AMATH 586 Homework 4

Cade Ballew #2120804

May 20, 2022

1 Problem 1

Consider the following heat equation with "linked" boundary conditions

$$\begin{cases} u_t = \frac{1}{2}u_{xx} \\ u(0,t) = su(1,t) \\ u_x(0,t) = u_x(1,t), \\ u(x,0) = \eta(x), \end{cases}$$

where $s \neq -1$. The MOL discretization with the standard second-order stencil can be written as

$$U'(t) = -\frac{1}{2h^2}AU(t) + \begin{pmatrix} \frac{U_0(t)}{2h^2} \\ 0 \\ \vdots \\ 0 \\ \frac{U_{m+1}(t)}{2h^2} \end{pmatrix}, \quad A = \begin{pmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & -1 & 2 & -1 & \\ & & \ddots & \ddots & \ddots \\ & & & -1 & 2 \end{pmatrix}.$$

If we enforce the boundary conditions via $U_0(t) = sU_{m+1}(t)$ and suppose that

$$\frac{U_1(t) - U_0(t)}{h} = \frac{U_{m+1}(t) - U_m(t)}{h},$$

we find that

$$\frac{U_1(t) - sU_{m+1}(t)}{h} = \frac{U_{m+1}(t) - U_m(t)}{h},$$

so

$$(1+s)U_{m+1}(t) = U_1(t) + U_m(t),$$

meaning that

$$\begin{aligned} U_{m+1}(t) &= \frac{1}{1+s}U_1(t) + \frac{1}{1+s}U_m(t), \\ U_0(t) &= \frac{s}{1+s}U_1(t) + \frac{s}{1+s}U_m(t). \end{aligned}$$

Then, we have from the MOL discretization that

$$\begin{aligned} U_1'(t) &= -\frac{1}{2h^2}(2U_1(t) - U_2(t)) + \frac{1}{2h^2}\left(\frac{s}{1+s}U_1(t) + \frac{s}{1+s}U_m(t)\right) \\ &= \frac{1}{2h^2}\left(\left(-2 + \frac{s}{1+s}\right)U_1(t) + U_2(t) + \frac{s}{1+s}U_m(t)\right). \end{aligned}$$

For $j = 2, \dots, m$,

$$U_j'(t) = -\frac{1}{2h^2}(-U_{j-1}(t) + 2U_j(t) - U_{j+1}(t)) = \frac{1}{2h^2}(U_{j-1}(t) - 2U_j(t) + U_{j+1}(t)).$$

Finally,

$$\begin{aligned} U_m'(t) &= -\frac{1}{2h^2}(-U_{m-1}(t) + 2U_m(t)) + \frac{1}{2h^2}\left(\frac{1}{1+s}U_1(t) + \frac{1}{1+s}U_m(t)\right) \\ &= \frac{1}{2h^2}\left(\frac{1}{1+s}U_1(t) + U_{m-1}(t) + \left(-2 + \frac{1}{1+s}\right)U_m(t)\right). \end{aligned}$$

Putting this back into matrix form, we find that

$$U'(t) = \frac{1}{2h^2}BU(t)$$

where

$$B = \begin{pmatrix} -2 + \frac{s}{1+s} & 1 & & & & \frac{s}{1+s} \\ 1 & -2 & 1 & & & \\ & 1 & -2 & 1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & 1 & -2 & 1 \\ \frac{1}{1+s} & & & & 1 & -2 + \frac{1}{1+s} \end{pmatrix}.$$

