OneNote for Windows 10

Home Insert Draw View Help

Shapes Ink to Shape Ink to Text

100 Days of ML

Day 33 - Working-...
Day 34 - Working...
Day 35 - Complete...
Day 36 - Handling...
Day 37 - Handling...
Day38-Missing Indi...
Day39 - KNN Impu...
Day40 - Iterative I...
Day 41 - Outliers in...
Day 42 - Outlier De...
Day 43 - Outlier de...
Day44 - Outlier De...
Day 45 - Feature C...
Day 46 - Curse of...
Day 47 - PCA
Day 48 - Simple Li...
Day 49 - Regressio...
Day 50 - Multiple L...
Day 51 - Gradient...

Add section

What is Gradient Descent?
The Plan
Intuition
Mathematical Formulation
Example
Code from Scratch
Visualization 1
Adding m into the mix
Code
Visualization 2
Effect of Learning Data
Effect of Loss Function
Effect of Data

Add page

What is Gradient Descent?

Thursday, May 20, 2021 1:46 PM

Linear Reg

Logistic Reg

Ishu

Deep Learning

OneNote for Windows 10

Home Insert Draw View Help

Shapes Ink to Shape Ink to Text

100 Days of ML

Day 33 - Working-...
Day 34 - Working...
Day 35 - Complete...
Day 36 - Handling...
Day 37 - Handling...
Day38-Missing Indi...
Day39 - KNN Impu...
Day40 - Iterative I...
Day 41 - Outliers in...
Day 42 - Outlier De...
Day 43 - Outlier de...
Day44 - Outlier De...
Day 45 - Feature C...
Day 46 - Curse of...
Day 47 - PCA
Day 48 - Simple Li...
Day 49 - Regressio...
Day 50 - Multiple L...
Day 51 - Gradient...

Add section

What is Gradient Descent?
The Plan
Intuition
Mathematical Formulation
Example
Code from Scratch
Visualization 1
Adding m into the mix
Code
Visualization 2
Effect of Learning Data
Effect of Loss Function
Effect of Data

Add page

The Plan

Thursday, May 20, 2021 1:46 PM

5 videos

1) Gradient Descent

2
3
4

Batch GD
SGD → SGDRegressor
MBGD

5

OneNote for Windows 10

Home Insert Draw View Help

Shapes Ink to Shape Ink to Text

100 Days of ML

Day 33 - Working-...
Day 34 - Working...
Day 35 - Complete...
Day 36 - Handling...
Day 37 - Handling...
Day38-Missing Indi...
Day39 - KNN Impu...
Day40 - Iterative I...
Day 41 - Outliers in...
Day 42 - Outlier De...
Day 43 - Outlier de...
Day44 - Outlier De...
Day 45 - Feature C...
Day 46 - Curse of...
Day 47 - PCA
Day 48 - Simple Li...
Day 49 - Regressio...
Day 50 - Multiple L...
Day 51 - Gradient...

Add section

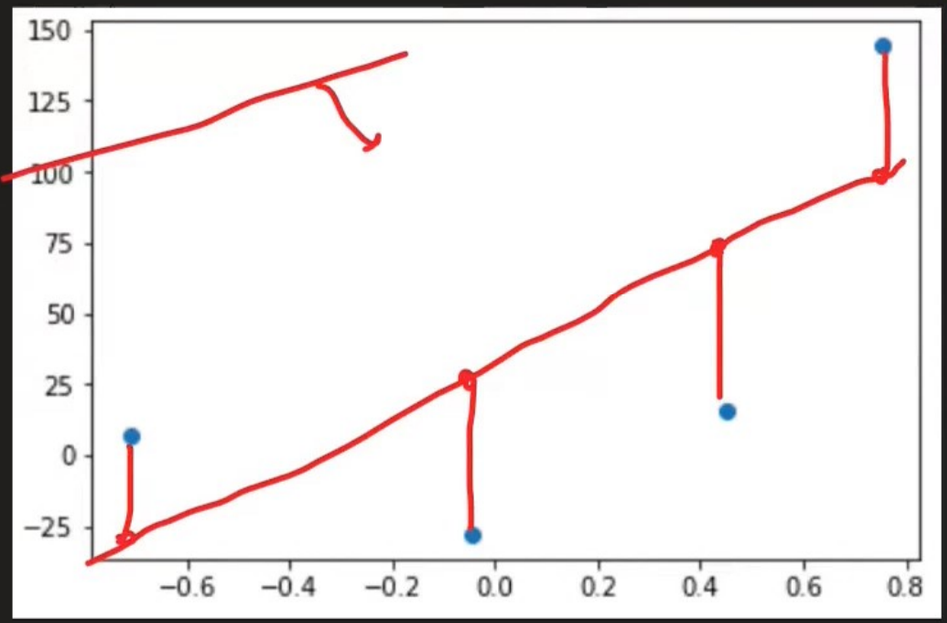
What is Gradient Descent?
The Plan
Intuition
Mathematical Formulation
Example
Code from Scratch
Visualization 1
Adding m into the mix
Code
Visualization 2
Effect of Learning Data
Effect of Loss Function
Effect of Data

