Georgia Journal of Science

Volume 80 *No. 2 Scholarly Contributions from the Membership and Others*

Article 5

2022

THE BEST LINEAR APPROXIMATION TO y= √x ON THE INTERVAL [0, b] USING THE MINIMAX ERROR

Hyounkyun Oh
Savannah State University, hoh@savannahstate.edu

THE BEST LINEAR APPROXIMATION TO y= √x ON THE INTERVAL [0, b] USING THE MINIMAX ERROR

Acknowledgements

This research project is proceeded with a sponsor of "NSF-TIP grant (171944- PI: Atean Agegnehu): Targeted Infusion Project: Developing a Minor in Applied Mathematics at Savannah State University"

THE BEST LINEAR APPROXIMATION TO $y = \sqrt{x}$ ON THE INTERVAL [o, b] USING THE MINIMAX ERROR

Hyounkyun Oh*, Josiah Riddick, Sujin Kim
Department of Mathematics
Savannah State University
Savannah, GA 31404
*corresponding author: hoh@savannahstate.edu

Abstract

This study discusses how to find the best linear approximation $p(x) = mx + n$ to a fundamental function $y = \sqrt{x}$ on the interval $[0, b]$, especially using the minimax error in Numerical Analysis. For this aim we employ two mathematical techniques: a) using the MATLAB code, positioning m and n values of the smallest maximum error on a broad range of m, and n value matrix in a rough scale and then repeatedly refining the regions in the smaller scales and b) finding three-point fitting line to a set of non-colinear three points. *We see that both results are successfully obtained and identical each other neglecting the error tolerance.*

Keywords: Best approximation, Minimax error, Three-point fitting line

I. Introduction

Approximation theory has always brought numerous applications to numerical issues appearing in various scientific and engineering disciplines. In mathematics, approximation theory is used to replace complicated functions with simpler functions, especially polynomial functions, or to attain the best fitting function to given data in a certain class. In this replacement, while the Taylor polynomial approximation emphasizes the local properties of the original function around a specific point, the Best Approximation or the Least Square method look for an alternative function which has the smallest error in a specified error type over a certain domain.

In this research we discuss how to find the best linear approximation $p(x) = mx + n$ to a fundamental function $y = \sqrt{x}$ on the interval $[0, \mathrm{b}]$ for some positive number $\mathrm{b} \in \mathbb{R}$, especially using the minimax error.

Mathematical Background: In mathematics, the best approximation algorithm with minimax error or minimax problem is a method to find a polynomial of degree n that minimizes the maximum error. The maximum error employs the maximum norm (or supremum norm or infinity norm) that assigns to a function $f: [a, b] \to \mathbb{R}$ on the interval $[a, b]$ the nonnegative number

$$\|f\|_\infty = \max_{a \le x \le b} |f(x)| = sup\{|f(x)|: x \in [a, b]\}.$$

Here we know that the supremum is actually the maximum for continuous functions because a continuous function on a closed bounded interval $[a, b]$ always attains its maximum on it. Then the minimax error for a polynomial approximation function $p(x)$ of degree n to a function $f(x)$ is defined as

$$E_n(f, p) = \min_{\deg(p) \le n} \|f - p\|_\infty = \min_{\deg(p) \le n} \left\{ \max_{a \le x \le b} |f(x) - p(x)| \right\}.$$

See the details in (Phillips 2003) and, (Burden *at al.* 2015). For example, when a function $f(x) = \sqrt{x}$ is defined on the interval $[0, 5]$, the approximation algorithm of degree 1 finds a linear polynomial $p(x) = mx + n$ of degree at most 1 in the absolute sense that minimizes the value

$$\max_{0 \le x \le 5} \left| \sqrt{x} - (mx + n) \right|.$$

Eventually, the problem is to find the m and n where a two-variable real valued absolute difference function g: $\mathbb{R}^2 \to \mathbb{R}$ defined by

$$g(m, n) = \max_{0 \le x \le 5} \{ |\sqrt{x} - (mx + n)| \}$$

attains its minimum. Since this minimax best approximation involves the absolute values, $g(m, n)$ may fail in differentiation at certain points and thus it is often compared to the Least Square Method using the L_2 norm

$$\|f\|_2 = \left(\int_a^b \left(f(x)\right)^2 dx \right)^{1/2}$$

in computation convenience.

