \quad k \geq 0, \tag{112}$$

(here we write $r(k)$ as r_k) and its associated homogeneous d.e. is

$$y_{k+n} + a_1 y_{k+n-1} + \cdots + a_n y_k = 0. \tag{113}$$

All of the theorems of the previous sections apply to these equations: Existence-uniqueness, general solutions, fundamental sets of solutions. Because of the importance of constructing the latter, we first focus on homogeneous d.e.'s, beginning with the quite simple case of first order d.e.'s.

Homogeneous First Order Difference Equations with Constant Coefficients

We write Eq. (113), with $n = 1$, as follows

$$y_{k+1} + a y_k = 0, \quad k \geq 0, \tag{114}$$

dropping the subscript from a_1 for convenience. Since $n = 1$, we need only specify y_0 in order to uniquely determine the entire solution. The d.e. can be rewritten as follows,

$$y_{k+1} = -a y_k \,, \quad k \geq 0 \,, \tag{115}$$

which is of the form

$$y_{k+1} = c y_k \,, \tag{116}$$

studied in Lecture 2. The solution of (115) is therefore

$$y_k = (-a)^k y_0 \,. \tag{117}$$

By allowing y_0 to be an arbitrary constant C, the general solution of (114) is therefore

$$y_k = C(-a)^k. \tag{118}$$

Since $n = 1$, this solution comprises the fundamental "set" of solutions.

Homogeneous Second Order Difference Equations with Constant Coefficients

We shall write Eq. (113), with $n = 2$, in the form

$$y_{k+2} + p y_{k+1} + q y_k = 0, \quad k \geq 0, \tag{119}$$

where $p = a_1$, and $q = a_2 \neq 0$ are constants. Motivated by Eq. (118), let us assume a solution of the form $y_k = m^k$ where m is to be determined. (Note: We do not have to include an additional constant "C" so that $y_k = Cm^k$. Why?) Substitution into (119) yields

$$m^{k+2} + pm^{k+1} + qm^k = 0 \tag{120}$$

or

$$m^k[m^2 + pm + q] = 0\,. \tag{121}$$

Eq. (121) is obviously satisfied by $m = 0$. This leads to the trivial solution $y_k = 0$ which is always a solution to homogeneous d.e.'s. The more interesting values of m are the non-zero roots m_1 and m_2 of the quadratic equation

$$\boxed{m^2 + pm + q = 0,} \tag{122}$$

which is called the **characteristic equation** associated with the d.e. in (119). We denote the roots as follows

$$m_{1,2} = -\tfrac{1}{2}p \pm \tfrac{1}{2}\sqrt{p^2 - 4q}. \tag{123}$$

(Note that the $m_i \neq 0$ since $q \neq 0$ by assumption.) There are three cases to consider.

Case 1: Two distinct real roots $m_1 \neq m_2$

This is the case when $p^2 - 4q > 0$. The two solutions, $y^{(1)}$ and $y^{(2)}$, corresponding to these roots are, respectively, $y_k^{(1)} = m_1^k$ and $y_k^{(2)} = m_2^k$. These two solutions form a fundamental set of solutions to (119) since

$$D = W_0 = \begin{vmatrix} y_0^{(1)} & y_0^{(2)} \\ y_1^{(2)} & y_1^{(2)} \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ m_1 & m_2 \end{vmatrix} = m_2 - m_1 \neq 0. \tag{124}$$

Thus, the general solution Y of (119) is given by

$$Y_k = C_1 m_1^k + C_2 m_2^k, \quad k \geq 0. \tag{125}$$

Example: The d.e.