2 Problem 2

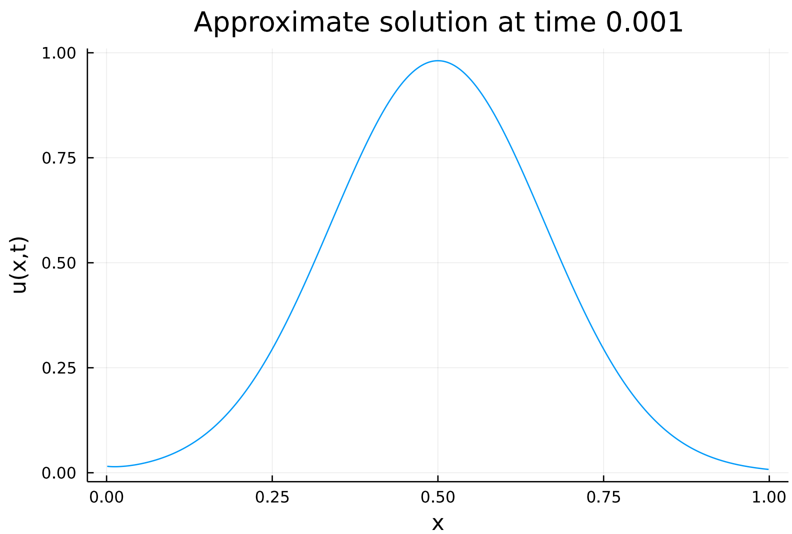
Applying backward Euler to the system from problem 1, we have the system

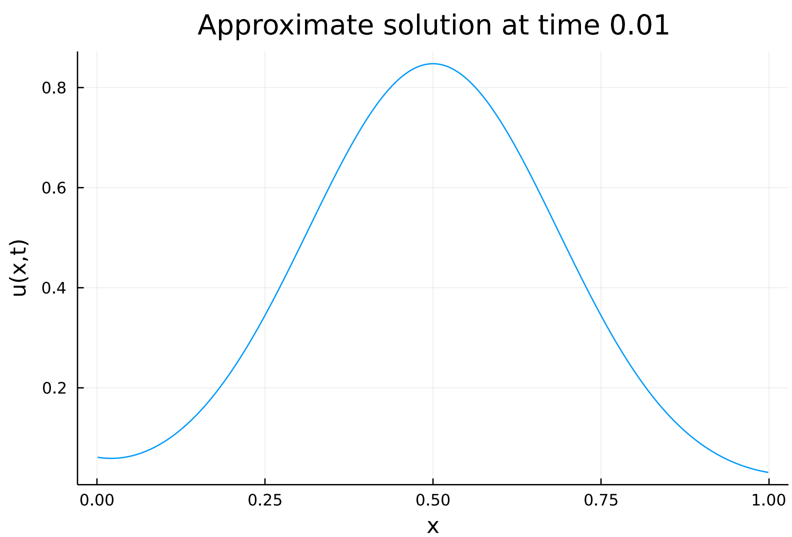
$$\left(I - \frac{k}{2h^2}B\right)U^{n+1} = U^n, \quad U^n = \begin{pmatrix} U_1^n \\ U_2^n \\ \vdots \\ U_m^n \end{pmatrix}. \tag{1}$$

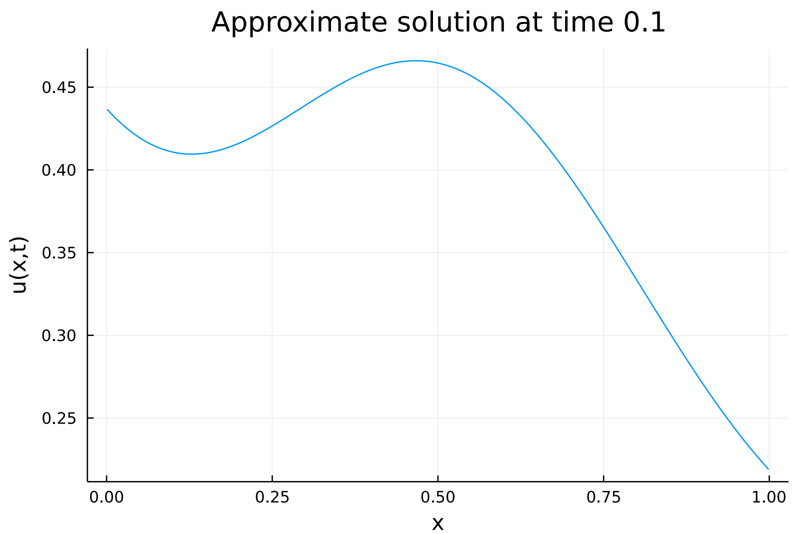
Using Julia, we solve this system with

$$\eta(x) = e^{-20(x-1/2)^2},$$

$k = h$ and $h = 0.001$ with $s = 2$. We plot the computed solution at times $t = 0.001, 0.01, 0.1$ as follows.


