SSB-30806: Modelling in Systems Biology
Week 1 Exercise Sheet 1 Solutions[1]

This practical will be split into three parts:

- writing ODEs using the "Law of Mass Action",
- methods of solving ODE equations analytically,
- numerically solving and ODE, and
- checking dimensions/units of an ODE and rescaling accodingly.

Solutions will be uploaded to Brightspace after the class has finished.

Part 1: Writing ODEs using the "Law of Mass Action"

Write the ODE models for the following sets of chemical reactions. Don't forget to rescale the ODEs appropriately when dimers and higher order complexes are involved!

i $\quad A + B \underset{k_r}{\overset{k_f}{\rightleftharpoons}} C \xrightarrow{k_d} \phi$

ii $\quad \phi \xrightarrow{3} A$

$\quad A + A \xrightarrow{k_f} B + C$

iii $\quad A \underset{k_r=0.1}{\overset{k_f=2}{\rightleftharpoons}} 2B$

iv $\quad 3A + B \xrightarrow{10} 2C$

v $\quad \phi \xrightarrow{k_{s1}} A$

$\quad \phi \xrightarrow{k_{s2}} B$

$\quad 2A + 3B \underset{k_r}{\overset{k_f}{\rightleftharpoons}} C + 2D$

$\quad C \xrightarrow{k_d} \phi$

[1] Material prepared by Dr Rob Smith: robert1.smith@wur.nl

Note that in parts *iv* and *v* we have a 4-th and 5-th order reaction (i.e. the number of molecules of A and B required for the reaction is equal to 4 or 5). These are not very realistic reactions as they imply that 5 things/molecules come together instantaneously — it is far more common that the reaction will happen in sequential steps, i.e. one A collides with one B and then a second B in separate reactions, rather than all molecules colliding at once.

Solutions:

i

$$\frac{dA}{dt} = k_r C - k_f AB$$
$$\frac{dB}{dt} = k_r C - k_f AB$$
$$\frac{dC}{dt} = k_f AB - k_r C - k_d C$$

ii

$$\frac{dA}{dt} = 3 - 2k_f A^2$$
$$\frac{dB}{dt} = \frac{dC}{dt} = k_f A^2$$

Note here that the dA/dt equation is a sum of the two equations and that only the second term/reaction is required to satisfy the relationship

$$\frac{1}{2}\frac{dA}{dt} = -\frac{dB}{dt} = -\frac{dC}{dt}.$$

Other reactions alter the dynamics of the system independently.

iii

$$\frac{dA}{dt} = 0.1B^2 - 2A$$
$$\frac{dB}{dt} = 4A - 0.2B^2$$

iv

$$\begin{aligned}\frac{dA}{dt} &= -30A^3B\\ \frac{dB}{dt} &= -10A^3B\\ \frac{dC}{dt} &= 20A^3B\end{aligned}$$

v

$$\begin{aligned}\frac{dA}{dt} &= k_{s1} + 2\big(k_rCD^2 - k_fA^2B^3\big)\\ \frac{dB}{dt} &= k_{s2} + 3\big(k_rCD^2 - k_fA^2B^3\big)\\ \frac{dC}{dt} &= k_fA^2B^3 - k_rCD^2 - k_dC\\ \frac{dD}{dt} &= 2\big(k_fA^2B^3 - k_rCD^2\big) \qquad (1)\end{aligned}$$

This questions uses the same process as in problem **ii**.

Part 2: Solving ODEs #1

We have already seen in lectures how to solve some 1-dimensional systems (1-dimensional as they contain 1 species) such that if we know $dx/dt = f(x(t))$ then we can calculate $x(t)$. Here, we will look at how the computer solves ODE systems using variants of the *Euler approximation*.

For this exercise you are free to use an Excel spreadsheet to store values and help with calculations.

