
CALCULUS, 2018-19

Bioinformatics degree

Lists of problems

Toni Guillamon (UPC), Àlex Haro (UB)

List 0. Preliminaries.

List 1. Exponential growth, basic modelling.

List 2. Zeroes, linear approximation, induction, extrema.

List 3. Optimization with modelling.

List 4. Taylor expansions and power series.

List 5. Interpolation and approximation.

List 6. Integration.

List 7. Differential equations.

Problems. List 0. Preliminaries.

1. *Geometric progressions.*

- Let $s_n = 1 + 1/2 + 1/4 + \dots + 1/2^n$. Do you think that $\lim_{n \rightarrow +\infty} s(n) = 2$ is true? Why?
- Let $s_n = 1 + 1.1 + 1.1^2 + \dots + 1.1^n$. Why $\lim_{n \rightarrow +\infty} s(n) = +\infty$?
- Let $s_n = 1 - 1/2 + 1/4 - 1/8 + \dots + \beta 1/2^n$. What is the most coherent expression for β ?
- Let $s_n = 1 + q + q^2 + \dots + q^n$ for $|q| < 1$. Compare s_n with $q s_n$. Using the previous comparison, compute $\lim_{n \rightarrow +\infty} s(n)$.

2. Let a sequence of real numbers $\{y_n\}_n$ be defined recurrently by $y_{n+1} = \lambda y_n - 1$, for $n \geq 0$. Find a general expression for y_n in terms of y_0 .

3. Discuss the limits $\lim_{x \rightarrow \pm\infty} \frac{x^n}{1 + x^n}$ in terms of $n \in \mathbb{N}$.

4. Consider the real function $g(s) = -s + \log(s)$.

(a) What is the domain of definition of g ?

(b) Compute

$$(b1) \lim_{s \rightarrow 0^+} g(s); \quad (b2) \lim_{s \rightarrow -\infty} g(s); \quad (b3) \lim_{s \rightarrow +\infty} g(s).$$

5. *Inequalities.* Find the minimal $N \in \mathbb{N}$ such that

$$(a) \frac{(0.1)^{N+1}}{N+1} < 10^{-4}; \quad (b) \left| \frac{(-1)^N}{(N+1)!} \right| < 10^{-4}.$$

6. Consider the real function $f(n) = \frac{1}{n} (\log(n-1) - \log(n+1))$.

(a) What is the domain of definition ($D(f)$) of f ?

(b) Prove that $f \leq 0$ for any $n \in D(f)$.

(c) Prove that $e^{f(n)} := \exp(f(n)) = \left(\frac{n-1}{n+1}\right)^{\frac{1}{n}}$.

7. Analyze the sign of $\psi(a, c) = 1 - c/a$ in terms of parameters a and c .

8. Solve the equation $x(1-x) = \frac{cx}{a+x}$ with respect to x . Observe that the solution is a function of parameters a and c . Describe the regions of parameters (a, c) where $(1-x)$ and $c/(a+x)$ have 1, 2 or 3 intersections, respectively.

9. Plot the functions $\sin(x)$, $\sin(\pi/2 - x)$, $3 \cos(x)$ and $\cos(3x)$.

Problems. List 1.

1. In 1950, a sample of carbon in the caves of Lascaux (France) gave an average of 0.97 disintegrations of ^{14}C per minute and gram, while for the wood of alive trees the measurements give 6.68 disintegrations of ^{14}C per minute and gram. We aim at estimating the age of the famous paintings in the cave.

- (a) The radioactive decay of ^{14}C in a death organism follows an exponential decay function, with a half-life of about 5580 years. Find the decay constant κ .

Let $y(t)$ be the average number of disintegrations of ^{14}C per minute and gram after a lapse of time t . This follows the law

$$y(t) = e^{-\kappa t} y_0,$$

where κ is the decay constant and y_0 is the average number at $t = 0$. In the present example, the half-time is $t_{1/2} = 5580$ years, from where we can estimate the decay constant:

$$y(t_{1/2}) = \frac{y_0}{2} = e^{-\kappa t_{1/2}} y_0 \Rightarrow \kappa = \frac{\log 2}{t_{1/2}} \simeq 0.12422 \times 10^{-3} \text{ year}^{-1}.$$

- (b) Estimate the age of the carbon sample found in 1950, taking into account that, in an alive tree, the consumption rate compensates the radioactive decay rate, and then, initially the average was 6.68 disintegrations of ^{14}C per minute and gram.

Let t_1 be the lapse of time until 1950, since the paintings were drawn. Then, $y_0 = y(0) = 6.68$, $y_1 = y(t_1) = 0.97$, and t_1 is unknown. From the decay law:

$$y_1 = e^{-\kappa t_1} y_0 \Rightarrow \frac{1}{\kappa} \log \left(\frac{y_0}{y_1} \right) \simeq 15534 \text{ years.}$$

Then, the paintings were drawn around 15600 years ago.

2. Detective Beckett is called to the scene of a crime where a dead body has just been found. She arrives on the scene at 10:23 pm and begins her investigation. Immediately, the temperature of the body is taken and is found to be 26.7°C . The detective checks the programmable thermostat and finds that the room has been kept at a constant 20°C for the past 3 days. After evidence from the crime scene is collected, the temperature of the body is taken once more one hour after the first one, and found to be 25.8°C . Assuming that the victim's body temperature was normal (37°C) prior to death, what time did the victim die?

- (a) To solve the problem, use Newton's Law of Cooling, that describes the cooling of a warmer object to the cooler temperature of the environment. Specifically we write this law as,

$$T(t) = T_e + (T_0 - T_e)e^{-k(t-t_0)},$$

where $T(t)$ is the temperature of the object at time t , T_e is the constant temperature of the environment, T_0 is the initial temperature of the object, at time t_0 , and k is a constant that depends on the material properties of the object.

The unknowns in the model are both t_0 (the initial time) and the constant κ . We know that at time $t_1 = 10:23 \text{ pm} \simeq 22.3833 \text{ h}$ the temperature of the victim's body is $T_1 = T(t_1) = 26.7^\circ\text{C}$, and at time $t_2 = 11:23 \text{ pm} \simeq 23.3833 \text{ h}$, the temperature is $T_2 = T(t_2) = 25.8^\circ\text{C}$. Also, we know that $T_e = 20.00^\circ\text{C}$ and we assume that $T_0 = T(0) = 37.00^\circ\text{C}$. The problem is, then, isolating t_0 and κ from the system:

$$\begin{aligned} T_1 - T_e &= (T_0 - T_e)e^{-k(t_1 - t_0)} \\ T_2 - T_e &= (T_0 - T_e)e^{-k(t_2 - t_0)} \end{aligned}$$

By dividing both equations we obtain

$$\frac{T_1 - T_e}{T_2 - T_e} = e^{\kappa(t_2 - t_1)} \Rightarrow \kappa = \frac{1}{t_2 - t_1} \log \left(\frac{T_1 - T_e}{T_2 - T_e} \right).$$

Since $t_2 - t_1 = 1 \text{ h}$, we obtain $\kappa = 0.144250$. From first equation we obtain

$$t_1 - t_0 = \frac{1}{\kappa} \log \left(\frac{T_0 - T_e}{T_1 - T_e} \right),$$

so that $t_1 - t_0 = 6.4548$ and $t_0 = 22.3833 - 6.4548 = 15.9285$. Hence, the crime was performed around 15:56 h = 3:56 pm.

- (b) Explain Newton's Law of Cooling, emphasizing the role of the parameters. What is the relation between the temperature and its rate of change?

The temperature decreases exponentially fast (with exponent κ), till reaching an equilibrium temperature T_e . Notice that

$$\lim_{t \rightarrow \infty} T(t) = \lim_{t \rightarrow \infty} T_e + (T_0 - T_e)e^{-k(t - t_0)} = T_e.$$

Notice also that T_0 is the temperature at initial time t_0 . The law is obtained by computing the derivative of $T(t)$ and relating it with $T(t)$:

$$\dot{T}(t) = -\kappa(T_0 - T_e)e^{-k(t - t_0)} = -\kappa(T(t) - T_e).$$

Hence, the rate of cooling at a time t is proportional to the difference of the temperature and the equilibrium temperature.

3. Scopolamine is a medication used in the treatment of motion sickness and postoperative nausea. Its half-life is 4.5 hours. In the following questions, present the results with 6 decimal digits.

- (a) If we administrate a dose of $A = 0.330 \text{ mg}$ every day, which is the amount of drug in the bloodstream just after the third administration? The amount of drug in the bloodstream y at a time t during one day is $y(t) = e^{-\kappa t} y_*$, where y_* is the initial amount. Since the half-life of the drug is $T = 4.5$, then

$$\kappa = \frac{\log 2}{T} \simeq 0.154033 \text{ h}^{-1}.$$

In one day, the amount of drug is decreased by a factor

$$\lambda = e^{-\kappa \cdot 24\text{h}} \simeq 0.0248031.$$

Hence, if at the beginning of the first day the amount of drug is $y_0 = A$, at the end it is λA . At the beginning of the second day, just after taking the dose A , the amount of drug is $y_1 = A + \lambda A$, and at the end of the second day, it is $\lambda(A + \lambda A)$. At the beginning of the third day, just after the third administration of drug, the amount is

$$y_2 = A + \lambda(A + \lambda A) = A(1 + \lambda + \lambda^2) \simeq 0.338388 \text{ mg}.$$

- (b) If the drug is administrated chronically on a daily basis, which will be the limiting amount of drug in the body just after each administration? By repeating the previous argumentation, the amount of drug at the beginning of the day $n + 1$, just after the $(n + 1)$ -th administration, is

$$y_n = A(1 + \lambda + \lambda^2 + \dots + \lambda^n) = A \frac{\lambda^{n+1} - 1}{\lambda - 1}.$$

Then, the limiting amount of drug after dose administration is

$$y_\infty = A(1 + \lambda + \lambda^2 + \dots) = \lim_{n \rightarrow \infty} A \frac{\lambda^{n+1} - 1}{\lambda - 1} = \frac{A}{1 - \lambda},$$

since $|\lambda| < 1$ (this is the well-known formula of the sum of all elements of a geometric series). In our problem, $y_\infty = 0.338393 \text{ mg}$.

- (c) If we wish to maintain a steady-state level of 0.5 mg, which would be the initial dose?

Since we wish $y_\infty = 0.5 \text{ mg}$, the initial dose should be

$$A = (1 - \lambda)y_\infty \simeq 0.487598 \text{ mg}.$$

4. Under a given set of assumptions, we can obtain the model

$$r = \ln(n)/a$$

for the fractional yearly population growth (i.e., $r = 0.01$ corresponds to a growth rate of 1% per year) as a function of the average number of children per adult (n) and the average age of the mother when a child is born (a).

- (a) Plot r as a function of n over the reasonable range $0.5 \leq n \leq 2$, with a values of 25, 30, and 35. These three curves should be plotted together on one set of axes. Use this figure to discuss the dependence of growth rate on number of children, focusing on how this dependence changes for different average reproduction ages.

Clearly, r depends on two variables, n and a ; that is, $r = r(n, a)$. To simplify the study of this 2-variable function (we have not studied them yet), this statement

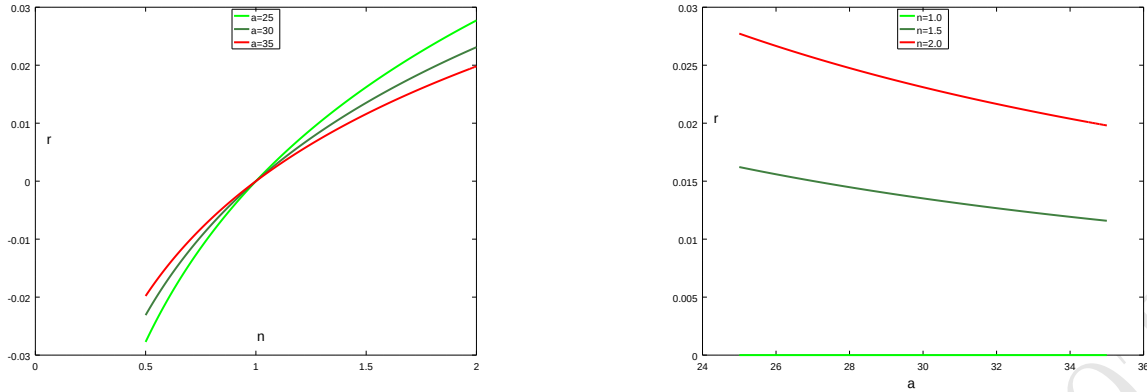


Figure 1: Graph of function $r = \log(n)/a$ (left) for $a = 25$ (green), 30 (blue), 35 (red), see Problem 4a, and (right) for $n = 1$ (green), 1.5 (blue), 2 (red), see Problem 4b.

suggests to look at the variable a as a parameter and plot the graph of $r(n; a)$ for different values of a , see Figure 1.

The plots consists essentially in plotting the logarithm, which is a basic elementary function, and noting that by increasing a we obtain higher values for $0 < n < 1$ (where $\log(n) < 0$) and lower values for $n > 1$ (where $\log(n) > 0$). In other words: take two positive parameters $a_1 < a_2$ (so, $1/a_1 > 1/a_2$). Then, if $n < 1$, $\log(n) < 0$ and $\log(n)/a_1 < \log(n)/a_2$; if $n > 1$, $\log(n) > 0$ and $\log(n)/a_1 > \log(n)/a_2$. Biologically speaking this means that higher (respectively, lower) reproductive ages lead to lower (respectively, higher) population growth rates in absolute values. The logarithm being increasing means that the higher the number of children per mother, the higher the population growth rates, which is a quite obvious argument.

- (b) Plot r as a function of a over the reasonable range $25 \leq a \leq 35$, with n values of 1, 1.5, and 2. These three curves should be plotted together on one set of axes. Use this figure to discuss the dependence of growth rate on average reproductive age, focusing on how this dependence changes for different average numbers of children.

Similarly, this statement suggests to look now at the variable n as a parameter and plot the graph of $r(a; n)$ (note that we reverse the order and use the semicolon to emphasize that n is treated as a parameter) for different values of n . It only consists of plotting $1/a$ and multiplying by $\log(n)$, which increases with n . The biological discussion is similar to that of (4a).

5. Consider a species that has a constant death rate of m individuals per capita per time unit.

- (a) Given a constant per capita death rate, survival of individuals in a cohort born at the same time is mathematically identical to maintenance of a drug in a patient who gets a single dose at time 0. Use this fact to find a formula for the fraction $y(t)$ of individuals alive at age t . In particular, find the fraction s of individuals alive at age 1 as a function of m .

Let $y(t; m)$ be the fraction of individuals alive at time t . Since the per capita death rate is m , then

$$y(t; m) = e^{-mt}.$$

(Notice that at time 0 all individuals of the cohort are alive.). At age 1, the fraction of alive individuals is $s(m) = y(1; m) = e^{-m}$.

- (b) Suppose that individuals of our hypothetical species mature at age 1 and reproduce at constant rate n until age 2. It can be shown that the birth rate necessary for the overall population to stay at the same size is

$$n = \frac{m}{e^{-m} - e^{-2m}}.$$

Combine this formula with the result of part (5a) to obtain an equation that determines the birth rate necessary for population maintenance as a function of the probability of surviving to maturity.

