Modeling Using Differential Equations

Your first model

Ordinary differential equations are frequently used for mathematical modeling in biology. Let us start with a simple model. Take a reversible reaction between A and B:

We want to know if we start with a specific amount of *A* and *B*, and then allow the reactions to happen, how much *A* and *B* will be there after time *t*. We also want to know how the concentrations of *A* and *B* change with time. In other words, we want to know the time traces of *A* and *B*.

We will use Ordinary Differential Equations (ODEs) to answer these questions. Using ODEs, we write equations that represent the rate of a particular process. For this example of reversible reaction:

Rate of change in concentration of *A*,

$$\frac{d[A]}{dt} = -k_1[A] + k_2[B]$$

Rate of change in concentration of *B*,

$$\frac{d[\mathrm{B}]}{dt} = k_1[A] - k_2[B]$$

Here, we have two dependent variables, *[A]* and *[B]*. They are called dependent as they are changing with time. Time *t* is the independent variable. *A* is converted into *B* with a

rate constant k_1. The rate constant for the backward reaction is k_2. These constants, k_1 and k_2, are called parameters of this system, and they do not change with time.

While writing such equations, be careful about the units of each term. The unit on the left-hand side of an equation must be the same as that of the right-hand side.

Also, note the signs (+ or −) used before each term in these equations. They represent the direction of the process. The forward process is reducing the concentration of *A*. Therefore, in the first ODE, the sign for $k_1.[A]$ is negative. The sign of $k_2.[B]$ is positive as it represents the backward reaction that increases A. In the next ODE, the corresponding signs are just the opposite, as this equation represents the change in the concentration of *B*.

We will now simulate this model using JSim. JSim is a software that allows us to simulate an ODE-based model using numerical methods. A primer on the use of JSim is provided in the Appendix. You can use any other software of your choice to simulate this model. Software that can be used for simulation of ODE-based models are listed in the Appendix.

For numerical simulation, we have to specify the values of the parameters in our model. Let, k_1 = 0.2 1/s and k_2 = 0.4 1/s. We have to specify the initial concentrations of *A* and *B*. Let these values were 10 M and 20 M, respectively. We will simulate the model for t = 12 s.

Use the JSim code given below and observe how the concentrations of *A* and *B* changes with time.

```
math revReaction{
      realDomain t ;
      t.min = 0;
      t.max = 12;
      t.delta = 0.1;

      // parameter values
      real k1 = 0.2;
      real k2 = 0.4;

     // Declare dependent variables
      real A(t), B(t);

      // Define initial condition
```

```
        when(t=t.min){A = 10;B = 20;}
        // ODEs
        A:t = -k1*A + k2*B;
        B:t = k1*A - k2*B;
}
```

The result of this simulation is shown in Figure 1.

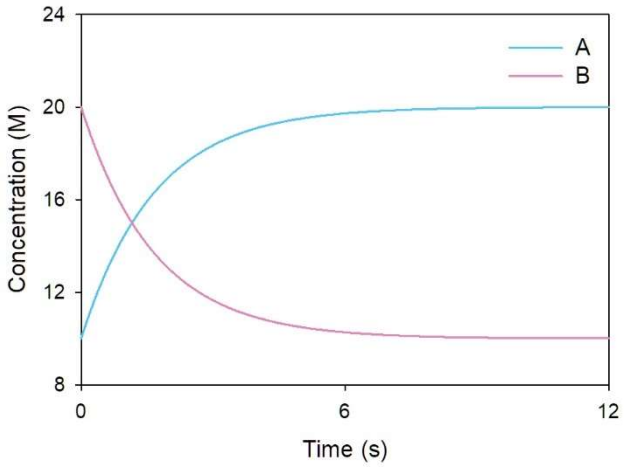

Figure 1: Temporal dynamics of A and B in the mathematical model for the reversible reaction. Initial conditions, $[A]$ = 10 M and $[B]$ = 20 M. k_1 = 0.2 1/s, k_2 = 0.4 1/s.

Ordinary differential equations

The differential equations that we have used for the reversible reaction are called Ordinary Differential Equations. Though called ordinary, there is nothing ordinary about these equations. Many exotic phenomena can be understood using ODEs.