Approximate solution at time 0.001
1.00
0.75
0.50
0.25
0.00
u(x,t)
0.00
0.25
0.50
0.75
1.00
x


Approximate solution at time 0.01
0.8
0.6
0.4
0.2
u(x,t)
0.00
0.25
0.50
0.75
1.00
x



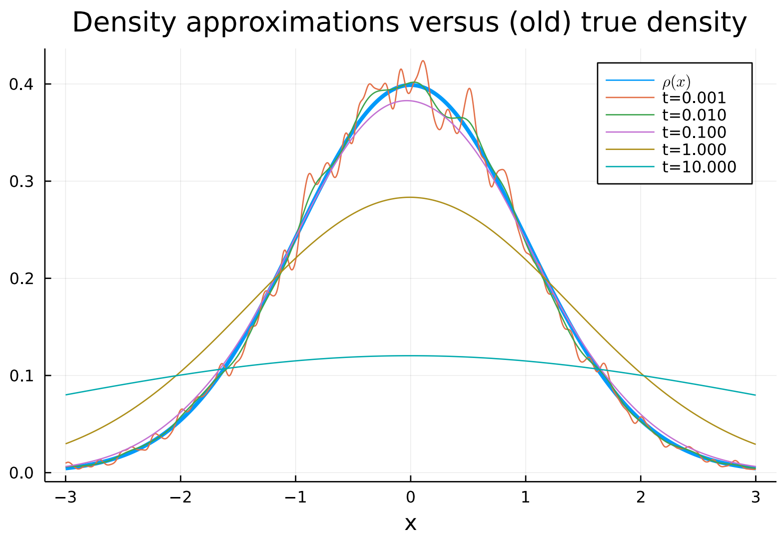
See Appendix A for the Julia code used to do this.

3 Problem 3

Consider the function

$$\frac{1}{N}\sum_{j=1}^{N}\frac{1}{\sqrt{2\pi t}}\exp\left(-\frac{(x-X_j)^2}{2t}\right), \quad t>0.$$

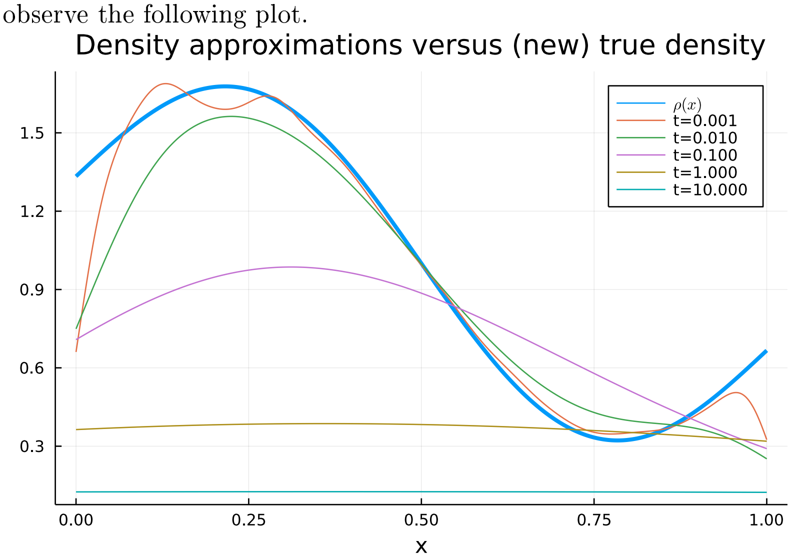
as an approximation to the density of data points $X_1, X_2, \dots, X_N, \dots$ each being a real number arising from a repeated experiment. Using Julia, we generate normally distributed data with $n = 10000$ and plot this function for $t = 0.001, 0.01, 0.1, 1, 10$ against the true probabilty density function for the data $\rho(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}$ in the following plot.