II. The Best Linear Approximation to $f(x) = \sqrt{x}$ on $[0, b]$

Finding the best linear approximation $p(x) = mx + n$ to a basic function $f(x) = \sqrt{x}$ on the general interval $[0, b]$ using the minimax error implies to find the straight line which is closest to the curve $f(x) = \sqrt{x}$ in the sense of maximum error. For example, letting $b = 5$ and $m = 0.4$ and $n = 0$, i.e. $p_1(x) = 0.4x$, the graphs of two functions and their absolute difference function $g_1(x) = |\sqrt{x} - 0.4x|$ are shown Figure 1.

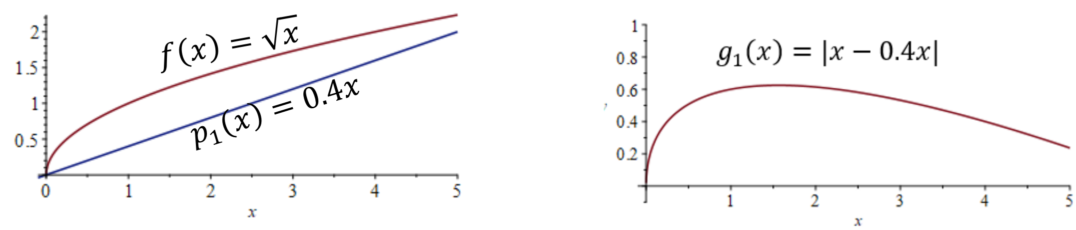

Figure 1. Graphs of $\boldsymbol{f(x) = \sqrt{x}}$ and $\boldsymbol{p_1(x) = 0.4x}$ (left) and their absolute difference function $\boldsymbol{g_1(x) = |\sqrt{x} - 0.4x|}$ (right)

Then, as shown in the right of Figure 1, the maximum possible error occurs at the critical point of $g_1(x)$, that is,

$$g_1'(x) = \frac{1}{2\sqrt{x}} - 0.4 = 0$$

which says $x = 1.5625$ and $g_1(1.5625) = 0.625$. That is, the maximum error of the approximation $p_1(x) = 0.4x + 0$ to the function $\boldsymbol{f(x) = \sqrt{x}}$ over the interval [0,5] is

$$\left\| \sqrt{x} - (0.4x + 0) \right\|_\infty = \|g_1\|_\infty = 0.625.$$

Again, if $m = 0.3577$ and $n = 0.39$, and if defining $p_2(x) = 0.3577x + 0.39$, shown as in Figure 2, the maximum error at the critical point is $g_2(1.9539) = 0.3089$ at $x = 1.9539$ from

$$g_2'(x) = \frac{1}{2\sqrt{x}} - 0.3577 = 0.$$

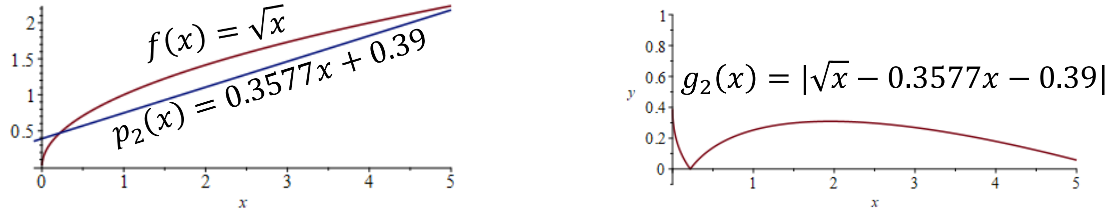

Figure 2. Graphs of $\boldsymbol{f(x)=\sqrt{x}}$ and $\boldsymbol{p_2(x)=0.3577x+0.39}$ (left) and their absolute difference function $\boldsymbol{g_2(x)=|\sqrt{x}-0.3577x-0.39|}$ (right)

On the other hand, at each end point the function value $g_2(0)=0.39$ and $g_2(5)=0.0576$, respectively and the maximum possible error then occurs at $x=0$ with $\|g_2\|_\infty=0.39$. Thus, $p_2(x)=0.3577x+0.39$ is considered as a better approximation than $p_1(x)=0.4x$ in minimax error sense and the minimax error for the linear functions must be $\mathrm{E}_1(f,p)\le 0.39$.

