Question 1 Given the quadratic equation $x^2+62.1x+1=0$. Find the approximation to each of the two solutions using **4 digit chopping arithmetic** and the appropriate equations for x_1 and x_2. Comment on the actual error.

Solution.

$$x^2 + 62.1x + 1 = 0 \tag{1}$$

Here first list the root finding equation of quadratic equation.

$$x_1 = \frac{-2c}{b + \sqrt{b^2 - 4ac}} \tag{2}$$

$$x_2 = \frac{-2c}{b - \sqrt{b^2 - 4ac}} \tag{3}$$

Here we first calculate the exact root of the equation

$$\begin{cases} x_1 = -0.016107237408969 \\ x_2 = -62.083892762597031 \end{cases} \tag{4}$$

Next use 4 digit chopping arithmetic, calculated the $b^2 - 4ac$

$$\begin{aligned} fl\left(b^2 - 4ac\right) &= (62.10)^2 - (4.000)(1.000)(1.000) \\ &= (3852.) \end{aligned} \tag{5}$$

Next calculate the $\sqrt{b^2 - 4ac}$ with (5)

$$fl(\sqrt{b^2 - 4ac}) = \sqrt{(3852.)} = 62.06 \tag{6}$$

Then calculate the $b + \sqrt{b^2 - 4ac}$,$b - \sqrt{b^2 - 4ac}$ with (6)

$$fl\left(b + \sqrt{b^2 - 4ac}\right) = 62.06 + 62.10 = 124.1 \tag{7}$$

$$fl\left(b - \sqrt{b^2 - 4ac}\right) = -62.06 + 62.10 = 0.04000 \tag{8}$$

Then calculate the value of $-2c$

$$fl(-2c) = -(2.000)(1.000) = -2.000 \tag{9}$$

Then calculate the $fl(x_1)$ and $fl(x_2)$ with (2), (3),(7), (8) and (9)

$$fl(x_1) = fl\left(\frac{-2c}{b + \sqrt{b^2 - 4ac}}\right) = \frac{-2.000}{124.1} = -0.01611 \tag{10}$$

$$fl\left(x_2\right) = fl\left(\frac{-2c}{b - \sqrt{b^2 - 4ac}}\right) = \frac{-2.000}{0.04000} = -50.00 \tag{11}$$

Finally compute the actual error of the solution by comparing (4) and (10),(11)

$$\begin{cases} x_1 - x_1^* = 0.7237 \times 10^{-5} \\ x_2 - x_2^* = -12.08389276 \end{cases} \tag{12}$$

Comment Here we can find that the root of the equation has a big difference in the size of numbers. Therefore it's clearer for as to campare the relative error instead of actual error. The relative error can be given that:

	actual value	measured value	absolute error	relative error
x_1	-0.016107237408967	-0.016120000000000	0.000012762591033	0.000792351333064
x_2	-62.083892762591034	-50.000000000000000	12.083892762591034	0.194638129551571

It can be clearly find that the relative of x_1 is acceptable. However, the relative error of the x_2 is more then 10%.

Through observation, we can find that in this question, the size of $b = 62$ is much larger than that of $a = 1$ and $c = 1$, which leads to the difference between b and $\sqrt{b^2 - 4ac}$ very tiny relative to themselves. and we find that the subtraction (8) gets a very inaccurate answer. Besides in equation (11), also the division by the small result of this subtraction. The inaccuracy that this combination produces, which also enlarge the absolute error of x_2.

Here we also apply another form of the root finding equation to check the result fi there will have any difference.

$$x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \tag{13}$$

By (6),it can be found that:

$$fl(-b + \sqrt{b^2 - 4ac}) = -0.04000$$

$$fl(-b - \sqrt{b^2 - 4ac}) = -124.1$$

Therefore:

$$fl(x_1) = -0.02000 \tag{14}$$

$$fl(x_2) = -62.05 \tag{15}$$

We skip the calculation process and look at the error directly:

	actual value	measured value	absolute error	relative error
x_1	-0.0161072374089670	-0.0200000000000000	0.00389276259103298	0.241677855251941
x_2	-62.0838927625910	-62.0500000000000	0.0338927625910372	0.000545918773499646

In this case, you can find that x_2 because it avoids the amplification of error caused by using a small number as a multiplier, a more accurate answer is obtained. However, under the same method, there is a big problem with x_1, which is also due to the small gap between B and $\sqrt{b^2-4ac}$ mentioned earlier.