The probability of surviving to maturity is obtained in part (5a): $s = e^{-m}$. Then, the birth rate necessary for population maintenance is

$$n = \frac{m}{e^{-m} - e^{-2m}} = \frac{-\log s}{s - s^2}.$$

- (c) Plot the resulting function $n(s)$. Choose a biologically reasonable range of s values for the plot. Discuss the biological significance of the result by comparing the birth rate requirements for animal species with different probabilities of survival to maturity.

The function to be plotted is $n(s) = \frac{-\log s}{s - s^2}$, in the domain $]0, 1[$. Notice that:

- $\lim_{s \rightarrow 0^+} \frac{-\log s}{s - s^2} = \frac{+\infty}{0^+} = +\infty$.
- $\lim_{s \rightarrow 1^-} \frac{-\log s}{s - s^2} = \frac{0}{0} = \lim_{s \rightarrow 1^-} \frac{-1/s}{1 - 2s} = 1$, where we have applied l'Hôpital rule.
- In order to study the monotonicity of $n(s)$ in the interval $]0, 1[$, consider the derivative $n'(s) = \frac{-1 + s + (1 - 2s) \log s}{s^2(1 - s)^2}$. Since the denominator is positive, we analyze the auxiliary function $g(s) = -1 + s + (1 - 2s) \log s$ in the interval $]0, 1[$. At the extrema of such interval, $\lim_{s \rightarrow 0^+} g(s) = -\infty$ and $g(1) = 0$, and we want to see that $g(s) < 0$ for any $s \in]0, 1[$. In fact, the function $g(s)$ is strictly increasing in $]0, 1[$, since its derivative

$$g'(s) = \left(-1 + \frac{1}{s}\right) + (-2 \log s)$$

is the sum of two positive addends. As a consequence of these arguments, $n'(s) < 0$ in $s \in]0, 1[$, and the function $n(s)$ is strictly decreasing.

The fact that $n(s)$ is strictly decreasing has the natural implication that the bigger is the probability of survival to maturity, s , the smaller is the birth rate necessary for population maintenance, n .

- (d) Use a sequence of s values that rapidly approaches 1 to conjecture the value of

$$\lim_{s \rightarrow 1^-} n(s).$$

We in fact know how to compute such a limit. It is an immediate consequence of l'Hôpital's rule. See above!

6. Influenza is an example of a disease that exists in multiple strains with frequent mutations. This means that individuals who recover from the disease are immune to the current strain, but not to the next strain to reach the population. Now suppose a new strain reaches a closed population. Under these circumstances, the standard model for such diseases can be used to predict the fraction d of individuals who contract the disease in terms of the basic reproductive number R_0 , a parameter that represents the average number of secondary infections due to one infected individual in a population of susceptibles. Values of R_0 can vary from 0 to about 20.6 with measles estimated at 12 – 18 and smallpox at 5 – 7, for instance. If $R_0 < 1$, then the disease can only infect a few individuals before it dies out. If $R_0 > 1$, then a disease epidemic occurs, and the model predicts that the fraction who gets the disease satisfies the equation

$$R_0 d + \log(1 - d) = 0.$$

- Solve the equation for R_0 and use a *guess-and-check* method to estimate the value of d for $R_0 = 5$. Independent of disease mortality, what does this suggest about the impact of smallpox on a Native American population newly exposed by explorers in the late fifteenth century?
- Plot d as a function of R_0 for the range $0 < R_0 < 5$. (Keep in mind that the given formula is not correct for all values of R_0 . Note that tabular data for two related variables can be plotted with either variable on the horizontal axis.)

We first analyze the setting of the problem. It is about finding the zeroes of the function

$$f(d) = R_0 d + \log(1 - d),$$

where R_0 is a positive parameter, and variable d belongs to the interval $[0, 1[$ (since d is a fraction of individuals). Obviously, $d_0 = 0$ is a zero (the trivial one), and we have to look for the other zeroes, if any (the non-trivial zeroes). In fact, the statement of the problem indicates that there should be only one non-trivial solution, for $R_0 > 1$.

Let us start by analyzing the function f . Notice that $f(0) = 0$ and $\lim_{d \rightarrow 1^-} f(d) = -\infty$.

Since $f'(d) = R_0 - \frac{1}{1-d}$, then the only critical point is $d_1 = 1 - \frac{1}{R_0}$. Two cases are in order:

- If $R_0 > 1$, then $d_1 \in]0, 1[$ and, from the previous information, it has to be the global maximum of the function, for which $f(d_1) > 0$ (In particular, $f(d_1) = R_0 - 1 - \log R_0$). Moreover, since $\lim_{d \rightarrow 1^-} f(d) = -\infty$, there is a value $d_2 < 1$, close to 1, such that $f(d_2) < 0$. As a consequence of Bolzano's theorem, there exists a solution of the equation $f(d) = 0$ in the interval $]d_1, d_2[$. But, since the function f is strictly decreasing at $]d_1, 1[$, then the function f has a unique non-trivial zero $d_* = d(R_0)$ in the interval $]d_1, 1[$. (Notice that there are not zeroes in the interval $]0, d_1[$.)
- If $0 < R_0 \leq 1$, then $d_1 \in]-\infty, 0]$. Therefore, the function $f(d)$ is strictly negative in $]0, 1[$, and there are not non-trivial zeroes.

In summary, for any $R_0 > 1$ there exists a unique $d_* = d(R_0) \in]0, 1[$ such that $f(d_*) = 0$. With the notation $d_* = d(R_0)$ we emphasize that the solution is a function of parameter R_0 , and this is the function demanded in part 6b of the problem.

In part 6a, we fix $R_0 = 5$. The guess-and-check method is the bisection method. Starting, for instance, from the interval $[a, b] = [0.8, 0.999]$ in which there is the zero, we repeatedly bisect such interval in such a way that the function is positive at the left extremum, and negative at the right extremum. After 8 steps we get an interval $[a, b] = [0.9931699219, 0.9927812500]$, so that $d_* \simeq 0.993$ with 3 correct digits. Note also that the application of Newton's method, taking the initial approximation 0.999, converges to the solution $d_* \simeq 0.993$ in 3 iterations.

The spreadsheet file `problem4.ods` contains more accurate results (use `libreoffice`, for instance). It is shown that after 31 steps (notice that $(b - a)/2^{31} \simeq 0.9 \cdot 10^{-10}$), we get a solution with 10 correct digits: $d_* \simeq 0.9930228463$. The application of Newton's method, taking the initial approximation 0.999, converges to the solution $d_* \simeq 0.9930228463$ in 6 steps.

Now, you should:

- Complete part 6b. Using the spreadsheet, it is easy now to perform the computations (the C2 box contains R_0). Notice that you probably need to change the initial interval or the initial seed, and that you have to increase or decrease the number of iterations of the algorithms. In the future, you will be able of writing a computer program to obtain the results in a very efficient way.
- The plotted function $d(R_0)$ has domain $[1, +\infty]$, with $d(1) = 0$, is strictly increasing, and $\lim_{R_0 \rightarrow +\infty} d(R_0) = 1$. Why? Notice that the equation is equivalent to $g(d) = R_0$, where

$$g(d) = \frac{-\log(1-d)}{d}.$$

Hence, the plotted function is the inverse function of g , g^{-1} . Observe that the graphs of the functions g and g^{-1} are symmetric with respect to the line $d = R_0$ in the $d-R_0$ plane.

- Play with other initial approximations (seeds) to apply Newton's method. When do they converge to the solution? The explanation is provided by the graph of the function f .
- Redo the problem using other equivalent formulations of the equation. For instance, the equation is equivalent to

$$e^{-R_0 d} - 1 + d = 0.$$

The domain of this function is the whole real line (although we are interested only in $[0, 1]$).

7. A resource x is consumed at the rate

$$f(x) = \frac{Ax}{B+x},$$

where A and B are positive parameters.

- (a) What is the maximum consumption rate in terms of A and/or B , achieved in the limit as $x \rightarrow \infty$?

We need to compute the limit of $f(x)$ when x tends to infinity and make sure that the function does not take a higher value for positive x .

On one side, observe that $f'(x) = \frac{AB}{(B+x)^2}$, which is positive for any $x > 0$ and so, f is increasing in $]0, +\infty[$. Notice also that $B+x$ never vanishes for positive x .

On the other side, $\lim_{x \rightarrow +\infty} f(x) = A$. This means that $f(x)$ is asymptotic to A , and this is the maximum consumption rate. Notice that we never reach this value; although we talk about *maximum consumption rates*, in mathematical language A is said to be the *supremum* of f (a *maximum* is a value that is taken for some value of the variable).

- (b) Suppose the actual consumption rate is half of the maximum from part (7a). What is the corresponding resource level x ?

If $f(x) = A/2$, then $\frac{Ax}{B+x} = \frac{A}{2}$ and, thus, $2Ax = A(B+x)$. Solving for x in terms of A and B , we obtain that $x = B$. This property gives a biological meaning to B , that can be interpreted as the resource level at which the consumption rate is half of the maximum.

CALCULUS, 2018-19-T2

Bioinformatics degree

Problems. List 2.

1. Solve the equation $x(1-x) = \frac{cx}{a+x}$ with respect to x . Observe that the solution is a function of parameters a and c . Describe the regions of the parameter space (a, c) where the functions $x(1-x)$ and $cx/(a+x)$ have 1, 2 or 3 intersections, respectively. [Problem 8, List 0]

In order to solve $x(1-x) = \frac{cx}{a+x}$, we first notice that $x = 0$ is a solution, so that we can consider $(1-x)(a+x) = c$ for $x \neq 0$. Expanding in powers of x , we get: $x^2 + (a-1)x - (a-c) = 0$. Solving this polynomial equation, we obtain two more solutions:

$$x_{\pm} = \frac{1-a \pm \sqrt{(a-1)^2 + 4(a-c)}}{2} = \frac{1-a \pm \sqrt{(a+1)^2 - 4c}}{2}.$$

Therefore: when $c < (a+1)^2/4$, the original equation has 3 solutions, when $c = (a+1)^2/4$, it has 2 solutions and when $c > (a+1)^2/4$, it has 1 solution.

2. Prove the existence of a real number c such that $c^5 - 4c + 1 = 7.21$.

By Bolzano's theorem, it is sufficient to find two values a and b such that $f(a) < 0$ and $f(b) > 0$, for $f(x) = x^5 - 4x + 1$. Since f is an odd degree polynomial, by taking a small enough and b big enough, we get the desired result. For instance, $a = -2$ and $b = 0$ satisfy $f(-2) = -23$ and $f(0) = 1$.

3. Let f and g be continuous functions such that $f(0) < g(0) < g(1) < f(1)$. Prove that there exists at least one point $c \in (0, 1)$ such that $f(c) = g(c)$.

If we consider the function $h(x) := f(x) - g(x)$, it is clear that $h(0) < 0$ and $h(1) > 0$ and so, by Bolzano's theorem (h is continuous), there exists a $c \in (0, 1)$ such that $h(c) = f(c) - g(c) = 0$ (that is, $f(c) = g(c)$).

4. Show that $g(x) = 2^{-x}$ has a unique fixed point on $[1/3, 1]$. Estimate the number of iterations of the bisection method required to achieve the fixed point with 10^{-4} accuracy.

A fixed point must satisfy $2^{-x} = x$. Notice that finding a fixed point is equivalent to find a zeros of $f(x) := 2^{-x} - x$. On one hand, we have that $f(1/3) = 2^{-1/3} - 1/3 = \frac{1}{\sqrt[3]{2}} - \frac{1}{3} \approx 0.46 > 0$ and $f(1) = \frac{1}{2} - 1 = -0.5 < 0$ and so, by Bolzano's theorem (f is continuous), there exists at least a zero of f in $(1/3, 1)$ (equivalently, a fixed point of g). To prove that it is unique, we need to prove that f is monotonous. Clearly, $f'(x) = -2^{-x} \ln(2) - 1 < 0$, which implied that f is always decreasing, which includes of course the interval $(1/3, 1)$.

5. Use the bisection method to find solutions for $x^3 - 7x^2 + 14x - 6 = 0$ on $[0, 1]$, with accuracy 10^{-2} .

Let $f(x) = x^3 - 7x^2 + 14x - 6 = 0$. Note that $f(0) = -6 < 0$ and $f(1) = 2 > 0$; therefore, based on the Intermediate Value Theorem, since f is continuous, there exists a value $p \in (0, 1)$ such that $f(p) = 0$. We take $a_0 = 0$, $b_0 = 1$, which satisfy $f(a_0) < 0$ and $f(b_0) > 0$. Let $p_0 = (a_0 + b_0)/2 = 0.5$. We have $f(p_0) = -0.6250 < 0$ (the same sign as $f(a_0)$); therefore, we take $a_1 = p_0 = 0.5$ and $b_1 = b_0 = 1$ and repeat: $p_1 = 0.75, \dots$ This

yields the following results for p_n and $f(p_n)$:

n	p_n	$f(p_n)$
0	0.5	-0.62500001
1	0.75000000	+0.98437502
2	0.62500000	+0.25976563
3	0.56250000	-0.16186524
4	0.59375000	+0.05404665
5	0.57812500	-0.05262376
6	0.58593750	+0.0010313

Note that, a priori, we know that the number of steps must be bigger than

$$\frac{\ln(L/\epsilon)}{\ln 2} - 1, \quad (1)$$

where L is the length of the initial interval and ϵ is the required precision.

The file `List2-5-2019.ods` contains the computations of this problem.

6. Consider the polynomial $p(x) = x^3 - ax - 1$, where a is a real parameter.

- (a) Find the value of a such that $p(x)$ has a root of multiplicity two (that is, a tangency of the graph of p with the axis $y = 0$).

A root of multiplicity two (tangency with the axis $y = 0$) arises when $p(x) = p'(x) = 0$. In this example, it requires:

$$\begin{aligned} x^3 - ax - 1 &= 0, \\ 3x^2 - a &= 0. \end{aligned}$$

Note that this is a system of nonlinear equations with unknowns x and a . From the second equation, we get that $x^2 = a/3$. If we substitute in the first term of the first equation, we obtain $\frac{a}{3}x - ax = 1$ and so, $x = -3/(2a)$. Thus, we have two relations between x and a , which lead to

$$\frac{a}{3} = \left(\frac{-3}{2a} \right)^2.$$

must hold. This implies that

$$a^3 = \frac{27}{4} \Rightarrow a = \frac{27}{\sqrt[3]{4}} \approx 1.88988.$$

- (b) For $a = 3$, find a bound for the number of iterations needed to achieve an approximation with accuracy 10^{-3} to the solution of $p(x) = 0$ on the interval $[1, 4]$ using the bisection method.

Using (1) with $L = 4 - 1 = 3$ and $\epsilon = 10^{-3}$, we know that the number of steps must be $n > \frac{\ln(3000)}{\ln 2} - 1 \approx 10.55$. Then, with 11 steps of the bisection method, we will find the root with the required accuracy. The file `List2-6-2019.ods` contains the computations of this problem.

- (c) Use the Newton-Raphson¹ method to approximate the same solution with the same accuracy starting with $x_0 = 4$. How many steps has it required?

