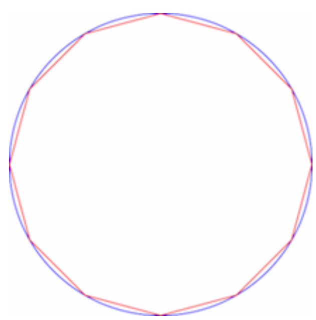
Theory Involved in Projects

Syed Sajid Husain Rizvi

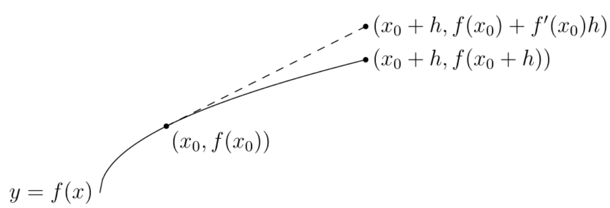
May 31, 2020

Euler's Method

Euler's method is used for approximating solutions to certain differential equations and works by approximating a solution curve with line segments.In the image below, the blue circle is being approximated by the red line segments.

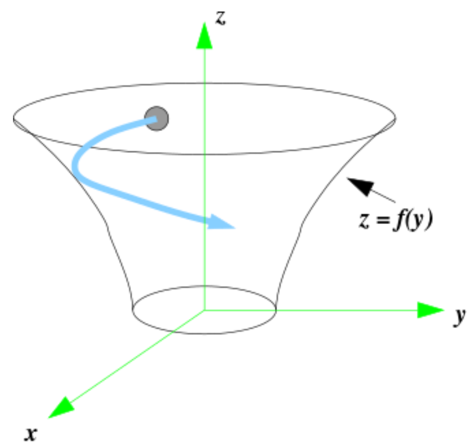


For a differential equation of the form $y' = f(x)$ where $y(x_0) = y_0$, a sequence of points x_n, y_n satisfying $x_{n+1} = x_n + h$ and $y_{n+1} = y_n + hf(x_n)$, where h is the step size, can be used to approximate $y(x)$ for some x.The idea is to choose h,the step size, so that one of the x_n will be the $y-$value to be estimated.



Gravity Well

Consider the following funnel-like surface. It is obtained by rotating the curve $z = f(y)$ around the the z-axis. For instance, $f(y) = \sqrt{y-1}$ would produce a surface like the following.



A ball is moving along the inner surface of the funnel. We shall ignore the radius of the ball and the friction of the surface. (Thus the ball is a point mass slipping, not rolling, on the funnel.) We know the initial position and velocity of the ball. We want to find out the path that the ball will follow. There are two forces acting on the ball: its weight and the reaction of the surface. The first works downwards, and so is

$$\begin{bmatrix} 0 \\ 0 \\ -mg \end{bmatrix}.$$

The reaction acts inwards along the normal to the surface at the current position of the ball. Let the current position of the ball be

$$\begin{bmatrix} t \\ y \\ f(u) \end{bmatrix},$$

where $u = \sqrt{t^2 + y^2}$. A little coordinate geometry shows that the normal lies along

$$\begin{bmatrix} -t \\ -y \\ u/f'(u) \end{bmatrix}.$$

So the reaction force is

$$\mathrm{R}\begin{bmatrix} -t \\ -y \\ u/f'(u) \end{bmatrix},$$

for some unknown function R of t. So we have the equation of motion:

$$\mathrm{m}\begin{bmatrix} x'' \\ y'' \\ z'' \end{bmatrix} = R\begin{bmatrix} -t \\ -y \\ u/f'(u) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ -mg \end{bmatrix}.$$

Thus, we have 3 equations in 3 unknowns: t,y, and R. Notice that z is a known function of t and y. To simplify the equations first find z" in terms of t,y and their derivatives. Then eliminate R to get two equations in two unknowns:

$$\begin{aligned} x'' &= -t\tilde{R} \\ y'' &= -y\tilde{R}, \end{aligned}$$

where

$$\tilde{R} = \frac{f'(u)\left(x'^2+y'^2-u'^2\right)/u+u'^2f''(u)+g}{u\left(f'(u)+\frac{1}{f'(u)}\right)}.$$

Newton Raphson

The Newton-Raphson method (also known as Newton's method) is a way to quickly find a good approximation for the root of a real-valued function f(x) = 0f(x)=0. It uses the idea that a continuous and differentiable function can be approximated by a straight line tangent to it.

How It Works

Suppose you need to find the root of a continuous, differentiable function $f(x)$, and you know the root you are looking for is near the point $x = x_0$. Then Newton's method tells us that a better approximation for the root is

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$$

$$x_n = x_{n-1} - \frac{f(x_{n-1}}{f'(x_{n-1}}$$

Maximum Likelihood Estimation

Maximum likelihood estimation (MLE) is a technique used for estimating the parameters of a given distribution, using some observed data. For example, if a population is known to follow a normal distribution but the mean and variance are unknown, MLE can be used to estimate them using a limited sample of the population, by finding particular values of the mean and variance so that the observation is the most likely result to have occurred.

