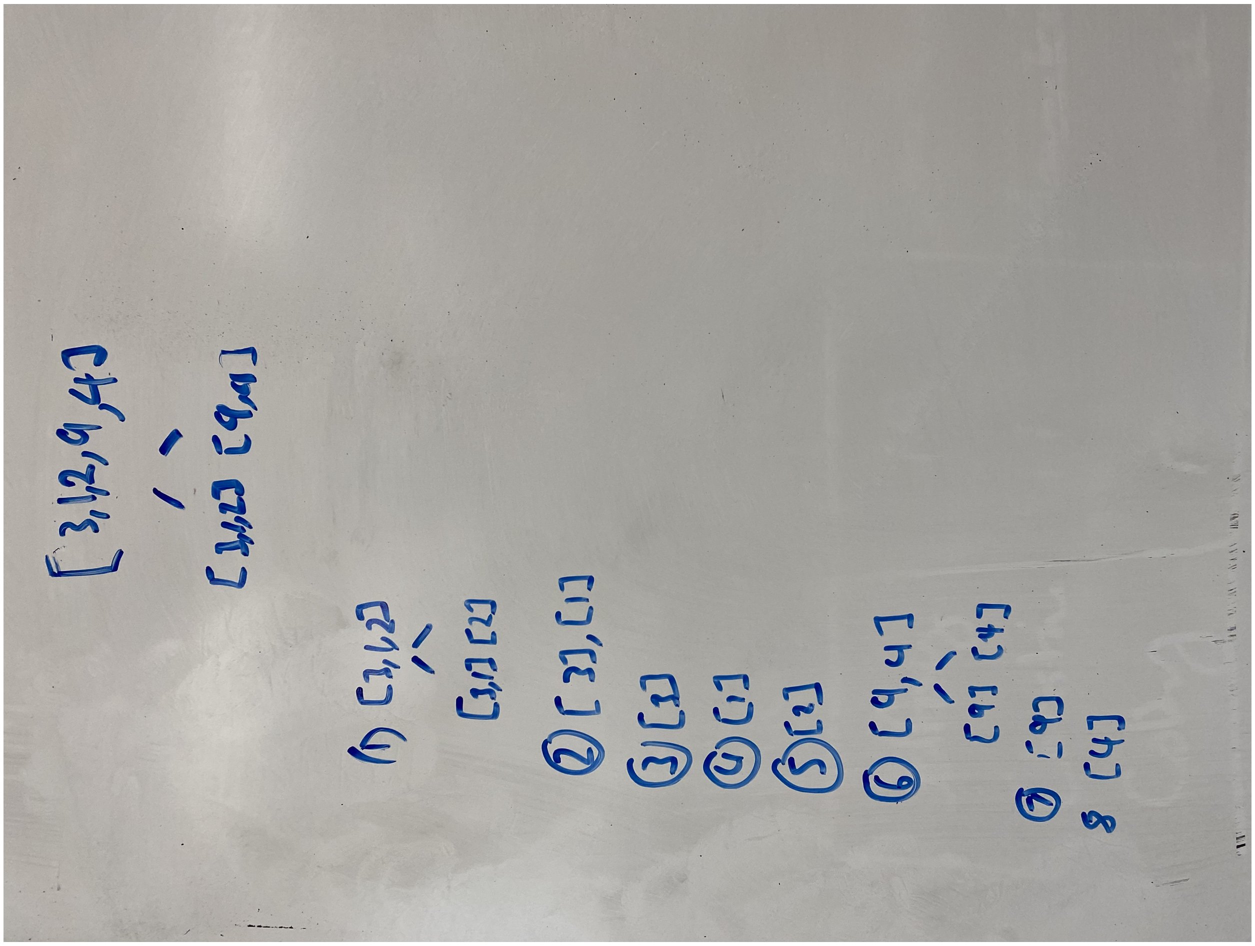
test answers

1)



2)

$$
\begin{gathered}
x^2 + x - 6 \\
x = 1 \\
f(x) = -4 \\
f'(x) = 3 \\
x- = \frac{f'(x)}{f(x)} = 1 - (-4/3) \\
x = 7/3
\end{gathered}
$$

$$f(x) = 49/9 + 7/3 - 6$$
$$f'(x) = 14/3 + 1$$
$$x- = \frac{f(x)}{f'(x)} = 7/3 - \frac{49/9 + 7/3 - 6}{14/3 + 1} = 103/51$$

approaches 2

3)

(1,2),(3,4),(5,5)

