Discrete (Difference) Equations

A discrete, or difference, equation expresses a relationship between the elements of a sequence, $\{y_n\}$, where $n \in \mathbb{N}_0 \equiv \{0, 1, 2, \ldots\}$. For example, the trivial difference equation

$$y_{n+1} = y_n, \tag{1}$$

has the solution $y_n = y_0$, which means that the sequence $\{y_n\}$ may be any constant sequence.

Many mathematical models are posed in the form of discrete equations. The (in)famous logistic map

$$y_{n+1} = r y_n (1 - y_n), \tag{2}$$

was introduced as a model for the growth of a population. Here, $y_n \in [0, 1]$ represents the (scaled) population in the n-th year, y_0 is the initial population; and $r > 0$ is a combined birth and death rate.

Discrete equations also arise when solving continuous models using numerical methods, a necessary task for all but the most simple models. Computers can only work with discrete data, so continuous equations must be discretised[1] before they can be solved numerically.

The continuous differential equation

$$\frac{\mathrm{d}y}{\mathrm{d}x} = 0; \tag{3}$$

has the simple solution $y = y(0)$, a constant function. A simple discretisation method is to use finite differences and approximate the derivative by the expression

$$\frac{\mathrm{d}y}{\mathrm{d}x}(x) \approx \frac{y(x+h) - y(x)}{h}, \quad \text{where } h \text{ is small.}$$

Now, we define the sequence $\{y_n\}$, such that $y_0 = y(0)$, $y_1 = y(h)$, $y_2 = y(h + h) = y(2h)$, ... $y_n = y(nh)$. Thus,

$$\frac{\mathrm{d}y}{\mathrm{d}x}(nh) \approx \frac{y(nh + h) - y(nh)}{h} = \frac{y_{n+1} - y_n}{h}.$$

The discrete approximation to the differential equation (3) is, therefore,

$$\frac{y_{n+1} - y_n}{h} = 0 \quad \Rightarrow \quad y_{n+1} - y_n = 0 \quad \Rightarrow \quad y_{n+1} = y_n,$$

our original example (1). In this simple example, the solutions to the discrete equation and continuous equation are exactly the same and independent of h, why?

In fact, there is a strong connection between difference equations and differential equations. Indeed a differential equation can always be approximated by a difference equation, can you see how?

[1]Discretisation is the general term used to describe the procedure of converting a continuous system into a set of discrete equations.

1 Linear difference equations

1.1 Equations with constant coefficients

A general m-th order difference equation with constant coefficients can be written in the form

$$a_m y_{n+m} + a_{m-1} y_{n+m-1} + \cdots + a_1 y_{n+1} + a_0 y_n = r_n, \tag{4}$$

where $\{a_k\}$, $k \in 0, \ldots, m$ are constants, and $r_n = r(n)$ is an arbitrary function. Note that the axis has been re-scaled so that the distance between points in the sequence $\{y_n\}$ is $h = 1$, the standard convention when working directly with difference equations.

The solution of equation (4) can be split into two separate procedures:

- Solution of the homogeneous equation ($r_n = 0$)
- Calculation of a particular solution
  (any function that satisfies the entire inhomogeneous equation)

The solution of the equation is the **sum** of the particular solution and the solution of the homogeneous equation, known as the complementary function. Note that this linear superposition (adding the two solutions together) is only possible because the difference equation is linear.

Homogeneous equations

1st order

A homogeneous difference equation is one in which the right-hand side is zero, *i. .e.* $r_n = 0$, in equation (4). Initially, the simplest case to consider is a first-order, linear, homogeneous difference equation with constant coefficients,

$$y_{n+1} + p y_n = 0, \tag{5}$$

where p is a constant.

We find the solution by an inductive procedure; rearranging equation (5) gives

$$y_n = -p y_{n-1}, \tag{6}$$

and because the equation (6) must be true for any n (> 1), then

$$y_{n-1} = -p y_{n-2}, \qquad \text{provided that } n > 2. \tag{7}$$

Combining equations (6) and (7) gives

$$y_n = (-p)^2 y_{n-2} = \cdots = (-p)^n y_0;$$

we have continued the inductive procedure until we reached y_0, which must be given as an initial condition Thus, the fundamental solution of the equation (5) is $y_0\,C^n$, where C is a constant that is the solution of the characteristic (or auxiliary) equation $C + p = 0$.