Formal Definition

Let $X_1, X_2, X_3, ..., X_n$ be a random sample from a distribution with a parameter θ. Suppose that we have observed $X_1 = x_1, X_2 = x_2, \cdots, X_n = x_n$. If X_i's are discrete, then the likelihood function is defined as

$$L(x_1, x_2, \cdots, x_n; \theta) = P_{X_1X_2\cdots X_n}(x_1, x_2, \cdots, x_n\theta).$$

If X_i's are jointly continuous, then the likelihood function is defined as

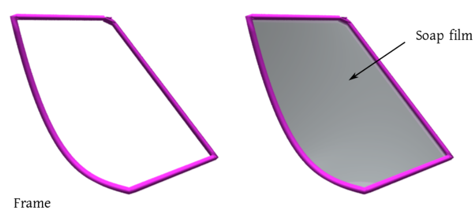
$$L(x_1, x_2, \cdots, x_n; \theta) = f_{X_1 X_2 \cdots X_n}(x_1, x_2, \cdots, x_n; \theta).$$

In some problems, it is easier to work with the log likelihood function given by

$$\ln L(x_1, x_2, \cdots, x_n; \theta).$$

Linear Equations Behind A Soap Film

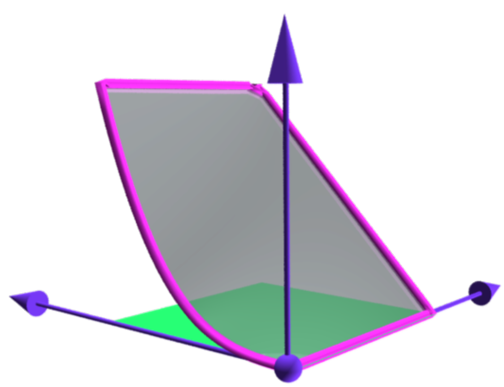
If you dip a wire frame in a soap solution, then a thin film of soap water will cling to it.



Given the shape of the frame, we want to find the natural shape of the film. This is an important question in architecture, where a structure must be given the most natural shape to reduce stress. It is known that a soap film will always occupy the position that minimises its elastic potential energy. In the next section we shall see how to express this mathematically.

Coordinate System

We imagine an xy-plane under the frame:



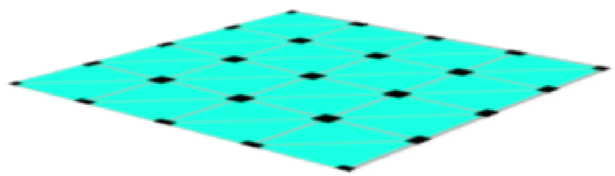
If the surface is given by a function $u(x, y)$,then its elastic potential energy is given by

$$E(u) = \int \int_R (u_x)^2 + (u_y)^2 \, dxdy.$$

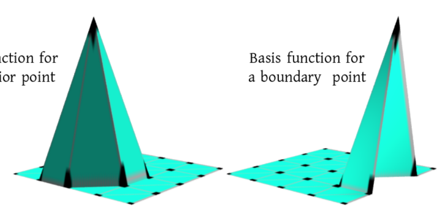
This is to be minimised subject to the boundary condition that the $u(x, y)$ must match the frame height at the boundary. In general, this is a difficult/impossible problem to solve analytically.

Numerical Solution

To proceed numerically, one starts with a triangulation of the base.



Then the aim is to find the value of $u(x, y)$ at the vertices. Let c_j denote the value of $u(x, y)$ at the j-th vertex. Since the target function involves u_x and u_y, we need to somehow approximate them using only the values at the vertices. For this we choose a set of basis functions, one for each vertex. It is constructed by "pulling up" the vertex to a height 1, while leaving all the other vertices at height 0. Here are two examples:



Notice that the graph of each basis function (ϕ_j) is a plane over each triangle (T_i). and hence we may write a basis function $\phi_j(x, y)$ as

$$\phi_j(x, y) = \alpha_{ij} + \beta_{ij}x + \gamma_{ij}y \quad \text{for}(x, y) \in T_i.$$

for suitable numbers $\alpha_{ij}, \beta_{ij} and \gamma_{ij}$. Also, notice that $\alpha_{ij}, \beta_{ij}, \gamma_{ij}$'s are zero if the j-th vertex is not part of T_i. Thus, most of these numbers are actually zero. Then we can approximate $u(x, y)$ by

$$u(x, y) = \sum_j c_j \phi_j(x, y).$$

Thus, the problem of finding $u(x, y)$ reduces to finding the finitely many numbers c_j's. Now

