2 Discrete Dynamical Systems (DDS)

2.1 Basics

A Discrete Dynamical System (DDS) models the change (or dynamics) of single or multiple populations or quantities in which the change occurs deterministically and on a discrete set of time values.

By breaking down the definition of a DDS we form the framework for our models. We first specify either the default time increment or spacing $\Delta t$ or specify particular moments of time: $t_0, t_1, t_2, \ldots, t_n, \ldots$. For uniform spacing we typically take $t_n = n\Delta t$. Then we let $X_n$ be the value (single or multiple/vector) of interest at time $t_n$. If we think of $X$ as a continuously varying quantity then $X_n = X(t_n)$, however, we won’t necessarily make that assumption. Because the change is deterministic, if we know the values $X_0, X_1, X_2, \ldots, X_n$ then we can determine $X_{n+1}$. To limit the complexity of the models we consider, we assume the future value only depends on a fixed number of prior values, thus we can take the model in general form as

$$X_{n+1} = f(n, X_n, X_{n-1}, X_{n-2}, \ldots, X_{n-k}).$$

(1)

The first argument ($n$) represents the possible dependence on the current time and could be replaced by $t_n$ or $(t_n, t_{n-1}, \ldots, t_{n-k})$. The specific form of the function $f$ is determined by the model mechanics that describe the dynamics of the population. There are classic functions used as models and they typically represent a family of models, with the parameters in the model determining the particular behavior. The parameters are determined by the specific circumstances we consider.

There are also expressions called Difference Equations which look much like a DDS. The only difference is that in a difference equation you describe the change from $X_n$ to $X_{n+1}$ instead of just the new value of $X_{n+1}$. So a typical difference equation looks like

$$X_{n+1} - X_n = g(n, X_n, X_{n-1}, \ldots, X_{n-k}).$$

(2)

Obviously, this could be written in the same form as a DDS and thus is not really a different object. From a modeling standpoint, it is sometimes easier to write out a DDS as a difference equation first, as we often understand the mechanisms that cause change in the values. For example, in the case of simple interest or geometric growth, we have that $X_{n+1} - X_n = aX_n$ where $a > 0$ is the interest or growth rate. This becomes the DDS $X_{n+1} = (1 + a)X_n$.

Another example: Suppose we are farming fish, so we have a pond in which the fish live and grow and we harvest a certain number of pounds of fish on a regular basis. So if $X_n$ is the number of pounds of fish in our pond at week $n$, the weight of the fish increases 20% each week, and we harvest 100 pounds of fish a week, then we can model this as $X_{n+1} - X_n = 0.2X_n - 100$ or $X_{n+1} = 1.2X_n - 100$.


2.2 Characterization

Our first task will be to look at different ways of classifying DDS into different categories or families. The first broad categories have to do with simple characteristics of the function $f$.  

\[\text{\copyright 2021} \text{AI4Ed, Inc.} \]
A DDS has **order** $k$ if $X_{n+1}$ depends only as far back as $X_{n+k}$. We'll commonly consider first-order systems where $X_{n+1}$ depends only on $X_n$.

To completely determine the values of $X_n$, $n = 0, 1, 2, \ldots$ in a order-$k$ DDS, we'll need to specify $k$ **initial values** $X_0, X_1, \ldots, X_{k-1}$.

A DDS is **autonomous** when the model depends on the current time only implicitly through the population values, i.e. there is no explicit mention of $n$ or $t_n$ in the model.

A DDS is **linear** if the function $f$ is a linear function of the arguments $X_n, \ldots, X_{n+k}$. Similarly if $f$ is affine, quadratic, etc. we say the same about the DDS.

A DDS is **non-homogeneous** if $f = g + h$ where $g$ is autonomous and $h$ depends only on $n$ or $t_n$. This distinction is primarily useful in developing solution techniques.

As we construct models, we can add to this list of characteristics.

### 2.3 Solution

Given a DDS, we would like to find a solution, but there are various types of solutions we can consider.