The iteration formula will be:

$$x_{n+1} = x_n - \frac{x_n^3 - 3x_n - 1}{3x_n^2 - 3}, \text{ with } x_0 = 4.$$

Again, the file `List2-6-2019.ods` contains the computations of this problem. Note that we get the same accuracy with only 5 iterations.

7. Let $f(x) = -x^3 - \cos x$ and consider the initial condition $x_0 = -1$. Use the Newton-Raphson's method to find a root of f with accuracy $\epsilon = 5 \cdot 10^{-7}$.

Let $f(x) = -x^3 - \cos x$ and $p_0 = -1$. We use Newton-Raphson's formula

$$x_{n+1} = x_n - \frac{-x_n^3 - \cos x_n}{-3x_n^2 + \sin x_n}, \text{ with } x_0 = -1.$$

The file `List2-7-2019.ods` contains the computations of this problem.

8. You are given a resistor with a resistance of 1200 ohms and a relative error of 5%. What is the absolute error in the resistance?

The relative error $\epsilon_r(x)$ is related with the absolute error $\epsilon_a(x)$ through $\epsilon_r(x) = \frac{\epsilon_a(x)}{x}$. In this case, $\epsilon_a(x) = 1200 \cdot 0.05 = 60$ ohms.

9. If you measure a cylinder with measurement errors of 1 mm both for the radius and for the height, which will be the absolute and relative errors of the estimated volume?

Denoting by r the radius of the cylinder and h its height, the actual volume will be $V = \pi r^2 h$ and the estimated volume $\bar{V} = \pi (r \pm \Delta r)^2 (h \pm \Delta h)$, with $\Delta r = \Delta h = 1$ mm (here, Δr and Δh stand for $\epsilon_a(r)$ and $\epsilon_a(h)$, respectively) according to the statement. If we only take into account the first order errors, then the expression of $\bar{V}$ can be approximated as:

$$\bar{V} = \pi (r^2 \pm 2r \Delta r + (\Delta r)^2) (h \pm \Delta h) \approx \pi (r^2 h \pm 2r h \Delta r \pm r^2 \Delta h).$$

Substituting $\Delta r = \Delta h = 1$ mm, we get:

$$\bar{V} \approx \pi (r^2 h \pm 2r h \pm r^2)$$

Then,

$$\epsilon_a(V) = |\bar{V} - V| \approx \pi (2r h + r^2); \quad \epsilon_r(V) \approx \frac{\epsilon_a(V)}{V} = \frac{\pi (2r h + r^2)}{\pi r^2 h} = \frac{2}{r} + \frac{1}{h} = 2\epsilon_r(r) + \epsilon_r(h).$$

¹Newton-Raphson's method to approximate solutions of $f(x) = 0$ consists of iterating an initial guess x_0 using the formula $x_n = x_{n-1} - f(x_{n-1})/f'(x_{n-1})$, for $n > 1$, up to $n = N$ such that $|f(x_N)| < \epsilon$ or $|x_N - x_{N-1}| < \epsilon$, for some prescribed accuracy ϵ .

10. We want to compute the quantity $Q := x \cos \theta$ for $x = 2 \text{ cm}$ and $\theta = 53^\circ$, with measurement errors $\Delta x = 0.2 \text{ cm}$ and $\Delta \theta = 2^\circ$. What is the absolute error of Q ?

$$\begin{aligned}\Delta Q &:= \Delta x |\cos \theta| + x |\sin \theta| \Delta \theta = 0.2 \cos 53^\circ + 2 \sin 53^\circ (2\pi/180)^2 \\ &\approx 0.120363 + 0.0557553 \approx 0.176118.\end{aligned}$$

11. Using mathematical induction, prove that the following statements are true for any natural number n :

(a) $1 + 3 + 5 + \cdots + (2n - 1) = n^2$.

We first test the equation for $n = 1$, which trivially gives $1 = 1$. Next, we assume that it is true up to some value $n \in \mathbb{N}$, and we prove to be true for $n + 1$. If we call $a_n := 1 + 3 + 5 + \cdots + (2n - 1)$, we have, applying the induction hypothesis, that

$$a_{n+1} = 1 + 3 + \cdots + (2n - 1) + (2n + 1) = n^2 + (2n + 1) = (n + 1)^2.$$

Therefore, we have proved that provided that $a_n = n^2$, then $a_{n+1} = (n + 1)^2$.

(b) $1^2 + 2^2 + 3^2 + \cdots + n^2 = \frac{1}{6} n(n + 1)(2n + 1)$.

For $n = 1$, the equation reads as $1^2 = \frac{1}{6} 1 \cdot 2 \cdot 3$, which is true. Next, we assume that it is true up to some value $n \in \mathbb{N}$, and we prove to be true for $n + 1$. If we call $a_n := 1^2 + 2^2 + 3^2 + \cdots + n^2$, we have, applying the induction hypothesis, that

$$\begin{aligned}a_{n+1} &= 1^2 + 2^2 + 3^2 + \cdots + n^2 + (n + 1)^2 = \frac{1}{6} n(n + 1)(2n + 1) + (n + 1)^2 \\ &= (n + 1) \frac{n(2n + 1) + 6(n + 1)}{6} = (n + 1) \frac{2n^2 + 7n + 6}{6} = (n + 1) \frac{(n + 2)(2n + 3)}{6}.\end{aligned}$$

The last expression is precisely the term $\frac{1}{6} n(n + 1)(2n + 1)$ after changing n by $n + 1$, so we have proved the equality for $n + 1$.

(c) $1^3 + \cdots + (n - 1)^3 < \frac{1}{4} n^4 < 1^3 + \cdots + n^3$.

Similar to previous ones.

(d) $3^{2n+1} + 2^{n+2}$ is divisible by 7.

For $n = 0$, the statement is $3^1 + 2^2$ is divisible by 7, which is obviously true. Now, assume the statement to be true up to some n , so that $3^{2n+1} + 2^{n+2} = 7k$ for some $k \in \mathbb{N}$. Let us write the expression for $n + 1$, work it out, and finally apply induction:

$$\begin{aligned}3^{2(n+1)+1} + 2^{(n+1)+2} &= 3^{2n+3} + 2^{n+3} = 9 \cdot 3^{2n+1} + 2 \cdot 2^{n+2} \\ &= 7 \cdot 3^{2n+1} + 2 \cdot 3^{2n+1} + 2 \cdot 2^{n+2} = 7 \cdot 3^{2n+1} + 2 \cdot 7k.\end{aligned}$$

Observe that the last expression is a multiple of 7 since $7 \cdot 3^{2n+1} + 2 \cdot 7k = 7(3^{2n+1} + 2k)$.

12. Find the global extrema of the function $f(x) = x^3 - 2x^2 + bx$ on the interval $[-2, 4]$ in terms of $b \in \mathbb{R}$.

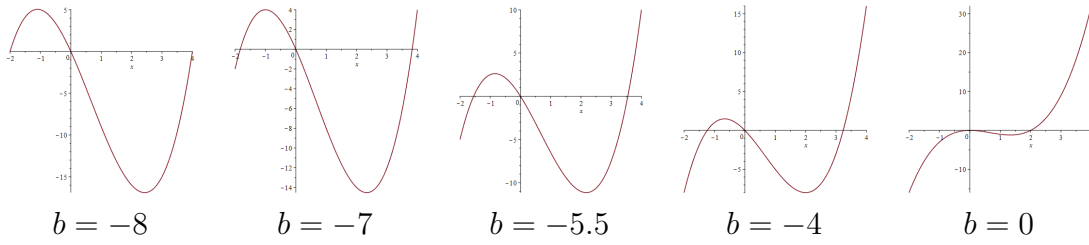


Figure 1: Graph of $f(x)$ from Problem 12, for $b \in \{-8, -7, -5.5, -4, 0\}$

We first compute the local extrema of the f in terms of b after its derivative $f'(x) = 3x^2 - 4x + b$. Clearly, $f'(x) = 0$ if and only if

$$x =: x_{\pm} = \frac{4 \pm \sqrt{16 - 12b}}{6}.$$

Since $16 - 12b \geq 0$ for $b \leq 4/3$, our first conclusion, then, is:

- For $b > 4/3$, $f'(x) > 0$ for all $x \in \mathbb{R}$, and so, the global minimum on $[-2, 4]$ is attained at $x = -2$ ($f(-2) = -16 - 2b$) whereas the global maximum is attained at $x = 4$ ($f(4) = 32 + 4b$).
- For $b \leq 4/3$, f has two local extrema, $x_{\pm}$. Given that $f''(x) = 6x - 4$, $f''(x_{\pm}) = 6 \frac{4 \pm \sqrt{16 - 12b}}{6} - 4 = \pm \sqrt{16 - 12b}$. Thus, f has a local maximum at $x = x_-$ and a local minimum at $x = x_+$. The first temptation is to state that these are, respectively, the global maximum and the global minimum, but this does not have to be true. Indeed, $x = x_-$ will be the global maximum only if $f(x_-) \geq \max\{f(-2), f(4)\}$ and $x = x_+$ will be the global minimum only if $f(x_+) \leq \min\{f(-2), f(4)\}$. Since $f(x_-) \geq f(-2)$ for all $b \leq 4/3$ and $f(x_+) \leq f(4)$ for all $b \leq 4/3$, it suffices to study two cases:
 - $f(x_-) \geq f(4)$. Tedious computations (see proof below) give the solution $b \leq -7$.
 - $f(x_+) \leq f(-2)$. Tedious computations (similar to the previous case) give the solution $b \leq -4$.

Summarizing,

- For $b \leq -7$, the global minimum is attained at $x = x_+$ and the global maximum at $x = x_-$.
- For $-7 \leq b \leq -4$, the global minimum is attained at $x = x_+$ and the global maximum at $x = 4$.
- For $-4 \leq b \leq +\infty$, the global minimum is attained at $x = -2$ and the global maximum at $x = 4$.

Check the results graphically in **python**. See also Figure 1 and the auxiliary file at <https://aula.esci.upf.edu/mod/resource/view.php?id=112014>.

Solving the equation $f(x_-) = f(4)$.

We first compute the value of f at the two points: $f(x_-) = 1/27 (-6b + 8) \sqrt{4 - 3b} +$

$2b/3 - \frac{16}{27}$ and $f(4) = 32 + 4b$.

Thus, $f(x_-) = f(4)$ implies

$$1/27 (-6b + 8) \sqrt{4 - 3b} + 2b/3 - \frac{16}{27} = 32 + 4b.$$

Multiplying the equation by 27 and simplifying terms, we get:

$$(-6b + 8)\sqrt{4 - 3b} = 880 + 90b. \quad (2)$$

After squaring this equation and simplifying, we get a polynomial equation of degree 3:

$$b^3 + 71b^2 + 1472b + 7168 = 0.$$

At this point, we have to study this polynomial. It can be seen that it only vanishes at $b = -32$ and $b = -7$. However, only for $b = -7$ equation (2) is satisfied (substituting $b = -7$ we get $250 = 250$); that is, the root $b = -32$ has appeared when we have squared (2), (substituting $b = -32$ we get $2000 = -2000$!). In summary, we have that $f(x_-) = f(4)$ if and only if $b = -7$. It is easy to check that $f(x_-) > f(4)$ and only if $b < -7$.

With similar computations we can check that $f(x_+) < f(-2)$ and only if $b < -4$.

13. Consider the family of functions $f(x) = x^4 - 4x^3 - 2bx^2$, where $b > 0$.

The problem is more interesting if we analyze all cases for b . We can think of b as a control parameter that we can tune to get a suitable answer of the model (in this case, to control the minimum of the function, for instance). The problem is also very easy if we take into account that, generically, a quartic polynomial with main term x^4 (hence, positive) has either 2 minima and 1 maximum in between, or just 1 minimum. By generic, we mean here the cases in which the critical points are nondegenerate (the first derivative vanishing at the critical point, but not the second derivative).

- (a) Find and classify the critical points of f . Which of them depend on the parameter b ? Are there any requirements that b must satisfy for these critical points to exist?

The critical points are the zeroes of $f'(x) = 4x^3 - 12x^2 - 4bx = 4x q(x)$, with $q(x) := x^2 - 3x - b$. The second derivative is $f''(x) = 12x^2 - 24x - 4b$. Then:

- $x_0 = 0$ is a critical point for any b . Since $f''(0) = -4b$, it is a maximum for $b > 0$ and a minimum for $b < 0$. When $b = 0$, the critical point is degenerate and, in fact, it is an inflection point (notice that $f''(0) = 0$ and $f'''(0) = -24 \neq 0$).
- The other critical points are the zeroes of $q(x)$. If $b > -9/4$ there are two different critical points

$$x_{\pm}(b) = \frac{3 \pm \sqrt{9 + 4b}}{2};$$

if $b = -9/4$ the critical point $x_{\pm}(-9/4) = 3/2$, is degenerate (in fact an inflection point), and no other critical points do exist for $b < -9/4$. In the notation, we have emphasized the dependence of the critical points on the parameter b .

We finally classify the critical points $x_{\pm}(b)$ for $b > -9/4$.

- If $b > 0$, $x_-(b) < 0 < x_+(b)$ and, since 0 is a maximum, then both $x_-(b)$ and $x_+(b)$ are minima.

- If $-9/4 < b < 0$, then $0 < x_-(b) < x_+(b)$ and, since 0 is a minimum, then $x_-(b)$ is a maximum and $x_+(b)$ is a minimum.

The parameter values $b_0 = -9/4$ and $b_1 = 0$ are bifurcations, in which the properties of the model (in this case the number of critical points or their classification) vary. All the information about the problem can be encoded in a bifurcation diagram, a plot on the $b - x$ plane indicating, for each b , the critical points 0 , $x_{\pm}$, and their characteristics. If you plot in the plane $b - x$ the line $x = 0$ and the curve $x^2 - 3x - b = 0$, you will easily obtain the bifurcation diagram of the problem.

Bifurcations at b_0 and b_1 are quite different, and mathematicians like to classify them. At $b_0 = -9/4$ two critical points are born (one minimum and one maximum for $b_0 \gtrsim -9/4$), and the bifurcation is referred to as *fold bifurcation* or *saddle-node bifurcation*. At $b_1 = 0$ two critical points interchange their extremal character, and the bifurcation is a *transcritical bifurcation*.

- (b) Suppose $x^* \neq 0$ is a critical point. Obtain a formula for $f''(x^*)$ that does not contain b and use it to design a test for determining whether x^* is a maximum or a minimum directly from the value of x^* (so that you won't have to evaluate f'' at x^*).

For this part, we recall that we have written the derivative of f as $f'(x) = 4xq(x)$ and that $q(x_{\pm}) = 0$ or, equivalently, $b = x_{\pm}^2 - 3x_{\pm}$. Then

$$f''(x_{\pm}) = 4(3x_{\pm}^2 - 6x_{\pm} - b) = 4(2x_{\pm}^2 - 3x_{\pm}) = 4x_{\pm}(2x_{\pm} - 3).$$

From this expression, we obtain similar conclusions as in part (13a):

- If $b > 0$, $x_-(b) < 0 < \frac{3}{2} < x_+(b)$, then $f''(x_-) > 0$ and $f''(x_+) > 0$, and both $x_{\pm}$ are minima.
 - If $-9/4 < b < 0$, then $0 < x_-(b) < \frac{3}{2} < x_+(b)$, then $f''(x_-) < 0$ and $f''(x_+) > 0$, and x_- is a maximum and x_+ is a minimum.
- (c) Suppose you measure a value x_0 with an absolute error of 4%. Use the linear approximation to estimate the maximum relative error of $f(x_0)$.