Over the coming weeks you are going to use computer algorithms generally referred to as *ODE solvers* to calculate the solution of our models. However, most solvers have been developed from a simple approximation derived by Euler. This approximation says

$$\frac{dx}{dt} \simeq \frac{x(t+\Delta t) - x(t)}{\Delta t} = f(x(t))$$

where Δt is the size of the time-step you wish to use. This means that to calculate the next time-point, $t + \Delta t$, in a time-series you just need to plug the existing numbers into the equation

$$x(t + \Delta t) = x(t) + \Delta t f(x(t)).$$

Given that $f(x(t)) = -2x(t)$ and $x(0) = 1$, you are asked to do three things:

i Solve $\frac{dx}{dt} = -2x$ to obtain the true solution $x_T(t)$.

ii Use the Euler approximation with $\Delta t = 0.2$. Calculate the values of $x_E(t)$ from $t = 0$ to $t = 1$. Compute the average error of the N points between x_T and x_E by summing the squared difference, i.e.

Error $= \frac{1}{N}\sum_{i=0}^{N}(x_T(t_i) - x_E(t_i))^2$.

Recall here that $\sum_{i=0}^{N} f_i$ means that you need to sum the expression f_i from 0 to N, i.e. the sum $= f_0 + f_1 + \ldots + f_{N-1} + f_N$.

iii Pick a smaller value of $\Delta t = 0.1$, compute a new time-series for $x_E(t)$ and the error. What does this tell you about the error between the Euler approximation and the true solution as $\Delta t \to 0$?

Solutions:

i

This equation can be rewritten as

$$\int \frac{dx}{x} = -2\int dt$$
$$\Rightarrow \ln(x) = -2t + C$$
$$x(t) = Ce^{-2t},$$

and using $x(0) = 1$ we find that $C = 1$, so $x_T(t) = e^{-2t}$.

ii + iii

By substituting the appropriate terms and values into the formulae you should get (to 4 decimal places)

t	$x_T(t)$	$x_E(t)$ $\Delta t = 0.2$	$\Delta t = 0.1$	Error $\Delta t = 0.2$	$\Delta t = 0.1$
0	1	1	1	0	0
0.1	0.8187		0.8		0.0004
0.2	0.6703	0.6	0.64	0.0049	0.0009
0.3	0.5488		0.512		0.0014
0.4	0.4493	0.36	0.4096	0.008	0.0016
0.5	0.3679		0.3277		0.0016
0.6	0.3012	0.216	0.2621	0.0073	0.0015
0.7	0.2466		0.2097		0.0014
0.8	0.2019	0.1296	0.1678	0.0052	0.0012
0.9	0.1653		0.1342		0.001
1	0.1353	0.0778	0.1074	0.0033	0.0008
			Average	0.0057	0.0012

Part 3: Solving ODEs #2

You have the following ODE system

$$\frac{dX}{dt} = \alpha - \beta X$$
$$\frac{dY}{dt} = \gamma X - \delta Y$$

where $X(0) = Y(0) = 0$. Using the Euler approximation, calculate the first three time-steps using $\Delta t = 1$, with $\alpha = 1$, $\beta = 0.1$, $\gamma = 2$, $\delta = 0.5$.

Now, try to solve the above equations analytically to obtain an expression for $Y(t)$. To help you with this, you can make use of the file *ode_system_solution.docx* from the Brightspace page. What is the error between the analytical solution and the Euler approximation of these early timepoints?

Solutions:

We will first use the Euler approximation to solve the first part of the problem. Converting from a 1-dimension to 2-dimension system is fairly straightforward, but the next time-step of $X(t+\Delta t)$ and $Y(t+\Delta t)$ depend on $X(t)$ and $Y(t)$. So you can get

$$
\begin{aligned}
X(t+\Delta t) &= X(t) + \Delta t\Big(1 - 0.1X(t)\Big) \\
Y(t+\Delta t) &= Y(t) + \Delta t\Big(2X(t) - 0.5Y(t)\Big)
\end{aligned}
$$