$$y_{k+2} - 3y_{k+1} + 2y_k = 0\,,$$

examined in the previous lecture, has the characteristic equation

$$m^2 - 3m + 2 = 0,$$

with roots $m_1 = 2$, $m_2 = 1$. The corresponding solutions are $y_k^{(1)} = 2^k$, $y_k^{(2)} = 1^k = 1$, so the general solution is

$$Y_k = C_1 2^k + C_2.$$

If we impose the initial conditions $y_0 = 1$, $y_1 = 2$, then

$$\begin{aligned} k = 0: &\qquad C_1 + C_2 = 1 \\ k = 1: &\qquad 2C_1 + C_2 = 2 \end{aligned}$$

with solution $C_1 = 1$, $C_2 = 0$. The particular solution with $y_0 = 1$ and $y_1 = 2$ is then

$$y_k = 2^k.$$

For the sake of safety, let's verify that this sequence satisfies the d.e.. Substitution into the LHS yields,

$$\begin{aligned} y_{k+2} - 3y_{k+1} + 2y_k &= 2^{k+2} - 3 \cdot 2^{k+1} + 2 \cdot 2^k \\ &= 2^k(4 - 6 + 2) \\ &= 0\,, \end{aligned}$$

which verifies that it is a solution – in fact, the unique solution to the given initial value problem.

Case 2: Equal real roots $m_1 = m_2$

This is the case when $p^2 - 4q = 0$, so that $m_1 = m_2 = m = -\frac{1}{2}p$. One solution to the d.e. is $y_k^{(1)} = m^k$. What about another, linearly independent one? It turns out that $y_k^{(2)} = km^k$ is another solution. We verify this by substitution into the d.e. (119):

$$\begin{aligned} y_{k+2} + py_{k+1} + qy_{k+2} &= (k+2)m^{k+2} + p(k+1)m^{k+1} + qkm^k \\ &= m^k\big[k(m^2 + pm + q) + m(2m + p)\big]. \end{aligned} \tag{126}$$

Since m is a root of the characteristic equation, the first term in the round brackets vanishes. The second term vanishes since $m = -\frac{1}{2}p$. Therefore the d.e. is satisfied.

The solutions $y_k^{(1)} = m^k$, $y_k^{(2)} = km^k$, form a fundamental set since

$$D = W_0 = \begin{vmatrix} y_0^{(1)} & y_0^{(2)} \\ y_1^{(1)} & y_1^{(2)} \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ m & m \end{vmatrix} = m \neq 0. \tag{127}$$

The reader may note that this case is reminiscent of the case of double roots for homogeneous differential equations with constant coefficients. In this case, the roots of the characteristic polynomial are $r_1 = r_2 = r$ so only one solution $y_1(t) = e^{rt}$ is obtained. A second, linearly independent solution is obtained by simply multiplying this solution by t, i.e., $y_2(t) = te^{rt}$.

Example: The d.e.

$$y_{k+2} - y_{k+1} + \tfrac{1}{4}y_k = 0$$

has as characteristic equation

$$m^2 - m + \tfrac{1}{4} = \left(m - \tfrac{1}{2}\right)^2 = 0.$$

Thus $m_1 = m_2 = \frac{1}{2}$. One solution is $y_k^{(1)} = \left(\frac{1}{2}\right)^k$. The second is $y_k^{(2)} = k\left(\frac{1}{2}\right)^k$. The general solution is

$$Y_k = C_1\left(\tfrac{1}{2}\right)^k + C_2 k\left(\tfrac{1}{2}\right)^k.$$

If we impose the initial conditions $Y_0 = 1$, $Y_1 = 1$, then

$$\begin{aligned} k = 0: &\qquad C_1 + 0 \cdot C_2 = 1 \\ k = 1: &\qquad \tfrac{1}{2}C_1 + \tfrac{1}{2}C_2 = 1 \end{aligned}$$

with solutions $C_1 = 1$, $C_2 = 1$. The particular solution with $y_0 = 1$, $y_1 = 1$ is therefore

$$y_k = \left(\tfrac{1}{2}\right)^k + k\left(\tfrac{1}{2}\right)^k = \tfrac{1+k}{2^k}, \quad k \geq 0.$$

The first few elements of this sequence are

$$y = (1, 1, \tfrac{3}{4}, \tfrac{1}{2}, \tfrac{5}{16}, \tfrac{3}{16}, \cdots)$$

From the formula, we see that $y_k \to 0$ as $k \to \infty$.