The last (outsider) question is just to recall that the derivative lets us to estimate the propagation of an uncertainty (error in a variable). In general, if a value x is given by a certain absolute error ε , one writes $x = \bar{x} \pm \varepsilon$. Then

$$f(x) = f(x \pm \varepsilon) \simeq f(x) \pm f'(x)\varepsilon.$$

That is, the absolute value of $f(x)$ is approximately $|f'(x)|\varepsilon$.

14. Suppose the number of prey animals eaten per predator per day is given by

$$y = \frac{x}{1 + hx}, \quad h > 0,$$

where x is the number of prey animals per unit area.

- (a) Determine the rate at which predation is changing with respect to time in terms of the quantities $x(t)$ and $x'(t)$.

We have to assume that both $x(t)$ and $y(t)$ change with time. Thus, $y(t) = \frac{x(t)}{1 + hx(t)}$ and so,

$$y'(t) = \frac{x'(t)(1 + hx(t)) - x(t)h x'(t)}{(1 + hx(t))^2} = \frac{x'(t)}{(1 + hx(t))^2}.$$

- (b) Suppose the prey population is growing at the rate rx , where $r > 0$. Use this additional information to determine the rate at which predation is changing with respect to time in terms of $x(t)$ alone.

The variation of the prey population can be written as $x'(t) = rx(t)$. Then, we can substitute into the previous relation:

$$y'(t) = \frac{rx(t)}{(1 + hx(t))^2}.$$

CALCULUS, 2018-19-T2

Bioinformatics degree

Problems. List 3. Optimization with modelling.

1. The gene responsible for *sickle cell disease* has two possible *alleles*, **A** and **a**. Individuals with the **aa** genotype have *sickle cell anemia*. Under natural conditions, people with sickle cell anemia did not live long enough to have children. Since the gene reduces fitness (adaptability), one would expect it to have been eliminated by natural selection. Indeed, people whose ancestors come from cold climates do not have the **a** allele, but this allele is not uncommon in parts of Africa. Individuals with the **Aa** genotype have a higher natural resistance to malaria than individuals with the **AA** genotype; where malaria is endemic, individuals with the **Aa** genotype have a longer life expectancy, which means they are likely to have more surviving children. Thus, the **a** allele has an indirect benefit along with its direct liability. To model the population dynamics of the sickle cell gene, let p be the fraction of **A** genes in the gene pool. We assume that the *fitness* of the **aa** genotype is 0, the fitness of the **Aa** genotype is 1, and the fitness of the **AA** genotype is a parameter w . According to the theory of genetics, the *overall fitness of the gene pool* is given by

$$F(p) = wp^2 + 2p(1-p) = 2p + (w-2)p^2, \quad p \in [0, 1].$$

- (a) Which fraction $p^* \in [0, 1]$ of **A** alleles gives the highest fitness?

Observe first that $F(0) = 0$ and $F(1) = w$. On the other hand, since $F'(p) = 2 + 2(w-2)p$, then $p^* = 1/(2-w)$ is the only critical point, which satisfies $F(p^*) = p^* = 1/(2-w)$. Moreover, $F''(p) = 2(w-2)$. If we restrict to the case $w \in [0, 2)$, then $F''(p) < 0$ and so, p^* is always a local maximum. It is important to note also that $F(p^*) \geq F(1)$ for any $w \in [0, 2)$ and $F(p^*) = F(1)$ only for $w = 1$ (you can prove it analyzing $F(p^*) - F(1) = \frac{1}{2-w} - w$).

Clearly, p^* is always positive for $w \in [0, 2)$, but $p^* \leq 1$ if and only if $w \in [0, 1]$. For $w < 1 \Rightarrow$ (relatively low fitness of the genotype **AA**), p^* will be the global maximum of F for $p \in [0, 1]$. The fraction $1-p = 11/(2-w) = (1-w)/(2-w)$ is the fraction of alleles of type **a**, and so the sickle cell trait remains (malaria becomes endemic) with the **aa**-probability being $(1-w)^2/(2-w)^2$.

On the other hand, for $1 < w < 2 \Rightarrow$ (relatively high fitness of **AA**), the global maximum of F in $[0, 1]$ is attained always at $p = 1$; thus, the model states the disappearance of the **a**-type alleles and so the eradication of the sickle cell trait which produces the malaria (this, all genotypes are **AA** and the individual is free of malaria).

- (b) How does p^* depend on the the fitness w of the **AA** genotype?

From the previous statement, we know that $p^* = 1/(2-w)$, which is the basic

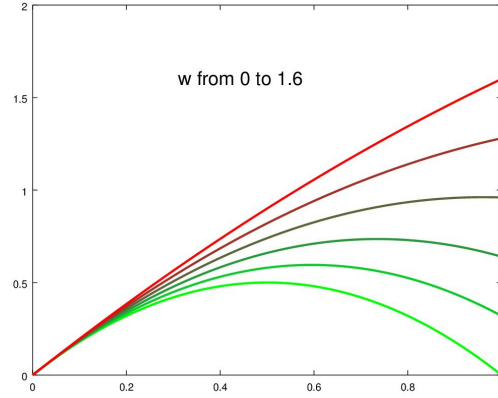


Figure 1: Graph of F for $w \in \{0, 0.32, 0.64, 0.96, 1.28, 1.6\}$ (graded from green to red).

elementary function represented in Figure 2.

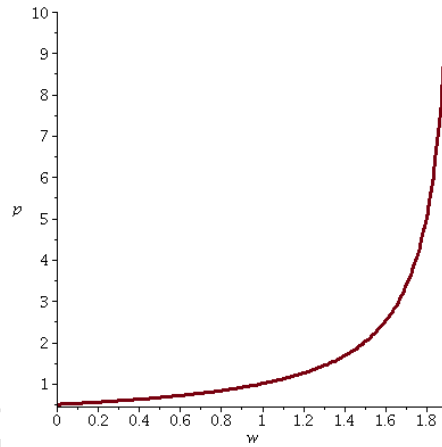


Figure 2: Critical point p^* in terms of w , for $w \in [0, 2)$.

2. The marginal value theorem (see last page of this document) determines the optimal time at which to move to a new patch. A more useful result would be the optimal level of remaining resources. We explore this issue in this problem. Let $f(t)$ be the net amount of resources collected at time t and let $x(t)$ be the size of the patch at time t . Assume that the patch is initially of unit size and that the initial resource deficit is an amount D ; that is, $f(0) = -D$, with $D < 1$.

- (a) Suppose the rate at which the forager obtains resources from a patch is proportional to the resource level of the patch; that is,

$$\frac{dx}{dt} = -kx, \quad k > 0,$$

where the parameter k represents the feeding rate of the forager. Use this equation, along with the information about the initial size, to determine the resource amount $x(t)$.

The equation $\frac{dx}{dt} = -kx$, is a differential equation, whose solution has been found previously in the course: $x(t) = x_0 \exp(-kt)$, where $x_0 = x(0)$. In this case, we are told that the patch is initially of unit size, so $x_0 = 1$. Then, the resource amount at a time t is given by $x(t) = \exp(-kt)$.

- (b) Explain why the quantity $f(t) + x(t)$ should be constant. Use this fact, the initial data, and the answer to part (2a) to obtain a formula for $f(t)$.

Clearly, resources are eliminated from the patch because they are collected. Then, $f(t) + x(t)$ must be constant. On the other hand, at time $t = 0$, we know that $x(0) = 1$ and $f(0) = -D$. Therefore, it holds that $f(t) + x(t) = 1 - D$. In particular,

$$f(t) = 1 - D - x(t) = 1 - D - \exp(-kt).$$

- (c) Apply the marginal value theorem to the function f from part (2b) to determine an equation relating the optimal time t^* and the parameters k and D . You will not be able to solve this equation by hand.

We have to maximize $f(t)/t$, that is, the amount collected relative to the time spent. Since f satisfies the hypothesis of the maximal value theorem, we know that this maximum has to be attained at a point t^* such that $f'(t^*) = f(t^*)/t^*$. Applying this to our particular function, we get that t^* solves the equation:

$$1 - D - \exp(-kt^*) - kt^* \exp(-kt^*) = 0. \quad (1)$$

As we can appreciate, it is not solvable explicitly, so we should proceed numerically (for instance, with the Newton-Raphson method)...provided that we know the numerical values of parameters D and k .

- (d) Use the result of part (2a) to eliminate t^* from the equation of part (2c). The result should be an equation that has only x^* and D .

We must take into account that $x^* = \exp(-kt^*)$ or, equivalently, $\log(x^*) = -kt^*$. Substituting these expressions conveniently into (1), we get that x^* solves the equation

$$1 - D - x^* + x^* \log(x^*) = 0. \quad (2)$$

- (e) Plot the optimal residual resource amount x^* against D , using the result of part (2d). Note that the equation does not need to be solved for x^* . Given any x^* between 0 and 1, you can calculate the corresponding value of D and plot the point in the Dx^* plane.

From (2), we can solve for D as $D = 1 - x^* + x^* \log(x^*)$. The function $D(x) := 1 - x + x \log(x)$ is strictly decreasing in the interval $]0, 1[$ since $D'(x) = \log(x)$, and $\lim_{x \rightarrow 0^+} D(x) = 1$ and $\lim_{x \rightarrow 1} D(x) = 0$ (this is left as an exercise). Then, as a consequence of Bolzano's theorem and the fact that D is strictly decreasing, for any $D \in]0, 1[$ there exists a unique solution $x^* \in]0, 1[$ of the equation $D = 1 - x + x \log(x)$,

that we refer to as $x^* = x^*(D)$. The plot for $x^* \in]0, 1[$ is given in Figure 3. The graph of the function $x^*(D)$ is the symmetric one with respect to the line $x^* = D$.

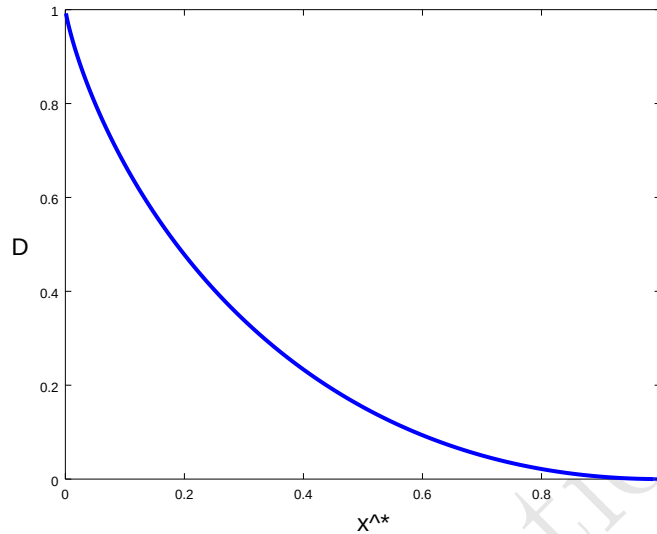


Figure 3: Graph of function $D = 1 - x^* + x^* \log(x^*)$, see Problem 2e.

- (f) Interpret the graph of part (2e), noting that D is the cost of travel between patches and x^* is the optimal amount of resource left at the time the forager moves to a new patch. Your discussion should note whether x^* increases or decreases with D and explain this result. It should also address the issues of how large D can be, what happens in the limit as D tends to 0, and what resource deficit amounts correspond to the limiting values $x^* = 0$ and $x^* = 1$.

We recall that x^* is the optimal size of the patch and that the resource deficit is D . The resource deficit can be also interpreted as the energy cost to go from one patch to another. The graph in Figure 3 shows that a “low-cost trip” strategy ($D = 0$) is linked to the decision to leave the patch (optimal time) when it is full ($x^* = 1$), so the strategy of this forager consists of picking a minimum amount from each patch and keeping travelling from one patch to another. Conversely, when $D \rightarrow 1^-$, we see that $x^* \rightarrow 0^+$. That is when moving from one patch to the other is highly expensive, then the strategy of the forager consisting of emptying each patch (leave 0 resources) and move to the following one.

- (g) The final result does not include the rate parameter k . Explain why it is reasonable that the relationship of x^* and D should not depend on k .

The relationship of x^* and D deals with optimal resource levels to leave the patch in terms of the difficulty of moving from one patch to the others. It is coherent, then, that this does not depend on the feeding rate of the forager. However, note that t^* depends on k since this parameter affects directly the time spend to reach the optimal resource level.

3. In a well-known optimal foraging paper, R. Zach reported data on the foraging behavior of crows. The crows break whelk shells by carrying them above a rocky area and then dropping them onto the rocks. Zach measured the probability of a shell breaking in a fall from a height of x meters and observed that the crows consistently dropped the whelks from a height of approximately 5.25 m. He explained these results as the solution of an optimal foraging problem. Zach did not use calculus to solve the problem, but we will do so in this problem.

It seems reasonable that the optimal strategy for the crows would be to minimize the product of the height and the average number of drops required for the shell to break. This assumption is based on the fact that the energy the crow expends on an upward flight is proportional to the total height. The product of the height per attempt and the number of attempts therefore represents the total energy cost of breaking a shell. Note, however, that the average number of drops required should be a function of the height of the drop. Thus, the function to be minimized is

$$f(x) = x n(x),$$

where we need a model for the function $n(x)$.

- (a) What properties should we expect a graph of $n(x)$ to have? Think in terms of slopes and limiting values.
- (b) Joe Mahaffy has proposed the model

$$n(x) = 1 + \frac{a}{x - b}$$

for the average number of drops needed to break a shell falling from a height of x . Mahaffy found the estimates $a \approx 16$ and $b \approx 1.2$ using Zach's data for shell breaking experiments. Plot the functions $n(x)$ and $f(x)$ with these parameter values. Your plot should only cover the range of x values that make sense (What happens if $x = 1$, for example?). What happens physically if x is outside of the reasonable range? Does Mahaffy's model have the right properties?

- (c) Determine the optimal dropping height for Mahaffy's model with the parameter values he obtained from Zach's data. How does the answer compare with what Zach observed? Does this answer support the model?
- (d) Determine the optimal dropping height for unspecified parameter values.
- (e) Suppose $b = 1.2$. Plot the optimal dropping height as a function of the parameter a .
- (f) Does a larger value of a indicate a harder surface or a softer one? How do you know?

We will do here the parametric analysis. The answers to particular values and plots are left to you.

The function $n(x)$ should (idealistically) satisfy the following properties:

- the larger is the height, the smaller is the number of drops to broke the shell;
- it should indicate that there are short distances for which the shell will never break;
- for large distances, the shell is broken at the first drop in time.

In mathematical terms:

- n is strictly decreasing;
- the domain of n is $]b, +\infty[$ for a certain $b > 0$;
- $\lim_{x \rightarrow +\infty} n(x) = 1$.

Notice that the properties imply that $\lim_{x \rightarrow b^-} n(x) \in]1, +\infty]$.