m-th order

Our result may be generalised to homogeneous difference equations of any order:

$$a_m y_{n+m} + a_{m-1} y_{n+m-1} + \cdots + a_1 y_{n+1} + a_0 y_n = 0. \tag{8}$$

The equation (8) can only be satisfied if y_{n+1} is equal to y_n multiplied by a constant, $y_{n+1} = C y_n$; in which case, each term in the equation is equal to any other term multiplied by a constant, *e. g.* $y_{n+2} = C y_{n+1} = C^2 y_n$, etc. Note that the constant will not necessarily be the same for each pair of terms. The condition is merely equation(5) in another guise, however, and we have already shown that the solution has the form $y_n = C^n$. On substitution of this result into equation (8), we obtain

$$a_m C^{n+m} + a_{m-1} C^{n+m-1} + \cdots + a_1 C^{n+1} + a_0 C^n = 0.$$

Dividing though by C^n,[2] we obtain the characteristic equation

$$a^m C^m + a_{m-1} C^{m-1} + \cdots + a_1 C + a_0 = 0.$$

The characteristic equation is an m-th order polynomial and will have m roots, each of which represents a different solution of the difference equation. The general solution may by found by a linear combination of these m solutions:

$$y_n = A_1 C_1^n + A_2 C_2^n + \cdots + A_m C_m^n,$$

where $\{A_k\}$ are constants and $\{C_k\}$ are the roots of the characteristic equation. Note that calculation of the constants $\{A_k\}$ requires m boundary conditions.

Aside

You will see in MT1222/MT1232 that the same general philosophy can be used to deduce the general solution of a linear homogeneous differential equation with constant coefficients.

$$a_m \frac{\mathrm{d}^m y}{\mathrm{d}x^m} + a_{m-1} \frac{\mathrm{d}^{m-1} y}{\mathrm{d}x^{m-1}} + \cdots + a_1 \frac{\mathrm{d}y}{\mathrm{d}x} + a_0 y = 0. \tag{9}$$

In this case the condition to be satisfied is that $\frac{\mathrm{d}y}{\mathrm{d}x} = Cy$, which implies that $y = \mathrm{e}^{Cx}$. Substitution of this form into equation (9) gives the result

$$a_m C^m \mathrm{e}^{Cx} + a_{m-1} C^{m-1} \mathrm{e}^{Cx} + \cdots + a_1 C \mathrm{e}^{Cx} + a_0 \mathrm{e}^{Cx} = 0.$$

We now divide through by e^{Cx} to obtain the same characteristic equation as in the case of the difference equation:

$$a^m C^m + a_{m-1} C^{m-1} + \cdots + a_1 C + a_0 = 0,$$

but the general solution of the differential equation is

$$y(x) = A_1 \mathrm{e}^{C_1 x} + A_2 \mathrm{e}^{C_2 x} + \cdots + A_m \mathrm{e}^{C_m x}.$$

[2]We note that the trivial sequence $y_n = 0$ is always a solution of a homogeneous equation, and we assume that $C \neq 0$, in order to find a non-trivial solution.

Examples

- Express the series $\{5, 10, 20, 40, \ldots\}$ as a difference equation and solve it to find an explicit expression for y_n.
  The difference equation is

$$y_{n+1} = 2y_n \quad \Rightarrow \quad y_{n+1} - 2y_n = 0.$$

  The characteristic equation is, therefore, $C - 2 = 0$, and so the general solution is

$$y_n = A_1 2^n$$

  The initial condition $y_0 = 5$, can be used to find $A_1 = 5$ and so

$$y_n = 5 \times 2^n.$$

Particular solutions

In order to calculate a particular solution, we use the "method of substitution", which is fancy way of saying that we determine the solution by working from a general (guessed) solution.

Example

- Find a particular solution of the difference equation

$$y_{n+1} + 2y_n = n. \tag{10}$$

  If the right-hand side is a polynomial in n, then we should try a general polynomial solution of the same order. In this case, we try a solution of the form $y_n = An + B$. On substitution into the equation we obtain

$$A(n+1) + B + 2An + 2B = n.$$

  We find the unknown values A and B by equating coefficients of powers of n:

$$n^1 : \quad 3A = 1, \qquad n^0 : A + 3B = 0,$$

  Solving these two simultaneous equations we find the solution

$$A = 1/3, \quad B = -1/9 \quad \Rightarrow \quad y_n = \frac{3n-1}{9}.$$

The complementary function of the equation (10) is $A_1(-2)^n$, because the characteristic equation is $C + 2 = 0$. The general solution of the difference equation (10) is, therefore

$$y_n = \frac{3n-1}{9} + A_1(-2)^n,$$

and the constant A_1 will depend on the initial condition.