$\begin{bmatrix}2\\4\\5\end{bmatrix}$ = $\begin{bmatrix}1 & 1\\1 & 3\\1 & 5\end{bmatrix} * \begin{bmatrix}a\\b\end{bmatrix}$

$\begin{bmatrix}1 & 1 & 1\\1 & 3 & 5\end{bmatrix} * \begin{bmatrix}2\\4\\5\end{bmatrix}$ = $\begin{bmatrix}1 & 1 & 1\\1 & 3 & 5\end{bmatrix} * \begin{bmatrix}1 & 1\\1 & 3\\1 & 5\end{bmatrix} * \begin{bmatrix}a\\b\end{bmatrix}$

x = $\begin{bmatrix}1 & 1 & 1\\1 & 3 & 5\end{bmatrix} * \begin{bmatrix}2\\4\\5\end{bmatrix}$

y = $\begin{bmatrix}1 & 1 & 1\\1 & 3 & 5\end{bmatrix} * \begin{bmatrix}1 & 1\\1 & 3\\1 & 5\end{bmatrix}$

y^-1 * x = $\begin{bmatrix}a\\b\end{bmatrix}$

4)

for 2, you can’t have 0,0

for 4, ln(1) = 0, which would be null if you do 1/0

for 1), it’s the same datapoint 3 times, which gives 1 equation 2 variables

5)

Doing #3

[(1,0.5),(2,0.4),(3,0.45)]

$$\begin{bmatrix} ln(\frac{1}{0.5} - 1) \\ ln(\frac{1}{0.9} - 1) \\ ln(\frac{1}{0.95} - 1) \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} * \begin{bmatrix} a \\ b \end{bmatrix} \tag{1}$$

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} * \begin{bmatrix} 0 \\ -2.2 \\ -2.94 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} * \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} * \begin{bmatrix} a \\ b \end{bmatrix} \tag{2}$$

x = $\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} * \begin{bmatrix} 0 \\ -2.2 \\ -2.94 \end{bmatrix}$

y = $\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} * \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}$

$y^{-1} * x = \begin{bmatrix} a \\ b \end{bmatrix}$

6)

now 2 and 4 is possible, but 1 still doesn’t work, as changing bounds can’t fix the 2 variables 1 equation dilemma.

2 : m = -1, M = 10

4 :m = 0. M = 2

7)

2 : [0,0),(1,0.5),(2,0.9)]

$$\begin{bmatrix} ln(\frac{10-0}{0--1}) \\ ln(\frac{10-0.5}{0.5--1}) \\ ln(\frac{10-0.9}{0.9--1}) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} * \begin{bmatrix} a \\ b \end{bmatrix} \tag{3}$$

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix} * \begin{bmatrix} 2.363 \\ 1.846 \\ 1.566 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix} * \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} * \begin{bmatrix} a \\ b \end{bmatrix} \tag{4}$$

x = $\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix} * \begin{bmatrix} 2.363 \\ 1.846 \\ 1.566 \end{bmatrix}$

y = $\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix} * \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}$

$y^{-1}x = \begin{bmatrix} a \\ b \end{bmatrix}$ == $\begin{bmatrix} 2.274 \\ -0.369 \end{bmatrix}$

y = -1 + $\frac{11}{1+e^{-0.369x+2.274}}$

8)