In particular, Mahaffy's model considers

$$n(x) = 1 + \frac{a}{x - b},$$

where $a, b > 0$ are parameters. This function satisfies the desired properties, and $\lim_{x \rightarrow b^-} n(x) = +\infty$. Since the only critical point of $f(x) = x n(x)$ at the domain $]b, +\infty[$ is $x_*(a, b) = b + \sqrt{ab}$, and $\lim_{x \rightarrow b^-} f(x) = +\infty$ and $\lim_{x \rightarrow +\infty} f(x) = +\infty$, then such a critical point is the absolute minimum of $f(x)$ (at $]b, +\infty[$).

Hence, $x_*(a, b) = b + \sqrt{ab}$ is the optimal dropping height, and the corresponding average of drops is $n_*(a, b) = n(x_*(a, b)) = 1 + \sqrt{\frac{a}{b}}$. Notice that both functions $x_*(a, b)$ and $n_*(a, b)$ are increasing with respect to a .

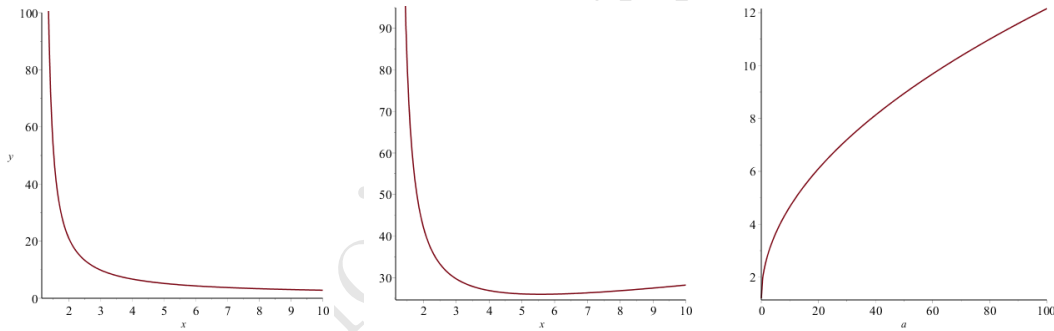


Figure 4: Graphs of functions $n(x)$, $f(x)$ and $x_*(a, 1.2)$, see Problem 3.

4. Suppose the biomass $x(t)$ of a plant changes according to the differential equation

$$\frac{dx}{dt} = r x - \frac{q x}{a + x}, \quad a, q, r > 0,$$

where the first term represents plant growth and the second represents herbivory. Define new variables X and T and parameter q by $X = x/a$, $T = rt$, and $Q = q/(ra)$.

- (a) Use the chain rule to obtain a differential equation for $X(T)$. [Hint: You can work with more than two factors, which is necessary here since you have a derivative formula for dx/dt along with algebraic relations between x and X and between t and T . Note that you must also use the algebraic formula to replace x by X ; your result should have no lowercase symbols.]

- (b) Note that X and T represent the plant biomass and time. If we want to study the effect of growth and herbivory on plant biomass, why might we prefer to study $X(T)$ rather than $x(t)$?

We will apply the chain's rule taking into account that X depends linearly on x , x depends on t (the differential equation is about this dependence), and t depends linearly on T .

$$\begin{aligned}\frac{dX}{dT} &= \frac{dX}{dx} \frac{dx}{dt} \frac{dt}{dT} = \frac{1}{a} \left(r x - \frac{q x}{a + x} \right) \frac{1}{r} \\ &\quad \frac{1}{a r} \left(r a X - \frac{q a X}{a + a X} \right) = X - \frac{Q X}{1 + X}.\end{aligned}$$

Notice that these changes of variables and parameters allow a simpler relation, that depends on a single parameter, Q .

CALCULUS, 2018-19-T3

Bioinformatics degree

Problems. List 4a. Taylor expansions.

1. We have the following information about a function f and their first three derivatives at $x = 0$:

x	f	f'	f''	f'''
0	2	-1	4	-3

Compute the polynomial of degree 3 that coincides with f and its first three derivatives at $x = 0$. $2 - x + 2x^2 - \frac{1}{2}x^3$.

2. Compute the Taylor polynomial of degree 4 of $f(x) = e^{(x^2-1)} + x$ around $x_0 = 1$. $p_4(x) = \frac{11}{6} - \frac{17x}{3} + 12x^2 - \frac{28x^3}{3} + \frac{19x^4}{6}$.

3. Consider the function $\sinh x := \frac{e^x - e^{-x}}{2}$ around $x_0 = 0$.

(a) Find an expression for the n -th derivative of $\sinh x$.

(b) Find the Taylor polynomial of degree n of the $\sinh x$ around $x_0 = 0$.

$$1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3 + \dots + \frac{1}{n!}x^n.$$

4. Compute the Taylor polynomial of degree n of $f(x) = \sin(x^2) + \cos(x^2)$ around $x = 0$. $1 + x^2 - \frac{1}{2}x^4$

5. Consider the function $f(x) = e^x \cos x$.

(a) Find the corresponding second order Taylor polynomial, $p_2(x)$, around $x_0 = 0$.

(b) Use $p_2(0.5)$ to approximate $f(0.5)$. Find an error bound for $|f(0.5) - p_2(0.5)|$ using the error formula and compare it with the actual error.

(c) Find an error bound for $|f(x) - p_2(x)|$ for $x \in [0, 1]$.

(d) How many terms should we add to compute $f(0.5)$ with a maximum error of 10^{-8} ?

The main method to compute the Taylor series of f around $x_0 = 0$ consists in computing the successive derivatives of f at $x_0 = 0$ and apply Taylor's formula. Notice that

$$\begin{array}{l|l} f^{(0)}(x) = e^x \cos x, & f^{(0)}(0) = 1; \\ f^{(1)}(x) = e^x (\cos x - \sin x), & f^{(1)}(0) = 1; \\ f^{(2)}(x) = -2e^x \sin x, & f^{(2)}(0) = 0; \\ f^{(3)}(x) = -2e^x (\sin x + \cos x), & f^{(3)}(0) = -2; \\ f^{(4)}(x) = -4e^x \cos x, & f^{(4)}(0) = -4; \\ \vdots & \vdots \end{array}$$

and we observe that the derivatives appear in blocks of size 4. Using the division rule, we can write any $k \geq 0$ as $k = 4q + r$, where $r \in \{0, 1, 2, 3\}$. The quotient q gives the block, and the residue r gives the element in the block. In summary:

$$f^{(k)}(x) = (-4)^q f^{(r)}(x), f^{(k)}(0) = (-4)^q f^{(r)}(0).$$

We are ready to complete the exercise.

(a) The second order Taylor polynomial is

$$p_2(x) = f(0) + f'(0)x + \frac{1}{2!}f''(0)x^2 = 1 + x.$$

(b) Given $x \in \mathbb{R}$, the (absolute value of the) error is

$$|f(x) - p_2(x)| = \frac{1}{3!}|f'''(\xi)||x|^3,$$

where ξ is a number between $x_0 = 0$ and x . Since we do not know the value of ξ , we can just bound the error, so we have to bound the absolute value of the third derivative between 0 and x . In the present exercise, $x = 0.5$ (in the next exercise, $x \in [0, 1]$). Since $f'''(0) = -2$, and $f'''(x)$ is decreasing in the interval $[0, \pi]$ (since its derivative is $f''''(x) = -4e^x \cos x$, which is negative in $[0, \pi]$), then an upper bound of $|f'''(\xi)|$ for $\xi \in [0, x]$ is $|f'''(x)|$ (for the two cases considered).

In particular for $x = 0.5$,

$$|f(0.5) - p_2(0.5)| \leq \frac{1}{6}|f'''(0.5)|0.5^3 \simeq 0.0932220 \leq 0.094.$$

(We only give two significant digits for the error.)

(c) Following the previous arguments, for $x \in [0, 1]$,

$$|f(x) - p_2(x)| \leq \frac{1}{6}|f'''(x)|x^3 \leq \frac{1}{6}|f'''(1)|1^3 \simeq 1.25202 \leq 1.3.$$

(d) In order to bound

$$e_n = |f(0.5) - p_n(0.5)| = \frac{|f^{(n+1)}(\xi)|}{(n+1)!}0.5^{n+1},$$

where $\xi \in]0, 0.5[$, we have to bound the absolute value of $|f^{(n+1)}(x)|$ in the interval $[0, 0.5]$. Using similar arguments as we did in exercise (b) we have, for $x \in]0, 0.5[$:

- $|f(x)| \leq f(0.5) \simeq 1.44689$, since $f(x)$ is increasing and positive;
- $|f'(x)| \leq f'(0) = 1$, since $f'(x)$ is decreasing and positive;
- $|f''(x)| \leq -f''(0.5) \simeq 1.58088$, since $f''(x)$ is decreasing and negative;
- $|f'''(x)| \leq -f'''(0.5) \simeq 4.47466$, since $f'''(x)$ is decreasing and negative.

We obtain that the sequence of errors bounds (using floating point notation)

$$\begin{aligned} e_0 &= 0.5, e_1 = 0.20, e_2 = 9.4e-02, e_3 = 1.6e-02, e_4 = 1.1e-03, \\ e_5 &= 1.4e-04, e_6 = 2.8e-05, e_7 = 2.3e-06, e_8 = 8.8e-08, e_9 = 6.9e-09, \dots \end{aligned}$$

so we have to sum up to order 9.

6. Using Taylor expansions, compute the numbers e and e^{-1} with absolute error 10^{-8} .
7. Let f be a function with bounded derivatives, more specifically, with $|f(x)| \leq 1$, $|f^{(1)}(x)| \leq 1$, $|f^{(2)}(x)| \leq 2$, $|f^{(3)}(x)| \leq 6, \dots$ and, in general, $|f^{(k)}(x)| \leq k!$. We approximate f by means of its Taylor polynomial of degree 5 around $x_0 = 0$. Give an bound of the truncation error for $x \in [-\frac{1}{2}, \frac{1}{2}]$. [0.016](#),
8. We know that a function $y = y(x)$ satisfies the two following equations:

$$\begin{cases} y' &= (x+1)y, \\ y(0) &= 2. \end{cases}$$

Compute the corresponding polynomial of order n around $x_0 = 0$.

CALCULUS, 2018-19-T3

Bioinformatics degree

Problems. List 4b. Power series.

9. Find the power series of the following functions and determine the corresponding interval of convergence.

(a) $f(x) = \frac{x}{1-x^2};$

(b) $f(x) = e^{x^3-1};$

(c) $f(x) = x^2 \sin x;$

(d) $f(x) = e^x \ln(1+x) + \cos x.$

(e) $f(x) = e^x \cos x.$

- (a) Using that

$$\frac{1}{1-y} = \sum_{n=0}^{\infty} y^n,$$

in the interval $] -1, 1[$, and using the change of variables $y = x^2$, we obtain

$$\frac{x}{1-x^2} = x \sum_{n=0}^{\infty} (x^2)^n = \sum_{n=0}^{\infty} x^{2n+1}$$

for $x^2 \in] -1, 1[$, that is for $x \in] -1, 1[$.

- (b) We use the power series of the exponential, that converges for any $x \in \mathbb{R}$, and obtain:

$$e^{x^3-1} = e^{-1} e^{x^3} = e^{-1} \sum_{n=0}^{\infty} \frac{1}{n!} x^{3n},$$

and the interval of convergence is $\mathbb{R}$.

- (c) We use the power series of the sinus, that converges for any $x \in \mathbb{R}$, and obtain:

$$x^2 \sin x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+3},$$

and the interval of convergence is $\mathbb{R}$.

- (d) We operate with the power series corresponding to the functions. Notice that the power series of $\log(1+x)$ converges in $] -1, 1[$, and those of $\exp x$ and $\cos x$ converge in $\mathbb{R}$, and as a result the whole power series converges in $] -1, 1[$. Then, since

$$\begin{aligned} e^x \log(1+x) &= \left(\sum_{n_1=0}^{\infty} \frac{1}{n_1!} x^{n_1} \right) \left(\sum_{n_2=1}^{\infty} \frac{(-1)^{n_2-1}}{n_2} x^{n_2} \right) \\ &= \sum_{n=1}^{\infty} \left(\sum_{n_1+n_2=n} \frac{1}{n_1!} \frac{(-1)^{n_2-1}}{n_2} \right) x^n \\ &= \sum_{n=1}^{\infty} \left(\sum_{k=1}^n \frac{(-1)^{k-1}}{(n-k)!k} \right) x^n, \end{aligned}$$

where in the last equality we write $k = n_2$ (that runs from 1 to n) and then $n_1 = n-k$. In summary,

$$e^x \log(1+x) + \cos x = \sum_{n=1}^{\infty} \left(\sum_{k=1}^n \frac{(-1)^{k-1}}{(n-k)!k} \right) x^n + \sum_{m=0}^{\infty} \frac{(-1)^m}{(2m)!} x^{2m}.$$

- (e) Using exercise 5, the power series of $f(x) = e^x \cos x$ is

$$\begin{aligned} e^x \cos x &= \\ &= \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k \\ &= \sum_{q=0}^{\infty} (-4)^q \left(\frac{f^{(0)}(0)}{(4q)!} x^{4q} + \frac{f^{(1)}(0)}{(4q+1)!} x^{4q+1} + \frac{f^{(2)}(0)}{(4q+2)!} x^{4q+2} + \frac{f^{(3)}(0)}{(4q+3)!} x^{4q+3} \right) \\ &= \sum_{q=0}^{\infty} (-4)^q \left(\frac{1}{(4q)!} + \frac{1}{(4q+1)!} x - \frac{2}{(4q+3)!} x^3 \right) x^{4q}. \end{aligned}$$

The convergence (and the equality between the function and the series) is in all the set of real numbers, since the Taylor series of both exponential and cosinus functions converge in the set of real numbers, and the series of $f(x) = e^x \cos x$ is just the product of the corresponding Taylor series. (This is, in fact, another method to compute the power series.)

10. Consider the function $f(x) = \arctan x$.

- Compute the series of $f(x)$. Hint: $f'(x) = \frac{1}{1+x^2}$.
- Compute the radius of convergence.
- Study the convergence of the series at the extreme values of the interval of convergence (you may study partial sums or apply Leibnitz's criterion).
- Considering that $\arctan(1) = \pi/4$ (and using Abel's theorem), find a numerical series to compute π .
- What can you obtain by considering $\arctan(1/\sqrt{3}) = \pi/6$ instead of $\arctan 1$?

- (a) We recall that $\frac{d \arctan x}{dx} = \frac{1}{1+x^2}$ (that is, $\arctan x$ is the antiderivative of $(1+x^2)^{-1}$), which suggests to depart from the knowledge of the power series of $\frac{1}{1+x^2}$:

$$\frac{1}{1+x^2} = \sum_{n=0}^{\infty} (-1)^n x^{2n}.$$

The power series of $\arctan x$ is then the antiderivative of the above power series. Using that $\frac{d}{dx} \left(\frac{(-1)^n}{(2n+1)} x^{2n+1} + C \right) = (-1)^n x^{2n}$ for $C \in \mathbb{R}$, we get that:

$$\arctan x = \sum_{n=0}^{\infty} \left(\frac{(-1)^n}{(2n+1)} x^{2n+1} + C \right).$$

Since $\arctan 0 = 0$, we determine that $C = 0$. Consequently,

$$\arctan x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)} x^{2n+1}.$$

- (b) The radius of convergence is $R = 1$. It can be deduced from the fact the radius of convergence of the series associated to $g(x) = 1/(1-x)$, which is also $R = 1$. Since $f'(x) = g(-x^2)$, then the series of $1/(1+x^2)$ will converge for $-1 < -x^2 < 1$, that is for $|x| < 1$. Finally, since this is the radius of convergence of the series for $f'(x)$, it is the same for $f(x)$.
- (c) By Leibniz's criterion, it also converges for $x = 1$ since it is an alternate series whose main coefficients ($|a_{2n+1}| = 1/(2n+1)$) are positive and form a decreasing sequence that tends to zero.