This gives the numbers

t	$X(t)$	$Y(t)$
0	0	0
1	1	0
2	1.9	2
3	2.71	4.8

Next, let's find the solution to the system and obtain an expression for $Y(t)$. Following the steps in the document available on Brightspace, assuming that $\lambda = 0$, you should find that

$$
X(t) = \frac{\alpha}{\beta}\Big(1 - e^{-\beta t}\Big).
$$

Now, we worry about $Y(t)$. By ignoring the first term of dY/dt we can approximate

$$
Y(t) = C_y(t)e^{-\delta t} + D_y
$$

which, upon substitution, gives us

$$
\begin{aligned}
&\frac{dY}{dt} + \delta Y = \gamma X \\
\Rightarrow\ &\frac{dC_y}{dt}e^{-\delta t} - \delta C_y e^{-\delta t} + \delta C_y e^{-\delta t} + \delta D_y = \frac{\gamma\alpha}{\beta}\Big(1 - e^{-\beta t}\Big) \\
\Rightarrow\ &\frac{dC_y}{dt} = -\frac{\gamma\alpha}{\beta}e^{(\delta-\beta)t} + \left(\frac{\gamma\alpha}{\beta} - \delta D_y\right)e^{\delta t}.
\end{aligned}
$$

We can now integrate this to get

$$
\begin{aligned}
C_y(t) &= \frac{-\gamma\alpha}{\beta(\delta-\beta)}e^{(\delta-\beta)t} + \frac{1}{\delta}\left(\frac{\gamma\alpha}{\beta} - \delta D_y\right)e^{\delta t} + K_y, \text{and} \\
Y(t) &= \frac{\gamma\alpha}{\delta\beta} - \frac{\gamma\alpha}{\beta(\delta-\beta)}e^{-\beta t} + K_y e^{-\delta t},
\end{aligned}
$$

where K_y is the integration constant. We can now use the initial condition to get

$$K_y = \frac{\gamma\alpha}{\beta(\delta-\beta)} - \frac{\gamma\alpha}{\delta\beta} = \frac{\gamma\alpha}{\delta(\delta-\beta)}.$$

Upon substitution of the parameter values, we have

$$Y(t) = 40 - 50e^{-0.1t} + 10e^{-0.5t}$$

and we can calculate the first three timepoints and their error

t	$Y_{Analytical}(t)$	$Y_{Euler}(t)$	Error
0	0	0	0
1	0.82	0	0.82
2	2.74	2	0.74
3	5.19	4.8	0.39

This shows that with a time-step of $\Delta t = 1$ our Euler approximation of the true solution was not very good!

Part 4: Checking units and rescaling ODEs

In the final slide of Wednesday's lecture, we had the following system

$$\begin{aligned}\frac{dX_1}{dt} &= \alpha R - \beta X_1 \\ \frac{dX_2}{dt} &= \alpha\gamma X_1 - \beta X_2\end{aligned}$$

where α is known to be 2 min^{-1}, β is 5 sec^{-1}, R and X_1 are measured in μM, and X_2 is measured in nM. You are also given that $R = 10\mu$M, $X_1(0) = 2\mu$M and $X_2(0) = 0$ nM. We will also assume that we want to model time, t, in minutes. Are the equations balanced such that the units on the left hand side of the equation match those on the right hand side?

Next, you are given that $\gamma = 1000$ nM/μM. With this scaling factor, how can the model be rewritten such that both equations have dimensions of nM/min on both the left and right hand side of the equals sign?

Finally, let's assume we have data measured every hour for 10 hours. This would require solving our system for 600 minutes as the rates are on the minute timescale. In these instances it can be useful to rescale the model from minutes to hours. This can allow for an easier comparison between model and data. Rescale the model so that the equations now represent nM / h instead of nM / min. What are the resulting equations?

Solutions:

For the first question, let's check whether the equations balance. To do this you need to check the units of each term. So we get

$$
\begin{aligned}
&\frac{dX_1}{dt} = \alpha R - \beta X_1 \\
\Rightarrow\ &\frac{\mu M}{min} = \frac{\mu M}{min} - \frac{\mu M}{sec} \\
&\frac{dX_2}{dt} = \alpha\gamma X_1 - \beta X_2 \\
\Rightarrow\ &\frac{nM}{min} = \frac{\mu M}{min}\frac{nM}{\mu M} - \frac{nM}{sec} \\
\Rightarrow\ &\frac{nM}{min} = \frac{nM}{min} - \frac{nM}{sec}
\end{aligned}
$$

and you can see that the left and right hand side of the equals signs do not match. Therefore our equations are inconsistent/incorrect.