Visually, it seems that $t = 0.01$ gives the best approximation to the true density as it is barely visible over the true density.
Now, consider data points which instead correspond to the density

$$\rho(x) = -\frac{2}{3}x + \frac{4}{3} + \frac{1}{2}\sin(2\pi x).$$

Generating data with the provided Julia code, we repeat the experiment and observe the following plot.



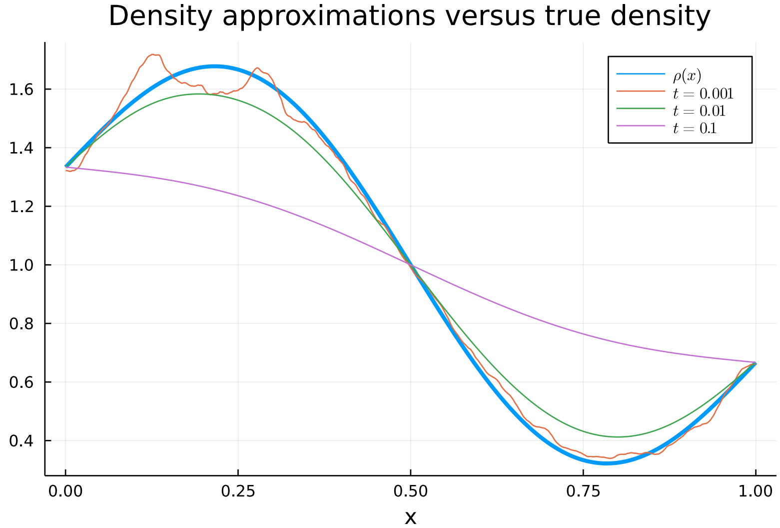
Now, $t = 0.001$ seems to produce the best approximation.
See Appendix A for the relevant Julia code.

4 Problem 4

Now, we again generate data $X_1, X_2, \ldots, X_N$ with density

$$\rho(x) = -\frac{2}{3}x + \frac{4}{3} + \frac{1}{2}\sin(2\pi x).$$

using the provided Julia code and find Y_i so that Y_i is the number of data points X_j that lie in the interval $[ih, (i+1)h) = [x_i, x_{i+1})$. We then set $U_i^0 = \frac{Y_i}{hN}$ to be the initial condition in the heat equation described in problem 1. We take $N = m$, $h = 0.0001$, $k = 10h$, $s = 2$ and use the Julia code from problem 2 to solve this system. The following plots the computed solution at $t = 0.001, 0.01, 0.1$ against the true density ρ.



The result looks quite similar to that of part 2 of problem 3 but does appear slightly better graphically specifically near $x = 1$.
See Appendix A for the relevant Julia code.

5 Problem 5

Let $A = I - \frac{k}{2h^2}B$ where B is defined in accordance with problem 1 and take $s > 0$. First, observe that

$$\left(\begin{pmatrix}1 & 1 & \cdots & 1\end{pmatrix} B\right)_1 = -2 + \frac{s}{1+s} + 1 + \frac{1}{1+s} = -1 + \frac{s+1}{1+s} = 0,$$

$$\left(\begin{pmatrix}1 & 1 & \cdots & 1\end{pmatrix} B\right)_m = \frac{s}{1+s} + 1 - 2 + \frac{1}{1+s} = \frac{s+1}{1+s} - 1 = 0,$$

and

$$\left(\begin{pmatrix}1 & 1 & \cdots & 1\end{pmatrix} B\right)_j = 1 - 2 + 1 = 0$$

for $j = 2, \ldots, m-1$. Thus,

$$\begin{pmatrix} 1 & 1 & \cdots & 1 \end{pmatrix} B = 0.$$

This implies that

$$\begin{pmatrix} 1 & 1 & \cdots & 1 \end{pmatrix} A = \begin{pmatrix} 1 & 1 & \cdots & 1 \end{pmatrix} I - \frac{k}{2h^2} \begin{pmatrix} 1 & 1 & \cdots & 1 \end{pmatrix} B = \begin{pmatrix} 1 & 1 & \cdots & 1 \end{pmatrix}.$$