Add page

Intuition

Thursday, May 20, 2021 1:46 PM

2 cols
4 rows



$L \rightarrow b^2$

$L(b)$ (b)

cgpa | lpa

OLS →

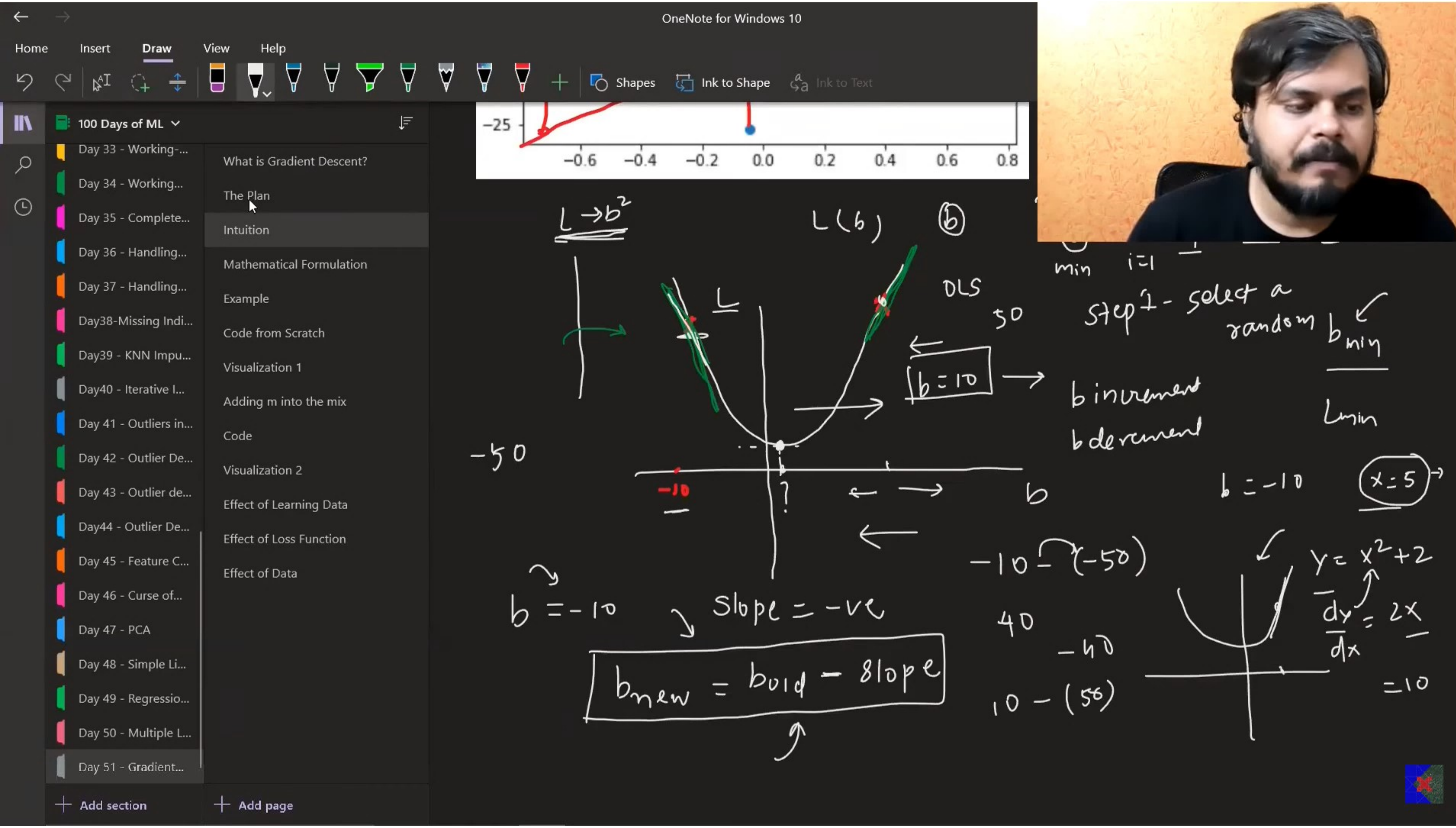
$\hat{y}_i = mx_i + b$

$m = 78.35$

$L = \sum_{i=1}^{4} (y_i - \hat{y}_i)^2$

$L = \sum_{i=1}^{4} (y_i - mx_i - b)^2$

$\min_{b} L = \sum_{i=1}^{4} (y_i - 78.35 * x_i - b)^2$



OneNote for Windows 10
Home
Insert
Draw
View
Help
Shapes
Ink to Shape
Ink to Text
100 Days of ML
Day 33 - Working-...
Day 34 - Working...
Day 35 - Complete...
Day 36 - Handling...
Day 37 - Handling...
Day38-Missing Indi...
Day39 - KNN Impu...
Day40 - Iterative I...
Day 41 - Outliers in...
Day 42 - Outlier De...
Day 43 - Outlier de...
Day44 - Outlier De...
Day 45 - Feature C...
Day 46 - Curse of...
Day 47 - PCA
Day 48 - Simple Li...
Day 49 - Regressio...
Day 50 - Multiple L...
Day 51 - Gradient...
Add section
What is Gradient Descent?
The Plan
Intuition
Mathematical Formulation
Example
Code from Scratch
Visualization 1
Adding m into the mix
Code
Visualization 2
Effect of Learning Data
Effect of Loss Function
Effect of Data
Add page
−25
−0.6
−0.4
−0.2
0.0
0.2
0.4
0.6
0.8
L → b²
L(b)
OLS
step 1 - select a random b