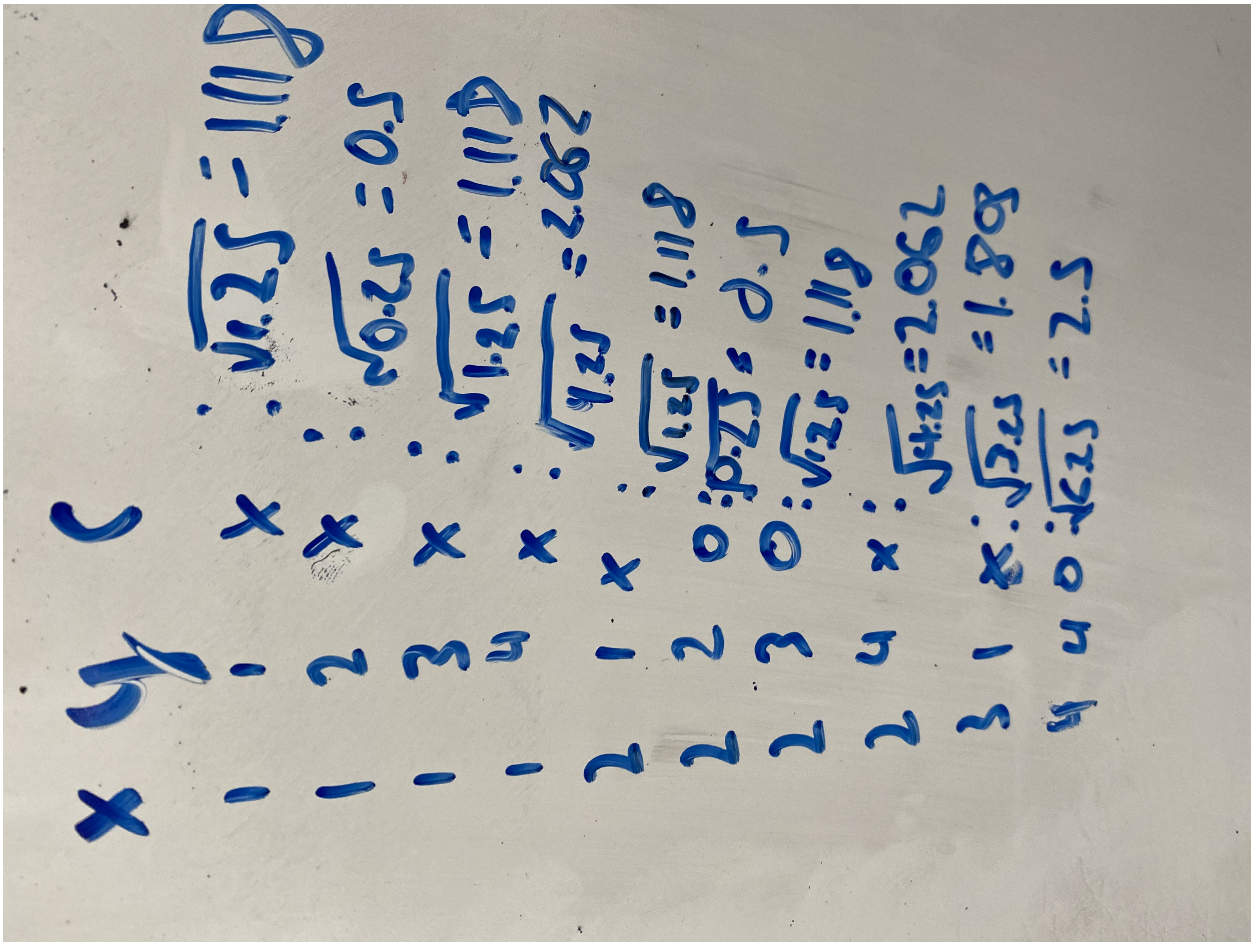
[(1,0,3),(2,0,6),(3,0,9),(4,0,7),(0,2,5),(0,3,6),(0,4,7),(3,3,0),(1,1,0)]

9)

An example of where interaction terms would be useful is with something like painting. Say for instance red flowers look nice, blue flowers look nice, yellow flowers look nice, so if you mix them all together red + blue + yellow = brown flowers must look nice, which isn't as good as pure ___ (i haven't seen a brown flower in real life so I can't talk, maybe they are superior in every way)

10)

the equation is $\sqrt{(1.5-x)^2+(2-y)^2}$, didn't have space/ ability to write in straight line to write all



a) For k =1, it would be split in between as there is an equal number of x and o's for 0.5

b) k = 3 would give x, since you have $0.5(x), 0.5(o), 1.118(x), 1.118(x), 1.118(o)$

c) k = 10, would go to x as there is 7 x, 3 o

11)

$$f(x) = x^2 + x - 6, f'(x) = 2x + 1 \tag{5}$$

$$\rho = 0.01, x = 1 \tag{6}$$

$$x- = 0.01(3) = 0.97 \tag{7}$$

$$x- = 0.01(2.94) = 0.9406 \tag{8}$$

going to 0.5

12)

$x^2 + (y-2)^2 + 3(z-1)^2, x_0 = (1,1,1), \rho = 0.01$

$f_x = 2x, f_y = 2(y-2), f_z = 6(z-1)$

$x = 1 - 0.01(f_x(1)) = 0.98$

$y = 1 - 0.01(f_y(2)) = 1.02$

z =$1 - 0.01(f_z(0)) = 1$

approaches 1

13)

a) False. Most times it's slower since more calls = more cpu power, and it's generally harder to debug

b)True

c)False. gradient descent tries to find the minimum, while Newton's tries to find f'(x) = 0, or where the slope is zero

d)False, wider graphs require more memory at one time (since you have more nodes width than height)

e)depends on the graphs, but it's usually similar

f)true

g)False, depending on if it's a left or right inverse

right is : $A^T(AA^T)^{-1}$

left is : $(AA^T)^{-1}A^T$