Using the fact that $AU^{n+1} = U^n$ for $n = 0, 1, \ldots$ inductively, we have that $A^n U^n = U^0$. We now have that

$$\begin{aligned} \begin{pmatrix} 1 & 1 & \cdots & 1 \end{pmatrix} A^n &= \left(\begin{pmatrix} 1 & 1 & \cdots & 1 \end{pmatrix} A \right) A^{n-1} \\ &= \begin{pmatrix} 1 & 1 & \cdots & 1 \end{pmatrix} A^{n-1} = \ldots = \begin{pmatrix} 1 & 1 & \cdots & 1 \end{pmatrix}, \end{aligned}$$

so we can left multiply both sides of $A^n U^n = U^0$ by $\begin{pmatrix} 1 & 1 & \cdots & 1 \end{pmatrix}$ to find that

$$\begin{aligned} \text{LHS} &= \begin{pmatrix} 1 & 1 & \cdots & 1 \end{pmatrix} A^n U^n = \begin{pmatrix} 1 & 1 & \cdots & 1 \end{pmatrix} U^n = \sum_j U_j^n, \\ \text{RHS} &= \begin{pmatrix} 1 & 1 & \cdots & 1 \end{pmatrix} U^0 = \sum_j U_j^0, \end{aligned}$$

so we can conclude that

$$\sum_j U_j^n = \sum_j U_j^0$$

for all $n = 0, 1, \ldots$.

Now, we assume that for $s > 0$, if y is a vector with non-negative entries and $\left(I - \frac{k}{2h^2} B\right) x = y$ then x has non-negative entries. This also means that if y is a vector with non-positive entries and $\left(I - \frac{k}{2h^2} B\right) x = y$, then x has non-positive entries as then $-x$ would have nonnegative entries, so $-y = \left(I - \frac{k}{2h^2} B\right)(-x)$ would have non-negative entries. This of course means that the same property holds for $A^k x = y$. Now, note that any vector u can be decomposed as $u = p + n$ where

$$p_j = \begin{cases} u_j, & u_j > 0 \\ 0, & u_j \leq 0, \end{cases}$$

$$n_j = \begin{cases} 0, & u_j > 0 \\ u_j, & u_j \leq 0, \end{cases}$$

for all j. Then, by definition,

$$\|A^k\| = \max_{\|u\|_1 = 1} \|A^k u\|_1.$$

If we consider an arbitrary $u \in \mathbb{R}^m$ and decompose it as $u = p + n$, then by the triangle inequality and our assumed property of A,

$$\begin{aligned}\|A^k u\|_1 = \|A^k(p+n)\|_1 \le \|A^k p\|_1 + \|A^k n\|_1 = h\sum_{j=1}^{M}|(A^k p)_j| + h\sum_{j=1}^{M}|(A^k n)_j| \\ = h\sum_{j=1}^{M}(A^k p)_j - h\sum_{j=1}^{M}(A^k n)_j = h\sum_{j=1}^{M}(A^k(p-n))_j\end{aligned}$$

Now, we use the property that $\sum_j U_j^n = \sum_j U_j^0$ by considering[1] $U^0 = A^k(p-n)$, $U^k = p - n$ which gives us that

$$\begin{aligned}\|A^k u\|_1 \le h\sum_{j=1}^{M}(p-n)_j = h\sum_{j=1}^{M}p_j + h\sum_{j=1}^{M}(-n)_j = h\sum_{j=1}^{M}|p_j| + h\sum_{j=1}^{M}|n_j| \\ = h\sum_{j=1}^{M}|(p+n)_j| = \sum_{j=1}^{M}|u_j| = \|u\|_1 = 1\end{aligned}$$

Because we can do this for any choice of u, we actually have that

$$\|A^k\| = \max_{\|u\|_1=1}\|A^k u\|_1 \le \max_{\|u\|_1=1}\|u\|_1 = 1.$$

Because we are able to bound our iteration matrix in this way, we can take $C_T = 1$ to conclude that the forward Euler iteration is Lax-Richtmyer stable in the 1-norm.

