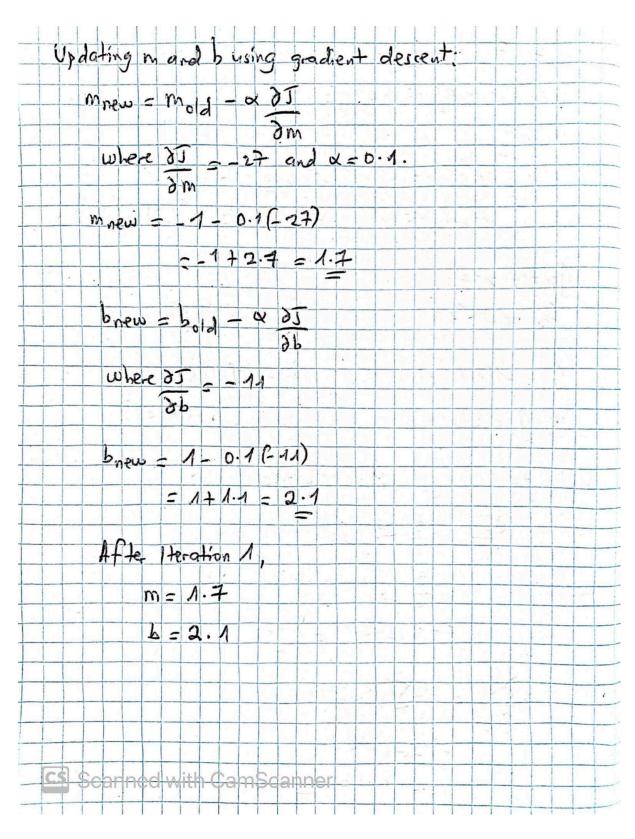
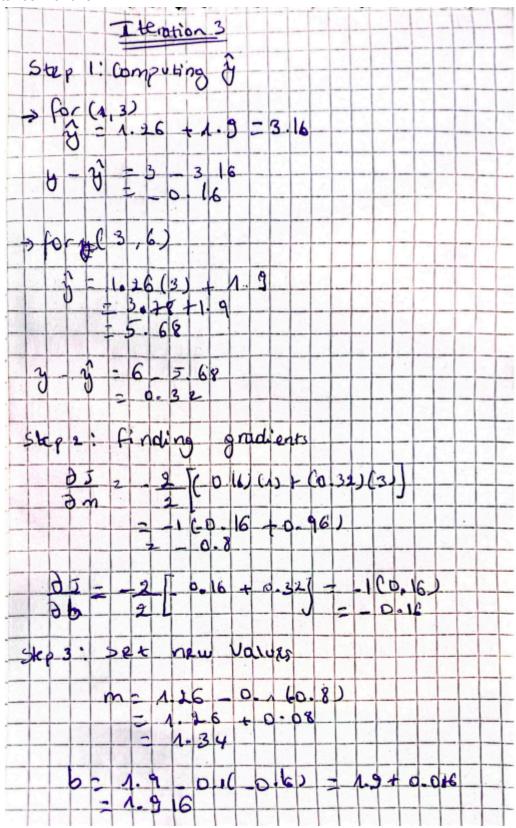
rart 3.	Gradient Descen	t Manual C	a last ofion	
Given	linear equation:	y=mx+b	30.	~
whe	e		2 2	
	Intial m= -1	Learning T	-1- ~	1
	Intrial in a			
	Intial b = 1	Given poil	Ats - (1,3) and (3,6
			1	
û- =	mx;+b			
3				
to-	point (1,3),	ter poin	+ (3,6)	
9; =	-1(-1)+1=0	9; = -1	(3) ta = -	2 18 4
11 1	0 0 7		(C2)	2
y;-y	= 3-0 - 3	9;-9;	- 6-(-2)	- 0
		1 1 2 2 3 3	-0 -1 N -1	
Using +	e Mean squared	error where n	=2:	
	3-	(1)	- 1 - 1 - 2	
	3T = -2 \(\)	41-4: 176		
	8m	×1) + (8×3)]	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	144
	= -2 (3	X1)+(8XS)-	- 1	(6) - (1)
	=-1 (34:	24/ = -27	4.25	
	25 = -25(y)	-41		
	76 7	3		
	2-2 (3	+8)		
Scan	ned with Camea			



Makuochi Prince

Iteration 2 de la
Step I: Calculate predicted values (q) for Point (1,3): y, = 1.7 x 1 + 2.1 = 3.8
For Point (3,6):
$\hat{y}_2 = 1.4 \times 3 + 2.1 = 4.2$
Step 2: Calculate errors (4-9)
for point (1,3): 1, -9, = 3-8.8 = -0.8
700 - point (3,6) 72 - 72 = 6 - 7.2 = -1.2
Step 3: Compute gradients
For m: $\frac{\partial J}{\partial m} = -2/2 \left[(-0.8 \times 1) + (-1.2 \times 3) \right] = -1(-0.8 \times 1) + (-1.2 \times 3) = -1(-0.8 \times 1) + (-1.2 \times 1) + (-1.2 \times 1) = -1(-0.8 \times 1) + (-1.2 \times 1) = -1(-0.8 \times 1)$
$\frac{1}{2}$ for $\frac{1}{2}$
New M: b M New = 1.7-0.1 × 4.4 = 1.7-0.44
1-26// New b: b_ New = 2.1-0.1×2 = 2.1-0.2=1.9
After Iteration 2
M = 1.26, $b = 1.9$

Hortance Irakoze



The values of m and b are moving towards reducing the error. Translating to our model learning and adjusting to the observations to better fit our data on each iteration.

Davy Ngamije

1 tention IV	Step 3: Update Values
5 tepri Compole ý	n = 1.34 - 0.1(0.064)
→ Por(1,3) 1 = N34 + 1.916	1 1 2 1.3336
9 = N3 4 1 9 16 = 3 9 56 4 - 9 = 0 9 56	6 = 1.916 - 0-1(0.1921 = 1.916 - 0.019
3 - for (3,6) = 1.34(3) 1.916	After 4 iterations
1.02 + 1.016 -5.936 2 - 2 - 0.064	m 2 1.3336
07 - 1 (CO, 256)(1)+(0)	-06W(3)
7 - 1 -0.256 +0.19	
CS Scanned with CamSca	anner () - 0. 192

Trend Observation:

m: -1
$$\rightarrow$$
 1.7 \rightarrow 1.26 \rightarrow 1.34 \rightarrow 1.3336 (stabilizing ~1.33)

b:
$$1 \rightarrow 2.1 \rightarrow 1.9 \rightarrow 1.916 \rightarrow 1.8968$$
 (stabilizing ~1.90)

Key Insight:

Parameters do not move toward (0,0).

Instead, they converge to values that minimize error for the given data.

By the 4th iteration, changes become very small ($\Delta m \approx 0.006$, $\Delta b \approx 0.02$), indicating near-convergence.

Final Line Equation:

y≈1.3336x+1.8968

y≈1.3336x+1.8968

This fits the data points (1,3) and (3,6) with minimal error (MSE ≈ 0.032).