$$u_x(x,y) = \sum_j c_j\beta_{ij} \text{ for } (x,y) \in T_i^\circ,$$

where T_j° denotes the interior of T_j. Similarly for $u_y(x,y)$. Hence we have

$$E(u) = \sum_i \int\int_{T_i^\circ} (\sum_j c_j\beta_{ij})^2 + (\sum_j c_j\gamma_{ij})^2 = \sum_i |T_i|\{(\sum_j c_j\beta_{ij})^2 + (\sum_j c_j\gamma_{ij})^2\},$$

where $|T_i|$ denotes the area of T_i (same as the area of T_i°). Thus,

$$E(u) = \underline{c}' M \underline{c},$$

where M is the NND matrix with (j,j')-th entry given by

$$m_{jj'} = \sum_i |T_i|(\beta_{ij}\beta_{ij'} + \gamma_{ij}\gamma_{ij'}).$$

Solving

Suppose that the last k of the c_j's are known frame heights. Partition $\underline{c}$ as $(\underline{c}_1, \underline{c}_2)$. Then $\underline{c}_2$ is known, and $\underline{c}_1$ is to be chosen to minimise $E(u)$. Let us partition M accordingly as

$$M = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix}.$$

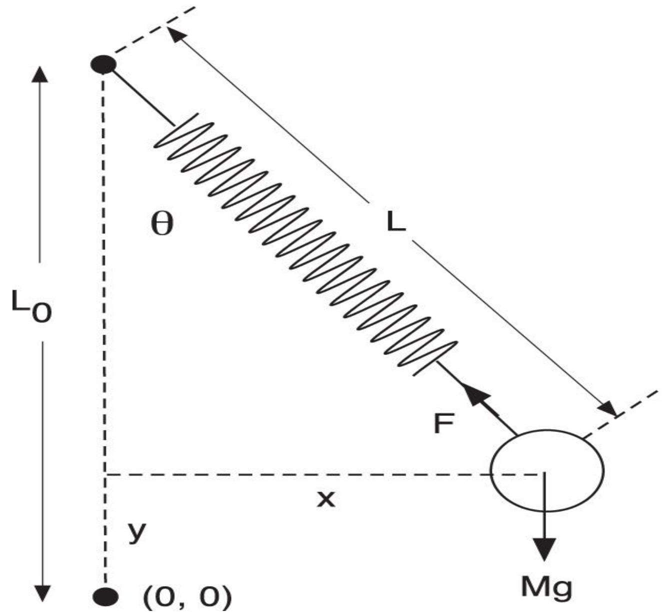
Then

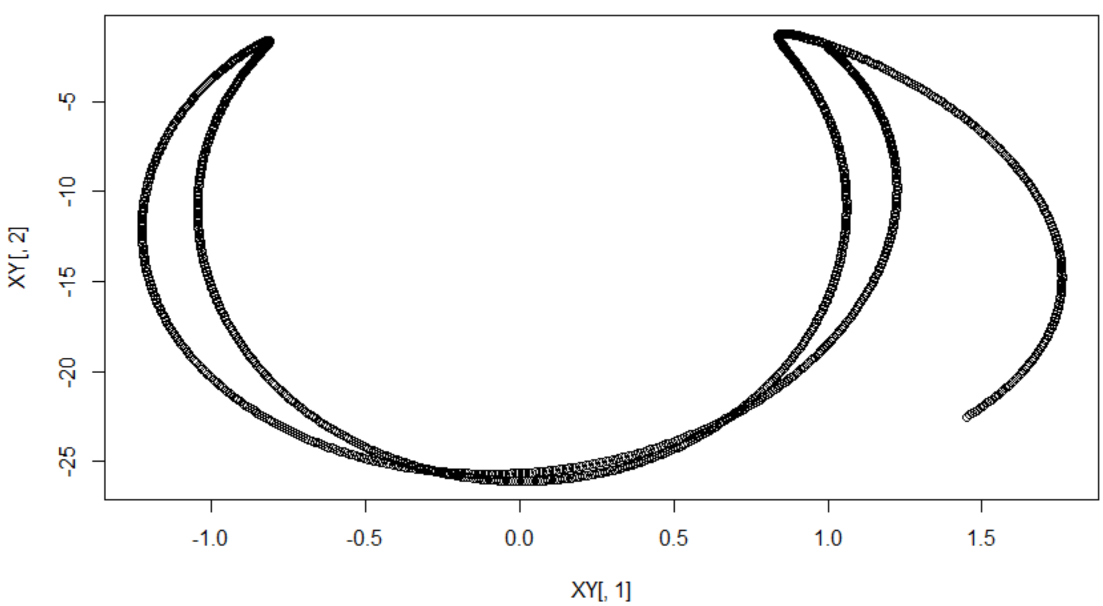
$$\underline{c}' M \underline{c} = \begin{bmatrix} \underline{c}_1' & \underline{c}_2' \end{bmatrix} \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \begin{bmatrix} \underline{c}_1 \\ \underline{c}_2 \end{bmatrix} = \underline{c}_1' M_{11} \underline{c}_1 + 2\underline{c}_1' M_{12} \underline{c}_2 + \underline{c}_2' M_{11} \underline{c}_2.$$

Differentiating w.r.t. $\underline{c}_1$, and equating to zero, we get

$$M_{11}\underline{c}_1 = -M_{12}\underline{c}_2.$$

It is a simple linear algebra exercise to show that this is always consistent. In fact, M_{11} will also be nonsingular (not easy to prove). So the problem always has unique solution.