Therefore, we can turn out two method to reduce the error:

1. We can transform the form of the equation to reduce the error of result in quadratic function to prevent the too small divisor.

$$x_{1,2} = \frac{-b \pm \sqrt{b^2-4ac}}{2a} = \frac{-2c}{b \pm \sqrt{b^2-4ac}}$$

2. We can add the digits when the two very close number applying subtraction.

■

Question 2 Given the quadratic equation $x^2+62.1x+1=0$. Find the approximation to each of the two solutions using **4 digit rounding arithmetic** and the appropriate equations for x_1 and x_2. Comment on the actual error.

Solution.

$$x^2+62.1x+1=0 \tag{16}$$

Here first list the root finding equation of quadratic equation.

$$x_1=\frac{-2c}{b+\sqrt{b^2-4ac}} \tag{17}$$

$$x_2=\frac{-2c}{b-\sqrt{b^2-4ac}} \tag{18}$$

Here we first calculate the exact root of the equation

$$\begin{cases} x_1=-0.016107237408969 \\ x_2=-62.083892762597031 \end{cases} \tag{19}$$

Next use 4 digit chopping arithmetic, calculated the b^2-4ac

$$\begin{aligned} fl\left(b^2-4ac\right) &= (62.10)^2-(4.000)(1.000)(1.000) \\ &= (3852.) \end{aligned} \tag{20}$$

Next calculate the $\sqrt{b^2-4ac}$ with (20)

$$fl(\sqrt{b^2-4ac})=\sqrt{(3852.)}=62.07 \tag{21}$$

Then calculate the $b+\sqrt{b^2-4ac}$,$b-\sqrt{b^2-4ac}$ with (21)

$$\begin{aligned} fl\left(b+\sqrt{b^2-4ac}\right) &= 62.07+62.10=124.2 \\ fl\left(b-\sqrt{b^2-4ac}\right) &= -62.07+62.10=0.03000 \end{aligned} \tag{22}$$

Then calculate the value of $-2c$

$$fl(-2c)=-(2.000)(1.000)=-2.000 \tag{23}$$

Then calculate the $fl(x_1)$ and $fl(x_2)$ with (17),(18),(),(23)

$$fl(x_1)=fl\left(\frac{-2c}{b+\sqrt{b^2-4ac}}\right)=\frac{-2.000}{124.2}=-0.01610 \tag{24}$$

$$fl\left(x_2\right)=fl\left(\frac{-2c}{b-\sqrt{b^2-4ac}}\right)=\frac{-2.000}{-0.03000}=-66.67 \tag{25}$$

Finally compute the actuall error of the solution by comparing (19) and (24),(25)

$$\begin{cases} x_1 - x_1^* = -0.7237 \times 10^{-5} \\ x_2 - x_2^* = 4.576007240 \end{cases} \tag{26}$$

Comment Similar with Q1, here list the relative error:

	actual value	measured value	absolute error	relative error
x_1	-0.0161072374089670	-0.0161000000000000	7.23740896702366e-06	0.000449326522187755
x_2	-62.0838927625910	-66.6700000000000	4.58610723740897	0.0738695180559353

Here it can be found that the x_2 has a relatively huge error, and the reason is similar with the previous problem, b and $\sqrt{b^2 - 4ac}$ too close and subtraction produce a small difference, and the error enlarged because of the tiny division. Here we also applied the similar method in question 1 that use the different form of the root finding equation.

$$x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \tag{27}$$

$$fl\left(-b + \sqrt{b^2 - 4ac}\right) = 62.07 - 62.10 = -0.03000$$
$$fl\left(-b - \sqrt{b^2 - 4ac}\right) = -62.07 - 62.10 = -124.2$$

$$fl(x_1) = -0.01500$$

$$fl(x_2) = -62.10$$

And the error can be given that:

	actual value	measured value	absolute error	relative error
x_1	-0.0161072374089670	-0.0150000000000000	-0.00110723740896702	0.0687416085610445
x_2	-62.0838927625910	-62.1000000000000	0.0161072374089670	0.000259443096948852