6 Problem 6

Consider the bi-infinite matrix

$$L = \begin{pmatrix} \ddots & \vdots & \vdots & \vdots & \iddots \\ \cdots & L_{-1,-1} & L_{-1,0} & L_{-1,1} & \cdots \\ \cdots & L_{0,-1} & L_{0,0} & L_{0,1} & \cdots \\ \cdots & L_{1,-1} & L_{1,0} & L_{1,1} & \cdots \\ \iddots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

and suppose the matrix L defines a bounded linear operator on $\ell^2(\mathbb{Z})$ via matrix-vector multiplication with

$$\ell^2(\mathbb{Z}) \ni V = \begin{pmatrix} \vdots \\ V_{-1} \\ V_0 \\ V_1 \\ \vdots \end{pmatrix}.$$

[1] Note that we use U^k instead of of U^n to avoid reusing variables.

Let e_k denote the kth unit vector and let S_ℓ denote the ℓth shift operator. Then, we let $j \in \mathbb{Z}$ and by basic matrix multiplication we see that

$$(Le_k)_j = L_{j,k}.$$

We can also see from the definition of the shift operator that

$$S_\ell e_k = e_{k+\ell},$$

so

$$(LS_\ell e_k)_j = (Le_{k+\ell})_j = L_{j,k+\ell},$$

and

$$(S_{-\ell}LS_\ell e_k)_j = (LS_\ell e_k)_{j+\ell} = L_{j+\ell,k+\ell}.$$

Because the unit vectors form a basis for $\ell^2(\mathbb{Z})$, we can decompose any $V \in \ell^2(\mathbb{Z})$ into a sum of unit vectors. $LV = S_{-\ell}LS_\ell V$ for any $V \in \ell^2(\mathbb{Z})$ if and only if $Le_k = S_{-\ell}LS_\ell e_k$ for every $k \in \mathbb{Z}$. Thus, by definition, L is shift-invariant if and only if

$$L_{j,k} = L_{j+\ell,k+\ell}$$

for every $j, k, \ell \in \mathbb{Z}$. However, this is precisely the definition of of a Toeplitz matrix as L is constant along diagonals, so $L_{i,j} = c_{i-j}$ for a sequence $(c_j)_{j=-\infty}^{\infty}$. Thus, we can conclude that L is shift-invariant iff $L_{i,j} = c_{i-j}$ for a sequence $(c_j)_{j=-\infty}^{\infty}$.

7 Problem 7

In the notation of Problem 2, for $s > 0$, we wish to establish that if y is a vector with non-negative entries and $\left(I - \frac{k}{2h^2}B\right)x = y$ then x has non-negative entries. We first establish notation by letting $A = I - \frac{k}{2h^2}B$ and $\alpha = \frac{k}{2h^2}$. We do this by way of Farkas' lemma; namely, we wish to show that there does not exist some $z \in \mathbb{R}^m$ such that $A^T z \geq 0$ and $z^T y < 0$ where $y \geq 0$ is given and The "$\geq$" symbol means that all components of the vector are nonnegative. If we can show this, then Farkas' lemma implies that there must exist $x \in \mathbb{R}^m$ such that $Ax = y$ and $x \geq 0$ for the same given y which is precisely what we wish to show in this problem.
To show our new statement, we note that because $y \geq 0$, at least one component of z must be strictly negative in order for $z^T y < 0$ to hold. Thus, it suffices to show that $A^T z$ must be strictly negative in at least one component when z is strictly negative in at least one component. We do this by considering i to be the index at which z_i is minimized; i.e., $z_i \leq z_j$ for all $j = 1, \ldots, m$. Of course, we must have that $z_i < 0$. We now show that $(A^T z)_i < 0$.