Elastic Pendulum Simulation

EULER'S METHOD

Syed Sajid Husain Rizvi | B.Stat | 31/05/2020

Problem

Simulate an elastic pendulum with spring constant γ, where tension in the spring is given by $T = \gamma\left(\sqrt{x^2 + y^2} - l\right)$ where l is the unstretched length of spring. Numerically solve these assuming that $l = 4, x(0) = 1, y(0) = -2, x'(0) = y'(0) = 0$. Animate to see if the solution looks natural.

Algorithm

- Write down the forces acting on the bob.
- Resolve the forces into x, y components
- Use Newton's Laws of Motion and form equations
- Apply Euler's Method to solve the problem numerically
- Tweak dt until the solution looks natural.

Calculations

$$a_x = -\frac{\gamma\left(\sqrt{x^2 + y^2} - l\right)x}{m\left(\sqrt{(x^2 + y^2\}}\right)}$$

$$a_{\mathrm{y}} = -\frac{\gamma\left(\sqrt{x^2 + y^2} - l\right)\mathrm{y}}{m\left(\sqrt{(x^2 + y^2\}}\right)} - g$$

$$v_{i+1} = v_i + dt \times a_i$$

$$\mathrm{p}_{i+1} = \mathrm{p}_i + dt \times p_i$$

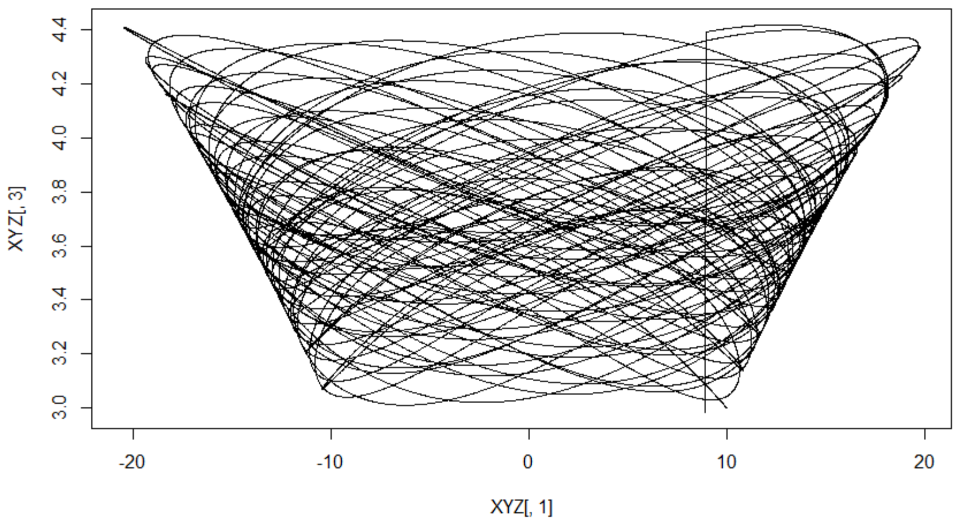
Here, p_i denotes the position of the pendulum at the i^{th} step.

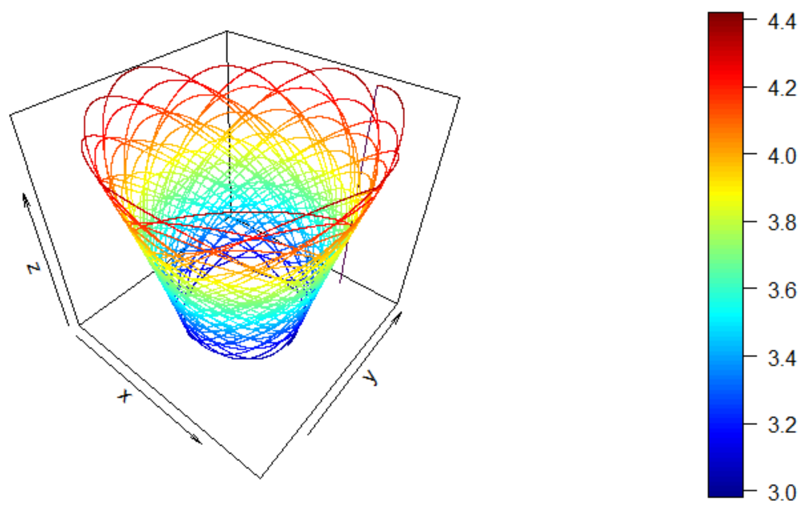
Using the above equations, we can keep on finding the successive values using Euler's Method.