An ODE will have only one independent variable. In our first model, the independent variable is time. However, in certain circumstances, a system may change with respect to more than one independent variable. For example, imagine the diffusion of some molecules in two-dimension. The concentration of the molecule will change with the position (x and y) and time. So, we have three independent variables x, y, and t.

In this case, we can not use ODEs. We have to use Partial Differential Equation (PDE). The following PDE is used to capture the dynamics of diffusion in two-dimension:

$$\frac{\partial C}{\partial t} = D\left(\frac{\partial^2 C}{\partial x^2} + \frac{\partial^2 C}{\partial y^2}\right)$$

Here, *C* is the concentration at a point in the two-dimensional space, and *D* is the diffusion coefficient.

In this book, we will restrict our discussion to ODEs. An ordinary differential equation is an equation involving the ordinary derivative of a function. We integrate this equation to get the function. For example, $\frac{dx}{dt} = t$ is an ODE, and it is a derivative of the function $x = f(t) = \frac{t^2}{2} + c$, where *c* is a constant.

All the ODEs discussed above are called first-order differential equations. The order of a differential equation is the order of the highest order derivative in the equation. For example, the following equation is of order 3:

$$\frac{d^3x}{dt^3} = k\frac{dx}{dt} + bx + c$$

In this book, we will deal only with first-order differential equations. Some of these equations would be linear ODEs and others nonlinear. A linear ODE has the following characteristics:

a) The dependent variable and its derivative should have the power of one.
b) There should be no product of the dependent variable and its derivative.
c) It should not involve any transcendental functions (like trigonometric, logarithmic functions) of the dependent variable or its derivative.

The general form of a linear first-order ODE is

$$\frac{dx}{dt} = a(t)x + b(t)$$

where, a(t) and b(t) are continuous functions.

On the other hand, examples of nonlinear ODE are:

$$\frac{dx}{dt} = ax^3 + b$$
$$x\frac{dx}{dt} = ax + b$$
$$\frac{dx}{dt} = \sin\left(a + x\right)$$

The following differential equation is a particular type of first-order linear equation, called a homogeneous equation:

$$\frac{dx}{dt} = bx$$

Note that this equation does not have a separate constant term. The generalized form of an n^{th} order linear homogenous ODE is:

$$a_n \frac{d^n x}{dt^n} + a_{n-1} \frac{d^{n-1} x}{dt^{n-1}} + \cdots + a_1 \frac{d^1 x}{dt^1} + a_0 x = 0$$

here, $a_n,\ a_{n-1}, \ldots, a_1, a_0$ are constants.

A special type of ODE is called an autonomous equation. In an autonomous ODE the only place where the independent variable appears is the derivative:

$$\frac{dx}{dt} = f\left(x\right)$$

Often, we will deal with more than one dependent variable. In such a situation, we will separate ODEs to capture the rate of change of each of these variables. Therefore, we will have a set of ODEs. Such a set of ODEs is called a system of ODE. The following is a system of two ODEs:

$$\frac{dx}{dt} = ax + by$$
$$\frac{dy}{dt} = cx$$

Check these ODEs carefully. The derivative of x depends on both x and y. That means to know x we have to know y, which itself changing with time. On the other hand, the derivative of y depends upon x. So, we have to know x in order to know y. Therefore, the behavior of x and y are linked. This type of system of ODEs is called a system of coupled ODEs.

Law of mass action

Most of the processes in cell biology are biochemical reactions. Therefore, we can use the rules of chemical kinetics to model those processes.

Take a generalized chemical reaction:

$$aA + bB \xrightarrow{k} cC + dD$$

The rate of this reaction is represented as:

$$\text{Rate} = -\frac{1}{a}\frac{d[A]}{dt} = -\frac{1}{b}\frac{d[B]}{dt} = \frac{1}{c}\frac{d[C]}{dt} = \frac{1}{d}\frac{d[D]}{dt} = k[A]^{a'}[B]^{b'}$$

Here, *[A]*, *[B]*, *[C],* and *[D]* are molar concentrations of free *A*, *B*, *C* and *D* respectively. *a*, *b*, *c,* and *d* are the stoichiometric constants. *a'* and *b'* are the orders of reactant *A* and *B,* respectively. The summation of the superscripts in the rate equation, $(a'+b')$, is called the overall order of the reaction.