Similar result with the question 1 that the different form of equation result different accuracy with different root, here it can be found that root x_2 got more accurate while the x_1 error increase. Therefore similar result could be given with question 1:

1. We can transform the form of the equation to reduce the error of result in quadratic function to prevent the too small divisor.

$$x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} = \frac{-2c}{b \pm \sqrt{b^2 - 4ac}}$$

2. We can add the digits when the two very close number applying subtraction.

■

Question 3
Construct the third Taylor polynomial about $x_0 = 1$ approximating $\ln(x)$. Use this polynomial to approximate $\ln(1.1)$. Then, using the truncation error (or remainder term) for this Taylor polynomial, bound the error of you approximation to $\ln(1.1)$. Comment on the actual error.

Solution.

Here the Taylor polynomial is given that

$$P_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k \tag{28}$$

The remainder equation is given that

$$R_n(x) = \frac{f^{(n+1)}(\xi(x))}{(n+1)!}(x - x_0)^{n+1} \tag{29}$$

Here, since the function is $f(x) = \ln(x)$, the derivative is given that

$$\begin{cases} f'(x) = \dfrac{1}{x} \\ f''(x) = -\dfrac{1}{x^2} \\ f^{(3)}(x) = \dfrac{2}{x^3} \\ f^{(4)}(x) = -\dfrac{2 \cdot 3}{x^4} \end{cases} \tag{30}$$

Here we input the derivative (30) into the Taylor polynomial (28):

$$\begin{aligned} P_3(x) &= f(1) + \frac{f'(1)}{1!}(x-1) + \frac{f''(1)}{2!}(x-1)^2 + \frac{f'''(1)}{3!}(x-1)^3 \\ &= 0 + 1(x-1) - \frac{1(x-1)^2}{2!} + \frac{2(x-1)^3}{3!} \end{aligned}$$

Here input $x = 1.1$ and approximate $\ln(x)$

$$\begin{aligned} P_3(1.1) &= 0.1 - \frac{1 \cdot 0.1^2}{2!} + \frac{2 \cdot 0.1^3}{3!} \\ &= 0.09533 \end{aligned}$$

Then the remainder function, $R_3(x)$ is given that

$$R_3(x) = \frac{f^{(4}(\xi(x))}{(4)!}(x-1)^{n+1} = -\frac{3!}{4!\xi^4}(x-1) \tag{31}$$

Where $1 \leq \xi \leq 1.1$,since $\xi^{-4} \leq 1$

$$|R_3(x)| \leq \frac{3!}{4!}\frac{1}{1^4}(1.1-1)^4 = 0.2500 \times 10^{-4} \tag{32}$$

Therefore we can conclude that the Now compare the approximation result and the exact value and generate the actual error:

$$|\ln(1.1) - P_3(1.1)| = 0.23154 \times 10^{-4} < 0.2500 \times 10^{-4} \tag{33}$$

Comment It could be found that the Taylor polynomial gives the approximation of ln(1.1) and the Truncation error give the bound of error, it could be found that the actual error is in the bound of the truncation error and close to the error bound, therefore it's a good approximation for third Taylor polynomial in this question.

■

Question 4
Given $f(x) = 2 - x^2 \sin(x)$

(a) Verify that the Bisection method can be applied to the function $f(x)$ on $[-1, 2]$.

(b) Using the error formula for the Bisection method find the number of iterations needed for accuracy 0.000001.

(c) Write a program by using MATLAB/Octave to determine an approximation to the root that is accurate to at least within 0.000001.

(d) Comment on the number of iterations in (b) and (c).

Solution.

(a) Here $f(x) - 2 - x^2 \sin(x)$ is continous function defined on $[-1, 2]$,

$$\begin{cases} f(-1) = 2 - (-1)^2 \sin(-1) = 2 - \sin(-1) = 2.8414 > 0 \\ f(2) = 2 - 2^2 \sin(2) = -1.6371 < 0 \end{cases} \tag{34}$$

The Intermediate Value Theorem implies that a number p exists in $[-1, 2]$ with $f(p) = 0$. Therefore the Bisection method can be applied to the function.