Conclusion

Keeping $m, \gamma = 1$, with $dt = 0.01$ gives a natural looking solution with $n = 1500$ steps.

For some other values of dt, the pendulum seems to move in a haphazard manner.





Gravity Well Simulation

SECOND ORDER TAYLOR METHOD

Syed Sajid Husain Rizvi | B.Stat | 31/05/2020

Problem

Use 2nd order Taylor method to simulate the Gravity Well for the following conditions

$$t(0) = 10, y(0) = 0, x'(0) = 0, y'(0) = 5$$

Take $g = 9.8$

Algorithm

- Find forces acting on the ball
- Resolve the forces into x, y, z components
- Assume initial coordinates and use the laws of motion
- Simplify and solve the differential equation using 2nd order Taylor's Method

Calculations

$$u = \sqrt{x^2 + y^2}$$

$$f(u) = \sqrt{u-1}$$

$$z = f(u)$$

$$u' = \frac{\mathrm{xx'} + yy'}{\sqrt{x^2 + y^2}}$$

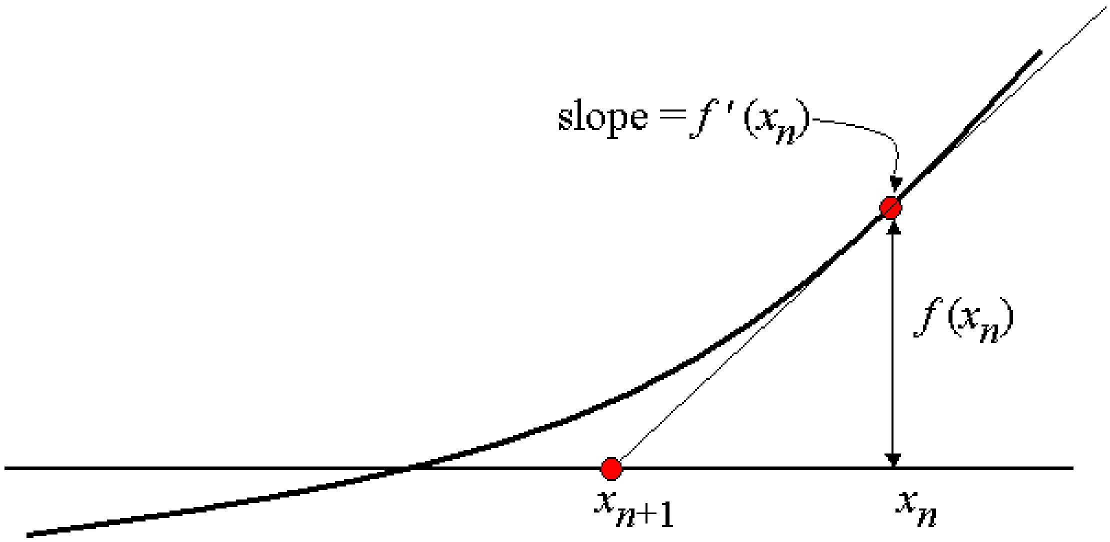
$$f'(u) = \frac{1}{2\sqrt{u-1}}$$

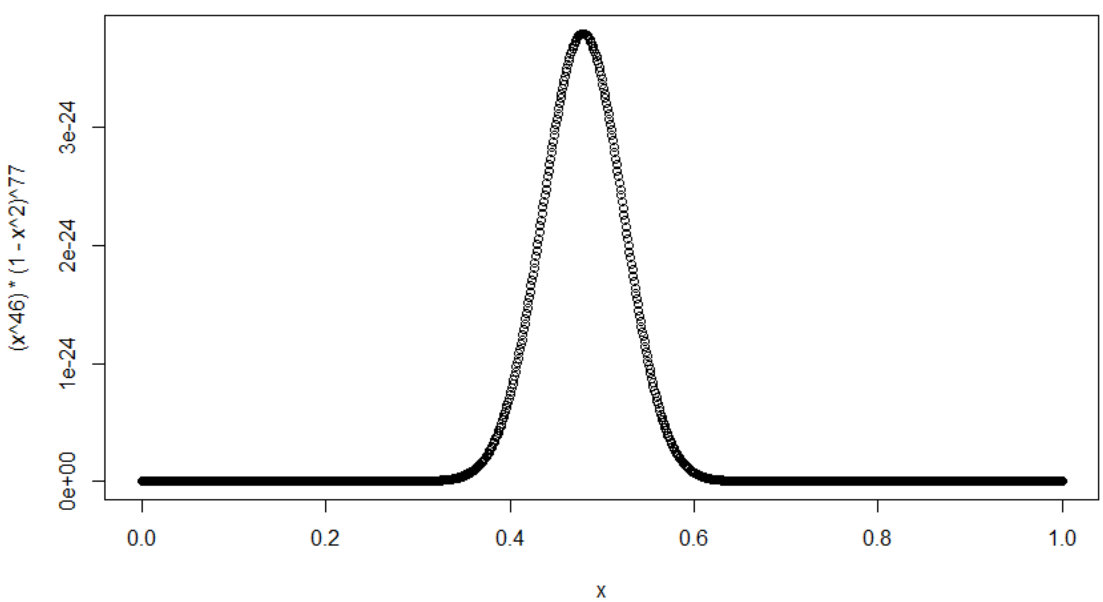
$$f''(u) = -\frac{1}{4(u-1)^{3/2}}$$

Using these we can find $\overline{R}$, and hence (x'', y'') given initial values of x, y, x', y'.

Conclusion

The results obtained by Taylor's Method with 100000 iterations with a time step of 0.01 seconds gives very precise results.





Estimation in Bio-Statistics

PROBABILITY OF RECEIVING AN ALLELE