The parameter k is the rate constant. Rate constants and reaction orders are measured by experiments. In some cases, an order of a reactant can be equal to its stoichiometric constant. For many biological systems, we do not have a clue about orders of reactions. For simplicity, stoichiometric constants are commonly used in place of orders.

The rate equation, given above, is based on the well-known Law of Mass Action that states (in brief) that the rate of a reaction is proportional to the product of concentrations of the reactants raised to their orders. Not all reactions, in living systems, follow the Law of Mass action. Most of the biochemical reactions are not elementary reactions and involve multiple steps that are often unknown. Sometimes, multiple known steps in a biochemical process are clubbed together to reduce the complexity of the model. Though the Law of Mass Action is not applicable, such processes are still modeled using ODE-based rate equations similar to the one we wrote earlier. Such ODE-based rate equations are also

used beyond chemical reactions, like in modeling population dynamics of organisms. In all cases, we formulate the ODEs, rationally, to represent the rates of the processes in the system.

The Law of Mass Action is valid only when large numbers of molecules of the reactants are present in a well-stirred container. These two pre-conditions, high concentrations, and spatial homogeneity assure that the probability of molecules colliding with each other, by diffusion, is proportional only to their concentrations. It is crucial to note that these two conditions must meet for any ODE-based model even when that is not associated with the Law of Mass Action.

Law of large number

Suppose we have a container of defined volume with N_A number of molecule A and N_B number of molecule B. Imagine these molecules as solid spheres that are jiggling in the container. When one molecule of A collides with one molecule of B, they fuse and create a molecule of C, another solid sphere.

When N_A and N_B are very large, there will be a collision between A and B at every moment somewhere in the container. Therefore, almost at every moment, a new C will be created. Though each collision is a discrete event, we will perceive the production of C as a continuous process. The rate of the process will depend upon the concentrations of A and B.

However, when N_A and N_B are very small, collisions will not happen at every moment, and we will be uncertain about the frequency of collision. As an observer, we may observe a collision right now and then may have to wait for an hour to see the next collision. Likewise, the third collision may happen within a few seconds after the second one. Therefore, the process of production of C will be discrete with uncertainty. This type of process with uncertainty is called a stochastic process. We cannot use the law of mass action and differential equations to model such processes.

The requirement of high concentration or number can be appreciated from another perspective. We use differential equations to model reactions. In such equations, concentrations of molecules are dependent variables. Dependent variables in a differential equation must be continuous, and the function must be smooth. That means the dependent variable must not change abruptly.

These conditions are met only when the molecules involved in a reaction are present at high concentrations. Consider a cell of 1 μm radius. Its volume would be approximately 4×10^{-15} lt. Suppose the concentration of a molecule, A, inside the cell is high, say 10 mM and it is degrading. Using Avogadro number, we know that 10mM of A in this small volume is equal to approximately 24000000 molecules of A. When one molecule of A degrades, the number of A decreases to 23999999 and the concentration of A changes to 9.99999958 mM. So the change in the concentration of A is small and smooth.

Now imagine the same cell, but the concentration of A is very low, say 10 nM. This concentration is equivalent to approximately 24 molecules. If one molecule of A degrades, its number reduces to 23, and the concentration changes to 9.58 nM. That is an abrupt change and we cannot consider the change in concentration as smooth. Therefore, we can not use a differential equation to model the degradation of A when its concertation is so low.

In general, ODE based models are used when concentrations of molecules are above 10 nM. For a cell of 10 μm diameter, this is equivalent to around 3000 molecules per cell. Concentrations of some molecules in a single cell may be lower than this limit. To circumvent this issue, one can imagine that an ODE-based model of a biochemical system does not represent a single cell but an ensemble of cells considered as a single unit. This approach matches with our experimental techniques. Most of the time, we do not measure the number of molecules in a single cell. Say, we want to measure the amount of phosphorylated Akt in human fibroblast cells treated with insulin. Usually, Western Blot is used for this purpose. Millions of human fibroblast cells, grown as a monolayer in a flask, are treated with insulin, and subsequently harvested together. These cells are lysed to analyze by SDS-PAGE and Western blot. Here, we do not measure the amount of phospho-Akt in one single cell but estimate an average amount of it in an ensemble of millions of cells.