For $x = -1$, the numerical series becomes

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)} (-1)^{2n+1} = \sum_{n=0}^{\infty} \frac{(-1)^{n+1}}{(2n+1)},$$

which also converges by the Leibniz's criterion. Observe that the both the arctan function and the arctan series are odd function and so, the series for $x = -1$ is the same than for $x = 1$ but changing sign.

- (d) Since the radius of convergence of the power series obtained in part (10a) has radius 1, see part (10b), and the numerical series for $x = 1$ obtained in part (10c) is convergent, by Abel's theorem, the series $\sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)} x^{2n+1}$ is continuous from the left at $x = 1$. As a consequence, we can approximate $\arctan 1$ by the partial sums $S_M := \sum_{n=0}^M \frac{(-1)^n}{(2n+1)}$; that is,

$$\frac{\pi}{4} = \arctan(1) \approx \sum_{n=0}^M \frac{(-1)^n}{(2n+1)} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots \pm \frac{1}{2M+1}.$$

Note that the series converges very slowly. In order to compute π with an error smaller than, say, 10^{-8} , we should add to around 200000000 terms (200 million terms!).

(e) Evaluating at $x = \frac{1}{\sqrt{3}}$, we obtain:

$$\frac{\pi}{6} = \arctan(1/\sqrt{3}) = \frac{1}{\sqrt{3}} \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)3^n}.$$

Note: The series converges much faster. To compute π with an error smaller than 10^{-8} , we should add only 15 terms.

11. Let us consider the series $f(x) = \sum_{n=0}^{\infty} \frac{1}{2^n} x^n$.

Compute the interval of convergence (analysing the convergence at the boundary), and compute the sum (using operational properties of series, including derivation and antiderivation).

Since $a_k = \frac{1}{2^k}$, the radius of convergence is

$$R = \lim_{k \rightarrow \infty} \left| \frac{a_k}{a_{k+1}} \right| = \lim_{k \rightarrow \infty} \left| \frac{2^{k+1}}{2^k} \right| = 2.$$

We consider now, the numerical series for $x = 2$ and $x = -2$:

$$\begin{aligned} f(2) &= \sum_{n=0}^{\infty} \frac{1}{2^n} 2^n = \sum_{n=0}^{\infty} 1 = +\infty; \\ f(-2) &= \sum_{n=0}^{\infty} \frac{1}{2^n} (-2)^n = \sum_{n=0}^{\infty} (-1)^n. \end{aligned}$$

Both numerical series are divergent: the first one for obvious reasons because the partial sums tend to infinity; the second one because the partial sums are $1, 0, 1, 0, \dots$, which is a non convergent sequence. Thus, the interval of convergence is $] -2, 2[$.

To compute the sum, we observe that $\frac{1}{2^n} x^n = \left(\frac{x}{2}\right)^n$ and so:

$$f(x) = \sum_{n=0}^{\infty} \frac{1}{2^n} x^n = \sum_{n=0}^{\infty} \left(\frac{x}{2}\right)^n = \frac{1}{(1 - x/2)} = \frac{2}{2 - x},$$

for $x \in] -2, 2[$

12. Let us consider the series $f(x) = \sum_{n=1}^{\infty} nx^n$.

Compute the interval of convergence (analysing the convergence at the boundary), and compute the sum (using operational properties of series, including derivation and antiderivation).

Since $a_k = k$, the radius of convergence is

$$R = \lim_{k \rightarrow \infty} \left| \frac{k}{k+1} \right| = 1.$$

We consider now, the numerical series for $x = 1$ and $x = -1$:

$$\begin{aligned} f(1) &= \sum_{n=1}^{\infty} n = +\infty; \\ f(-1) &= \sum_{n=1}^{\infty} n(-1)^n. \end{aligned}$$

Both numerical series are divergent: the first one for obvious reasons because the partial sums tend to infinity; the second one because the partial sums are $-1, 1, -1, 1, -1, \dots$, which is a non convergent sequence. Thus, the interval of convergence is $] -1, 1[$. To compute the sum, we observe that nx^{n-1} is the derivative of x^n and we use this fact accordingly:

$$\begin{aligned} f(x) &= \sum_{n=1}^{\infty} nx^n = x \sum_{n=1}^{\infty} nx^{n-1} = x \frac{d}{dx} \left(\sum_{n=1}^{\infty} x^n \right) \\ &= x \frac{d}{dx} \left(\frac{1}{1-x} - 1 \right) = x \frac{1}{(1-x)^2} = \frac{x}{(1-x)^2}. \end{aligned}$$

for $x \in] -1, 1[$

13. The state equation of a certain gas follows the *Dieterici's equation*:

$$P(V - nb) = nRTe^{-an/RTV}.$$

Find the first terms of the corresponding virial series. Compute the Boyle's temperature in terms of a and b (see the notes around the Van der Waals model given in Lecture05.) We have to express the compressibility factor $Z = PV/nRT$ in terms of the molar density $\rho = n/V$, and then find the power series with respect to ρ . To do so:

$$Z = \frac{PV}{nRT} = \frac{nRTe^{-an/RTV}}{(V - nb)} \frac{V}{nRT} = \frac{e^{-an/RTV}V}{(V - nb)} = \frac{e^{-a\rho/RT}}{1 - b\rho},$$

and then by multiplying the power series of $e^{-(a/RT)\rho}$ and $\frac{1}{1-b\rho}$, we get

$$Z = \sum_{n=0}^{\infty} \left(\sum_{k=0}^n \frac{(-1)^k (a/RT)^k b^{n-k}}{k!} \right) \rho^n$$

where the series converges for $\rho \in [0, b^{-1}[$.

The first virial coefficient is the term for $n = 1$,

$$B_1(T) = b - a/RT,$$

and, hence, Boyle's temperature is $T_B = a/Rb$.

CALCULUS, 2018-19-T3

Bioinformatics degree

*Problems. List 5. Interpolation and approximation.*¹

1. () Compute the Lagrange polynomials and the interpolation polynomial for the following datasets obtained from exact values of a function f :

- (a) $f(-1) = 5$, $f(2) = -2.5$, $f(3) = -3$.
- (b) $f(-3) = -4$, $f(-2) = -6$, $f(1) = 0$, $f(3) = 14$.
- (c) $f(0) = 0.50$, $f(2) = 0.35$, $f(4) = 0.15$, $f(6) = 0.05$.

Compute as well the corresponding divided differences schemes and check that you obtain the same interpolation polynomials.

2. () Consider the following table of values of the function $f(x) = e^x$:

x	0.0	0.2	0.4	0.6
$f(x)$	1.0000	1.2214	1.4918	1.8221

- (a) Estimate $\sqrt[3]{e}$ using both linear, cubic polynomial and cubic splines () interpolation.
- (b) Provide an error bound for the linear and the cubic polynomial interpolation and check it by comparing the results with the direct numerical computation of $\sqrt[3]{e}$.

3. Consider the function $f(x) = \frac{1}{1+x^2}$.

- (a) () Compute the interpolation polynomial using the values of f at $x = -4$, $x = -2$, $x = 0$, $x = 2$ and $x = 4$, both using the Lagrange method and the divided differences method.
- (b) () Compare graphically the function and the interpolation polynomial.
- (c) () If we add a new point to the dataset, say $x = 6$, which additional (minimal) modification do we need to perform in both methods?

4. () The position $x = x(t)$ of a moving particle at four times is $x(0) = 0$, $x(1) = 1.5$, $x(2) = 3.5$, $x(3) = 5.25$.

- (a) Compute the interpolation polynomial.

¹We use the images ,  and  (by-hand computations, calculator and computer, respectively) to give an orientation about the methodology to follow in each problem. Computer simulations refer to `python` scripts, that is, using tools from `Lab020-2019-Interpolation.py` for interpolation and from `Lab030-2019-ExponentialDecay.py` for approximation.

(b) Estimate the position, velocity and acceleration of the particle at time $t = 2.5$.

5. () Intracellular recordings of the membrane voltage v of a neuron gives the following data during 0.6 ms:

t (ms)	0.0	0.3	0.5	0.6
$v(t)$ (mV)	-50	-48	-46	-44

- (a) Find the interpolation polynomial $p(t)$, and give a prediction for the membrane voltage at $t = 0.4$ ms.
- (b) Find an approximation for the time t^* at which the membrane voltage crosses the threshold $v = -45$ mV, by applying one step of Newton method to the equation $p(t) + 45 = 0$ from an initial approximation $t_0 = 0.55$.
- (c) Find an approximation for the time t^* by interpolating the inverse table (t versus v instead of v versus t).

6. () In order to obtain information about an individual's glucose tolerance, recordings of glucose levels are performed in clinical practice to control diabetes. It is common practice to use simple summary measures, such as fasting² value, two-hour value or the area under the curve. Table shows OGTTs (Oral Glucose Tolerance Tests) with glucose measurements over two hours recorded for healthy pregnant women in their first trimester (averages over population), after ingesting a 75 g preparation of glucose.

t (min)	0	10	20	30	40	50	60	70	80	90	100	110	120
x (mmol/l)	4.00	5.02	5.56	5.77	5.75	5.56	5.27	4.91	4.52	4.12	3.73	3.34	2.98

- (a) Adapt the code in `Lab020-2019-Interpolation.py` to compute the interpolating polynomial of the table subset containing only the measurements at $t = 0, 30, 60, 90$ and 120 minutes.
- (b) Give a prediction of the glucose levels at $t = 40$ and $t = 50$ and compare the result with the actual values given in the full table.
- (c) Adapt the code in `Lab020-2019-Interpolation.py` to compute both the interpolating polynomial $P_{12}(t)$ and the interpolating spline $S_{12}(t)$ of the full table and compare them, both visually and numerically (you can compute $\frac{1}{M} \sum_{j=0}^M |P_{12}(t_j) - S_{12}(t_j)|$, where $t_j = 120j/M$ and M is a “sufficiently” large integer).

7. () In Lecture01, we treated the modeling of drug clearance (see also Example 1.1.1 in [Ledder]) with data obtained from a hypothetical clinical analysis in order to test the effects of two tablets of acetaminophen in one subject's body. We assume here that the results of the analysis are those shown in the following table:

Time (h)	0	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0
Acetaminophen (mg)	650.0	552.3	471.5	406.7	347.1	297.6	254.9	219.8	184.2

²dejú (cat), ayuno (cast), jeûne (fr)

We aim at testing the exponential decay behavior

$$y(t) = e^{-\kappa t} A \quad (1)$$

of the drug clearance in the subject's body.

- (a) Use the code from `Lab030-2019-ExponentialDecay.py` to plot the points of the table in the space “Time-Acetaminophen” $t - y$.
- (b) Apply logarithms to make the expression (1) linear in terms of t .
- (c) Use the code from `Lab030-2019-ExponentialDecay.py` to plot the points of the table in the space $t - \log y$. Write the corresponding output into a file called “FigureLog.png”.
- (d) Assuming $A = 650.0$ mg (initial value), we aim at minimizing the *residual sum of squares*:

$$RSS(\kappa) := \sum_{i=0}^8 (\log(A/y_i) - \kappa t_i)^2,$$

where (t_i, y_i) for $i = 0, \dots, 8$ are the values of the table, in terms of κ . Prove (🔪) that the minimum must satisfy

$$\kappa = \frac{\sum_{i=0}^8 t_i \log(A/y_i)}{\sum_{i=0}^8 t_i^2}.$$

- (e) Use the code from `Lab030-2019-ExponentialDecay.py` to compute the optimal κ for the given data set. Plot the points of the table and the best-fitting curve both for in the $t - \log y$ and the $t - y$ plane.
- (f) Using the above computations, make a prediction about the amount of acetaminophen in the body after 100 minutes.
- (g) Using the tools provided in `Lab020-2019-Interpolation.py`, make predictions about the amount of acetaminophen in the body after 100 minutes using both a suitable interpolation polynomial of degree 3, the interpolation polynomial of the entire table and cubic splines.

CALCULUS, 2018-19-T3

Bioinformatics degree

Problems. List 6. Integration

1. Compute the following indefinite integrals.

a) $\int \frac{dx}{(2-3x)^2}$ b) $\int \frac{x}{(4x^2+9)^2} dx$ c) $\int \frac{x}{\sqrt{x^2+1}} dx$

d) $\int x e^x dx$ e) $\int e^x \sin x dx$ f) $\int \tan x dx$

g) $\int \sin x \cos^3(x) dx$ h) $\int \frac{dx}{x^2+2x+2}$ j) $\int \frac{\ln x}{x} dx$

a) $\frac{1}{3} (2-3x)^{-1} + C, C \in \mathbb{R};$

b) $-\frac{1}{8} (4x^2+9)^{-1} + C, C \in \mathbb{R};$

c) $\sqrt{x^2+1} + C, C \in \mathbb{R};$

d) $(x-1)e^x + C, C \in \mathbb{R};$

e) $\frac{1}{2}e^x (\sin(x) - \cos(x)) + C, C \in \mathbb{R};$

f) $-\ln(\cos(x)) + C, C \in \mathbb{R};$

g) $-\frac{1}{4} \cos^4 x + C, C \in \mathbb{R};$

h) $\arctan(x+1) + C, C \in \mathbb{R};$

j) $\frac{1}{2} \ln^2 x + C, C \in \mathbb{R}.$

2. Compute the following integrals using the Fundamental Theorem of Calculus.

a) $\int_1^4 \sqrt{x} dx$ b) $\int_2^0 \frac{dx}{(x+1)^2}$ c) $\int_{-1}^1 (x-2)^2 dx$

d) $\int_1^\infty \frac{dx}{x^2}$ e) $\int_0^{+\infty} \frac{dx}{(9x^2+16)}$ f) $\int_0^1 \frac{dx}{(9x^2-16)}$

a) $14/3;$

b) $-2/3;$

c) $26/3;$

d) $1;$

e) $\pi/24;$

f) $-\frac{1}{12} \operatorname{arctanh} \frac{3}{4} = \frac{1}{24} \ln \left(\frac{3x-4}{3x+4} \right).$

3. Represent the bounded region defined by the graphs of the functions $f(x)$ and $g(x)$, and compute its area.

- (a) $f(x) = \sqrt{x}$, $g(x) = x^2$.
- (b) $f(x) = 5 - x^2$, $g(x) = 3 - x$.
- (c) $f(x) = 8$, $g(x) = x^2 + 2x$.