Syed Sajid Husain Rizvi | B.Stat | 30/05/2020

Problem

A certain trait in rabbits is controlled by a pair of alleles, a and A. Each rabbit receives one of these from the father and another from the mother. Thus, the possible pairs are $aa, aA\ and\ AA$(order is immaterial). The probability that a parent gives an a to the offspring is p (unknown). So, the probability of an A is $q = 1 - p$. The father's contribution is independent of the mother's, and so the probabilities of $aa, aA\ and\ AA$ in the offspring are, respectively, $p^2, 2pq\ and\ q^2$. Our aim is to estimate p. Unfortunately, it is impossible to detect the pair an offspring has. It is only possible to detect if an offspring has at least one A, $i.e.$, whether aa or aA, AA. The probabilities are, respectively, p^2 and $q^2 + 2pq$. In a random sample of 100 offsprings, only 23 are without any A. The probability of this is

$$L(p) = p^{46}(q^2 + 2pq)^{77}$$

The value of $p \in (0,1)$ for which this is the maximum is called the **maximum likelihood estimator** of p. Find it.

Calculations

$$L(p) = p^{46}(q^2 + 2pq)^{77}$$

$$L'(p) = 2p^{45}(1 - p^2)^{76}(23 - 100p^2)$$

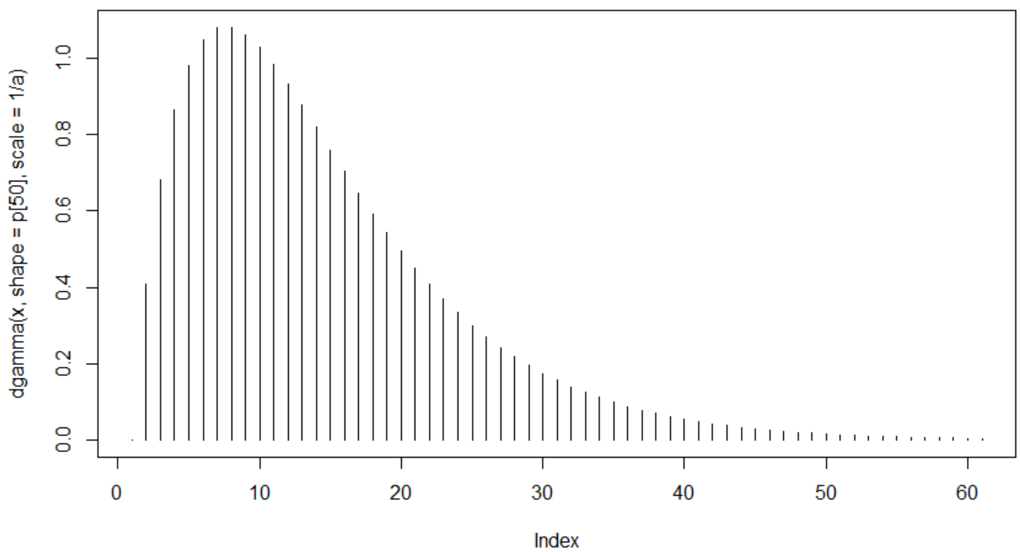
$$L''(p) = 2p^{44}(1 - p^2)^{75}(19900p^4 - 100p^2 - 9131p^2 + 1035)$$

