CHALMERS, GÖTEBORGS UNIVERSITET

EXAM for
COMPUTATIONAL BIOLOGY A

COURSE CODES: **FFR 110, FIM740GU, PhD**

Time:	March 15, 2018, at $14^{00} - 18^{00}$
Place:	Johanneberg
Teachers:	Kristian Gustafsson, 070-050 2211 (mobile), visits once around 15^{00}
Allowed material:	Mathematics Handbook for Science and Engineering
Not allowed:	any other written material, calculator

Maximum score on this exam: 12 points (need 5 points to pass).
Maximum score for homework problems: 18 points (need 7 points to pass).
CTH $\geq$15 grade 3; $\geq$20 grade 4; $\geq$25 grade 5,
GU $\geq$15 grade G; $\geq$ 23 grade VG.

---

1. Short questions [3 points] For each of the following questions give a concise answer within a few lines per question.

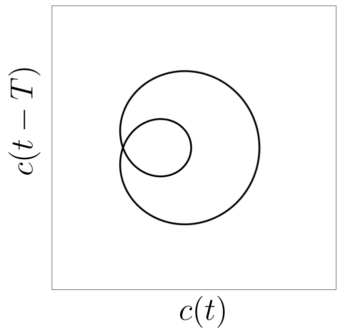
a) When we analyze growth models we often use dimensionless units. Explain what the advantage of using dimensionless units is.

Solution
It is instructive to convert to dimensionless units in problems that are governed by many dimensional parameters, for example in typical growth models. Using dimensionless units allows us to identify the minimal number of dimensionless parameters. It also allows for easier comparison between different dynamical quantities and different parameters. In dimensionless form the magnitude of all dimensionless parameters can be directly related to the values of other parameters or to unity (this simplifies numerical simulations or investigation of limiting cases). In contrast, dimensional parameters must be compared to all other combinations of parameters with the same dimension.

b) Consider a system with a single time delay T for the concentration c:

$$\dot{c} = f(c(t), c(t-T))\,.$$

The delay embedding below shows an example of $c(t-T)$ against $c(t)$. But in one-dimensional dynamical systems without time delay, the existence and uniqueness theorem states that trajectories cannot cross. Explain why the curves may cross in the delay embedding below.



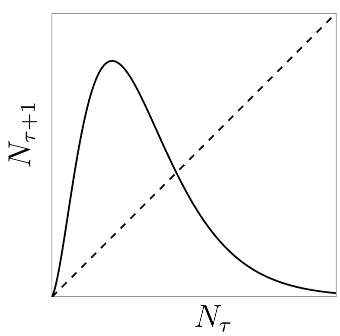
Solution

In the delay embedding the change in $c(t)$ at any time t is determined by not only $c(t)$, but also $c(t-T)$. Assuming that two curves of different history intersects at $(c(t), c(t-T))$, then at the next time step both curves take the same value $c(t+\delta t)$, but the value of $c(t-T+\delta t)$ may differ. Therefore the curves may move to different locations in the delay embedding at $t+\delta t$ even though they were at the same location at t. To have a unique evolution of two curves, they must be equal for the entire interval from $t-T$ to t.

c) Consider a discrete growth model for a single species of population size N:

$$N_{\tau+1} = F(N_\tau)\,.$$

The figure below shows a particular choice of the map F (solid line) and the curve $N_{\tau+1} = N_\tau$ (dashed line). The scales of the axes are equal. Classify all fixed points in this system with respect to their stability and whether they show oscillations.