(b) Here apply the error function:

$$|p_n - p| \leq \frac{b-a}{2^n} \tag{35}$$

Here $b = 2$, $a = -1$,and accuracy 1.00×10^{-6}, input to (35)

$$|p_N - p| \leq 2^{-N}(2 - (-1)) = 3 \times 2^{-N} < 10^{-6}$$

$$\begin{aligned} 3 \times 10^6 &< 2^N \\ N &> \log_2(3 \times 10^6) \approx 21.5165 \\ N &\geq 22 \end{aligned}$$

Therefore 22 times of iterations need for accuracy 0.000001.

(c) The .m file is attached with file

Here is the list of the Iteration (1), the iteration stops at 22 times with accuracy 0.000001.

(d) Here it is found that the time of iterations is exactly the number of turn out by error function.This is because the criterion here used to test the accuracy is same as the error function used in (b)(35).

Here we find that it required much work to do to enhance the accuracy with the Bisection Method.

n	a_n	b_n	p_n	Error
1	-1	2	0.500000000000000	1.50000000000000
2	0.500000000000000	2	1.25000000000000	0.750000000000000
3	1.25000000000000	2	1.62500000000000	0.375000000000000
4	1.25000000000000	1.62500000000000	1.43750000000000	0.187500000000000
5	1.25000000000000	1.43750000000000	1.34375000000000	0.0937500000000000
6	1.34375000000000	1.43750000000000	1.39062500000000	0.0468750000000000
7	1.39062500000000	1.43750000000000	1.41406250000000	0.0234375000000000
8	1.41406250000000	1.43750000000000	1.42578125000000	0.0117187500000000
9	1.41406250000000	1.42578125000000	1.41992187500000	0.00585937500000000
10	1.41992187500000	1.42578125000000	1.42285156250000	0.00292968750000000
11	1.41992187500000	1.42285156250000	1.42138671875000	0.00146484375000000
12	1.42138671875000	1.42285156250000	1.42211914062500	0.000732421875000000
13	1.42138671875000	1.42211914062500	1.42175292968750	0.000366210937500000
14	1.42175292968750	1.42211914062500	1.42193603515625	0.000183105468750000
15	1.42193603515625	1.42211914062500	1.42202758789063	9.15527343750000e-05
16	1.42202758789063	1.42211914062500	1.42207336425781	4.57763671875000e-05
17	1.42207336425781	1.42211914062500	1.42209625244141	2.28881835937500e-05
18	1.42207336425781	1.42209625244141	1.42208480834961	1.14440917968750e-05
19	1.42207336425781	1.42208480834961	1.42207908630371	5.72204589843750e-06
20	1.42207908630371	1.42208480834961	1.42208194732666	2.86102294921875e-06
21	1.42208194732666	1.42208480834961	1.42208337783813	1.43051147460938e-06
22	1.42208337783813	1.42208480834961	1.42208409309387	7.15255737304688e-07

Table 1: Iteration List of Bisection Method

Question 5 Given $f(x) = 2 - x^2 \sin(x)$ has a solution in the interval $[-1, 2]$.

(a) Sketch the graph for $f(x)$ on interval $[-10, 10]$

(b) Apply five steps of Newton's Method with initial guess $x_0 = -1$ to attempt to find this root.

(c) Apply five steps of Newton's Method with initial guess $x_0 = 2$ to attempt to find this root.

(d) Compare the results of (b) and (c), explain what happens.

Solution.

(a) The Diagram generated by MatLab (1)

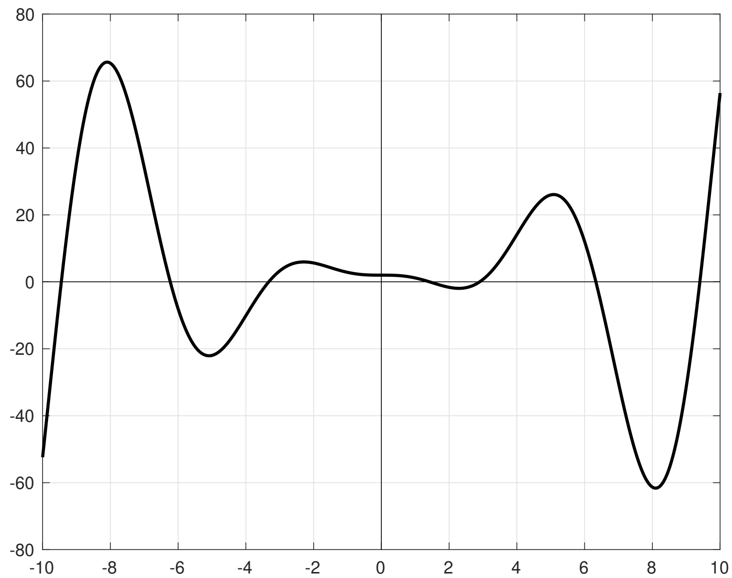


Figure 1: (a) $f(x) = 2 - x^2 \sin(x)$

(b) Here first find the iteration method.