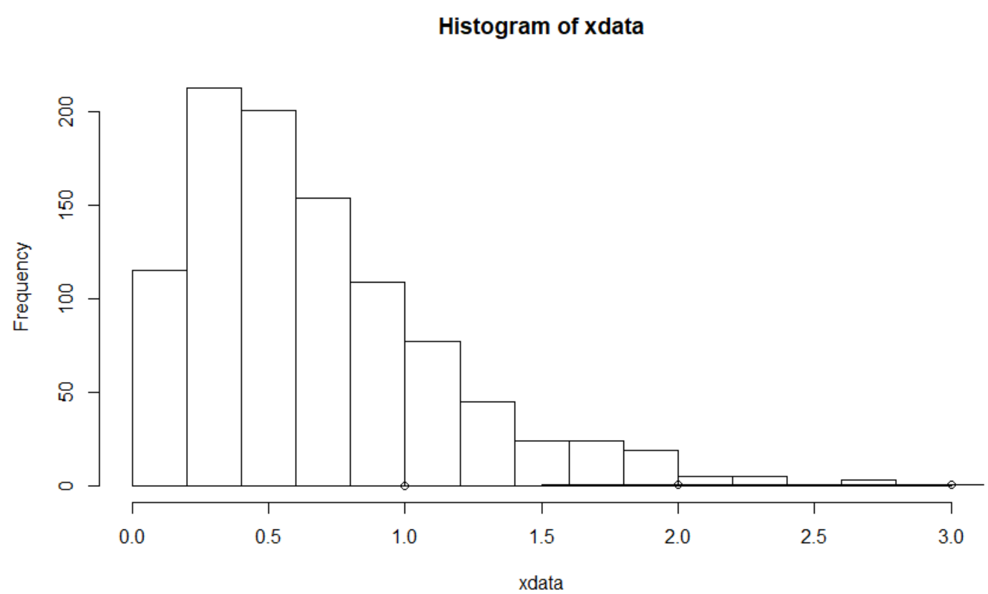
We need to find the positive root of $L'(p)$ using Newton-Raphson method.

$$p_{n+1} = p_n - L'(p)/L''(p)$$

Conclusion

The value of p converges to 0.4795832 which is a very good approximation to the analytic solution of $p = \sqrt{23}/10$. Since, $L''(0.4695832) < 0$, the value indeed maximises $L(p)$.





Fitting a Distribution

ESTIMATING UNKNOWN CONSTANTS BY MLE

Syed Sajid Husain Rizvi | B.Stat | 30/05/2020

Problem

The file data.txt has $n = 996$ random numbers that are generated from the density

$$f(x; p, a) = \frac{a^p}{\Gamma(p)} x^{p-1} e^{-ax}, \quad x > 0$$

for unknown constants $p, a > 0$. The principle of maximum likelihood estimation suggests estimating p, a by maximizing

$$L(p, a) = \prod_{i=1}^{n} f(x_i; p, a)$$

where $x_1, \dots, x_n$ are the data in the file. Perform this estimation, and check your answer graphically by overlaying the graph of $f(x_i; p, a)$ on the histogram of the data

Algorithm

- Obtain the expression for $L(p, a)$
- Take logarithm on both sides and equate the partial derivates w.r.t. p, a with 0
- Numerically solve those two equations for values of p, a
- Compare the histogram and the graph of $f(x_i; p, a)$