First there is the **analytic** solution. In this case we seek a function $\hat{X}$ such that $X_n = \hat{X}(t_n)$ or $X_n = \hat{X}(n)$. For most DDS, it is impossible to determine an analytic solution. However, in the case of autonomous linear systems, it is possible to construct a solution and we'll discuss that technique later. Also, for some non-autonomous systems one can find a solution using Z-transforms, which we'll not discuss. For non-linear systems there are not any general solution techniques.

A more accessible type of solution although inherently less accurate is the **qualitative** solution. In this case, we look for a description of the behavior of the solution under various scenarios. The description doesn't say when something happens or tell much about the quantities of the change, but usually indicates when the values are increasing or decreasing, and possible what the asymptotic behavior is (i.e the tendencies as $t_n \to \infty$).

Finally, the format of a DDS, naturally leads to **computed** solutions. This can be programmed or implemented in a spreadsheet, all one needs is values for the parameters and the initial values. Typically for a computed solution we compute several solutions changing the parameters and initial values, to explore the dependency of the results on the data for the problem.

#### 2.3.1 Analytic Solution Method

For a homogeneous, autonomous, linear, first-order DDS, we have $X_{n+1} = aX_n$. To solve this, I like the **human computer** method, which is to take $X_0$ and then sequentially compute $X_1, X_2, \ldots$ by hand until I recognize the pattern. In this case we have $X_1 = aX_0$, $X_2 = a^2X_0$, $X_3 = a^3X_0$, and the apparent pattern is $X_n = a^nX_0$. (We can use induction to prove that this is the solution to the DDS if necessary). Note that if $|a| > 1$ then $|X_n| \to \infty$ as $n \to \infty$ and if $|a| < 1$ then $|X_n| \to 0$ as $n \to \infty$. If $a < 0$ then the solution oscillates. These observations, though simple, are important for the next section.

For a non-homogeneous, autonomous, affine, first-order DDS, we have $X_{n+1} = aX_n + b$. There are two approaches to solving this problem. The first is the human computer method, in which we get $X_1 = aX_0 + b$, $X_2 = a(aX_0 + b) + b$, $X_3 = a(a(aX_0 + b) + b) + b$, etc. Note that in this method it is often helpful to not simplify the expression too much, and this is especially true in the case
when you are using real numbers. Maybe we see a pattern now, so we can guess the solution is

\[ X_n = a^n X_0 + (a^{n-1} + a^{n-2} + \ldots + a^2 + a + 1)b. \]

This is not bad, but we can use some algebra to simplify. Since the coefficient of \( b \) is a geometric series we can use the property that \( \sum_{k=0}^{n} r^k = (1 - r^{n+1})/(1 - r) \) when \( r \neq 1 \) to simplify this solution to \( X_n = a^n X_0 + (1 - a^n)/(1 - a) b \) when \( a \neq 1 \). If \( a = 1 \), we can easily recognize the solution to be \( X_n = X_0 + nb \).

A second approach is to use the result for the homogeneous problem and figure out what else is needed due to the \( b \)-term. So we know part of the solution should be \( Ca^n \), because if \( b = 0 \) then the solution is \( X_0 a^n \). Now to figure out the part that depends on \( b \) we just use guess and check to find some expression which fits the DDS. Since \( b \) is a constant, it makes sense to at least try \( X_n = D \) for all \( n \) where \( D \) is a constant. Then we have \( D = aD + b \) or \( D = b/(1 - a) \). We then have a general solution \( X_n = Ca^n + b/(1 - a) \) and can use the value of \( X_0 \) to determine \( C \). Thus for \( n = 0 \) we have \( X_0 = C + b/(1 - a) \) or \( C = X_0 - b/(1 - a) \). Thus our solution is \( X_n = (X_0 - b/(1 - a))a^n + b/(1 - a) \). This solution matches the solution using the other method.

This second form is somewhat more useful as we can observe that if \( |a| < 1 \) then \( X_n \to b/(1 - a) \) as \( n \to \infty \).

For some non-homogeneous, affine or linear DDS, we can solve non-autonomous versions by either using a similar method or using Z-transforms. It is, however, easy to write down a pleasant looking problem that is impossible to solve. So basically, we’ve exhausted all possible analytic solution methods for DDS. Sometimes you can figure something out about a DDS by using the human computer method and some lucky observations, but that is not typical.