b = 10
b increment
b decrement
b = −10
Slope = −ve
b_new = b_old − slope
−10 − (−50)
40
10 − (50)
y = x² + 2
dy/dx = 2x
x = 5

OneNote for Windows 10

Home Insert Draw View Help

Shapes Ink to Shape Ink to Text

100 Days of ML

- Day 33 - Working-...
- Day 34 - Working...
- Day 35 - Complete...
- Day 36 - Handling...
- Day 37 - Handling...
- Day38-Missing Indi...
- Day39 - KNN Impu...
- Day40 - Iterative I...
- Day 41 - Outliers in...
- Day 42 - Outlier De...
- Day 43 - Outlier de...
- Day44 - Outlier De...
- Day 45 - Feature C...
- Day 46 - Curse of...
- Day 47 - PCA
- Day 48 - Simple Li...
- Day 49 - Regressio...
- Day 50 - Multiple L...
- Day 51 - Gradient...

Add section

- What is Gradient Descent?
- The Plan
- Intuition
- Mathematical Formulation
- Example
- Code from Scratch
- Visualization 1
- Adding m into the mix
- Code
- Visualization 2
- Effect of Learning Data
- Effect of Loss Function
- Effect of Data

Add page

b = -10 → Slope = -ve

$b_{new} = b_{old} - slope$

$b_{new} = b_{old} - \eta \, slope$

1

$b_{new} = -9.5 - (0.01 \times -40) = -9.5 + 0.4 = -9.1$

When to stop

1) diff b_{old} b_{new} => 0.0001

2) Iteration => 1000, 100, ...