Solution

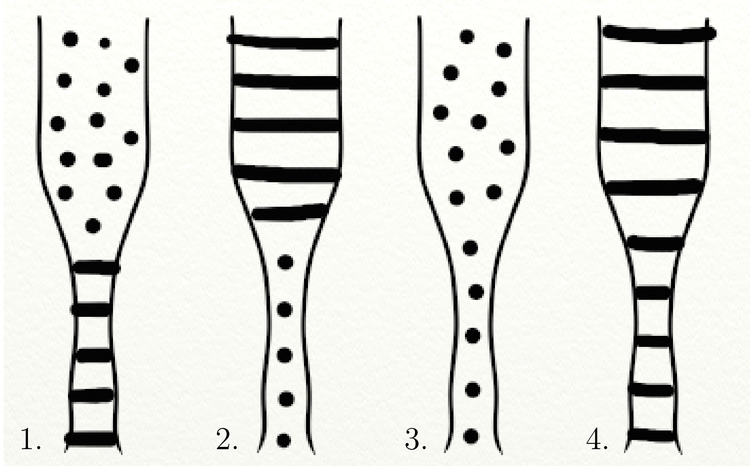
The system has two fixed points where $N_{\tau+1}$ intersects N_τ. They are both unstable because the slopes of the map at the fixed points are larger than unity. The slope at the fixed point $N^* = 0$ is positive and there is therefore no oscillations in the vicinity of $N^* = 0$. The slope evaluated at the second fixed point is negative and therefore it exhibits oscillations in its vicinity.

d) Explain what Brownian motion is. How does it differ from population diffusion of biological species?

Solution

Lecture notes 6.1

e) Which (could be more than one) of the patterns below are consistent with being formed in a Turing instability (diffusion driven instability)? Explain your answer.



Solution

Answer: 1,3,4. When the domain size becomes smaller in the horizontal direction, the values of the allowed wave numbers in the horizontal direction becomes either smaller or remain constant. This is consistent with pattern 1, but inconsistent with pattern 2. For patterns 3 and 4 the wave number of the pattern does not change significantly, which is also consistent with patterns being formed in a Turing instability.

f) Explain what a metapopulation is. What is the rescue effect?

Note: The theory for this problem is not covered in the course this year

Solution

A metapopulation is a collection of unstable local populations that are distributed on a number of spatially separated patches.

The population at isolated patches eventually goes extinct due to stochastic fluctuations. Adding migration over a critical threshold, parts of the populations from patches with large population could migrate to a patch with small population to rescue it from extinction.

g) Consider a number N of coupled oscillators with phases $\theta_1, \theta_2, \ldots \theta_N$ with the following time evolution (Kuramoto model)

$$\dot{\theta}_i = \omega_i + \frac{K}{N}\sum_{j=1}^{N} \sin(\theta_j - \theta_i)\,. \tag{1}$$

Explain the assumptions made to derive Eq. (1).

Solution

A general ansatz of the dynamics for interaction between N oscillators that oscillate with constant angular velocities in the absence of the interaction is:

$$\dot{\theta}_i = \omega_i + \sum_{j=1}^{N} \Gamma_{ij}(\theta_j - \theta_i)\,.$$

Here interactions only depends on the relative phases between oscillators. Starting from this general form the Kuramoto model assumes that there is full network connectivity of equal weight and that the interactions take the form $\Gamma_{ij}(\theta_j - \theta_i) = K/N \sin(\theta_j - \theta_i)$.

h) This problem is not part of the course material this year.
Discuss and contrast the effects of measurement noise and dynamical noise on a time series which is generated by linear (Malthus) decay (negative growth rate).

Solution
Lecture notes 14.2

2. A model for competition with Neanderthals [2.5 points] Construct and analyze a mathematical model for competition between Neanderthal man (with population size N) and Early modern man (with population size E). Assume that the population sizes are large enough for a continuous model to apply. Assume further that no reproduction occurs between the two species. Assume that both species show linear (Malthus) growth for small population sizes, with equal birth rate, b, and with different death rates. The death rate of Neanderthal man is d (assume that $d < b$). Due to better survival rate, the death rate of Early modern man is sd, where s is a parameter taking a value in the range $0 < s < 1$.
Include competition for resources into the model. Assume that both species are equally competitive, i.e. assume that the per capita growth rates for both species are reduced proportional to the total number of competing individuals.

a) Using the assumptions above derive growth equations for the population sizes N and E.

