Dugga 2024-03-06

TBMT19 / TBMT37

Write your Dugga-ID on all pages and your answers in Swedish or English. There are in total 5 questions, each worth 3 points. You need at least 2 points per question or 12 points in total to pass. Good luck! /William

1 Model formulation and model parts

You have the following model:

$$\frac{dA}{dt} = -v1 + v2 \qquad v1 = k1 \cdot A$$
$$\frac{dB}{dt} = v1 - v2 - v3 \qquad v2 = k2 \cdot B \cdot u$$
$$\frac{dC}{dt} = v3 - v4 \qquad v3 = k3 \cdot B$$
$$\frac{dD}{dt} = v4 \qquad v4 = k4 \cdot C \cdot A$$
$$[k1, k2, k3, k4] = [1, 2, 4, 1]$$
$$[A(0), B(0), C(0), D(0)] = [1, 0, 0, 1]$$
$$u = 10$$

(a) Give an interaction graph which represents the model. (1 point)

(b) List the model states and the model parameters, and explain what the difference between states and parameters is? (1 point)

(c) Make the reaction $B \rightarrow C$ saturated with respect to B. (1 point)

2 Model simulation

(a) Use the Euler forward method to simulate the following model, with a step-length of 0.1. What are the values of A and B in t = 0.1, and t = 0.2? (2 points)

$$
\begin{aligned}
d/dt(A) &= -v1 + v2 \qquad & v1 &= k1 \cdot A \cdot u \\
d/dt(B) &= \quad v1 - v2 \qquad & v2 &= k2 \cdot B
\end{aligned}
$$

$$
\begin{aligned}
&[k1, k2] = [4, 1] \\
&[A(0), B(0)] = [20, 30] \\
&u = 1
\end{aligned}
$$

(b) Why do we need to numerically simulate systems biology models? (1 point)

3 Parameter estimation

(a) You have a model simulation ($\hat{y}$) of four different time-points for a given parameter set (θ):

$$\hat{y}(\theta) = [12, 15, 23, 30]$$

You also have corresponding experimental data (y) with measurement uncertainties (SEM):

$$y = [8, 17, 21, 33], \qquad SEM = [2, 1, 4, 3]$$

What is the cost for the parameter set θ? (1 points)

(b) What are the input/output to the optimization algorithms? (1 point)

(c) What is the difference between *local* and *global* optimization algorithms? (1 point)

4 Statistical tests

Assume that we again have the same model simulation and data as in question 3 above.

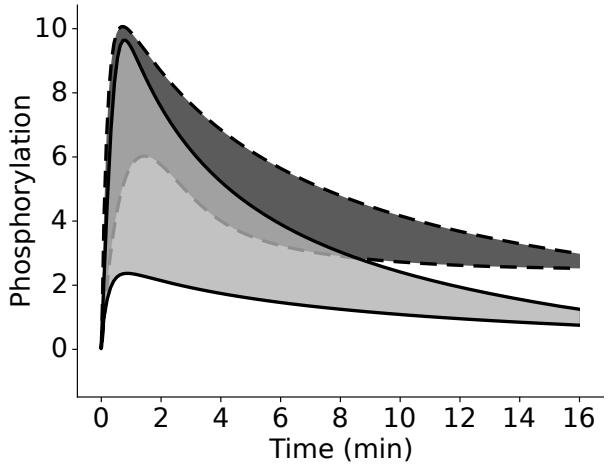
(a) If the χ^2-limit is 9.48, should you reject the model simulation given the parameter set θ? What is the conclusion of the test? (1 point)
If you did not calculate the cost, assume that the cost is 8.5.

(b) What is the null hypothesis of a Whiteness-test, and what do you conclude if you *reject* the model simulation with the Whiteness-test? (1 point)

(c) Based on your conclusion from 4(a), should you reject the *model*? (1 point)

5 Predictions and experimental design

Assume that you have the following model predictions from two different models. The first model is represented by the light gray area with ~~striped~~ solid edges, and the second model is represented by the dark gray area with ~~solid~~ striped edges. The predictions correspond to the phophorylation of a protein over time.



(a) If you could ask an experimentalist to measure the phosphorylation at only one time-point, which time-point would you pick? What could be the potential outcomes when you get the data? Motivate your answer! (2 points)

(b) Would there be any benefit of measuring in t = 2? (1 point)