Most of the biology that we know today is based on cellular ensemble data. Such measurements rely on the assumption that all cells in an ensemble behave similarly. However, single-cell experiments have shown that this assumption breaks in many cases, and the ensemble average can be misleading. Such a situation requires an entirely different approach that we will discuss in stochastic simulations of biological systems.

Your second model

We will model the spread of an infectious disease in a population using ODE. To make the model simple (though a bit unrealistic) we make the following assumptions: (a) the

disease spreads when an infected person comes in contact with an uninfected one, (b) no one gets cured, (c) everybody can come in contact with everyone, (d) the total population is large, and (e) the total population remains constant over time (no death or birth).

Assumptions (c) and (d) are made to make the model suitable for modeling using ODEs. The society is considered as a homogenous well-stirred tank, where everyone can come in contact with very one. Assumption (d) is equivalent to having a reaction with a high concentration of reactants.

Assumptions (b) and (e) make the equation simple. In real life, there will be birth, death, and even people will get cured of the disease. You may consider that rates of death and birth are the same, so that the total population remains constant. Often that is a reasonable assumption. We make similar assumptions in modeling many cellular processes. For example, while modeling a signaling network in a cell, you may assume that the total amount of some proteins remains constant.

With these assumptions, we can compare the problem of the spread of infection with a chemical reaction. In a chemical reaction, a product is formed when reactants interact with each other. Similarly, the infection spreads when an uninfected person comes in direct or indirect contact with an infected person.

In chemical kinetics, we measure the change in the concentration of a molecule. Here, we will measure the change in size of the infected population. Suppose a fraction, x, of the population is infected with a disease. So, (1-x) fraction of the population is free of the disease. Using the metaphor of a chemical reaction, we can write the following ODE:

$$\frac{dx}{dt} = rx\left(1-x\right)$$

Here r is the rate constant for the spread of the disease. In this system, with time, the number of infected people will change. This temporal change in the size of the infected population can be estimated by integrating this ODE.

Rearrange the ODE to integrate it:

$$\frac{dx}{dt} = rx(1-x)$$

$$\frac{dx}{x(1-x)} = rt$$

Let at $t = 0$, $x = x_0$. This is the initial condition. We need to specify the initial condition to solve any ODE based problem. Integrating both sides with this initial condition, we get:

$$\int_{x_0}^{x} \frac{dx}{x(1-x)} = r\int_0^t t$$

$$\int_{x_0}^{x} \frac{dx}{x} + \int_{x_0}^{x} \frac{dx}{(1-x)} = r\int_0^t t$$

$$\left[\ln(x)\right]_{x_0}^{x} - \left[\ln(1-x)\right]_{x_0}^{x} = rt$$

$$\ln\left(\frac{x}{x_0}\right) - \ln\left(\frac{1-x}{1-x_0}\right) = rt$$

$$\frac{x(1-x_0)}{x_0(1-x)} = e^{rt}$$

$$x = \frac{1}{1+\left(\frac{1}{x_0}-1\right)e^{-rt}}$$

So, by integration with the initial conditions, we have got x as a function of time. This function gives the time evolution of x. Suppose, in the beginning, 2% of the population was infected, and the rate constant for spread of infection is 0.5 per day. We can now estimate the size of the infected population after 20 days:

$$x = \frac{1}{1+\left(\frac{1}{x_0}-1\right)e^{-rt}} = \frac{1}{1+\left(\frac{1}{0.02}-1\right)e^{-0.5\times 20}} = 0.9977 = 99.77\%$$

Let us try to understand the change in x with time as per the function obtained by integration. As time increases, e^{-rt} decreases, thereby decreasing the denominator of the function. Therefore, with time, x increases. When t tends to infinity ($t \to \infty$), x tends to one ($x \to 1$).

It is expected as with time, more and more people are infected. Eventually, everyone in the population is infected. However, this dynamics is observed only when $0 < x_0 < 1$.

When no one is infected to start with ($x_0 = 0$), there will be no spread of infection. When everyone is infected ($x_0 = 1$), there will be no further spread too.

If we know the numerical values of x_0 and r, we can also plot x with respect to t. that will provide us a graphical representation of time evolution of x or, in other words, the spread of infection with time. In Figure 2, we have plotted the dynamics of spread of infection for $x_0 = 0.02$ and $r = 0.5$.

The dynamics of the spread of the disease is clearer from this figure. Notice the slope of the curve in the period of zero to five. The slop is very shallow. Initially, a small fraction of the population was infected. Therefore, the rate of spread of the disease was also low. With time, the size of the infected population increases so is the rate of infection. Eventually, the curve tends to saturate as the majority of the population is already infected.

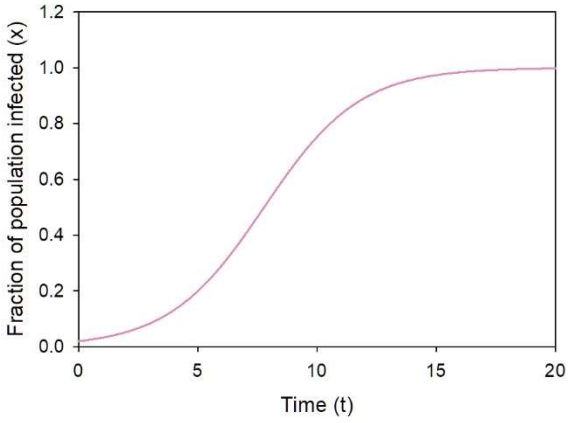

Figure 2: The dynamics of the spread of infection in a population. The initial value of x is $x_0 = 0.02$ and the rate constant $r = 0.5$.

You can also simulate this model in JSim using the following code:

```
math INF{
      realDomain t;
      t.min = 0;
      t.max = 20;
      t.delta = 0.1;

      // parameter values
      real r = 0.5;
```

```
        // the dependent variable
        real x(t);

        // initial condition
        when(t=t.min){x = 0.02;}

        // ODE
        x:t = r*(1-x)*x;
}
```

Nondimensionalization of equations

Let us have a system of ODEs representing a biochemical reaction involving two molecules.

$$\frac{d[X]}{dt} = K_1[Y]([X_T]-[X]) - d_1[X]$$
$$\frac{d[Y]}{dt} = K_2[X]([Y_T]-[Y]) - d_2[Y]$$

Imagine that X and Y are active forms of two enzymes that control each other. [X] and [Y] are molar concentrations of these two active enzymes, respectively. This model has six parameters. Two of the parameters are [X_T] and [Y_T] - the total concentrations (active and inactive) of these two enzymes. K_1, K_2, d_1, and d_2 are rate constants with appropriate unites- $M^{-1}s^{-1}$, $M^{-1}s^{-1}$, s^{-1}, and s^{-1}, respectively.

We will try to rearrange the terms in these ODEs so that both the equations become dimensionless. To do so, let us assume, [X] = X^*x, such that X^* is a constant with unit M, and x is a dimensionless variable representing the concentration of X. Similarly, for Y we assume [Y] = Y^*y. For time t, assume $t = T\tau$. T is a constant with sec as its unit, and τ is the dimensionless time.

Now, rewrite our original ODEs in terms of these new variables. First, use the chain rule to rewrite the derivatives,

$$\frac{d[X]}{dt} = \frac{d[X]}{d\tau}\cdot\frac{d\tau}{dt} = \frac{1}{T}\frac{d[X]}{d\tau} = \frac{X^*}{T}\frac{dx}{d\tau}$$
$$\frac{d[Y]}{dt} = \frac{d[Y]}{d\tau}\cdot\frac{d\tau}{dt} = \frac{1}{T}\frac{d[Y]}{d\tau} = \frac{Y^*}{T}\frac{dy}{d\tau}$$

We can rewrite the original ODEs as,

$$\frac{X^*}{T}\frac{dx}{d\tau} = K_1 Y^* y\left([X_T] - X^* x\right) - d_1 X^* x$$
$$\frac{Y^*}{T}\frac{dx}{d\tau} = K_2 X^* x\left([Y_T] - Y^* y\right) - d_2 Y^* y$$

Still, the ODEs have dimensions. To make these ODEs dimensionless, we need to rearrange the terms,

$$\frac{dx}{d\tau} = K_1 T Y^* y\left(\frac{[X_T]}{X^*} - x\right) - d_1 T x$$
$$\frac{dx}{d\tau} = K_2 T X^* x\left(\frac{[Y_T]}{Y^*} - y\right) - d_2 T y$$

Now check dimensions of both sides of these ODEs. You will find that these are dimensionless. Note that we have not defined the new constants, X^*, Y^*, and T yet. We will define them in such a way that the number of parameters is reduced.

