

- Take analytical derivative of sigmoid function (on a paper with pencil, take a photo of your calculation and attach it to pdf report)

Sigmoid

$$G(x) = \frac{1}{1+e^{-x}}$$

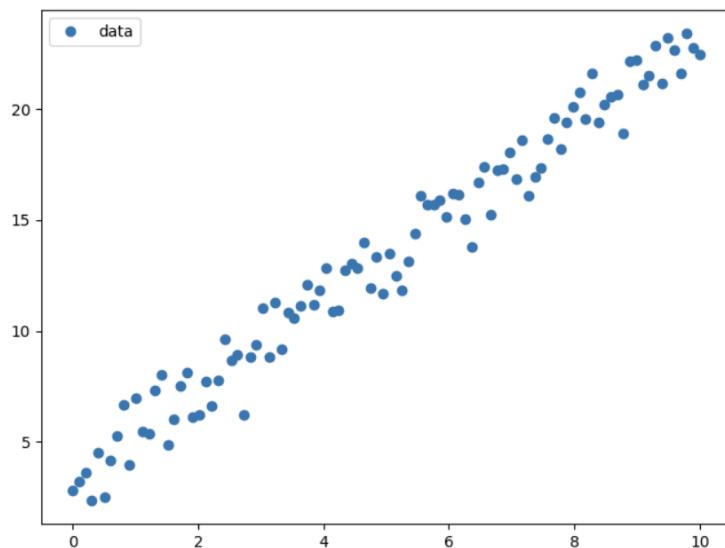
$$G'(x) = \frac{e^{-x}}{(1+e^{-x})^2} = G(x) \cdot \frac{e^{-x}}{1+e^{-x}} = G(x) \left(\frac{1+e^{-x}}{1+e^{-x}} - \frac{1}{1+e^{-x}} \right) = G(x)(1-G(x))$$

- Experiments with demo code (gradient descent)

- Vary learning rate
- Vary epochs
- Plot MSE over training (over epochs for specific learning rate)

Вхідні дані: x - рівномірно розподілені на проміжку від 0 до 10;

y - розраховуються за формулою $y = 2x + 3,5$ з додаванням шуму (випадкові числа з нормальногого розподілу).



Експеримент 1

Експеримент 2

Експеримент 3

Експеримент 4

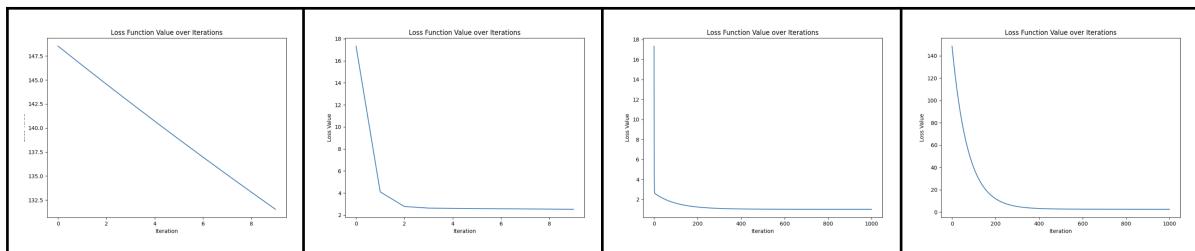
Швидкість навчання (η):

0,0001	0,01	0,01	0,0001
--------	------	------	--------

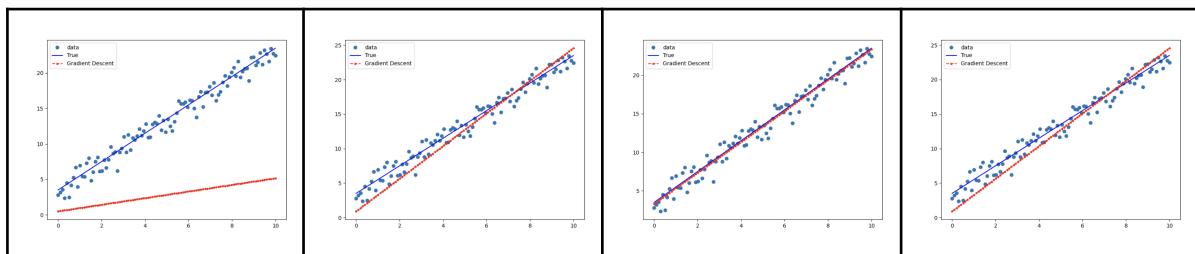
Кількість ітерацій (epoch):

10	10	1000	1000
----	----	------	------

Графік значень функції похибок:



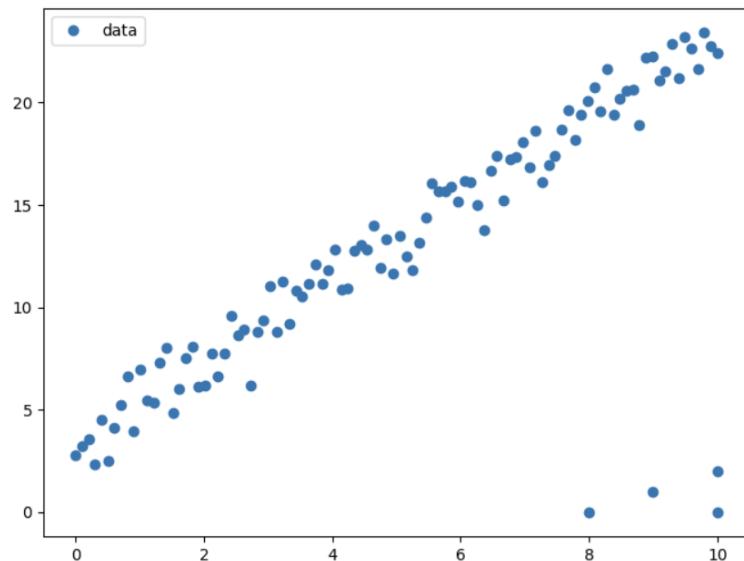
Регресійна модель:



Знайдені значення параметрів:

$m = 0.466298463$	$m = 2.365872592$	$m = 2.003222509$	$m = 2.363812529$
$b = 0.501789827$	$b = 0.912882162$	$b = 3.324892038$	$b = 0.912268326$

Вхідні дані: x - рівномірно розподілені на проміжку від 0 до 10;
 y - розраховуються за формулою $y = 2x + 3,5$ з додаванням
шуму (випадкові числа з нормальногого розподілу);
викиди (outlier): (10,0), (9,1), (10,2), (8,0).



Експеримент 5

Експеримент 6

Експеримент 7

Експеримент 8

Швидкість навчання (lr):

0,0001

0,01

0,01

0,0001

Кількість ітерацій (epoch):

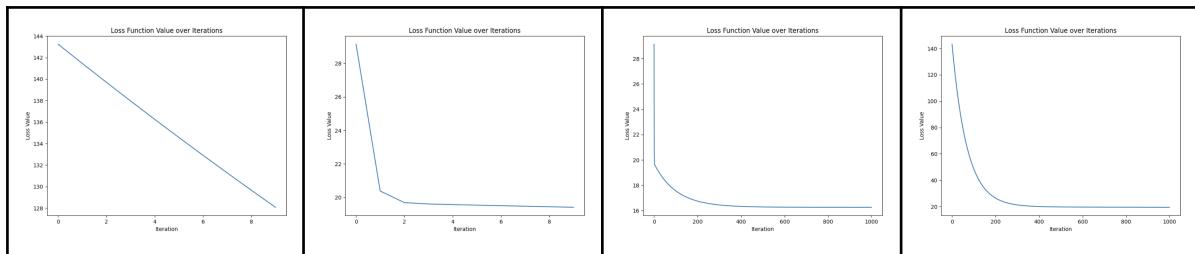
10

10

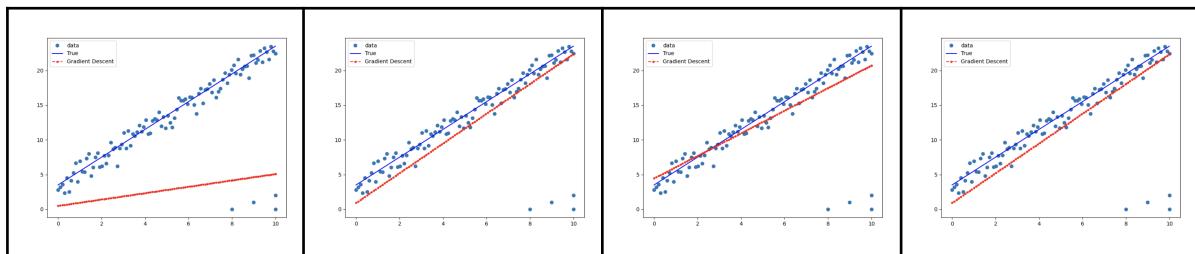
1000

1000

Графік значень функції похибок:



Регресійна модель:



Знайдені значення параметрів:

$$m = 0.458905055 \\ b = 0.500728060$$

$$m = 2.142669098 \\ b = 0.927759149$$

$$m = 1.626148177 \\ b = 4.457588623$$

$$m = 2.141472671 \\ b = 0.927152783$$

Висновки:

В експериментах 1 та 5 ($lr = 0,0001$, epoch = 10), де значення параметрів m та b підібрані не найкращим чином, модель немає достатньої кількості ітерацій та значень lr , щоб наблизитися до оптимальних значень. В експерименті 3 ($lr = 0.01$, epoch = 1000, без аутлаерів) досягнуто точних параметрів регресійної моделі. Це показує, що відповідне значення lr і достатня кількість ітерацій сприяють точному навчанню моделі. Знайдені параметри інших експериментів близькі до оптимальних, але ними не є. Також наявність аутлаерів у даних негативно вплинули на процес навчання моделі, збільшивши значення функції втрат та знизвивши точність моделі.

Загалом, експерименти демонструють важливість належного налаштування параметрів, контролю аутлаєрів і вибору оптимальних значень lr та epoch для досягнення найкращих результатів навчання моделі.

3. Make one forward and backward steps for $L = (2a + b)(c - d)$, where
 a, b, c, d are arbitrary numbers

$L = (2a + b)(c - d)$

If $a = 2, b = 3, c = -1, d = -2$, then:

1. Forward step:

$$L = (2 \cdot 2 + 3)(-1 - (-2)) = (4 + 3)(-1 + 2) = 7 \cdot 1 = 7.$$

2. Backward step:

$$\frac{dL}{da} = 2 \cdot (c - d) = 2; \quad \frac{dL}{dc} = 2a + b = 7$$

$$\frac{dL}{db} = c - d = 1; \quad \frac{dL}{dd} = -(2a + b) = -7.$$