epochs

0.01 0.1, 0.001

$b_{new} = -10 + (0.01 \times 50)$

$= -10 + 0.5 = -9.5$

$b_{new} - b_{old} = 0.0001$

$b_{new} - b_{old} = 0$

OneNote for Windows 10

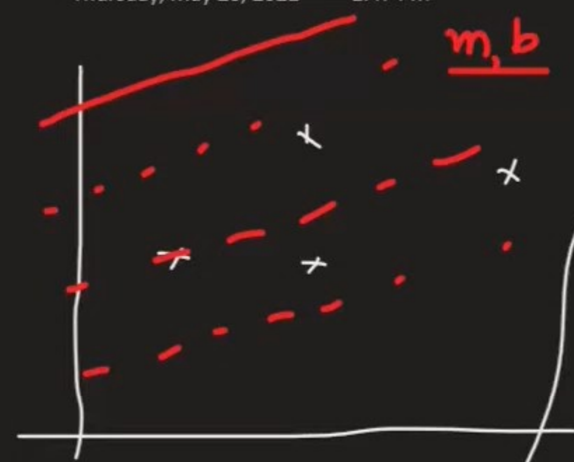
Home | Insert | Draw | View | Help

Shapes | Ink to Shape | Ink to Text

Mathematical Formulation

Thursday, May 20, 2021 1:47 PM

m, b



m = 78.35

Step → Start with a random b = b

i = 0 n, r,

1000, 100

$\eta = 0.01$

b = 0

for i in epochs:

$$b_{new} = b_{old} - \eta * slope \quad (b = 0)$$

$$L = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

$$\frac{dL}{db} = \frac{d}{db}\left(\sum_{i=1}^{n} (y_i - \hat{y}_i)^2\right) = 2\sum_{i=1}^{n} (y_i - mx_i - b)(-1)$$

b = 0

$$\frac{d}{db}\sum_{i=1}^{n} (y_i - mx_i - b)^2 \qquad -1$$

$$slope = -2\sum_{i=1}^{n} (y_i - mx_i - b)$$

$$= -2\sum_{i=1}^{n} (y_i - mx_i - 0)$$

OneNote for Windows 10

Home Insert Draw View Help

Shapes Ink to Shape Ink to Text

100 Days of ML

- Day 33 - Working-...
- Day 34 - Working...
- Day 35 - Complete...
- Day 36 - Handling...
- Day 37 - Handling...
- Day38-Missing Indi...
- Day39 - KNN Impu...
- Day40 - Iterative I...
- Day 41 - Outliers in...
- Day 42 - Outlier De...
- Day 43 - Outlier de...
- Day44 - Outlier De...
- Day 45 - Feature C...
- Day 46 - Curse of...
- Day 47 - PCA
- Day 48 - Simple Li...
- Day 49 - Regressio...
- Day 50 - Multiple L...
- Day 51 - Gradient...

Add section

- What is Gradient Descent?
- The Plan
- Intuition
- Mathematical Formulation
- Example
- Code from Scratch
- Visualization 1
- Adding m into the mix
- Code
- Visualization 2
- Effect of Learning Data
- Effect of Loss Function
- Effect of Data

Add page

Step → Start with

for i in epochs (10)

$$L = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

$$\frac{dL}{db} = \frac{d}{db}\left(\sum_{i=1}^{n} (y_i - \hat{y}_i)^2\right) = 2\sum_{i=1}^{n} (y_i - mx_i - b)(-1)$$

$$\frac{d}{db}\sum_{i=1}^{n} (y_i - mx_i - b)^2 \quad \text{Slope} = -2\sum_{i=1}^{n} (y_i - mx_i - b)$$

$$= -2\sum_{i=1}^{n} (y_i - 78.35 \times x_i - 0)$$

$b = 0$

$i = 1$

epoch

$$b_{new} = b_{old} - \eta \cdot \text{slope}_{b=b_{old}}$$

slope (b=0)

b_{new}

$b_{new} = b_{old} - \eta \times \text{slope}\ (b=0)$

OneNote for Windows 10

Home Insert Draw View Help

Shapes Ink to Shape Ink to Text

100 Days of ML

- Day 33 - Working-...
- Day 34 - Working...
- Day 35 - Complete...
- Day 36 - Handling...
- Day 37 - Handling...
- Day38-Missing Indi...
- Day39 - KNN Impu...
- Day40 - Iterative I...
- Day 41 - Outliers in...
- Day 42 - Outlier De...
- Day 43 - Outlier de...
- Day44 - Outlier De...
- Day 45 - Feature C...
- Day 46 - Curse of...
- Day 47 - PCA
- Day 48 - Simple Li...
- Day 49 - Regressio...
- Day 50 - Multiple L...
- Day 51 - Gradient...