In general, the approximation functions position as one of pictures in Figure 3. That is, they meet at either no point or one point or two points. The graphs of corresponding

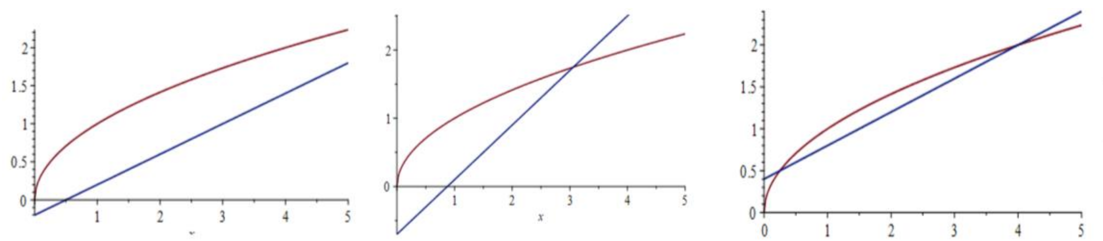

Figure 3. Possible three cases of $\boldsymbol{f(x)=\sqrt{x}}$ and its linear approximation

absolute difference functions are also shown in Figure 4. Thus, the maximum possible error $\|g\|_\infty$ occurs at the critical point of the regular difference function $\sqrt{x}-p(x)$ or the end point of domain. The points where the differentiation fails does not have any effect on maximum error and thus it is reasonable to neglect the non-differentiable points.

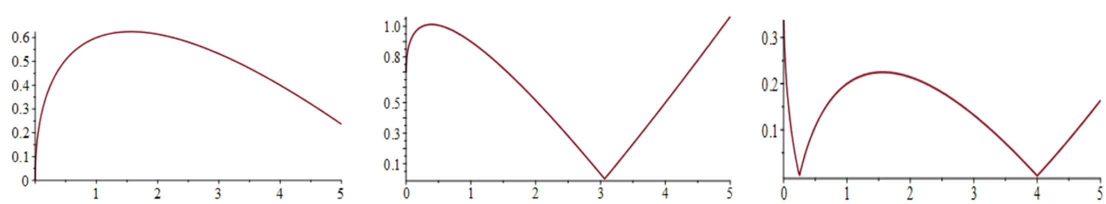

Figure 4. Three possible cases of patterns of absolute difference function g(x)

Now letting the difference function $g(x)=\sqrt{x}-mx-n$, we have the critical point

$$g'(x)=\frac{1}{2\sqrt{x}}-m=0 \quad \Leftrightarrow \quad x=\frac{1}{4m^2}. \tag{1}$$

So, it is enough to compare $g(0), g(b)$ and $g\left(\frac{1}{4m^2}\right)$ to get the maximum error. That is, the maximum error for the approximation $p(x) = mx + n$ is one of followings:

$$|g(0)| = |n| \tag{2a}$$

$$|g(b)| = \left|\sqrt{b} - mb - n\right| \tag{2b}$$

$$\left|g\left(\frac{1}{4m^2}\right)\right| = \left|\frac{1}{2m} - \frac{1}{4m} - n\right| = |\frac{1}{4m} - n| \tag{2c}$$

Approach 1: Utilization of MATLAB loops

To find the better approximation than $p_2(x) = 0.3577x + 0.39$ on [0, 5], that is, to find a linear function with the maximum error less than 0.39, we try to test the maximum error for $p(x) = mx + n$ on a rectangular range of $0.1..21.1$ for m and $-10..10$ for n by 0.1-unit scales. Since the original function $f(x) = \sqrt{x}$ is an increasing function, with reasonable intuition we skip to test the negative values in slope m. Table 1 shows a MATLAB code for this purpose.