Answers: Dugga 2024-03-06

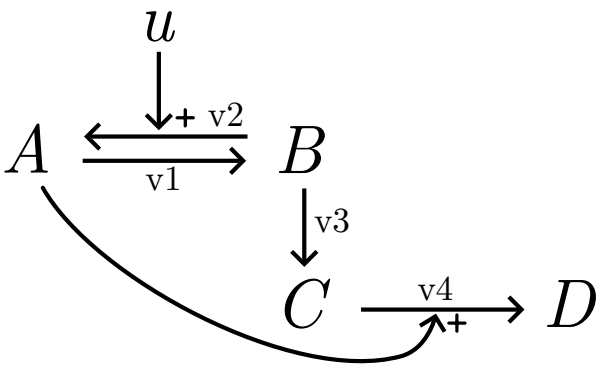
1 Model formulation and model parts

You have the following model:

$$\frac{dA}{dt} = -v1 + v2 \qquad v1 = k1 \cdot A$$
$$\frac{dB}{dt} = v1 - v2 - v3 \qquad v2 = k2 \cdot B \cdot u$$
$$\frac{dC}{dt} = v3 - v4 \qquad v3 = k3 \cdot B$$
$$\frac{dD}{dt} = v4 \qquad v4 = k4 \cdot C \cdot A$$
$$[k1, k2, k3, k4] = [1, 2, 4, 1]$$
$$[A(0), B(0), C(0), D(0)] = [1, 0, 0, 1]$$
$$u = 10$$

(a) Give an interaction graph which represents the model. (1 point)

Answer:
The interaction graph should look something like this:



(b) List the model states and the model parameters, and explain what the difference between states and parameters is? (1 point)

Answer:
The model states are: A, B, C, and D. The model parameters are: k1, k2, k3, and k4. The initial conditions are also often considered to be parameters: A(0), B(0), C(0), and D(0).

The difference between states and parameters is that the states are the variables that change over time, while the parameters are constants that do not change over time.

(c) Make the reaction $B \rightarrow C$ saturated with respect to B. (1 point)

Answer:
The reaction $B \rightarrow C$ corresponds to the reaction rate v3. To make v3 saturated with respect to B we can update the the reaction rate to:

$$v3 = Vmax \cdot \frac{B}{K_m + B} \tag{1}$$

2 Model simulation

(a) Use the Euler forward method to simulate the following model, with a step-length of 0.1. What are the values of A and B in t = 0.1, and t = 0.2? (2 points)

$$\begin{aligned} d/dt(A) &= -v1 + v2 \qquad v1 = k1 \cdot A \cdot u \\ d/dt(B) &= \quad v1 - v2 \qquad v2 = k2 \cdot B \end{aligned}$$

$$\begin{aligned} &[k1, k2] = [4, 1] \\ &[A(0), B(0)] = [20, 30] \\ &u = 1 \end{aligned}$$

Answer:
In the Euler forward method, we start with the initial values of the states and calculate the values of the ODEs at the initial time-point. We then take a small step in the direction of the ODEs, and use the new values to calculate the ODEs at the next time-point. This is repeated until the time-frame asked for by the user has been simulated.

The Euler forward method can be written as:

$$x(t + \Delta t) = x(t) + d/dt(x(t)) \cdot \Delta t$$

For this model, to calculate the values of A and B in t = 0.1, and t = 0.2, we start by calculating the values for t = 0.1 using the following steps:

$$v1(0) = k1 \cdot A(0) \cdot u = 4 \cdot 20 \cdot 1 = 80$$
$$v2(0) = k2 \cdot B(0) = 1 \cdot 30 = 30$$

$$d/dt(A(0)) = -v1(0) + v2(0) = -80 + 30 = -50$$
$$d/dt(B(0)) = v1(0) - v2(0) = 80 - 30 = 50$$

$$A(0.1) = A(0) + d/dt(A(0)) \cdot \Delta t = 20 - 50 \cdot 0.1 = 15$$
$$B(0.1) = B(0) + d/dt(B(0)) \cdot \Delta t = 30 + 50 \cdot 0.1 = 35$$

Then we calculate the values for t=0.2 using the following steps:

$$v1(0.1) = k1 \cdot A(0.1) \cdot u = 4 \cdot 15 \cdot 1 = 60$$
$$v2(0.1) = k2 \cdot B(0.1) = 1 \cdot 35 = 35$$

$$d/dt(A(0.1)) = -v1(0.1) + v2(0.1) = -60 + 35 = -25$$
$$d/dt(B(0.1)) = v1(0.1) - v2(0.1) = 60 - 35 = 25$$

$$A(0.2) = A(0.1) + d/dt(A(0.1)) \cdot \Delta t = 15 - 25 \cdot 0.1 = 12.5$$
$$B(0.2) = B(0.1) + d/dt(B(0.1)) \cdot \Delta t = 35 + 25 \cdot 0.1 = 37.5$$

So, the values of A and B in t = 0.1, and t = 0.2 are:

$$A(0.1) = 15, \quad B(0.1) = 35$$
$$A(0.2) = 12.5, \quad B(0.2) = 37.5$$

(b) Why do we need to numerically simulate systems biology models? (1 point)

Answer:
In systems biology, we often have complex models with many states and reactions, and it is often not feasible to solve these models analytically. Therefore, we need to use numerical simulation to solve the models and obtain the model simulations.