2 Non-linear difference equations

In a linear difference equation, every term of the equation contains at most one of the elements of the sequence $\{y_n\}$ and the elements occur only "as themselves" *i.e.* they are not raised to any power (other than one) and are not an argument to a function. In a non-linear difference equation, all these restrictions are lifted.

Linear	Non-linear
$y_{n+1} = y_n,$	$y_{n+1} = y_n^2,$
$y_{n+2} = y_{n+1} + y_n$	$y_{n+2} = y_{n+1}\, y_n,$
$y_{n+2} = 4y_{n+1} + 3y_n$	$y_{n+2} = 4\mathrm{e}^{y_n},$

As you might expect, methods of solution for the two different types of equation are very different and the solutions themselves exhibit very different properties:

Linear	Non-linear
Standard methods for finding analytic solutions, see §1	No standard method for finding analytic solutions.
Unique solutions,	Non-uniqueness — multiple solutions
Linear Superposition — sum of solutions is also a solution.	Complex behaviour, bifurcations, chaos

There is no standard method for finding analytic solutions to non-linear difference equations. Indeed, it is an area, where numerical experiments are as important as theory in understanding the system. You may wish to try and write programs that calculate the sequences $\{y_n\}$ for non-linear difference equations.

2.1 Fixed points of difference equations

A simple technique that can be used to obtain a great deal of information about non-linear difference equations is to use a fixed-point analysis. The idea is to find particular points for which the solution is fixed (does not change), in other words, $y_{n+1} = y_n, \forall n$. The analysis is not restricted to non-linear difference equations; in fact, the second problem on Worksheet 1 included a fixed point analysis.

Example

- Find the fixed points of the difference equation

$$y_{n+1} = y_n^2. \tag{11}$$

At a fixed point $y_{n+1} = y_n$, so we are looking for values of y_n such that

$$y_n = y_n^2 \quad \Rightarrow \quad y_n^2 - y_n = 0, \quad \Rightarrow \quad y_n(y_n - 1) = 0,$$

and the equation has two fixed points at $y_n = 0$ and 1. That these solutions are fixed points is easy to verify on substitution into equation (11).[3]

The fixed points of a difference equation are rather special because if any sequence ever "visits" one of the fixed points, it will always remain there. We might expect, therefore, that any solution will eventually reach a fixed point of the system. Unfortunately, things aren't quite that simple. For example, if there is more than one fixed point, can we say which one the system will reach first? The answer **must** depend on where the sequence starts — the initial condition. If the sequence starts at a fixed point, we know that it will remain there, but what about when it starts nearby? How near is near? What do we mean by near anyway? In order to answer these types of questions, we can perform a **linear stability analysis** of the fixed points.

2.2 Linear stability analysis of fixed points

Definition of linear stability

The first step is to define what we mean by the term linear stability. The idea is remarkably simple and is best (always) explained through a simple analogy. Consider a rigid pendulum moving freely under the action of gravity. The pendulum is a mechanical system that consists of a rigid bar with one end fixed to a rigid support. The bar is free to rotate about the fixed end (the pivot).

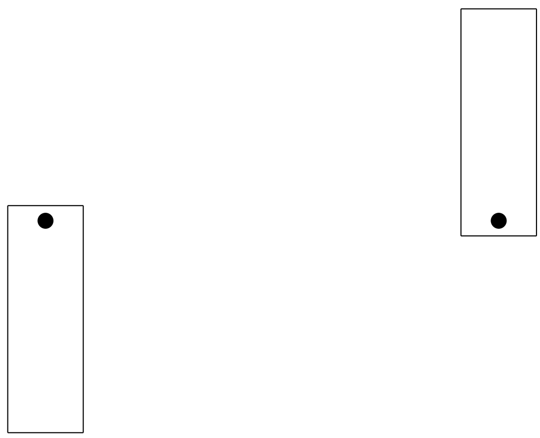

Figure 1: A rigid pendulum has two fixed points: the free end can be either directly above or directly below the pivot (assuming that gravity acts vertically downwards.)

There are, in fact, two possible fixed points of the system, corresponding to the states when the end of the pendulum is either directly above or directly below the pivot. If you were to place the pendulum (carefully) in either of these positions, it would remain there for all time. The idea of a linear stability analysis is to see what happens if we displace the pendulum slightly from either of these positions, *i. e.* give it a flick with your finger.

A fixed point is said to be **linearly stable** if the system **moves back** to the fixed point after it has been slightly perturbed. In the context of our mechanical analogy, the pendulum hanging below the pivot is a linearly stable fixed point.

[3]Note that we could also regard ∞ as a fixed point of the system, although in order to prove this we must use the substitution $z_n = 1/y_n$ and show that $z_n = 0$ is a fixed point of the resulting difference equation.

A fixed point is said to be **linearly unstable** if the system **moves away** from the fixed point after it has been slightly perturbed. The free end of the pendulum balanced vertically above the pivot is a linearly unstable fixed point. If the pendulum is displaced, it will swing around the pivot, oscillate about the lower fixed point and eventually settle there.

Analysis

In mathematical terms, we must consider what happens when

$$\boxed{y_n = Y + \tilde{y}_n,} \tag{12}$$

where Y is a fixed point of the system and $\tilde{y}_n \ll 1$ is a small perturbation to the system. The idea of the analysis is to find an explicit expression for $\tilde{y}_n$. If $|\tilde{y}_n| \to 0$ when $n \to \infty$, then the system moves back to the fixed point at Y and we conclude that the fixed point is linearly stable. The word **linear** in the expression linear stability analysis refers to the following important assumption:

> In a linear stability analysis we assume that the perturbation $\tilde{y}_n$ is so small that we can neglect all terms of the form $\tilde{y}_n^m$ where $m \geq 2$. That is, we keep only terms linear in $\tilde{y}_n$.

Example

- Recall that the difference equation $y_{n+1} = y_n^2$ (11) has two fixed points at $y_n = 0$, 1. Analyse the linear stability of these fixed points.

  We substitute the ansatz (12) into the equation (11) to obtain

  $$Y + \tilde{y}_{n+1} = (Y + \tilde{y}_n)^2 = Y^2 + 2Y\tilde{y}_n + \tilde{y}_n^2.$$

  Now, Y is a fixed point of equation (11), so $Y = Y^2$ and we have

  $$\tilde{y}_{n+1} = 2Y\tilde{y}_n + \tilde{y}_n^2 \approx 2Y\tilde{y}_n,$$

  when we neglect all non-linear terms, because $\tilde{y}_n \ll 1$. The governing linear equation for the perturbation, $\tilde{y}_n$, is, therefore,

  $$\tilde{y}_{n+1} = 2Y\tilde{y}_n,$$

  which has the solution $\tilde{y}_n = (2Y)^n\,\tilde{y}_0$.

  Fixed point at $y_n = 0$

  When $Y = 0$, we find that $\tilde{y}_n = 0$ and we conclude that the fixed point is **linearly stable**. This is a particularly easy case; and it's easy to see that provided $\tilde{y}_n < 1$, the system will always return to the fixed point at $y_n = 0$.

  Fixed point at $y_n = 1$

  Here, $Y = 1$, and we find that $\tilde{y}_n = 2^n\,\tilde{y}_0$. Hence, $|\tilde{y}_n| \to \infty$, as $n \to \infty$ and we conclude that the fixed point is **linearly unstable**. In fact, if $\tilde{y}_n < 0$, the system will be attracted to the fixed point at $y_n = 0$; whereas if $\tilde{y}_n > 0$, the system will grow without bound.

2.3 Bifurcations

When a difference equation involves a parameter, then the behaviour of the system can change, sometimes dramatically, as the parameter varies. Consider the difference equation:

$$y_{n+1} = y_n^2 + r, \tag{13}$$

which is the same as the example (11) only with the addition of a single parameter r. The fixed points of the equation occur when

$$y_n = y_n^2 + r \quad \Rightarrow \quad y_n^2 - y_n + r = 0 \quad \Rightarrow \quad y_n = \frac{1}{2} \pm \sqrt{\frac{1}{4} - r}.$$

Figure 2 shows the locations of the fixed points plotted as functions of the parameter r.

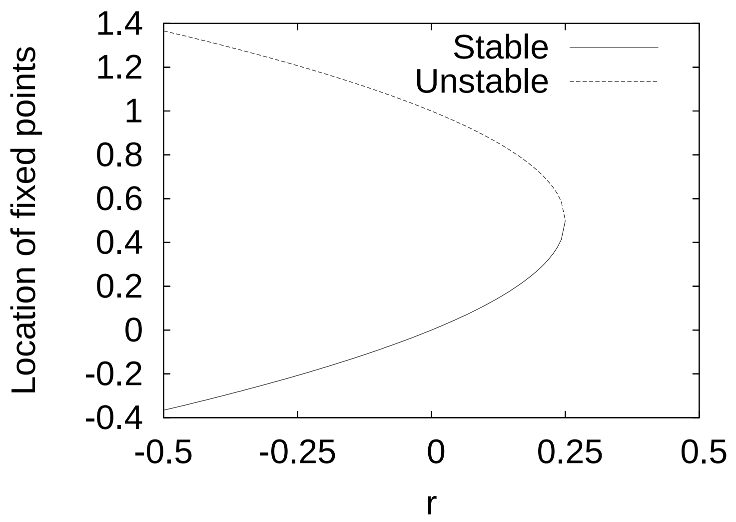

Figure 2: The fixed points of the difference equation $y_{n+1} = y_n^2 + r$, as a function of r.

Notice that there is a very abrupt change in the character of the system when $r = \frac{1}{4}$. If $r < \frac{1}{4}$, there are **two** fixed points (when $r = 0$, the fixed points are located at 0 and 1, as found earlier). If $r > \frac{1}{4}$, there are **no** fixed points. We say that $r = \frac{1}{4}$ is a bifurcation point. In general, a bifurcation point is a point where there is a change in the stability of the fixed points of the system. The bifurcation shown in Figure 2 is called a saddle-node bifurcation, or turning point.

2.4 Does the sequence remain bounded?

The sequence generated by the difference equation can sometimes remain bounded for a certain range of parameters and initial conditions. For example, consider the tent map:

$$y_{n+1} = r\left(\frac{1}{2} - \left|y_n - \frac{1}{2}\right|\right). \tag{14}$$

If we let the function $F(x) = \frac{1}{2} - |x - \frac{1}{2}|$, see Figure 3, then the RHS of the tent map is merely $rF(y_n)$.

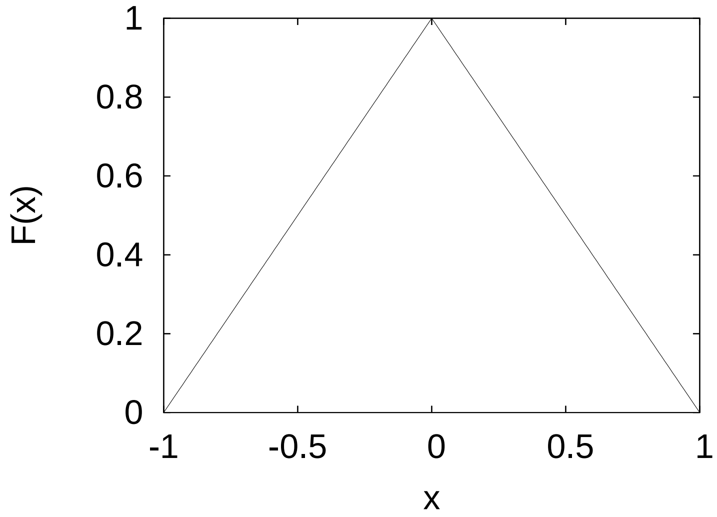

Figure 3: The function $F(x) = \frac{1}{2} - |x - \frac{1}{2}|$ plotted as a function of x.

Example

- For what values of r does the sequence generated by the difference equation (14) remain in the range $[-1, 1]$ if $y_0 \in [-1, 1]$?

  From Figure 3, we see that

  $$\max_{x\in[-1,1]} F(x) = 1, \quad \text{at } x = 0 \quad \text{and} \quad \min_{x\in[-1,1]} F(x) = 0, \quad \text{at } x = -1, 1.$$

  If $y_n \in [-1, 1]$, then

  $$y_{n+1} \leq 1, \text{ if } \max_{y_n\in[-1,1]} rF(y_n) \leq 1 \quad \Rightarrow \quad r \leq 1.$$

  Similarly, if $y_n \in [-1, 1]$, then

  $$y_{n+1} \geq -1, \text{ if } \min_{x\in[-1,1]} rF(x) \geq -1 \quad \Rightarrow \quad r \geq -1.$$

  Therefore, if $r \in [-1, 1]$ the sequence generated by (14) remains in the range $[-1, 1]$ when it starts in $[-1, 1]$.