Observations

$$\overline{x} = 0.676100416405622$$

$$\overline{ln(x)} = -0.670380719903167$$

Numerical Solution

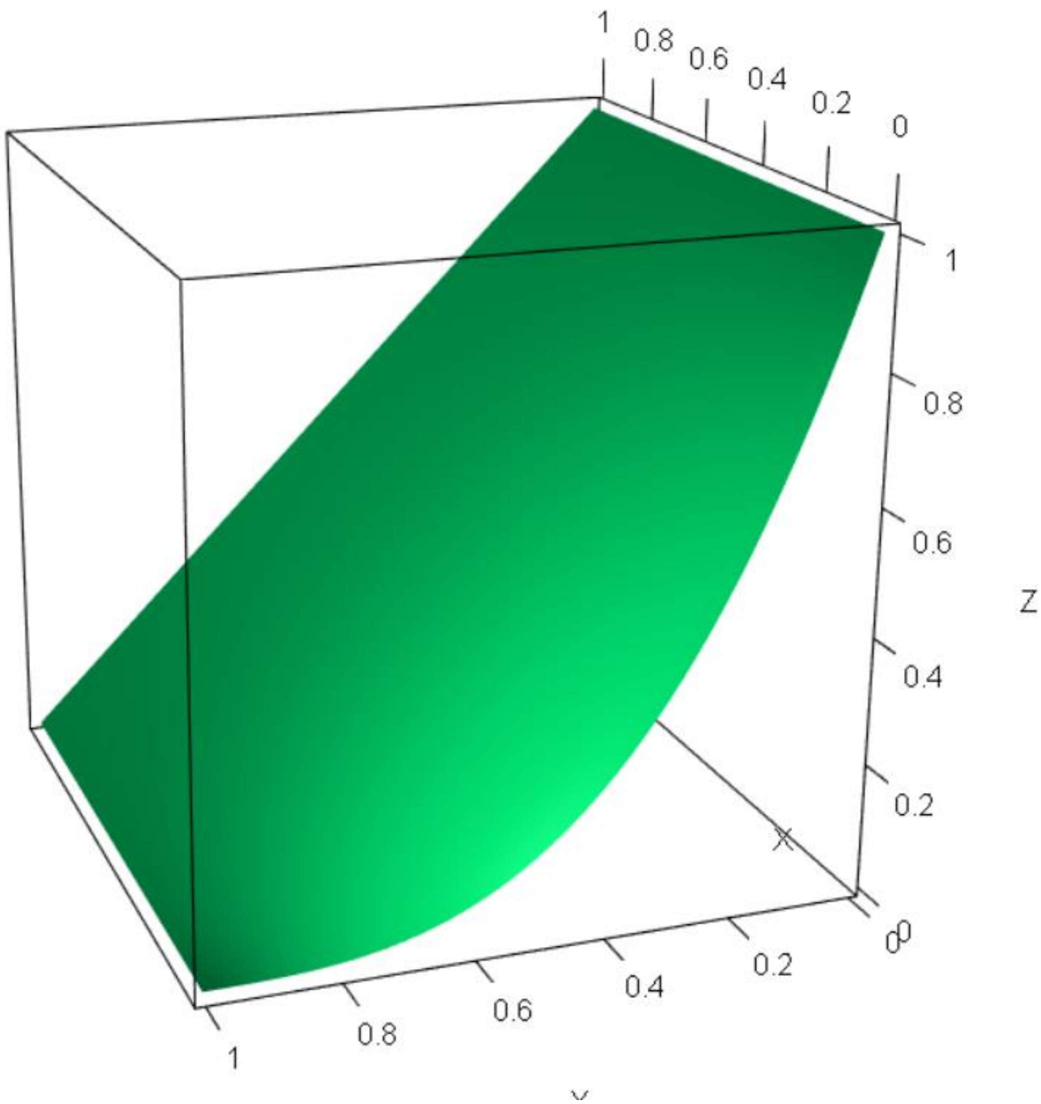
Equating the partial derivatives with 0 gives,

$$p = a\overline{x}$$

$$\psi(p) - ln(p) + ln(\overline{x}) - \overline{ln(x)} = 0$$

The second equation is well behaved but has no closed form solution. Hence, we used Newton-Raphson method to solve it.

Upon solving we get, $p = 1.94244969540639$ and $a = 2.8730195223560$



Soap Film Simulation

FINDING SHAPE OF SOAP FILM WITH GIVEN FRAME

Syed Sajid Husain Rizvi | B.Stat | 28/05/2020

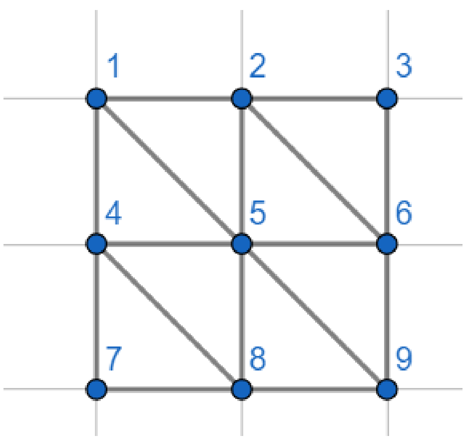
Problem:

Find the natural shape of the soap film where the base is the unit square, and the boundaries are the graphs of the functions 0, x, 1 and x^3. Use 50 subdivisions for each side.

Algorithm

- Start with a triangulation of the base.
- Numerate the interior points and then the exterior points
- Choose a set of basis functions
- Find the potential energy, E(u)
- Find M and c such that E(u)= cMc'
- Minimize E(u)
- Partition c as (c1:c2) where c2 is known.
- Partition M accordingly as well.
- Find c1 from M11c1+M12c2=0 (M11 is invertible)

Observations:



Let n denote the number of subdivisions.

Pulling up the 5th point to a height of 1. We have the following planes:

Points	Coefficient of X	Coefficient of Y
1,4,5	50	0
1,2,5	0	-50
2,6,5	-50	-50
6,9,5	-50	0
8,9,5	0	50
4,5,5	50	50

- Since the problem is symmetric, just using these 6 planes we can find every other plane. Eg: Plane formed by 4,8,7(height 1) is translation of the plane formed by 2,6,5(height 1).

Since, $m_{jj'} = \sum_i |T_i| \left(\beta_{ij}\beta_{ij'} + \gamma_{ij}\gamma_{ij'}\right)$

- For two consecutive points j,j' on the same horizontal line (like 4 & 5) $m_{jj'} = -2n^2|T_i|$
- For two consecutive points j,j' on the same vertical line (like 5 & 8) $m_{jj'} = -2n^2|T_i|$
- For equal j,j' $m_{jj} = 8n^2|T_i|$
- For all other values of j,j $m_{jj'} = 0$

We can create M matrix using the above observations

Shape: