Discrete derivative

Jose Manuel Perez, Ernesto Caceres

September 6, 2023

Abstract

Using the discrete derivative we will analyze the margin of error of three different methods: the forward discrete derivative method, the backward discrete derivative method, and te central discrete derivative method. This paper will explain how these methods work and calculate the error thorugh the infinate norm. The aim is to find which method offers a better approximation.

1 Introduction

2 Analysis

The derivative is the rate of change of a function with respect to one of its variables. The formal definition of the derivative of a function with one variable x is.

$$f'(x) = \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x}$$

For the discrete approximation of the derivative, we will assume that we want to evaluate the derivative for a discrete range $[0,1]$, such that N is the number of subdivisions of that range. $h = \frac{1}{N}$ represents the spacing between each subdivision, meaning the discrete form of Δx. As we increase the number of subdivisions N, h dicreases. The smaller h is, the closer our discrete approximation will be to the continuous derivative. There are three main methods: the forward approximation, the backward approximation and the central approximation.

2.1 The forward method

$$D_h^{f,1}(u)_j = \frac{u(x_{j+1}) - u(x_j)}{h}$$

where $u(x_j)$ is the image of the function we are differentiating at the j^{th} point. The spacing between each point is h and thus $u(x_{j+1}) = u(x_j + h)$. Since for this method we only consider inner points of the range $[0,1]$, $x_1 = h, x_2 = 2h, \ldots, x_n = nh$ and $1 \leq j \leq N-1$.

2.2 The backward method

following the same reasoning, we can obtain the backward method.

$$D_h^{b,1}(j) = \frac{u(x_j) - u(x_{j-1})}{h}$$

2.3 The central method

Using the arithmetic mean of the forward and the backward method:

$$D_h^{c,1}(j) = \frac{D_1^f + D_1^b}{2}$$

$$D_h^{c,1}(j) = \frac{u(x_{j+1}) - u(u_{x-1})}{2h}$$

3 A priori errror

3.1 Taylor series

To calculate the a priori error we can use Taylor series. The Taylor expansion of a function f centered at a is defined as:

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f'''(a)}{3!}(x-a)^3 + \dots$$

For our discrete analysis of the derivative we assume $u(x)$ has 3 continuous derivatives in the analyzed domain and we center the Taylor approximation around x_j.

$$u(x_{j+1}) = u(x_j + h) = u(x_j) + u'(x_j)h + \frac{u''(x_j)h^2}{2} + \frac{u'''(x^*)h^3}{6} \quad , x_j \leq x^* \leq x_j + h$$

$$u(x_{j-1}) = u(x_j + h) = u(x_j) - u'(x_j)h + \frac{u''(x_j)h^2}{2} - \frac{u'''(x^*)h^3}{6} \quad , x_j \leq x^* \leq x_j + h$$

3.2 The infinity norm

The infinity norm measures how large the vector is by the magnitude of its largest entry.

$$||\nu||_{h,\infty} = \max(|x_1|, \dots, |x_n|) = \max_{j=1,\dots,n} |x_j|$$

For the different methods, we can use the infinity norm to calculate the error factor.

3.2.1 Forward method error

Using the Taylor expansion of $u(x_{j+1})$:

$$D_h^{f,1}(j) = \frac{u(x_{j+1}) - u(x_j)}{h}$$

$$= \frac{1}{h}[u(x_j) + u'(x_j)h + \frac{u''(x_j)h^2}{2} + \frac{u'''(x^*)h^3}{6} - u(x_j)]$$

$$= u'(x_j) + \frac{u''(x_j)h}{2} + \frac{u'''(x^*)h^2}{6}$$

$$\implies |D_h^{f,1}(u_j) - u'(x_j)| = |\frac{u''(x_j)h}{2} + \frac{u'''(x^*)h^2}{6}|$$

$$\implies ||D_h^{f,1}(u_j) - u'(x_j)||_{h,\infty} \leq Ah \quad A \in \mathbb{N}$$

3.2.2 Backward method error

Using the Taylor expansion of $u(x_{j-1})$:

$$D_h^{b,1}(j) = \frac{u(x_j) - u(x_{j-1})}{h} =$$

$$= \frac{1}{h}[u(x_j) - u(x_j) + u'(x_j h - \frac{u''(x_j)h^2}{2} + \frac{u'''(x^*)h^3}{6}]$$

$$= u'(x_j) - \frac{u''(x_j)h}{2} + \frac{u'''(x^*)h^2}{6}$$

$$\implies |D_h^{b,1}(u_j) - u'(x_j)| = |-\frac{u''(x_j)h}{2} + \frac{u'''(x^*)h^2}{6}|$$

$$\implies ||D_h^{b,1}(u_j) - u'(x_j)||_{h,\infty} \leq Ah \quad A \in \mathbb{N}$$

3.2.3 Central method error

Using the Taylor expansion of $u(x_{j+1})$ and $u(x_{j-1})$:

$$D_h^{c,1}(j) = \frac{u(x_{j+1}) - u(u_{x-1})}{2h}$$

$$= \frac{1}{2} \cdot [u'(x_j) + \frac{u''(x_j)h}{2} + \frac{u'''(x^*)h^2}{6} + u'(x_j) - \frac{u''(x_j)h}{2} + \frac{u'''(x^*)h^2}{6}]$$

$$= u'(x_j) + \frac{h^2}{12} \cdot [u'''(x^f) + u'''(x^b)]$$

$$\implies |D_h^{c,1}(u_j) - u'(x_j)| = |\frac{h^2}{12} \cdot [u'''(x^f) + u'''(x^b)]|$$

$$\implies ||D_h^{c,1}(u_j) - u'(x_j)||_{h,\infty} \leq Ah^2 \quad A \in \mathbb{N}$$

3.2.4 Analysis

Since the rate of convergence for the central approximation method of the derivative is one order lower than the convergence rate of the forward and backward methods, it follows that $\lim_{h\to 0} e_c(u) \leq \lim_{h\to 0} e_f(u) = \lim_{h\to 0} e_b(u)$.

We can obtain a mathematical expression for the exponent rate of convergence α that we can test empirically for the the three different methods. We know that: $e_j \approx Ah_j^\alpha$ and $e_{j+1} \approx Ah_{j+1}^\alpha$ for $j = 1, 2, \ldots, N$; where $A \in \mathbb{N}$ and where e is the error calculated with the infinity norm. Thus:

$$\frac{e_j}{e_{j+1}} \approx (\frac{h_j}{j_{j+1}})^\alpha$$

$$\implies \alpha \approx \frac{\log(\frac{e_j}{e_{j+1}})}{\log(\frac{h_j}{h_{j+1}})}$$

4 Numerical experiments

4.1 tables

Table 1

Calculation of the greatest error e and exponent rate of convergenceα for the function x^2 using the forward, backward, and central approximation methods.

h	$e_f(u)$	$e_b(u)$	$e_c(u)$	α_f	α_b	α_c
0.25	0.25	0.25	0	0	0	0
0.125	0.125	0.235	0	1	1	NaN
0.0625	0.0625	0.0625	0	1	1	NaN
0.03125	0.03125	0.03125	0	1	1	NaN

Table 2

Calculation of the greatest error e and exponent rate of convergenceα for the function x using the forward, backward, and central approximation methods.

h	$e_f(u)$	$e_b(u)$	$e_c(u)$	α_f	α_b	α_c
0.25	0	0	0	0	0	0
0.125	0	0	0	NaN	NaN	NaN
0.0625	0	0	0	NaN	NaN	NaN
0.03125	0	0	0	NaN	NaN	NaN

TABLE 3

Calculation of the greatest error e and exponent rate of convergenceα for the function e^x using the forward, backward, and central approximation methods.

h	$e_f(u)$	$e_b(u)$	$e_c(u)$	α_f	α_b	α_c
0.25	0.28813	0.24389	0.022121	0	0	0
0.125	0.15638	0.14387	0.006252	0.88168	0.7614	1.823
0.0625	0.081488	0.078163	0.0016628	0.94036	0.88024	1.9107
0.03125	0.041599	0.040741	0.00042884	0.97006	0.94	1.9551

TABLE 4

Calculation of the greatest error e and exponent rate α for the function $sin(10x)$ using the forward, backward, and central approximation methods.

h	$e_f(u)$	$e_b(u)$	$e_c(u)$	α_f	α_b	α_c
0.25	9.3944	10.405	6.0936	0	0	0
0.125	5.9726	5.9355	2.4068	0.65346	0.80987	1.3402
0.0625	3.0716	3.0851	0.63836	0.95936	0.94405	1.9147
0.03125	1.5552	1.5484	0.16195	0.98188	0.99452	1.9789

5 Finate difference

5.1 The heat equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$$

5.1.1 The explicit method

To discretize this equation we can use the forward method on the derivative over time and the central method on the second derivative over x:

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} = \frac{u_{j-1}^n - 2u_j^n + u_{j+1}^n}{h^2} + error$$

where $u(x_j, t_n) = u_j^n$. Thus, our initial condition is $u(x, 0) = u_j^0$. Solving for u_j^{n+1}:

$$u_j^{n+1} = u_j^n + \frac{\Delta t}{h^2}(u_{j-1}^n - 2u_j^n + u_{j+1}^n)$$

Since we only know the initial state of the material. meaning $u(x,t)$ when $t = 0$; $u(x, 0) = u_j^0$ is known for any j, $0 \leq j \leq 1$. If we let $n = 0$ we have:

$$u_j^1 = u_j^0 + \frac{\Delta t}{h^2}(u_{j-1}^0 - 2u_j^0 + u_{j+1}^0)$$

Encoding u_j^t as matrices, we have:

$$\vec{u}^{n+1} = (I + \frac{\Delta t}{h^2} D_c^2)\vec{u}^n$$

where I is the identity matrix, D_c^2 corresponds to the central approximation for the second derivative, and $\vec{u}^n = u_j^n, u_{j+1}^n, \ldots, u_N^n$, meaning the values of all the vertices at a time n. For the same reason that u is known when $t = 0$, $\vec{u}^0$ is known.

This formula turns out to be recursive, because we depend upon $\vec{u}^n$ to know $\vec{u}^{n+1}$, where we know the intial value of $\vec{u}^0$. To obtain a closed formula we can expand the recursion:

$$\vec{u}^1 = (I + \frac{\Delta t}{h^2} D_c^2)\vec{u}^0$$

$$\vec{u}^2 = (I + \frac{\Delta t}{h^2} D_c^2)\vec{u}^1 = (I + \frac{\Delta t}{h^2} D_c^2)(I + \frac{\Delta t}{h^2} D_c^2)\vec{u}^0 = (I + \frac{\Delta t}{h^2} D_c^2)^2\vec{u}^0$$

$$\vec{u}^3 = (I + \frac{\Delta t}{h^2} D_c^2)^3\vec{u}^0$$

$$\vdots$$

$$\vec{u}^n = (I + \frac{\Delta t}{h^2} D_c^2)^n\vec{u}^0$$

This last equation corresponds to the explicit method and describes all $u(x, t)$ given $u(x, 0)$.

5.1.2 The implicit method

To discretize the heat equation we will again use the forward method on the derivative over time and the central method on the second derivative over x. However, the change in x will this time depend on $n + 1$ instead of n :

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} = \frac{u_{j-1}^{n+1} - 2u_j^{n+1} + u_{j+1}^{n+1}}{h^2} + error$$

As per in the explicit method, using matrices and the central method of the second derivative we now have:

$$\vec{u}^n = (I - \frac{\Delta t}{h^2} D_c^2)\vec{u}^{n+1}$$

Solving for u^{n+1} we get a similar recurive expression as which uses the inverse instead:

$$\vec{u}^{n+1} = (I - \frac{\Delta t}{h^2} D_c^2)^{-1}\vec{u}^n$$

Similrly, expanding the recursion we get the closed formula

$$\vec{u}^n = (I - \frac{\Delta t}{h^2} D_c^2)^{n^{-1}}\vec{u}^0$$

This equation corresponds to the implicit method and describes all $u(x, t)$ given $u(x, 0)$.

5.1.3 The Crank-Nicholson method

The Crank-Nicholson method employs the arithmetic mean between the explicit and the implicit method.

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} = \frac{1}{2h^2}(u_{j-1}^{n+1} - 2u_j^{n+1} + u_{j+1}^{n+1} + u_{j-1}^n - 2u_j^n + u_{j+1}^n)$$

Solving for $\vec{u}^{n+1}$ as done with the previous 2 method we get:

$$\vec{u}^n = (I - \frac{\Delta t}{h^2} D_c^2)^{n^{-1}} (I + \frac{\Delta t}{h^2} D_c^2)^n \vec{u}^0$$

5.1.4 The theta method

We can generalize these 3 defined method as a linear interpolation depending on $0 \leq \theta \leq 1$:

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} = (1-\theta)\frac{u_{j-1}^{n+1} - 2u_j^{n+1} + u_{j+1}^{n+1}}{h^2} + \theta \frac{u_{j-1}^n - 2u_j^n + u_{j+1}^n}{h^2} + error$$

which yields:

$$\vec{u}^n = (I - (1-\theta)\frac{\Delta t}{h^2} D_c^2)^{n^{-1}} (I + \theta \frac{\Delta t}{h^2} D_c^2)^n \vec{u}^0$$

If we let $\theta = 0$ we would be using the explicit method. If $\theta = 1$ it will be equal to toe implicit method. If $\theta = 1/2$, it is the Crank-Nicholson method.

5.1.5 Calculation of error

Using the general theta method we can calculate the error.

5.1.6 Error truncation

$$\tau^n = \frac{u_j^{n+1} - u_j^n}{\Delta t} - \frac{u_{j-1}^n - 2u_j^n + u_{j+1}^n}{h^2}$$

$$u_j^{n+1} = u(x_j, t_n + \Delta t) = u(x_j, t_n) + \Delta t \frac{\partial u}{\partial t}(x_j, t_n) + \frac{(\Delta t)^2}{2}\frac{\partial^2 u}{\partial t^2}(x_j, t_n^*)$$

$$= u(x_j, t_n) + \Delta t \frac{\partial u}{\partial t}(x_j, t_n) + c\Delta t^2 \quad \text{where} \quad c \leq ||\frac{\partial^2 u}{\partial t^2}||_\infty \leq \infty$$

Thus:

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} = \frac{\partial u}{\partial t}(x_j, t_n) + c\Delta t$$

$$u_{j\pm 1}^n = u(x_j \pm h, t_n) = u(x_j, t_n) \pm h\frac{\partial u}{\partial x}(x_j, t_n) + \frac{h^2}{2}\frac{\partial^2 u}{\partial x^2}(x_j, t_n) \pm \frac{h^3}{6}\frac{\partial^3 u}{\partial x^3}(x_j, t_n) + \frac{h^4}{24}\frac{\partial^4 u}{\partial x^4}(x_j^{*,\pm}, t_n)$$

$$= u(x_j, t_n) \pm h\frac{\partial u}{\partial x}(x_j, t_n) + \frac{h^2}{2}\frac{\partial^2 u}{\partial x^2}(x_j, t_n) \pm \frac{h^3}{6}\frac{\partial^3 u}{\partial x^3}(x_j, t_n) + c_0 h^4 \quad \text{where} \quad c_0 \leq ||\frac{\partial^4 u}{\partial x^4}||_\infty \leq \infty$$

Thus:

$$u_{j-1}^n + u_{j+1}^n = 2u_j^n + h^2 \frac{\partial^2 u}{\partial x^2}(x_j, t_n) + c_0 h^4$$

Thus:

$$\frac{u_{j-1+}^n - 2u_j^n + u_{j+1}^n}{h^2} = \frac{\partial^2 u}{\partial x^2}(x_j, t_n) + c_0 h^2$$

Similarly:

$$\frac{u_{j-1+}^{n+1} - 2u_j^{n+1} + u_{j+1}^{n+1}}{h^2} = \frac{\partial^2 u}{\partial x^2}(x_j, t_{n+1}) + c_1 h^2$$

from the heat equation we have that:

$$= \frac{\partial u}{\partial t}(x_j, t_{n+1}) + c_1 h^2$$

Expanding using Taylor:

$$= \frac{\partial u}{\partial t}(x_j, t_n) + \Delta t \frac{\partial^2 u}{\partial t^2}(x_j, t_n^*) + ch^2$$

$$= \frac{\partial u}{\partial t}(x_j, t_n) + c(\Delta t + h^2)$$

Knwoing that:

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} = \frac{\partial u}{\partial t}(x_j, t_n) + \frac{\Delta t}{2}\frac{\partial^2 u}{\partial t^2}(x_j, t_n) + c_0(\Delta t)^2$$

$$\frac{u_{j-1+}^n - 2u_j^n + u_{j+1}^n}{h^2} = \frac{\partial u}{\partial t}(x_j, t_n) + c_0 h^2$$

$$\frac{u_{j-1+}^{n+1} - 2u_j^{n+1} + u_{j+1}^{n+1}}{h^2} = \frac{\partial u}{\partial t}(x_j, t_n) + c_1(\Delta t + h^2)$$

and that:

$$\tau_j^n = \frac{u_j^{n+1} - u_j^n}{\Delta t} - (1-\theta)\frac{u_{j-1}^n - 2u_j^n + u_{j+1}^n}{h^2} - \theta \frac{u_{j-1}^{n+1} - 2u_j^{n+1} + u_{j+1}^{n+1}}{h^2}$$

$$\tau_j^n \leq |\frac{\partial u}{\partial t}(x_j, t_n) + \frac{\Delta t}{2}\frac{\partial^2 u}{\partial t^2}(x_j, t_n) + c_0(\Delta t)^2 - (1-\theta)[\frac{\partial u}{\partial t}(x_j, t_n) + c_0 h^2] - \theta[\frac{\partial u}{\partial t}(x_j, t_n) + c_1(\Delta t + h^2)]|$$

It follows that:

$$\tau_j^n = \Delta t(\frac{1}{2} - \theta)\frac{\partial^2 u}{\partial t^2}(x_j, t_n) + c(\Delta t^2 + h^2)$$

When $\theta = 1$, the spcefic case of the theta method is the explicit method. Since Δt dominates over Δt^2, the explicit method the error is:

$$\tau_j^n \leq c(\Delta t + h^2)$$

When $\theta = 0$, the spcefic case of the theta method is the implicit method. Since Δt dominates over Δt^2, the implicit method the error is:

$$\tau_j^n \leq c(\Delta t + h^2)$$

When $\theta = 1/2$, the spcefic case of the theta method is the Crank-Nicholson method. Since in such case the term Δt is cancelled, the Crank-Nicholson error is:

$$\tau^n \leq c(\Delta t^2 + h^2)$$

5.1.7 Method with error approximation

Given the heat equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$$

we approximate the left hand side of the equation using the forward method and we approximate the right hand side of the equation using the theta method.

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} = (1-\theta)\frac{u_{j-1}^{n+1} - 2u_j^{n+1} + u_{j+1}^{n+1}}{h^2} + \theta\frac{u_{j-1}^{n} - 2u_j^{n} + u_{j+1}^{n}}{h^2} + \tau_j^n$$

Given that $\vec{u}^0$ is known, solving for $\vec{u}^n$ and discretizing we get that:

$$\vec{u}^n = (I - (1-\theta)\frac{\Delta t}{h^2}D_c^2)^{n^{-1}}(I + \theta\frac{\Delta t}{h^2}D_c^2)^n\vec{u}^0$$

When solving for the error we get that:

$$\tau_j^n \leq \Delta t(\frac{1}{2} - \theta)\frac{\partial^2 u}{\partial t^2}(x_j, t_n) + c(\Delta t^2 + h^2)$$

There are 3 special cases of θ worth analyzing:

- $\theta = 1$. This is known as the explicit scheme. The recursive discrete form is give by $\vec{u}^{n+1} = (I + \frac{\Delta t}{h^2}D_c^2)\vec{u}^n$ and the closed forumala is given by $\vec{u}^n = (I + \frac{\Delta t}{h^2}D_c^2)^n\vec{u}^0$. The error is given by $\tau_j^n \leq c(\Delta t + h^2)$.

  The pros of this method is that is that the matrix $(I - \frac{\Delta t}{h^2}D_c^2)$ is easy to compute and guaranteed to exist. However, the convergance of this method depends on the ration between the length of the timesteps Δt and the spacing between each point h. The non convergance of the method is known as the CFL condition, which occurs when $\Delta t \leq ch^2$ where c is a constant obtained experimentally.

- $\theta = 0$. This is known as the implicit scheme. The recursive discrete form is given by $\vec{u}^{n+1} = (I - \frac{\Delta t}{h^2}D_c^2)^{-1}\vec{u}^n$ and the closed forumla is given by $\vec{u}^n = (I - \frac{\Delta t}{h^2}D_c^2)^{n^{-1}}\vec{u}^0$. The error is given by $\tau_j^n \leq c(\Delta t + h^2)$.

  The pros of this method is that it converges for all values of Δt unconditionally, meaning that it does not depend on the ratio $\Delta t, h$ to converge. However, the inverse of $(I - \frac{\Delta t}{h^2}D_c^2)$ can be expensive to compute and not guaranteed to exist.

- $\theta = 1/2$. This is known as the Crank-Nicholson scheme. The recursive discrete form is given by $\vec{u}^{n+1} = (I - \frac{\Delta t}{2h^2}D_c^2)^{-1}(I - \frac{\Delta t}{2h^2}D_c^2)\vec{u}^n$ and the closed forumla is given by $\vec{u}^n = (I - \frac{\Delta t}{2h^2}D_c^2)^{-n}(I - \frac{\Delta t}{2h^2}D_c^2)^n\vec{u}^0$. The error is given by $c(\Delta t^2 + h^2)$.

  The pros of this method is that it also converges for all vaues of Δt unconditionally. Besides, the error shows a better convergence in comparison to the other methods. The only problem is again the computation of the inverse of $(I - \frac{\Delta t}{h^2}D_c^2)$, which is expensive and not guaranteed to exist.

5.1.8 Tables

Table 5

Calculation of the greatest error e and exponent rate of convergence α for the function $u(x,t) = e^{-t\pi^2} \cdot \sin(\pi x)$ when $t = T = 1$ using the explicit, implicit and Crank-Nicholson schemes.

h	$e_f(u)$	$e_b(u)$	$e_c(u)$	α_f	α_b	α_c
0.125	∞	-	$1.6560 \cdot 10^{-3}$	-	$3.9551 \cdot 10^{-5}$	-
0.0625	∞	-	$4.1579 \cdot 10^{-4}$	1.9938	$1.3296 \cdot 10^{-5}$	1.5727
0.03125	∞	-	$1.3306 \cdot 10^{-4}$	1.6438	$3.5574 \cdot 10^{-6}$	1.9021
0.015652	∞	-	$5.1660 \cdot 10^{-5}$	1.3649	$9.042 \cdot 10^{-7}$	1.9761
0.019531	∞	-	$2.2597 \cdot 10^{-5}$	1.193	$2.2698 \cdot 10^{-7}$	1.9941

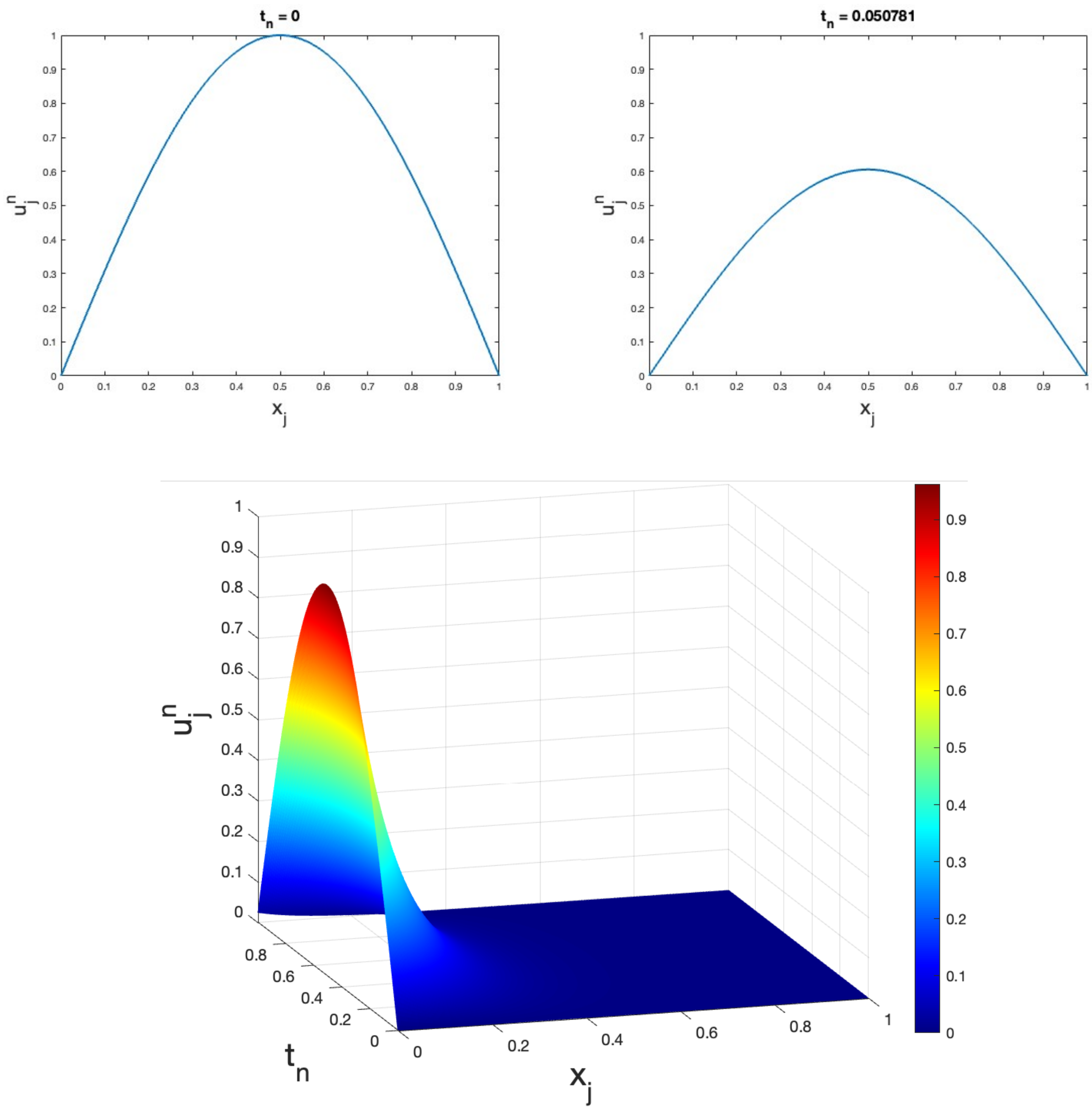

Figure 1: Function over time.