$$x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)} \tag{36}$$

Here the function is given that

$$f(x) = 2 - x^2 \sin(x) \tag{37}$$

$$f'(x) = -2x \sin(x) - x^2 \cos(x) \tag{38}$$

Therefore input (37) and (38) into (36),the function is given that:

$$x_{i+1} = x_i - \frac{2 - x_i^2 \sin(x_i)}{-2x_i \sin(x_i) - x_i^2 \cos(x_i)} \tag{39}$$

In this section we set the initial guess, $x_0 = -1$:

$$x_1 = x_0 - \frac{2 - x_0^2 \sin(x_0)}{-2x_0 \sin(x_0) - x_0^2 \cos(x_0)} = 0.2781$$

$$x_2 = x_1 - \frac{2 - x_1^2 \sin(x_1)}{-2x_1 \sin(x_1) - x_1^2 \cos(x_1)} = 8.994$$

$$x_3 = x_2 - \frac{2 - x_2^2 \sin(x_2)}{-2x_2 \sin(x_2) - x_2^2 \cos(x_2)} = 9.475$$

$$x_4 = x_3 - \frac{2 - x_3^2 \sin(x_3)}{-2x_3 \sin(x_3) - x_3^2 \cos(x_3)} = 9.403$$

$$x_5 = x_4 - \frac{2 - x_4^2 \sin(x_4)}{-2x_4 \sin(x_4) - x_4^2 \cos(x_4)} = 9.402$$

(c) In this section we set the initial guess, $x_0 = 2$:

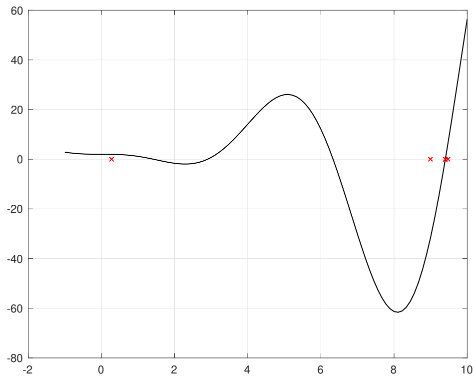
$$x_1 = x_0 - \frac{2 - x_0^2 \sin(x_0)}{-2x_0 \sin(x_0) - x_0^2 \cos(x_0)} = 1.170$$

$$x_2 = x_1 - \frac{2 - x_1^2 \sin(x_1)}{-2x_1 \sin(x_1) - x_1^2 \cos(x_1)} = 1.445$$

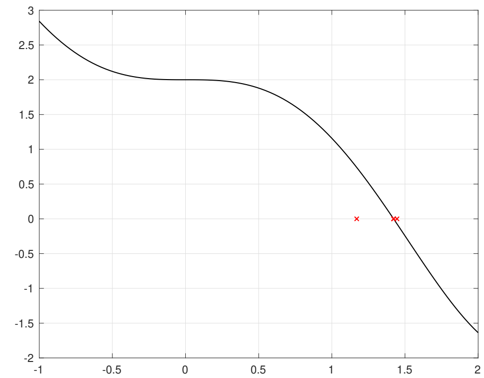
$$x_3 = x_2 - \frac{2 - x_2^2 \sin(x_2)}{-2x_2 \sin(x_2) - x_2^2 \cos(x_2)} = 1.422$$

$$x_4 = x_3 - \frac{2 - x_3^2 \sin(x_3)}{-2x_3 \sin(x_3) - x_3^2 \cos(x_3)} = 1.422$$

$$x_5 = x_4 - \frac{2 - x_4^2 \sin(x_4)}{-2x_4 \sin(x_4) - x_4^2 \cos(x_4)} = 1.422$$



(a) 5(b)



(b) 5(c)

Figure 2: Iteration Point of 5(b) and 5(c)

(d) Here we first plot the graph of the iteration point generated by 5(b) and 5(c), (2):

Therefore it is easy to find that in 5(b),the initial guess leads to the root jump to another region, out of the required range. Which is one of the drawback of the Newton's Method: In some cases where the function is oscillating and has a number of roots, one may choose an initial guess close to a root. However, the guesses may jump and converge to some other root.