### 2.3.2 Qualitative Solution Methods

For qualitatively describing either the very short-term or very long-term behavior of the solution of a DDS, we use basically two observations: (1) if \( X_{n+1} - X_n > 0 \) then \( X_{n+1} > X_n \) (and similarly for \( < \)) and (2) if \( X_{n+1} = aX_n \) then if \( |a| > 1 \) \( |X_n| \) grows (to \( \infty \)) and if \( |a| < 1 \) then \( |X_n| \) tends towards \( 0 \).

We’ll start a with first order, autonomous DDS: \( X_{n+1} = f(X_n) \). It may be helpful to rewrite this as a Difference Equation: \( X_{n+1} - X_n = g(X_n) \) where \( f(x) = x + g(x) \). We’ll assume that \( f \) and \( g \) are continuous and differentiable as much as we want. Then, using the difference form, we see that for some values of \( X_n \), \( X_{n+1} - X_n > 0 \) and for some \( X_{n+1} - X_n < 0 \). For example if \( g(x) = ax(K - x) \) then \( X_{n+1} - X_n > 0 \) when \( 0 < X_n < K \) and \( X_{n+1} - X_n < 0 \) when \( X_n < 0 \) or \( X_n > K \). So we identify intervals, such that if \( X_n \) is in a certain interval we know whether \( X_{n+1} > X_n \), i.e. it increases or \( X_{n+1} < X_n \) it decreases. We can also make a conjecture about what happens over more iterations, if assume the values stay in the same interval. For example above, we have the population increasing if \( 0 < X_n < K \), so if we assume that the values stay in that interval for a while, then the population will increase during that time. Or we can imagine a scenario where the population will always increase, approaching the upper limit of the interval as \( n \to \infty \).

These boundary points between the intervals have a special property, they are fixed points for the DDS. A fixed point is a value \( p \) such that if \( X_n = X_{n-1} = X_{n-2} = \ldots = X_0 = p \) then \( X_{n+1} = p \). Or, equivalently, if \( X_n = p \) for all \( n \) is a solution of the DDS. We can find fixed points by (1) solving \( p = f(p) \) or (2) solving \( g(p) = 0 \) or (3) (for some but not all fixed points), calculating \( p \) as \( \lim_{n \to \infty} X_n \).
Once we have fixed points, we can categorize them into two broad types: attractors and repellers. If a fixed point is an attractor, then if a DDS starts near the fixed point it stays near or gets closer to the fixed point. If it is a repeller then if a DDS starts near the fixed point then it moves away from the fixed point. In simplest terms, if \( p \) is the fixed point and \(|X_0 - p| < \epsilon\) for some small epsilon, then if \(|X_1 - p| < |X_0 - p|\) \( p \) is an attracting fixed point, and if \(|X_1 - p| > |X_0 - p|\) \( p \) is a repelling fixed point.

As an example, consider our good friend, \( X_{n+1} = aX_n \). The only fixed point is \( p = 0 \) and it is attracting if \(|a| < 1\) and repelling if \(|a| > 1\). What if \(|a| = 1\)? In this case \( p = 0 \) is neither attracting nor repelling.

We can make the process of analyzing a fixed point even simpler by using linearization. Let \( p \) be a fixed point for \( X_{n+1} = f(X_n) \), i.e. \( p = f(p) \), and set \( X_n = p + E_n \). Then \( p + E_{n+1} = f(p + E_n) \approx f(p) + f'(p)E_n \) (using just the first two terms of the Taylor expansion). Simplifying, we get \( E_{n+1} \approx f'(p)E_n \). Since \( f'(p) \) is constant, this is a familiar DDS and we know that it has 0 as a fixed point and that fixed point is attracting if \( |f'(p)| < 1 \) and repelling if \( |f'(p)| > 1 \). Now in this case if \( |f'(p)| = 1 \) then we don’t know if it is repelling or attracting as it depends on the higher order terms in the Taylor expansion (which we’ve conveniently ignored).