First, note that under our notation, we can write

$$A^T = \begin{pmatrix} 1+\alpha\left(2-\frac{s}{1+s}\right) & -\alpha & & & & -\alpha\frac{1}{1+s} \\ -\alpha & 1+2\alpha & -\alpha & & & \\ & -\alpha & 1+2\alpha & -\alpha & & \\ & & \ddots & \ddots & \ddots & \\ & & & -\alpha & 1+2\alpha & -\alpha \\ -\alpha\frac{s}{1+s} & & & & -\alpha & 1+\alpha\left(2-\frac{1}{1+s}\right) \end{pmatrix}.$$

First consider the case where $i = 1$. Then,

$$\begin{aligned} (A^T z)_1 &= z_1 + \alpha\left(2 - \frac{s}{1+s}\right) z_1 - \alpha z_2 - \alpha\frac{1}{1+s} z_m \\ &\le z_1 + \alpha\left(2 - \frac{s}{1+s}\right) z_1 - \alpha z_1 - \alpha\frac{1}{1+s} z_1 \\ &= z_1 + \alpha\left(1 - \frac{s+1}{1+s}\right) z_1 = z_1 < 0. \end{aligned}$$

In the case where $i = m$,

$$\begin{aligned} (A^T z)_m &= -\alpha\frac{s}{1+s} z_1 - \alpha z_{m-1} + z_m + \alpha\left(2 - \frac{1}{1+s}\right) z_m \\ &\le -\alpha\frac{s}{1+s} z_m - \alpha z_m + z_m + \alpha\left(2 - \frac{1}{1+s}\right) z_m \\ &= z_m + \alpha\left(1 - \frac{s+1}{1+s}\right) = z_m < 0. \end{aligned}$$

Finally, if $i = 2, \ldots, m-1$,

$$(A^T z)_i = -\alpha z_{i-1} + (1+2\alpha) z_i - \alpha z_{i+1} \le -\alpha z_i + (1+2\alpha) z_i - \alpha z_i = z_i < 0.$$

Thus, no matter which $i = 1, \ldots m$ minimizes z_i, $(A^T z)_i < 0$. QED.

8 Appendix A

The following Julia code is used for Problem 2.

```
using LinearAlgebra, Plots, Printf, SparseArrays, LaTeXStrings

η = x -> exp(-20*(x-1/2)^2)

h, k = 0.001, 0.001
s = 2.
T = 0.1 #final time

x = 0:h:1
x = x[2:end-1] #remove BC
m = length(x)
```

```
B = spdiagm(-1 => ones(m-1), 0 => -2*ones(m), 1 => ones(m-1))
B[1,1] += s/(1+s)
B[1,end] += s/(1+s)
B[end,1] += 1/(1+s)
B[end,end] += 1/(1+s)

A = I - (k/(2h^2))*B

u₀ = η.(x)
n = convert(Int64,ceil(T/k))
U = zeros(m,n+1)
U[:,1] = u₀
t = zeros(n+1)
t[1] = 0
for i = 2:n+1
    t[i] = t[i-1] + k
    U[:,i] = A\U[:,i-1]
end

ind₁ = t.≈ 0.001
p1 = plot(x,U[:,ind₁], label=false)
xlabel!("x")
ylabel!("u(x,t)")
title!("Approximate solution at time 0.001")
savefig(p1, "p2_1.pdf")
display(p1)

ind₂ = t.≈ 0.01
p2 = plot(x,U[:,ind₂], label=false)
xlabel!("x")
ylabel!("u(x,t)")
title!("Approximate solution at time 0.01")
savefig(p2, "p2_2.pdf")
display(p2)

ind₃ = t.≈ 0.1
p3 = plot(x,U[:,ind₃], label=false)
xlabel!("x")
ylabel!("u(x,t)")
title!("Approximate solution at time 0.1")
savefig(p3, "p2_3.pdf")
display(p3)
```