Here we can plot the graph to illustrate the procedure of the root jumping in 5(b),(3). Newton's Method could be described in figure that we choose one point and found the tangent line on $f(x)$, the focus of the tangent and X axis will be the result and the initial value of the next iteration. Here we hope the focus point can be closer to the root we want.

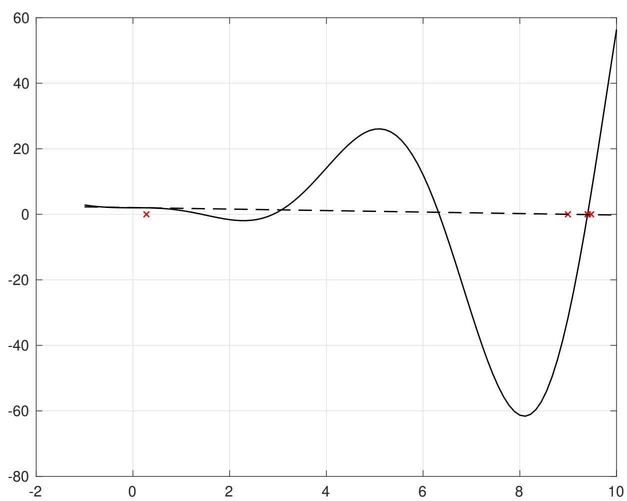
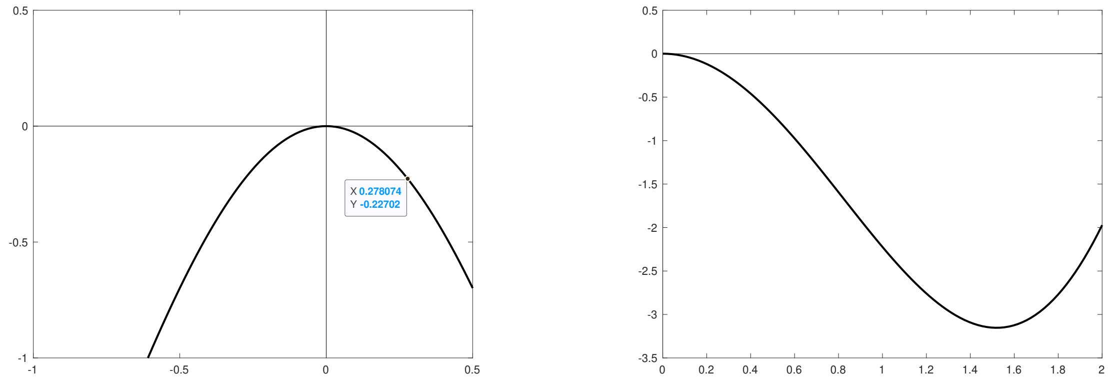


Figure 3: The procedure of the root jumping in 5(b)



(a) 5(b)

(b) 5(c)

Figure 4: Function first derivative (slope) diagram

However in 5(b), when the second times of the iteration leads to an abnormal focus with the x axis. This is because in the local region, the slope is too small and turn to be flat (4), which leads the focus locate in second iteration far from the required region, at 8.994375358445748, and then the further iteration leads to converge to another root around 9.4.

In 5(c) we start at $x_0 = 2$, here by plotting the graph of the derivative, we can find that the slope is enough to let the result converge to the root without jumping. (4)

To prevent the Root jumping problem when applying the Newton's Method, several tips could be applied to overcome:

(a) When the function is oscillating or have many roots, The Method of False Position may be a better choice. It has strong stability, but it lags behind in efficiency.

(b) Before iteration, check whether there is a phenomenon that the slope approaches zero near the selected initial point, which will lead to serious deviation of the iterative results in Newton method.

(c) Choose the point closer to the unknown root as the initial point, which will improve the stability and efficiency of the model to a certain extent.

■