For an example, consider \( X_{n+1} = X_n + a(K - X_n)X_n \). For this \( f(x) = x + a(K - x)x \). It has two fixed points 0 and \( K \). Since \( f'(x) = 1 + a(K - 2x) \). Then \( f'(0) = 1 + aK \), and, since we’ll assume \( a > 0 \) and \( K > 0 \), \(|f'(0)| > 1\) thus \( p = 0 \) is a repelling fixed point. Next \( f'(K) = 1 - aK \). This case is slightly more complicated as \(|f'(K)| < 1\) only if \( aK < 2 \). If \( aK > 2 \), \(|f'(K)| > 1\). Thus \( p = K \) is and attractor if \( aK < 2 \) and a repeller if \( aK > 2 \).

Another way to study this behavior is through the use of a cobweb plot. This graphical method is not as precise as the above calculations, but can give a better understanding of the overall dynamics of the system and how it might change as we modify the model. For a cobweb plot you start with the graph of the curve \( y = f(x) \) and the line \( y = x \). The first observation is that these curves intersect at fixed points. The second observation is that if for a value \( x \), \( f(x) > x \), i.e. the curve \( y = f(x) \) is above the line \( y = x \) then the values are increasing. So if \( f(X_n) > X_n \) then \( X_{n+1} > X_n \). Similarly if \( f(x) < x \) then the values are decreasing. To actually make the cobweb, start with a value \( X_0 \) near one of the fixed points, then mark the point \((X_0, f(X_0))\) (note: \( X_1 = f(X_0) \)). Next, move horizontally to the curve \( y = x \) and mark that point \((X_1, X_1)\), then vertically to the point \((X_1, f(X_1))\) (on the curve \( y = f(x) \)). Then you repeat this process over and over again: move horizontally to \( y = x \), move vertically to \( y = f(x) \). If you look at the sequence of \( x \)-values that you plot, they are the sequence \( X_0, X_1, X_2, \ldots \). Thus if the plot moves towards the fixed point \((p, f(p))\), that represents an attractor, and if the plot moves away, that is a repeller.

Some final comments: qualitative solution methods work best with first order, autonomous, single population systems. We can extend some of the results to higher order, like the idea of a fixed point. Often non-autonomous systems don’t have fixed points, so most of what we said doesn’t work. We can, and will, extend much of what we did to multiple populations. In general qualitative methods are good for describing general model behavior and thus are good to keep in mind during model development. They are also useful for identifying potentially interesting behavior in a model that we can then use computational methods to resolve further.
2.3.3 Computational Solution Methods

When we have a complex, fully populated model or we just want to see some examples of how the model behaves, nothing beats a computational method. It is also extremely convenient that the form of a DDS naturally leads to computing values. We’ll consider first order DDS of a single population, but everything we do can easily be extended to both higher-order systems and systems with multiple populations.

For this section, we write our DDS as \( X_{n+1} = f(n, X_n; a) \), \( X_0 \) given, where \( a \) represents all the parameters that go into our model. The fundamental computational element is the calculation of \( X_1, X_2, \ldots \) given values for \( X_0 \) and \( a \), and the basic process is some type of loop. For example, in MATLAB, the code for the DDS \( X_{n+1} = aX_n \) would look like:

```matlab
a = X(1) = % MATLAB starts at 1
for n = 1:100
    X(n+1) = a*X(n);
end
```

In other procedural languages (like C, FORTRAN, etc.) the code would be similar. This can also be done in a spreadsheet, where you put the values of the parameter(s) at the top, then the initial value in the cell below. Then in the next cell, you write the formula for the DDS, using the cell above, and the parameters at the top. Then by copying this formula to all the cells below, you compute the desired values.

Once you have produced a sequence of values, you usually want to process them in some way. This may include graphing the results, determining extreme values, determining time to some critical event, etc. What you look at can be related to the question you are trying to answer or to questions that came up about how the model functions.

Once you have your basic process created and have some means to process the results, then the next phase is to look at consequence of variations in the initial value and the parameter(s). For example if you first did some qualitative analysis and found some fixed points, you would look at the results for different initial values around the fixed points. Another possibility is that in building your model you did not find a specific value for some of the parameters but only a range. Then you could look at different results for different values of the parameters. Even if you think you know your initial value and your parameters, you still want to vary them to see how sensitive the results are to changes in their value(s).