It is straightforward to define X^* and Y^*. If we consider $X^* = [X_T]$ and $Y^* = [Y_T]$, then both $[X_T]/X^*$ and $[Y_T]/Y^*$ are equal to one. In reality, we may not know the exact total concentrations of X and Y in a cell. So by nondimensionalization, we get rid of these two potentially unknown terms. Further, in experiments, usually, we measure the fraction of an enzyme present in its active form, not the absolute concentration. By our definitions, x and y represent the fractions of two enzymes in active forms.

It is difficult to define T in the same fashion. It is present in multiple places in both the ODEs, along with different terms. So we club it with other constants and rewrite the ODEs,

$$\frac{dx}{d\tau} = \alpha_1 y(1-x) - \beta_1 x$$
$$\frac{dy}{d\tau} = \alpha_2 x(1-y) - \beta_2 y$$

This is our system of dimensionless ODEs with a reduced number of parameters. Here, $\alpha_1 = K_1 T Y^*$, $\alpha_2 = K_2 T X^*$, $\beta_1 = d_1 T$, and $\beta_2 = d_2 T$. Also, $0 \le x, y \le 1$.

Through this nondimensionalization, we have achieved two goals. The original system of ODEs had six parameters. The dimensionless system has only four parameters. Usually, these parameter values are unknown. So we either guess the values or estimate those from data. Therefore, a reduction in the number of parameters is a boon for us.

Secondly, all the terms in the ODEs are now just pure numbers. Therefore, this system of ODEs is a generalized one that can capture the dynamical behavior in any reference frame of units.

Exercises

1. Identify the linear and nonlinear ODEs:

$$\frac{dy}{dx} = \frac{y}{(1+y)} - by$$

$$\frac{dy}{dt} = \ln(y) - y$$

$$\frac{dx}{dt} = cx - t$$

2. Which of the following are not autonomous ODEs?

$$\frac{dy}{dx} = vy - kx$$

$$\frac{d^2x}{dt^2} = kx^3 + x - c$$

$$\frac{dy}{dt} = \frac{aty}{c+y}$$

3. Integrate the following ODEs, for the initial condition, $t = 0$, x = x_0

$$\frac{dx}{dt} = \frac{x}{a+x}$$

$$\frac{dx}{dt} = \frac{x^2}{a+x^2}$$

$$\frac{dx}{dt} = bx - c$$

4. The simplest ODE-based model for the growth of bacteria is:

$$\frac{dn}{dt} = rn$$

Here, n is the number of bacteria, and r is the rate constant for the growth of bacteria. For the initial condition, $t = 0$, $n = N$, calculate the function that will give the change in the number of bacteria with time.

5. The model in question 4 is very unrealistic as we have not considered the death of bacteria. Modify this model to include death. Consider that the rate of death depends upon the number of live bacteria. With the initial condition, $t = 0$, $n = N$, calculate the function that will give the change in the number of bacteria with time.
6. Write a system of ODEs to represent the dynamics of the following chemical reaction.

$$X + 2Y \underset{r}{\overset{k}{\rightleftharpoons}} 3P$$

7. The concentration of a protein in a cell is 25 μM. If the volume of the cell is 10^{-12} cc, calculate the number of copies of this protein in the cell.
8. Bionumbers is a curated web source for numbers relevant in cell and molecular biology. Look into this database to find – a) the average number of mRNA in a yeast cell, b) average concentration of a protein in a yeast cell, and c) the average number of a receptor on a human cell surface.
9. The model of population growth in question 4 is called the Malthusian model, named after Thomas Robert Malthus. The following is a non-Malthusian model considering that the competition among individuals for resources constrains growth. Derive the function for the population growth for this model.

$$\frac{dx}{dt} = rx - \frac{r}{k}x^2$$

10. Soluble ligands (L) bind to cell surface receptors (R) to create the receptor-ligand complex (C). This is a reversible process. The following ODE represents the temporal dynamics of complex formation.

$$\frac{d[C]}{dt} = k_1\left([R_T] - [C]\right)L - k_2[C]$$

$[R_T]$ is the total concentration of the receptor, free plus ligand-bound, and is constant. Reduce the number of parameters in this equation by creating a few meaningful dimensionless quantities.