The following Julia code is used for Problem 3.

```
using Random, Plots, LaTeXStrings, Printf

Random.seed!(123)

n = 10000
X = randn(n)
x = -3:0.001:3

ρ = x -> exp(-x^2/2)/√(2π)

function density_approx(x, t, X)
    N = length(X)
    eval = zeros(length(x))
    for j = 1:N
       eval += @. exp(-(x-X[j])^2/(2t))/√(2π*t)
    end
    eval ./= N
end

p1 = plot(x, ρ.(x), label=L"\rho(x)", linewidth=3)
```

```
xlabel!("x")
title!("Density approximations versus (old) true density")

for t = [0.001 0.01 0.1 1. 10.]
    plot!(x, density_approx(x, t, X),label=@sprintf("t=%2.3f",t))
end
display(p1)
savefig(p1,"density_approx.pdf")

# repeating with new density
function prand(m)
    p = x -> -(2.0/3)*x .+4.0/3 .+ .5sin.(2*pi*x)
    B = 1.7
    out = fill(0., m)
    for j = 1:m
        u = 10.
        y = 0.
        while u >= p(y)/ B
            y = rand()
            u = rand()
        end
        out[j] = y
    end
    out
end


x = 0:0.001:1 #new domain where density should be nonzero
ρ = x -> -2x/3+4/3+sin(2π*x)/2
X = prand(n)
p2 = plot(x, ρ.(x), label=L"\rho(x)", linewidth=3)
xlabel!("x")
title!("Density approximations versus (new) true density")

for t = [0.001 0.01 0.1 1. 10.]
    plot!(x, density_approx(x, t, X),label=@sprintf("t=%2.3f",t))
end
display(p2)
savefig(p2,"density_approx_new.pdf")
```

The following Julia code is used for Problem 4.

```
using LinearAlgebra, Plots, SparseArrays, LaTeXStrings, Random, UncertainData

Random.seed!(123)

function prand(m)
    p = x -> -(2.0/3)*x .+4.0/3 .+ .5sin.(2*pi*x)
    B = 1.7
    out = fill(0., m)
    for j = 1:m
        u = 10.
        y = 0.
        while u >= p(y)/ B
            y = rand()
            u = rand()
        end
        out[j] = y
    end
    out
end

h, s = 0.0001, 2.
k = 10*h
T = 0.1 #final time
```

```
x = 0:h:1
m = length(x)-2
N = m

#build Yᵢ
X = prand(N)
bins = bin(x[2:end], X, ones(length(X)))
Y = sum.(bins)

x = x[2:end-1] #remove BC

#build initial condition
u₀ = Y./(h*N)

#code from problem 2
B = spdiagm(-1 => ones(m-1), 0 => -2*ones(m), 1 => ones(m-1))
B[1,1] += s/(1+s)
B[1,end] += s/(1+s)
B[end,1] += 1/(1+s)
B[end,end] += 1/(1+s)

A = I - (k/(2h^2))*B

n = convert(Int64,ceil(T/k))
U = zeros(m,n+1)
U[:,1] = u₀
t = zeros(n+1)
t[1] = 0
for i = 2:n+1
    t[i] = t[i-1] + k
    U[:,i] = A\U[:,i-1]
end

ρ = x -> -2x/3+4/3+sin(2π*x)/2
p1 = plot(x, ρ.(x), label=L"\rho(x)", linewidth=3)
xlabel!("x")
title!("Density approximations versus true density")

ind₁ = t.≈ 0.001
plot!(x,U[:,ind₁], label=L"t=0.001")

ind₂ = t.≈ 0.01
plot!(x,U[:,ind₂], label=L"t=0.01")

ind₃ = t.≈ 0.1
plot!(x,U[:,ind₃], label=L"t=0.1")
savefig(p1, "p4.pdf")
display(p1)
```