Solution
One growth model that satisfies the assumptions above is the following:

$$\dot{N} = N(b - \frac{N+E}{K} - d)$$
$$\dot{E} = E(b - \frac{N+E}{K} - sd)$$

b) Analyze your model by finding its biologically relevant fixed points and their linear stability. Discuss the possible long-term behaviours of the system.

Solution
Fixed points:

$(N_1^*, E_1^*) = (0, 0)$ (unstable node)
$(N_2^*, E_2^*) = (0, K(b - ds))$ (stable node)
$(N_3^*, E_3^*) = (K(b - d), 0)$ (saddle)

Long-term behaviour: By sketching the phase portrait it becomes evident that for initial conditions with finite population sizes the neanderthals goes extinct (the dynamics is attracted to the stable node).

c) By rewriting your growth model in terms of logarithmic derivatives:

$$\frac{\mathrm{d}}{\mathrm{d}t}\ln E = \dots$$
$$\frac{\mathrm{d}}{\mathrm{d}t}\ln N = \dots$$

it should be straightforward to find a solution for N in terms of E and t. Write down the relation you find between N, E and t.

Solution

First rewrite

$$d_t \ln N = b - \frac{N+E}{K} - d$$
$$d_t \ln E = b - \frac{N+E}{K} - sd\,.$$

Subtract the equations

$$d_t \ln(N/E) = -d + sd$$
$$\Rightarrow \ln(N/E) = (-d+sd)t + \ln(N_0/E_0)$$
$$\Rightarrow N/E = N_0/E_0 e^{(-d+sd)t}\,.$$

Relation between population sizes

$$N(t) = E(t)\exp[dt(s-1)]$$

d) Historical data from the time of Neanderthals show that the lifetime of an individual (both Neanderthal man and Early modern man) was approximately 35 years and that the time to extinction, $T_{\text{ext.}}$, of the Neanderthals upon contact with Early modern man was $T_{\text{ext.}} \approx 10500$ years.

Use these historical data in your model to roughly estimate the parameter s. Since the population never reaches zero in a continuous system, you can approximate $T_{\text{ext.}}$ as the time where N/E reaches five percent of its initial value (it may be helpful to approximate five percent by $0.05 \approx e^{-3}$). Interpret and discuss the value of s you find.

Solution

Death rate $d = 1/35\,\mathrm{y}^{-1}$. Time to extinction $T_{\text{ext}} = 10500\,\mathrm{y}$. Solve $N/E = 0.05N_0/E_0 \approx e^{-3}N_0/E_0$:

$$\frac{N(t)}{E(t)} = \frac{N_0}{E_0}\exp[dT_{\text{ext.}}(s-1)] = \exp[-3]\frac{N_0}{E_0}$$
$$\Rightarrow dT_{\text{ext.}}(s-1) = -3$$
$$\Rightarrow s = 1 - \frac{3}{dT_{\text{ext.}}} = 1 - \frac{105}{10500} = \frac{99}{100}$$

gives $s = 0.99$. This value lies very close to unity, i.e. a very slight advantage of Early modern man drew the Neanderthals extinct according to our model.

3. Harvesting strategies [2 points] A simple continuous growth model with harvesting is given by

$$\dot{N} = rN\left(1 - \frac{N}{K}\right) - Y(N)\,.$$

Here N is the size of the population, r and K are positive constants, and the yield $Y(N)$ denotes removal rate of the population due to harvesting. A good harvesting strategy (choice of $Y(N)$) is a strategy that gives a large yield Y in the long run, while also allowing the system to quickly recover from perturbations.

a) Consider the case $Y(N) = EN$, where $0 < E < r$ is a constant. Determine the long-term yield of the system. What is the maximal long-term yield?