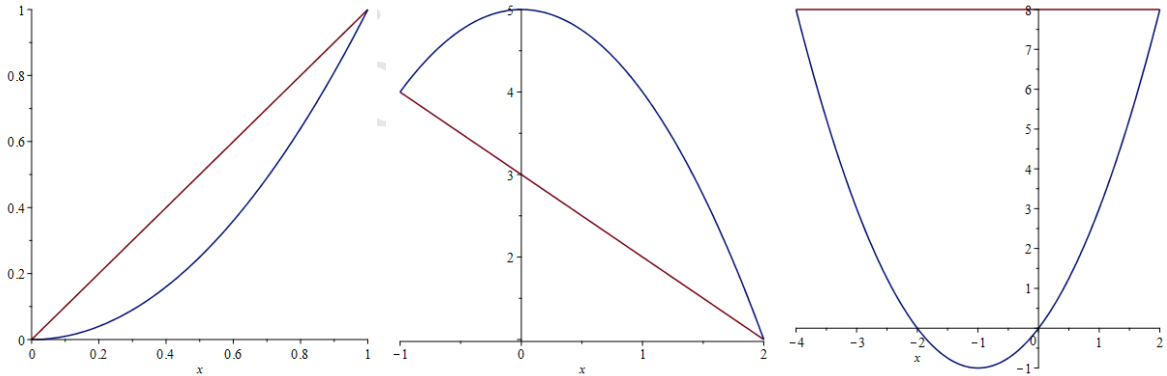
Let us call A the bounded area between the two functions. The procedure consists of:

- (a) Looking for the intersection points, that is, solving $f(x) = g(x)$. We call $S := \{x_1, x_2, \dots\}$ the set of solutions.
- (b) Localizing those x_i such that $f(x) - g(x)$ is bounded between x_i and x_{i+1} . We call $J \subset S$ the set of subscripts satisfying this boundedness condition.
- (c) Integrating $\int_{x_i}^{x_{i+1}} |f(x) - g(x)| dx$ for all $i \in J$. The bounded area is then

$$A = \sum_{i \in J} \int_{x_i}^{x_{i+1}} |f(x) - g(x)| dx.$$

In practice, we usually have $S = \{x_1, x_2\}$, $J = \{1\}$ and so, $A = \int_{x_1}^{x_2} |f(x) - g(x)| dx$.

- a) $f(x) = g(x) = 0 \Leftrightarrow x \in \{0, 1\}$. Then, $A = \int_0^1 (x - x^2) dx = \frac{1}{6}$;
- b) $f(x) = g(x) = 0 \Leftrightarrow x \in \{-1, 2\}$. Then, $A = \int_{-1}^2 (5 - x^2 - (3 - x)) dx = \frac{9}{2}$;
- c) $f(x) = g(x) = 0 \Leftrightarrow x \in \{-4, 2\}$. Then, $A = \int_{-4}^2 (8 - (x^2 + 2x)) dx = 36$.



Graphical representation of the functions in Problem 3, (a)-(b)-(c) from left to right.

4. A bacterium is moving along a straight line (assume it is the x -axis) on a Petri dish with velocity $v(t) = 4(1 - t)(2 - t)$ $\mu\text{m/s}$. Without loss of generality, we establish that the initial position is $x(t) = 0$ μm .

- (a) Compute the position of the bacterium after 4 seconds.
- (b) Compute the distance covered by the bacterium during this 4-seconds period.
- (c) Repeat the above two questions in the case that we know that $x(0) = 0$ μm , $v(0) = 8$ $\mu\text{m/s}$ and the acceleration is $a(t) = (1 + t)^{1/2}$ $\mu\text{m/s}^2$.

(a) Since $v = \frac{dx}{dt}$, then

$$x(4) - x(0) = \int_0^4 v(t) dt = \left[\frac{4}{3} t^3 - 6t^2 + 8t \right]_0^4 = \frac{64}{3} - 0 = \frac{64}{3} \mu m.$$

Then, $x(4) = x(0) + \frac{64}{3} = \frac{64}{3} \mu m.$

(b) Notice that the distance does not need to coincide with the difference of positions because the direction of movement may change during the experiment. In particular, note that in our example, $v(t) < 0$ in $(1, 2)$. Therefore, the distance covered by the bacterium is:

$$d := \int_0^4 |v(t)| dt = \int_0^1 v(t) dt - \int_1^2 v(t) dt + \int_2^4 v(t) dt = \frac{64}{3} \mu m.$$

(c) Since $a = \frac{dv}{dt}$, then

$$v(4) - v(0) = \int_0^4 a(t) dt = \left[\frac{2}{3} (1+t)^{3/2} \right]_0^4 = -\frac{2}{3} + \frac{10}{3} \sqrt{5} \mu m/s.$$

Then, $v(4) = v(0) - \frac{2}{3} + \frac{10}{3} \sqrt{5} = \frac{22}{3} + \frac{10}{3} \sqrt{5} \mu m/s.$

In this case, $v(t) > 0$ in $[0, 4]$, and so,

$$d = x(4) - x(0) = \int_0^4 \frac{2}{3} (1+t)^{3/2} dt = \left[\frac{4}{15} (1+t)^{5/2} \right]_0^4 = \frac{20\sqrt{5}}{3} - \frac{4}{15} \mu m.$$

Since $x(0) = 0$, the value of the position $x(4)$ coincides with the value of the computed distance.

5. Practice the concept of integral in the following questions:

(a) A runner's respiration rate during a 30-min training run is $r(t)$ breaths per minute. Construct an integral to calculate the total number of breaths over the course of the run.

In a time interval of dt minutes, the number of breaths is $r(t) dt$; thus, in 30 minutes, the total number is

$$\int_0^{30} r(t) dt.$$

(b) Suppose $t = \int_0^Y f(y) dy$ represents the time required for Y grams of a toxin to be eliminated from an organism's bloodstream. What does $f(y)$ represent? In particular, what does $f(Y/2)$ represent?

By matching units, we guess that $f(y)$ must be the time required to eliminate 1 gram of the toxin when the current quantity is y grams. Another way to make a derivative of the equality assuming that $t = t(Y)$:

$$\frac{dt}{dY} = f(Y).$$

Indeed, if $t(Y)$ is invertible (which is true under the logical assumption that f is positive), then we can invert the relationship and state that:

$$\frac{dY}{dt} = \frac{1}{f(Y(t))}.$$

Therefore, $1/f(y)$ can be thought of as the rate of elimination of toxin.

- (c) The population density of a plant species is $f(x)$ individuals per square meter, where x is the distance from a river, with $f(x) \geq 0$ for $x \leq 100$ and $f(x) = 0$ for $x \geq 100$. Construct an integral to calculate the number of plants along a section of river of length L .

First of all, notice that the density does not depend on the position ($\in [0, L]$) along the shore. In each strip of area $dx \times L$, then, the number of plants will be $f(x) dx L$. Therefore, the total number of plants will be (using that $f(x) = 0$ for $x \geq 100$):

$$L \int_0^{100} f(x) dx.$$

- (d) The population density of a plant species is $f(r)$ individuals per square meter, where r is the distance from the center of a circular lake of radius 100m. Assume the density falls to 0 at a distance of 100m from the lake. Construct an integral to calculate the total number of plants.

First of all, notice that the density does not depend on the angle ($\theta \in [0, 2\pi]$) along the concentric circles. The area each infinitesimal annular ring between the circle of radius r and the circle of radius $r + dr$, with $r \in [100, 200]$, is $\pi(r + dr)^2 - \pi r^2 = \pi(r dr + dr^2)$, which can be approximated by $\pi r dr$ if we neglect dr^2 (second order infinitesimal). Thus, the number of plants in each infinitesimal annular ring is $f(r) \pi r dr$. Finally, the total number of plants is:

$$\pi \int_{100}^{200} r f(r) dr.$$

6. Write the sum $S(k) := \int_{-1}^0 (1+x)e^{-kx} dx + \int_0^1 (1-x)e^{-kx} dx$, where $k \neq 0$, as a single integral by the change of variables $y = 1+x$ in the first integral. Compute $S(k)$.

We first use the substitution $y = 1+x$ in the first definite integral. In particular, $dy = dx$, $x = -1 \Rightarrow y = 0$ and $x = 0 \Rightarrow y = 1$:

$$I_1(k) := \int_{-1}^0 (1+x)e^{-kx} dx = \int_0^1 y e^{-k(y-1)} dy = e^k \int_0^1 y e^{-ky} dy.$$

To solve this integral, we integrate by parts taking $u = y$ and $dv = e^{-ky} dy$. Thus, $du = dy$, $v = e^{-ky}/(-k)$ and:

$$\begin{aligned} S_1(k) &:= e^k \left([y e^{-ky}/(-k)]_0^1 - \int_0^1 e^{-ky}/(-k) dy \right) = e^k \left(\frac{e^{-k}}{-k} + \frac{1}{k} [e^{-ky}/(-k)]_0^1 \right) \\ &= e^k \left(\frac{e^{-k}}{-k} - \frac{1}{k^2} (e^{-k} - 1) \right) = \left(\frac{-1}{k} - \frac{1}{k^2} (1 - e^k) \right) = \frac{e^k - k - 1}{k^2}. \end{aligned}$$

The second integral is solved with similar arguments, thus obtaining:

$$\begin{aligned} S_2(k) &:= \int_0^1 (1-x)e^{-kx} dx = \int_0^1 e^{-kx} dx - \int_0^1 x e^{-kx} dx = \frac{1 - e^{-k}}{k} + \frac{e^{-k} k + e^{-k} - 1}{k^2} \\ &= \frac{e^{-k} + k - 1}{k^2}. \end{aligned}$$

Joining both results, we get that

$$S(k) = S_1(k) + S_2(k) = \frac{e^k - k - 1}{k^2} + \frac{e^{-k} + k - 1}{k^2} = \frac{e^k + e^{-k} - 2}{k^2}.$$

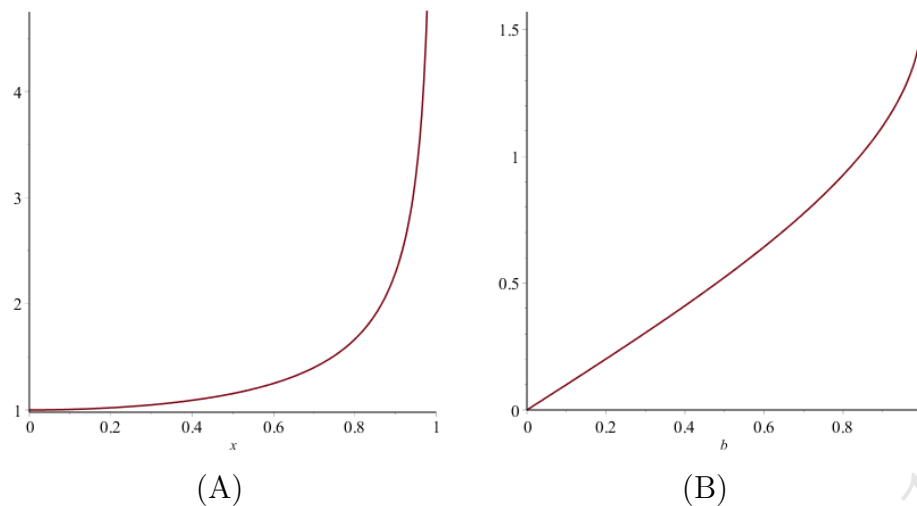


Figure 1: (A) Plot of $\frac{1}{\sqrt{1-x^2}}$ for $x \in [0, 1)$. (B) Plot of $g(b)$.

7. Consider the function $f(x) = \frac{1}{\sqrt{1-x^2}}$.

- (a) Plot $f(x)$ on the interval $0 \leq x < 1$. Can you tell from the graph whether the area under this curve is finite?

As we can observe in Figure 7a (A), the function has a vertical asymptote at $x \rightarrow 1^-$. This raises some doubts about the finiteness of the area under the graphic.

- (b) Use the table of basic primitives to compute $g(b) = \int_0^b f(x) dx$ for $0 \leq b \leq 1$.

$$g(b) = \int_0^b f(x) dx = [\arcsin(x)]_0^b = \arcsin b.$$

- (c) Plot the function $g(b)$ defined in (7b).

See Figure 7a(B).

- (d) What is $\int_0^1 f(x) dx$?

$$\int_0^1 f(x) dx = g(1) = \arcsin 1 = \frac{\pi}{2}.$$

8. In Lecture 8 we have studied the *area under the (glucose) curve* (AUC) as a measure to diagnose gestational diabetes. In particular, we have shown that experimental data have a good fit with functions of type $y(t) = (a + bt) \exp(-ct)$, where $a, b, c > 0$, t is the time in minutes and y is the level of glucose in mmol/l.

- (a) Give a general formula for the AUC $\int_0^{120} y(t) dt$ in terms of parameters a , b and c . Using the values $a = 4.0014$, $b = 0.20549$ and $c = 0.018858$, obtain the result given in Slide 7 of Lecture 8. What is the average level of glucose of this subject?

The integral can be reduced to integrals already computed in Problem 6 since

$$\int_0^{120} y(t) dt = \int_0^{120} (a + bt) \exp(-ct) dt = a \int_0^{120} \exp(-ct) dt + b \int_0^{120} t \exp(-ct) dt.$$

Therefore,

$$\int_0^{120} y(t) dt = \frac{-a}{c} [e^{-ct}]_0^{120} + \frac{-b}{c^2} [(ct + 1) e^{-ct}]_0^{120} \approx 571.7743117.$$

- (b) Show that the average level of glucose tends to zero if we increase the time interval indefinitely. Hint: use that $\int_0^{+\infty} y(t) dt := \lim_{M \rightarrow \infty} \int_0^M y(t) dt$ is finite.

Using the last formula provided we see that

$$\int_0^{+\infty} y(t) dt := \lim_{M \rightarrow \infty} \int_0^M y(t) dt = \frac{-a}{c} \lim_{M \rightarrow \infty} [e^{-ct}]_0^M + \frac{-b}{c^2} \lim_{M \rightarrow \infty} [(ct + 1) e^{-ct}]_0^M.$$

Since $\lim_{M \rightarrow \infty} e^{-cM} = 0$ (recall that $c > 0$) and $\lim_{M \rightarrow \infty} M e^{-cM} = 0$,

$$\int_0^{+\infty} y(t) dt = \frac{a}{c} + \frac{b}{c^2}.$$

Numerical integration

Next exercises are computational and have been studied in the Lab sessions. See the `python` and `excel` files in the Lab's folder.

9. Compute a power series expansion of $\int_0^2 \sin(x^2) dx$. How many terms do we need to sum in order to get an error smaller than 10^{-6} .
10. The error in computing numerically a definite integral by using the trapezoidal rule with step h is approximately proportional to h^2 . In fact, it can be proved the following asymptotic formula:

$$\lim_{h \rightarrow 0^+} \frac{I(f) - I_T(f, h)}{h^2} = C,$$

where $C = C_T(f)$ is a constant that depends on f (in fact, on its second derivative). Hence: $I(f) \simeq I_T(f, h) + C h^2$.

- (a) Applying the trapezoidal rule with steps $2h$ and h , obtain an approximation of the asymptotic constant C , and of the error $I(f) - I_T(f, h)$.
- (b) From the previous approximation, guess how many steps n we would need to compute the integral with an error smaller than a given tolerance ε ?
- (c) Apply the results to the example $\int_0^2 \sin(x^2) dx$, explained in Lecture 9, and guess the number of steps for getting an error smaller than 10^{-6} .

Note: You can make some of the computations by hand, but in order to analyze all results in the table of results, you may use a spreadsheet, such as `Excel` or `LibreOffice`.