Table 1: MATLAB code for the maximum error on a rectangular range of m and n

```
b = 5;

A = zeros(201,201);
i = 0;

for m = 0.1:0.1:21.1
        i = i+1;
        j = 0;
        for n = -10:0.1:10
            j = j+1;
            B = [abs(n), abs(sqrt(b)-m*b-n), abs(1/(4*m)-n)];
            A(i,j) = max(B);    %max error matrix
        end
end

minimum = min(min(A));
[x,y] = find(A==minimum);
m = 0.1+(x-1)*0.1
n = -10+(y-1)*0.1
```

The output of the code then shows that the smallest maximum error occurs at $m_3 = 0.4$ and $n_3 = 0.3$ with $\|g_3\|_\infty = 0.325$. Again, we examine the maximum error at a zoomed-in rectangular range for $m = 0.3..0.5$ and $n = 0.2..0.4$ by 0.001-unit scales. This trial gives the smallest maximum error occurs at $m_4 = 0.447$ and $n_4 = 0.280$ with $\|g_4\|_\infty = 0.280$. Following output with 0.00001 scale says that $p_5(x) = 0.44721x + 0.27951$ with

$\|g_5\|_\infty = 0.2795114888$ and this procedure is repeatedly done until the error tolerance $|\|g_n\|_\infty - \|g_{n+1}\|_\infty| < \epsilon$, that is, the difference between max errors gets small enough.

In the similar way for $b = 10$, we can still use the same code in Table 1 changing b value at the top of the code. Then on the rectangular range of $m = 0.1..21.1$ and $n = -10..10$ by 0.1-unit distance we have $m_1 = 0.3$ and $n_1 = -10 + 104 \cdot 0.1 = 0.4$ with $\|g_1\|_\infty = 0.43333$. In the first zoom-in rectangular range of $m = 0.2..0.4$ and $n = 0.3..0.5$ with 0.001-unit distance it gives $m_2 = 0.316$ and $n_2 = 0.396$ with $\|g_2\|_\infty = 0.396$. In the second zoom-in rectangular range of $m = 0.315..0.0.317$ and $n = 0.395..0.397$ with 0.00001-unit distance we have the better approximation $p_3(x) = 0.31623x + 0.39527$ with the maximum error $\|g_3\|_\infty = 0.39529383$. For other b values, the same procedures can be applied, repeatedly.

Approach 2: Fitting line to a set of non-colinear three points

Again, let's start with $f(x) = \sqrt{x}$ over the interval $[0, 5]$. Finding the linear approximation $p(x) = mx + n$ is the same to put a straight pen on the graph in which the pen is the least away from the curve, if possible. Since the biggest deviations occur at $x = 0$ or $x = 5$ or at a critical point in $(c, \sqrt{c})$ for some c in (0,5), the problem is eventually how to put the pen closely from three non-collinear points $(0,0)$, $(5, \sqrt{5})$ and $(c, \sqrt{c})$. See the left picture in Figure 5. In these graphs the length of two side arrows represent the magnitude of maximum error. Then we see that the smallest deviation occurs when the errors are exactly the same at all these three points. That is, the approximation must be parallel to the line connecting the two end points $(0,0)$ and $(5, \sqrt{5})$ which says the slope of $\frac{\sqrt{5}}{5}$. It looks more obviously when the left figure is rotated clockwise at the angle $\tan^{-1}(\sqrt{5})$ as shown in the right figure of Figure 5.