Solution

Rewrite as a logistic equation

$$\dot{N} = rN\left(1 - \frac{N}{K}\right) - EN = (r-E)N\left(1 - N\frac{r}{K(r-E)}\right)\,.$$

This system has the stable steady state $N^* = K(r-E)/r$. The long-term yield is thus

$$Y(N) = EK(r-E)/r\,.$$

The maximal yield is $Y_{\max} = rK/4$, obtained by choosing $E = r/2$.

b) Redo the analysis in subtask a) for the case $Y(N) = DN^2$, where D is a positive constant.

Solution

Rewrite as a logistic equation

$$\dot{N} = rN\left(1 - \frac{N}{K}\right) - DN^2 = rN\left(1 - N\frac{DK + r}{rK}\right)\,.$$

This system has the stable steady state $N^* = Kr/(DK + r)$. The long-term yield is thus

$$Y(N) = D(Kr/(DK + r))^2\,.$$

The maximal yield is $Y_{\max} = rK/4$, obtained by choosing $D = r/K$.

c) Discuss which of the two harvesting strategies in subtasks a) and b) is better. Explain which properties of this strategy makes it better. To come to a definite conclusion, it may be a good idea to, for each of the two strategies, consider the recovery time due to linear perturbations.

Solution

The two harvesting strategies have the same maximal yield.

Close to the stable fixed point $N^* = K$ in the logistic model, the time scale of recovery due to linear perturbations is

$$\tau = -[f'(N^*)]^{-1} = \frac{1}{r},$$

i.e. τ depends only on the linear growth rate.

For the harvesting strategy $Y = EN$, the linear growth rate is $r - E$ and the time scale becomes $\tau_1 = 1/(r - E)$.

For the harvesting strategy $Y = DN^2$, the linear growth rate is r and time scale becomes $\tau_2 = 1/r$.

Since $0 < r < E$, the second strategy recovers quicker from perturbations for any parameter combinations. For example, at the maximum yield we have $\tau_1 = 2/r$ and $\tau_2 = 1/r$. Since the second harvesting strategy has a shorter time scale of recovery, we expect it to be a better strategy within our model.

This makes sense because the effect of the second strategy is solely to reduce the stable steady state of the system by modifying the effective carrying capacity, while keeping the linear growth and stability intact. The first strategy on the other hand modifies both the carrying capacity and reduce the growth rate and stability of the system.

4. Spirals in reaction diffusion equations [2 points]

a) Give three examples where reaction-diffusion processes are of importance in mathematical models of biological systems.

Solution
E.g. Population migration, infection outbreaks, chemical reactions, BZ reaction, pattern formation and morphogenesis, neural activity in brain, electrical patterns in heart, signalling patterns of slime mold.

b) Typical solutions to reaction-diffusion equations are travelling waves and spiral waves. A typical ansatz for the phase of a simple rotating spiral is

$$\phi(r, \theta, t) = \Omega t \pm m\theta + \psi(r), \tag{2}$$

where ϕ is the phase, t is time, r and θ are the radial and angular coordinates in a polar coordinate system, Ω is a constant parameter, m is a positive integer, and $\psi(r)$ is some function of r.

Explain what is meant by a 'phase' and a 'wave front'.

Give interpretations of Ω, m and $\psi(r)$ in Eq. (2).

Solution

The phase is an angular coordinate that characterises the state of an oscillator. A wave front consists of regions of constant phase.

The parameters Ω (angular velocity), m (number of arms), and $\psi(r)$ (radial dependence of spiral arm) describe the spiral shape of the wave fronts.

c) Sketch the wave fronts for the following set of parameters at $t = 0$:

$$\psi(r) = r\,, \quad m = 2$$

d) Discuss how Eq. (2) could be applied in the context of reaction diffusion equations.

Solution

The form (2) can be used as an ansatz in a reaction-diffusion equation of dimension larger than one (similar to the travelling wave ansatz in order to numerically find rotating spiral solutions.)