Let's start with the first problem to hand. We are given that $\gamma = 1000$ nM/μM and we want the ODEs to have dimensions/units of nM/min on both sides of the equals sign. To solve this problem we need to rescale X_1. We can do this by using the transformation that

$$\tilde{X}_1 = \gamma X_1$$

where $\tilde{X}_1$ has units of nM.

This transformation implies that

$$
\begin{aligned}
&\frac{d\tilde{X}_1}{dt} = \frac{d(\gamma X_1)}{dt} = \gamma\frac{dX_1}{dt} \\
\Rightarrow\ &\frac{d\tilde{X}_1}{dt} = \alpha\gamma R - \beta\gamma X_1.
\end{aligned}
$$

Now we can notice that, on the right hand side of the equals sign, we can replace γX_1 with $\tilde{X}_1$, and we can also replace γR with a new variable $\tilde{R}$ that has units of nM too. This gives us

$$\frac{d\tilde{X}_1}{dt} = \alpha\tilde{R} - \beta\tilde{X}_1$$

with $\tilde{X}_1(0) = 2000$ nM, and $\tilde{R} = 10000$ nM. However, β is still given in units of 1/sec when we want 1/min. The easiest way to solve this is to spot that

$$60\beta \sim \frac{sec}{min}\frac{1}{sec} \sim \frac{1}{min}$$

so using $\tilde{\beta} = 60\beta$ will generate an ODE with units of min throughout:

$$\frac{d\tilde{X}_1}{dt} = \alpha\tilde{R} - \tilde{\beta}\tilde{X}_1.$$

The last change we need to make is to replace X_1 in the second equation to give

$$\begin{aligned}\frac{dX_2}{dt} &= \alpha\gamma\frac{\tilde{X}_1}{\gamma} - \tilde{\beta}X_2\\ \Rightarrow \frac{dX_2}{dt} &= \alpha\tilde{X}_1 - \tilde{\beta}X_2\end{aligned}$$

The last part of the question asks us to rescale our ODEs from minute to hour timescale. The way to solve this problem is to rescale time such that

$$\tilde{t} = \frac{t}{60} \sim \frac{min}{min.h^{-1}} = h$$

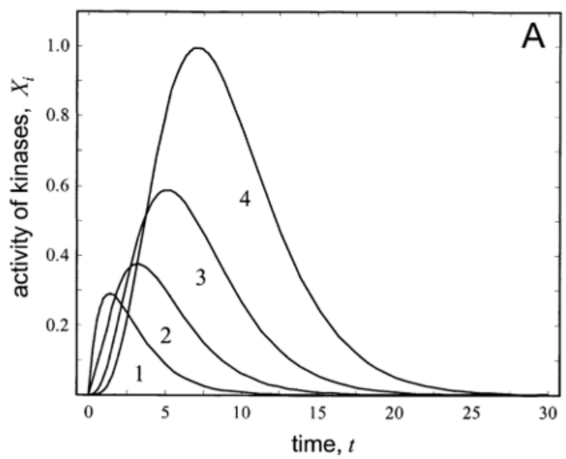
i.e. $\tilde{t}$ has units of h. We can now rewrite our ODEs in terms of $\tilde{t}$ using this transformation. So we get

$$\begin{aligned}\frac{d\tilde{X}_1}{d\tilde{t}} &= \frac{d\tilde{X}_1}{d\frac{t}{60}} = 60\frac{d\tilde{X}_1}{dt}\\ \Rightarrow \frac{d\tilde{X}_1}{d\tilde{t}} &= 60\alpha\tilde{R} - 60\tilde{\beta}\tilde{X}_1\\ \frac{dX_2}{d\tilde{t}} &= \frac{dX_2}{d\frac{t}{60}} = 60\frac{dX_2}{dt}\\ \Rightarrow \frac{dX_2}{d\tilde{t}} &= 60\alpha\tilde{X}_1 - 60\tilde{\beta}X_2.\end{aligned}$$

By rescaling time from min to h, our ODEs have "sped up"!