3 Parameter estimation

(a) You have a model simulation ($\hat{y}$) of four different time-points for a given parameter set (θ):

$$\hat{y}(\theta) = [12, 15, 23, 30]$$

You also have corresponding experimental data (y) with measurement uncertainties (SEM):

$$y = [8, 17, 21, 33], \qquad SEM = [2, 1, 4, 3]$$

What is the cost for the parameter set θ? (1 points)

Answer:
The cost for the parameter set θ is calculated as follows:

$$\text{residuals} = y - \hat{y} = [8, 17, 21, 33] - [12, 15, 23, 30] = [-4, 2, -2, 3]$$

$$v(\theta) = \sum_{t=1}^{N} \frac{(y(t) - \hat{y}(t, \theta))^2}{SEM(t)^2} = \frac{(-4)^2}{2^2} + \frac{2^2}{1^2} + \frac{(-2)^2}{4^2} + \frac{3^2}{3^2}$$

$$= \frac{16}{4} + \frac{4}{1} + \frac{4}{16} + \frac{9}{9} = 4 + 4 + \frac{1}{4} + 1 = 9.25$$

(b) What are the input/output to the optimization algorithms? (1 point)

Answer:
The input is a start guess of the parameter values, and the output is the optimized (better but not necessarily optimal) parameter values.

(c) What is the difference between *local* and *global* optimization algorithms? (1 point)

Answer:
Local optimization algorithms only search in directions which strictly gives an improvement of the cost, while global optimization algorithms can search in directions which gives a temporary increase of the cost. In other words, local optimization algorithms can only walk downhill in the cost landscape, while global optimization algorithms can also walk uphill sometimes.

Typically, local optimization algorithms are faster than global optimization algorithms, but global optimization algorithms are more likely to find better parameter values.

4 Statistical tests

Assume that we again have the same model simulation and data as in question 3 above.

(a) If the χ^2-limit is 9.48, should you reject the model simulation given the parameter set θ? What is the conclusion of the test? (1 point)
If you did not calculate the cost, assume that the cost is 8.5.

Answer:
The χ^2-limit is 9.48, and the cost is 9.25 (or 8.5). Since the cost is smaller than the χ^2-limit, we cannot reject the model simulation given the parameter set θ. The conclusion of the test is that we accept the model for now, and must continue with more tests (e.g. predictions).

(b) What is the null hypothesis of a Whiteness-test, and what do you conclude if you *reject* the model simulation with the Whiteness-test? (1 point)

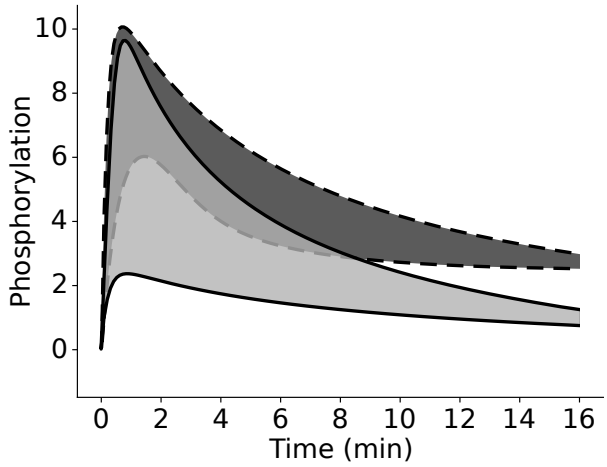
Answer:
The null hypothesis of a Whiteness-test is that the residuals are not correlated. If we reject the model simulation with a Whiteness-test, this means that the residuals are *too* correlated.

(c) Based on your conclusion from 4(a), should you reject the *model*? (1 point)

Answer:
No. Since we did not reject the model simulation (given θ) with the χ^2-test in 4(a), we have at least one parameter set which gives simulations that agrees with the data ”good enough”. Therefore, we cannot reject the model, and we tentatively accept the model and need to do further tests (such as making predictions).

5 Predictions and experimental design

Assume that you have the following model predictions from two different models. The first model is represented by the light gray area with solid edges, and the second model is represented by the dark gray area with striped edges. The predictions correspond to the phosphorylation of a protein over time.



(a) If you could ask an experimentalist to measure the phosphorylation at only one time-point, which time-point would you pick? What could be the potential outcomes when you get the data? Motivate your answer! (2 points)

Answer:
Long answer: I would pick the time-point where both of the two model predictions have a low uncertainty, and where they differ the most from each other. Given these predictions, the best point would be at t = 16, because here the models predictions are the most different from each other, and the prediction uncertainty is low. Since the predictions differ between the models, we will when we do the measurement get one of three possible outcomes: 1) the data agrees with the prediction from the light gray model, and we therefore reject the dark gray model, 2) The data agrees with the dark gray model prediction, and we reject the light gray model, or 3) the data does not agree with either of the model predictions, and we reject both models.

Short answer (equally correct): Pick t=16 because this is where the model prediction uncertaintes are small and differ the most. If the experimental data has a reasonable uncertainty, we will be able to reject at least one of the models.

(b) Would there be any benefit of also measuring in t = 2? (1 point)

Answer:
Yes! Since both of the models have a high uncertainty in the prediction at t = 2, we can use the experimental data to reject some of the simulations which lie outside the new data, and thus reduce the uncertainty of the model predictions. Also, if the data ends up being outside both model uncertaintes, we should reject both models.