5. Disease spreading in large but finite populations [2.5 points]
Assume that a population consists of N (constant in time) individuals. Each individual is either infected by, or susceptible to a disease. Assume that recovered individuals once again become susceptible (SIS model).

a) In the lecture notes and the hand-ins a Master equation was derived that describes the probability ρ to observe n infected individuals. Discuss what it means that this Master equation has a 'quasi-steady state'.

Solution

Bernhard's lecture notes, Section 4

An approximate solution for the quasi-steady state that is valid for large N can be found by an ansatz

$$\rho(I) = \exp\left[-NS_0(I) - S_1(I) - 1/NS_2(I) - \ldots\right]\,,$$

where $I = n/N$ is the ratio of infected individuals. To lowest order in N^{-1}, the dynamics of $I(t)$ and $p(t) = S_0'(I)$ was shown to follow Hamilton's equations

$$\begin{aligned} \dot{I} &= \beta I(1-I)e^{p} - \gamma I e^{-p} \\ \dot{p} &= -\beta(1-2I)(e^{p}-1) - \gamma(e^{-p}-1)\,. \end{aligned} \tag{3}$$

Here β and γ are positive parameters.

b) A disease is said to be endemic if it can sustain a finite number of infected individuals in the long run. Find a condition on the parameters β and γ for which the disease described by Eq. (3) is endemic in the limit $N \to \infty$ (corresponding to $p \to 0$).

Solution

When $p = 0$, we have the dynamics

$$\dot{I} = I[\beta - \gamma - I\beta]$$

This reaches a non-zero steady state if $\beta > \gamma$, which is therefore the condition for the disease to be endemic.

c) In the endemic limit found in subtask b), find all biologically relevant fixed points of Eq. (3) that lies on either of the axes $I = 0$ or $p = 0$ and determine their stability. To speed up this calculation, it may be helpful to first evaluate the trace of the stability matrix (Jacobian) for general points (I, p) and to think about how the flow behaves along the axes $I = 0$ and $p = 0$.

Solution

For $p = 0$ we have the fixed points:

$(I_1^*, p_1^*) = (0, 0)$

$(I_2^*, p_2^*) = (1 - \gamma/\beta, 0)$

When $I = 0$ we get

$$\begin{aligned}
&\beta e^p + \gamma e^{-p} = +\gamma + \beta \\
&e^{2p} - \frac{\gamma + \beta}{\beta} e^p + \frac{\gamma}{\beta} = 0 \\
&\Rightarrow e^p = 1 \text{ (case } p = 0 \text{ already treated) or } e^p = \frac{\gamma}{\beta}
\end{aligned}$$

We have the fixed point:

$(I_1^*, p_1^*) = (0, \log(\frac{\gamma}{\beta}))$

Either by explicit evaluation, or by using that the dynamics is Hamiltonian and hence volume preserving, the trace of the stability matrix is zero. The eigenvalues becomes

$$\lambda_{\pm} = \pm\sqrt{-4 \det \mathbb{J}}\,.$$

Assuming $\det \mathbb{J} \neq 0$, we either have a saddle point or a center. Since $\dot{p} = 0$ when $p = 0$, and since $\dot{I} = 0$ when $I = 0$, the flow moves along the axes and therefore centers are ruled out. All the fixed points must therefore be saddle points. It can be noted that for the case $\det \mathbb{J} = 0$ both eigenvalues are zero, but this is not consistent with a sketch of the flow along the axes $I = 0$ or $p = 0$ in the endemic limit ($\beta > \gamma$).

d) In the endemic limit found in subtask b), the dynamics (3) has one additional biologically relevant fixed point (I^*,p^*) with $I^* > 0$ and $p^* < 0$. You can assume that this fixed point is a center.

Contrast the case $p = 0$ (corresponding to $N \to \infty$) and $p > 0$ (corresponding to large but finite N). In particular, discuss the implications of a finite population size for endemic diseases. It could be helpful to sketch the phase portrait for the dynamics in Eq. (3) for non-negative values of I.

Solution

Bernhard's lecture notes, p. 36