Part 5: Using generative AI for graphical interpretation

One possible usage of generative AI is to help us interpret and make conclusions from graphs. For this purpose, we are going to use one of the graphs of our kinase network from Wednesday's lecture. To do this, take a screenshot of the graph below and use the prompt "Please interpret this graph for me" in your favourite generative AI tool.



Compare the output you obtain with what you learnt from lectures.

Rob's interpretation: the figure the activity of four active kinase protein's that are denoted X_i. The activity of X_i is determined from phosphorylation cycle i that take place in a sequential order: i.e. phosphorylation cycle 1 leads to activation of X_1 that then triggers activation of phosphorylation cycle 2 and increases in X_2 activity, and so on. Due to the step wise amplification, the peaks of X_i activity increase in amplitude and are delayed across time with the peak of X_4 being higher and delayed relative to the peak of X_1. After the peak has been passed, the activity of the phosphorylated protein decreases as the initial triggering input is removed from the system.

How does the explanation from generative AI compare with this explanation?

Solution:

Using the graph and prompt in the free version of ChatGPT, I obtained the following explanation:

This graph shows the activity of kinases X_i as a function of time t, with four distinct curves labeled 1, 2, 3, and 4.

Key Observations:

- General Trend:

  The graph shows that the kinase activity starts at a low value, increases over time, peaks, and then decreases again.

  Each curve represents a different instance of kinase activity, with the peak activity occurring at different times and at different magnitudes for each curve.

- Time to Peak:

  The time at which the peak occurs seems to decrease as the curve number increases (i.e., curve 4 peaks at the earliest time, and curve 1 peaks at the latest). This could suggest that the system represented by the graph experiences faster kinase activation as it progresses through different conditions or stimuli.

- Peak Height:

  The peak activity decreases from curve 1 to curve 4. Curve 1 reaches nearly 1.0, the maximum value for kinase activity, while curve 4 peaks at a lower value, around 0.5. This suggests a decrease in the maximum response or efficiency as time progresses or with increasing stimuli or conditions.

- Decay Phase:

  After reaching the peak, all the curves decline, and the kinase activity returns toward 0. This indicates that the kinases are eventually deactivated or that their activity is reduced due to some regulatory mechanism. The decay is slower for curves 1 and 2 and faster for curves 3 and 4, suggesting varying rates of deactivation or removal of the stimulus.

Possible Interpretations:

- Different Stimulus or Conditions: The different curves may represent different experimental conditions or stimulus intensities. The faster activation and decay in curves 3 and 4 could indicate that a higher stimulus leads to a quicker response, whereas curves 1 and 2 may represent a slower, more sustained response to a weaker stimulus.
- System Kinetics: If this graph is describing a biological or biochemical system (like kinase activation in a signaling pathway), these curves could indicate how the system responds to different initial conditions or perturbations. This could be related to feedback loops or other regulatory mechanisms that control the activity of kinases.
- Rate of Deactivation: The faster decay for curves 3 and 4 might suggest faster turnover or deactivation of kinases in response to a stronger or more immediate regulatory feedback.

In summary, this graph suggests that the kinase activity follows a time-dependent pattern that is influenced by the initial conditions or stimuli, with the system exhibiting both activation and deactivation phases. The overall behavior hints at a feedback-regulated process that differs in speed and intensity across the different conditions.

Rob's comment on this answer: to be fair to ChatGPT, it has picked up on some key ideas from the graph alone. For example, the use of "kinase" on the y-axis has led to discussion on kinase chains and that different activation speeds may be related to differing stimuli of phosphorylation cycles. The general trend section also appears accurate, although ChatGPT has clearly got confused by the labelling of curves where the number order has been reversed! I can only assume that this is because the "1" label has been assumed to be linked to the increase of curve 4 rather than the decay of curve 1 which we can observe.