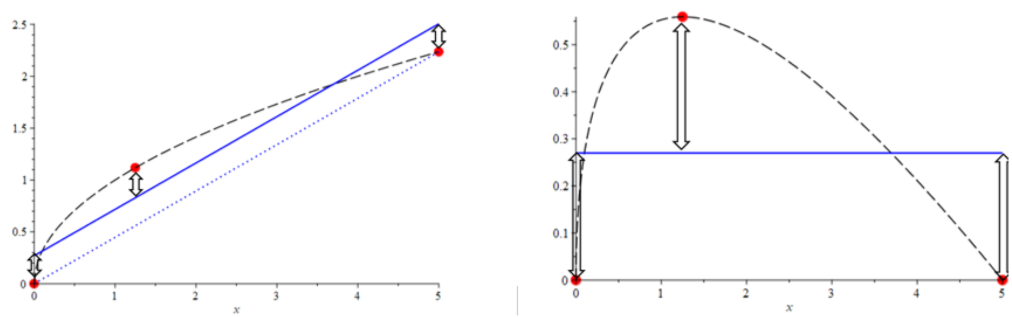

Figure 5. Best fitting line to non-colinear three points

From the Eq. (1), the critical point occurs at $c = \frac{1}{4m^2} = \frac{5}{4}$ for the function $g(x) = \sqrt{x} - mx - n$ and so three points are $(0,0)$, $(5, \sqrt{5})$ and $(\frac{5}{4}, \frac{\sqrt{5}}{2})$ on the curve. Since the errors are equivalent at these three points $x = 0$, $x = \frac{5}{4}$ and $x = 5$, setting Eqs. (2a) and (2a) to be equivalent in Eq. (3)

$$|n| = \left|\frac{1}{4m} - n\right| \iff n = \frac{1}{4\left(\frac{\sqrt{5}}{5}\right)} - n \tag{3}$$

gives $n = \frac{5}{8\sqrt{5}}$, that is, $p(x) = \frac{\sqrt{5}}{5}x + \frac{\sqrt{5}}{8}$ which is identical to $p_5(x) = 0.44721x + 0.27951$ in the result of the first approach.

In the same way, over the interval $[0, \mathrm{b}]$ the linear function which has the smallest maximum error from three non-collinear points (0,0), $(\mathrm{b}, \sqrt{\mathrm{b}})$ and $(\mathrm{c}, \sqrt{\mathrm{c}})$ for some point c in $(0, \mathrm{b})$ positions parallel to the line passing through two end points (0,0), and $(\mathrm{b}, \sqrt{\mathrm{b}})$. In other words, the approximation must be in the form of $p(x) = \frac{\sqrt{b}}{b}x + n$. Then the fact that the errors at $x = 0$ and at the critical point $\mathrm{c} = \frac{1}{4\mathrm{m}^2} = \frac{b}{4}$ equal:

$$|n| = \left|\frac{1}{4m} - n\right| \iff n = \frac{1}{4\left(\frac{\sqrt{b}}{b}\right)} - n \tag{4}$$

says $n = \frac{\sqrt{b}}{8}$. Therefore, the best linear approximation must be

$$p(x) = \frac{\sqrt{b}}{b}x + \frac{\sqrt{b}}{8} \tag{5}$$

and the minimax error is $|n| = \frac{\sqrt{b}}{8}$ for this case.

As another confirmation, the linear approximation to the function $f(x) = \sqrt{x}$ over the interval $[0, 10]$ is $p(x) = \frac{\sqrt{10}}{10}x + \frac{\sqrt{10}}{8} \approx 0.31623x + 0.39528$ which justifies the result $p_3(x) = 0.31623x + 0.39527$ and its error $\|g_3\|_\infty = 0.39529383$ in method 1.

III. Conclusion

As expressed, in order to find the best linear approximation of $f(x) = \sqrt{x}$ on the interval [0, b] in the minimax error, we examined two techniques: a) using the MATLAB code, positioning m and n values of the smallest maximum error on a broad range of m, and n value matrix in a rough scale and then repeatedly refining the selected regions in the smaller scales , and b) finding three-point fitting line to different three points not on the same line. We see that both experiments successfully attain the approximation

$$p(x) = \frac{\sqrt{b}}{b}x + \frac{\sqrt{b}}{8}$$

neglecting the tolerance with the minimax error $E_1(f, p) = \frac{\sqrt{b}}{8}$.

REFERENCES

[1] Burden, R. L., J. D. Faires and A. M. Burden. 2015. Numerical Analysis. 10 ed., Cengage Learning.

[2] Phillips, G. M. 2003. Interpolation and Approximation by Polynomials. Springer