11. The asymptotic expansion of the error in the approximation of a definite integral by using the trapezoidal rule is of the form

$$I(f) - I_T(f, h) = C_2 h^2 + C_4 h^4 + C_6 h^6 + \dots$$

- (a) Show that, by applying the trapezoidal rule with steps h and $h/2$, we can eliminate $C_1 h^2$ in the previous formula and get a new approximation of $I(f)$ whose error is approximately proportional to h^4 .
- (b) Show that, if we have applied the trapezoidal rule with steps h , $h/2$ and $h/4$, we can repeat the previous argument and get an approximation whose error is approximately proportional to h^6 . Notice that we could repeat the procedure if we had more initial approximations.
- (c) Apply the results to the example $\int_0^2 \sin(x^2) dx$, explained in Lecture 9.

Hint: Read the **Wikipedia** entries for *Richardson's extrapolation* and *Romberg's method*.

12. Adapt the two previous exercises to Simpson's method, having into account that the asymptotic expansion of the error is, in this case, of the form

$$I(f) - I_S(f, h) = C_4 h^4 + C_6 h^6 + \dots$$

13. Apply all the previous methodologies and results to compute $\int_0^1 \frac{\sin x}{x} dx$.

CALCULUS, 2018-19-T3

Bioinformatics degree

Problems. List 7. Differential equations.

1. Assuming that $y = y(x)$, solve the following linear differential equations.

(a) $x y' - 2y = 0$.

(b) $x y' - 2y = x^2$.

(c) $y' - y \tan x = 1$.

(a) $y = C x^2$.

(b) $y = C x^2 + x^2 \ln x$.

(c) $y = \tan x + C \sec x$.

2. Find the particular solution of the following differential equations.

(a) $y' \cos^2 x + y - 1 = 0$, $y(0) = 5$.

(b) $y' + y \tan x = \sec x + \cos x$, $y(0) = 1$.

(a) $y = 1 + 4e^{-\tan x}$.

(b) $y = \sin x + (x + 1) \cos x$.

3. Study the bifurcations of equilibria and plot the bifurcation curves in terms of the parameter r in the following cases:

(a) $\dot{x} = r x - x^2$ (*transcritical bifurcation* at $r = 0$).

(b) $\dot{x} = r + x^2$ (*saddle-node bifurcation* at $r = 0$).

(c) $\dot{x} = r x - x^3$ (*pitchfork bifurcation* at $r = 0$).

In each case, determine the stability character of each branch of the bifurcation diagram.

[Done in class.](#)

4. In Lecture 10-11, we have studied the logistic growth model

$$\dot{x} = r x (1 - x/K), \quad r > 0, K > 0,$$

where x stands for the size of a biological population, r is the birth rate and K is the carrying capacity (e.g., due to the competition for food).

(a) Prove that $x = 0$ and $x = K$ are the only equilibria and determine their stability.

- (b) Consider the logistic growth model with linear harvesting¹ introduced in Lecture 10-11 which assumes the harvesting to be modelled as $p x$, where $p > 0$ is a parameter:

$$\dot{x} = r x (1 - x/K) - p x, \quad r > 0, K > 0, p > 0. \quad (1)$$

Prove that the population can be preserved if we take $p < r$ and plot (qualitatively) the curve of equilibria versus p , for $p \in [0, 2r]$ (assuming r and K fixed). What does it happen, in biological terms, if $p > r$?

- (c) Match the bifurcation described in (4b) with one of the bifurcations studied in Problem 3.

(a) Clearly, $f(x) = r x (1 - x/K)$ is a polynomial of degree 2 in x whose two unique roots are $x = 0$ and $x = K$. On the other hand, $f'(x) = r (1 - 2x/K)$ and so, $f'(0) = r$ and $f'(K) = -r$. Then, for $r < 0$ (respectively, $r > 0$), $x = 0$ is a stable (resp., unstable) equilibrium and $x = K$ is an unstable (resp., stable) equilibrium. For $r = 0$, both equilibria coincide and $x = 0$ is a semistable equilibrium.

(b) This problem is very similar to problem (3a), so we are going just to sketch it. Here, only $x \geq 0$ is biologically meaningful, and r and K are “static” parameters, while the varying parameter is $p \geq 0$, that controls harvesting (or describes emigration). The fixed points of (1) are the zeroes of $f(x, p) = r x (1 - x/K) - p x$, which are 0 and $x_p = K(1 - p/r)$. Three cases are presented:

- $0 \leq p < r$: 0 is unstable and $x_p > 0$ is stable. The population goes to x_p when time increases.
- $p = r$: $0 = x_p$ is a semistable fixed point. The population goes to 0.
- $p > r$: $x_p < 0$ is unstable (no biological meaning) and 0 is stable. The population is extinguished.

(c) The bifurcation at $p = r$ is a transcritical bifurcation.

5. The growth of cancerous tumors can be modeled by the Gompertz law $\dot{N} = -a N \log(b N)$, where $N(t)$ is proportional to the number of cells in the tumor, and $a, b > 0$ are parameters. The predictions of this simple model agree surprisingly well with data on tumor growth, as long as N is not too small.

(a) Use phase line analysis to determine the stability of the equilibria.

(b) Give an explicit solution of the model assuming that $N(0) = e/b$ and show that the corresponding orbit tends to $N = 1/b$.

A simple analysis of the graph of $f(N) = -a N \log(b N)$ reveals that the only fixed points are 0 (unstable) and $x_1 = \frac{1}{b}$ (stable). It is also easy to draw the phase line curves in the x line. Notice that in this case f' is not defined at $x = 0$ and $\lim_{N \rightarrow 0^+} f'(N) = +\infty$; besides, $f'(x_1) = -a < 0$.

A more quantitative analysis let us to solve directly the equation, by using separation of

¹Also known as Schaeffer model, often used to model the impact of harvesting on a fish population.

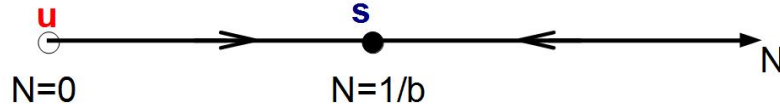
variables. A variant consists in considering that the time t to go from N_0 to N :

$$\begin{aligned} t &= \int_{N_0}^N \frac{1}{f(N)} dN = \int_{N_0}^N \frac{1}{-aN \log(bN)} dN \\ &= \int_{x_0}^x \frac{-1}{ax} dx = \left[\frac{-1}{a} \log |x| \right]_{x_0}^x = \frac{-1}{a} \log \left| \frac{x}{x_0} \right| \\ &= \frac{-1}{a} \log \left| \frac{\log(bN)}{\log(bN_0)} \right| = \frac{-1}{a} \log \frac{\log(bN)}{\log(bN_0)}, \end{aligned}$$

where in second line we have performed the change of variable $x = \log(bN)$, and $x_0 = \log(bN_0)$, and in the last equality we use that both $\log(bN)$ and $\log(bN_0)$ have the same sign (why?). Isolating N from this equality we get the Gompertzian growth

$$N(t) = \frac{1}{b} (bN_0)^{e^{-at}}.$$

Notice that, for any initial population $N_0 > 0$, we get $\lim_{t \rightarrow +\infty} N(t) = \frac{1}{b}$, as we knew from the phase line analysis, but with this quantitative analysis we know how the evolution goes to the stationary state $\frac{1}{b}$, see figure.

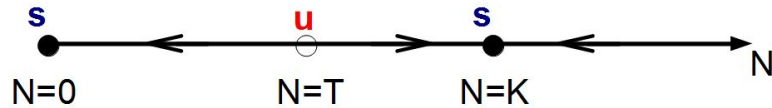


6. Use phase line analysis to determine the stability of the equilibria for the model

$$\frac{dN}{dt} = -N \left(1 - \frac{N}{T} \right) \left(1 - \frac{N}{K} \right), \quad 0 < T < K,$$

where the parameter K has the same biological meaning that it has in the logistic model. Assuming that N is a population, explain the biological meaning of T .

After the phase line analysis, we observe that T is the minimum population to grow (that is, a threshold for survival), since if the initial population is $N_0 < T$, then it goes to extinction, see figure.



7. Consider the *integrate and fire model*, introduced in Lecture 10-11, which is governed by equation

$$C \frac{dv}{dt} = -g_L (v - V_L) + I,$$

under the following rule: if v exceeds a certain value V_{thr} , then v is reset to a lower value V_{reset} . Typically, $V_{thr} \approx -30\text{mV}$ and $V_{reset} \approx -70\text{mV}$, while g_L is the order of 1 and $V_L \approx -65\text{mV}$. In Lecture 10-11, we proved that $\lim_{t \rightarrow +\infty} v(t) = V_L + I/g_L$. Assuming that $V_L + I/g_L > V_{thr}$, prove that the solution of the differential equation will be periodic with period

$$T = \frac{-C}{g_L} \log \left(\frac{V_{thr} - V_L - I/g_L}{V_{reset} - V_L - I/g_L} \right).$$

We are asked to compute the time to go from V_{res} to V_{thr} , which is

$$T = \int_{V_{res}}^{V_{thr}} \frac{C}{-g_L(v - V_L) + I} dv = \int_{V_{res}}^{V_{thr}} \frac{\frac{-C}{g_L}}{v - V_L - \frac{I}{g_L}} dv = \left[\frac{-C}{g_L} \log \left| v - V_L - \frac{I}{g_L} \right| \right]_{V_{res}}^{V_{thr}},$$

and leads to the result. Notice that the denominator does not vanish in the interval of integration, since $V_L + \frac{I}{g_L} > V_{thr}$ (do a phase line analysis to justify that).

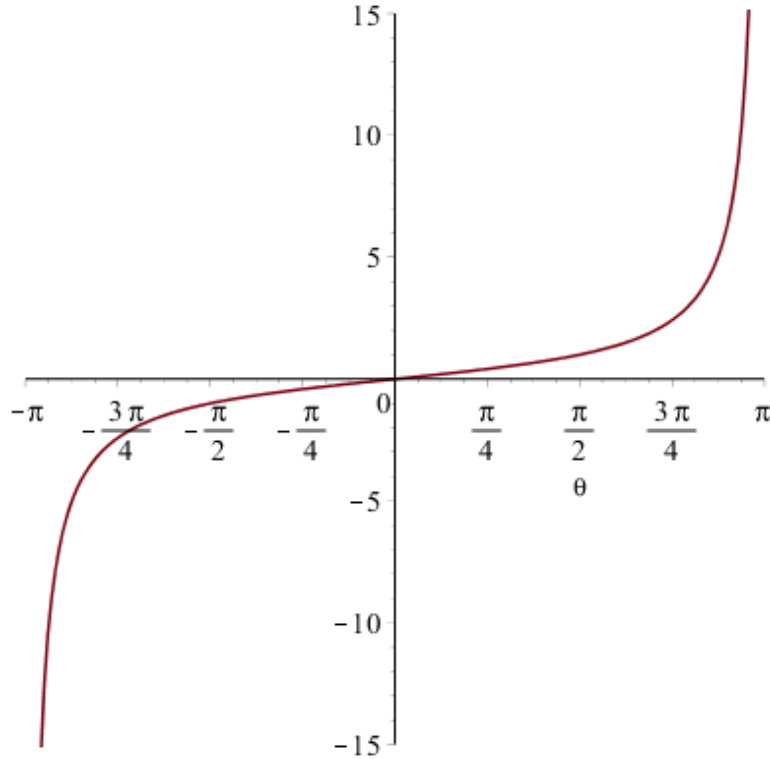
Note: We can also obtain T by finding the solution $v = v(t)$ of the ODE with $v(0) = V_{res}$, and then finding T so that $v(T) = V_{thr}$. The solution of the ODE can be done by different methods, such as separation of variables (which is close to what we have done above), transforming into a linear homogeneous equations (doing the change of variable $x = v - V_L - \frac{I}{g_L}$ we are lead to equation $\dot{x} = -\frac{g_L}{C}x$), or the variation of constant method. The last method works also for I depending on t .

8. The θ -model is a 1D model of neuronal dynamics that assumes the angle θ to be the phase along an oscillation. It is written as

$$\frac{d\theta}{dt} = 1 - \cos \theta + (1 + \cos \theta) I, \quad \theta \in [-\pi, \pi] \quad (2)$$

and θ is represented as angle on a circle. Moreover, when $\theta = \pi$, the neuron is supposed to reach the “spike” threshold.

- (a) Plot the function $x = \tan(\theta/2)$ for $\theta \in [-\pi, \pi]$.



- (b) Prove that equation (2) can be brought into equation

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

by means of the change of variables $x = \tan(\theta/2)$. From Problem 3, observe that the model undergoes a saddle-node bifurcation at $I = 0$.

Applying chain's rule and using equation (2), we have:

$$\frac{dx}{dt} = \frac{dx}{d\theta} \frac{d\theta}{dt} = (1 + \tan^2(\theta/2)) \frac{1}{2} (1 - \cos \theta + (1 + \cos \theta) I,). \quad (3)$$

We need to express $\cos \theta$ in terms of x . For that, we use trigonometric identities:

$$\cos \theta = \cos^2(\theta/2) - \sin^2(\theta/2) = \frac{1 - \tan^2(\theta/2)}{1 + \tan^2(\theta/2)} = \frac{1 - x^2}{1 + x^2},$$

and so

$$1 - \cos \theta = \frac{2x^2}{1 + x^2}; \quad 1 + \cos \theta = \frac{2}{1 + x^2}.$$

Substituting them in (3), we obtain:

$$\frac{dx}{dt} = (1 + x^2) \frac{1}{2} \left(\frac{2x^2}{1 + x^2} + \frac{2}{1 + x^2} I, \right) = x^2 + I. \quad (4)$$

- (c) Consider $I > 0$ and take an initial condition $x(0) = x_0$. Prove that the time to reach the spike threshold can be expressed as

$$T_{spike} = 1/2 \left(-2 \arctan \left(\frac{x(0)}{\sqrt{I}} \right) + \pi \right) \frac{1}{\sqrt{I}}.$$

In particular, if we reset $x(t)$ to $-\infty$ ($\theta = -\pi$) the *total transit time* to reach $+\infty$ ($\theta = \pi$) is

$$T_{per} = \frac{\pi}{2\sqrt{I}}.$$

Spike threshold is reached when $\theta = \pi$, so $x = \tan \frac{\pi}{2} = +\infty$. Since $\frac{dx}{dt} = x^2 + I$,

$$\int_{x_0}^{+\infty} \frac{dx}{x^2 + I} = \int_0^{T_{spike}} dt$$

since $x(0) = x_0$ and the spike time is determined by $\lim_{t \rightarrow T_{spike}^-} x(t) = +\infty$.

We solve the integral in x , whose primitive is $\frac{1}{\sqrt{I}} \arctan \frac{x}{\sqrt{I}}$ and substitute by $x \rightarrow +\infty$ and x_0 . Therefore,

$$T_{spike} = \frac{1}{\sqrt{I}} \left(\lim_{x \rightarrow +\infty} \arctan \frac{x}{\sqrt{I}} - \arctan \frac{x_0}{\sqrt{I}} \right) = \frac{1}{\sqrt{I}} \left(\frac{\pi}{2} - \arctan \frac{x_0}{\sqrt{I}} \right),$$

as we wanted to prove.

Finally, setting $x(0) = 0$, we easily get $T_{per} = \frac{\pi}{2\sqrt{I}}$.