Add section

- What is Gradient Descent?
- The Plan
- Intuition
- Mathematical Formulation
- Example
- Code from Scratch
- Visualization 1
- Few Discussions
- Adding m into the mix
- Code
- Visualization 2
- Effect of Learning Data
- Effect of Loss Function
- Effect of Data

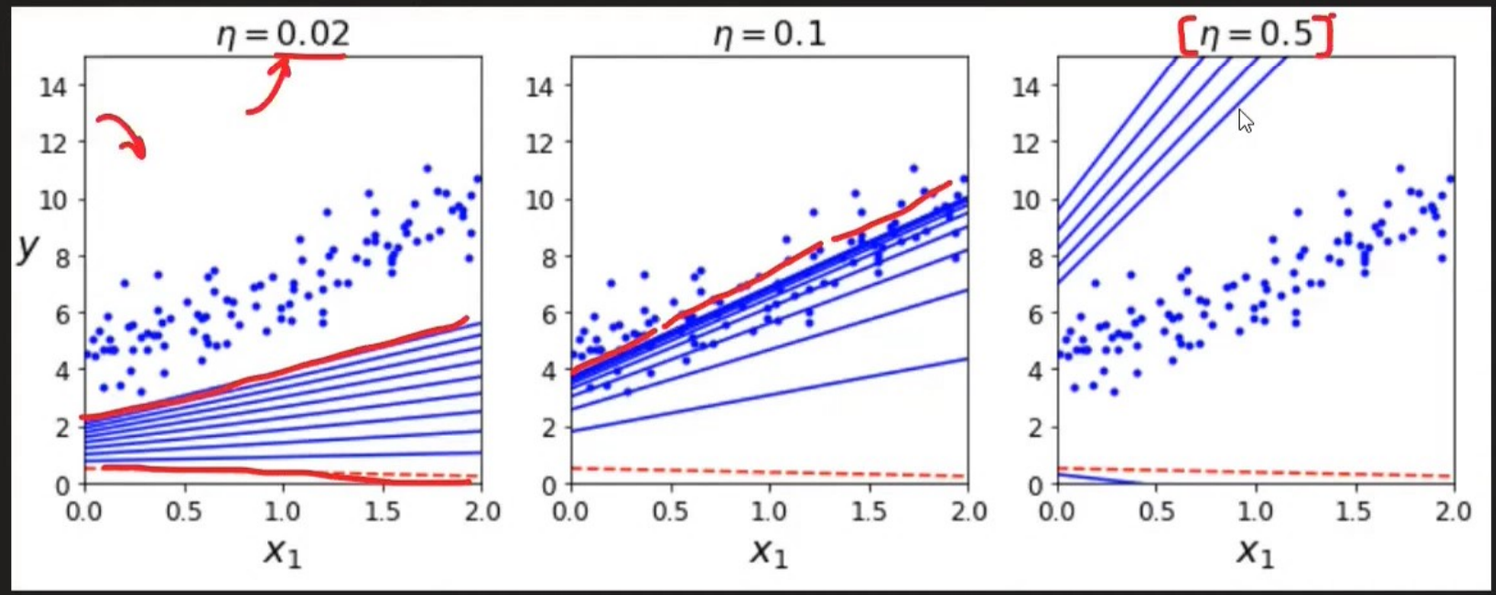
Add page

Few Discussions

Thursday, May 20, 2021 4:47 PM

1. Effect of Learning rate
2. The universality of Gradient Descent

epoch = 10



OneNote for Windows 10

Home Insert Draw View Help

Shapes Ink to Shape Ink to Text

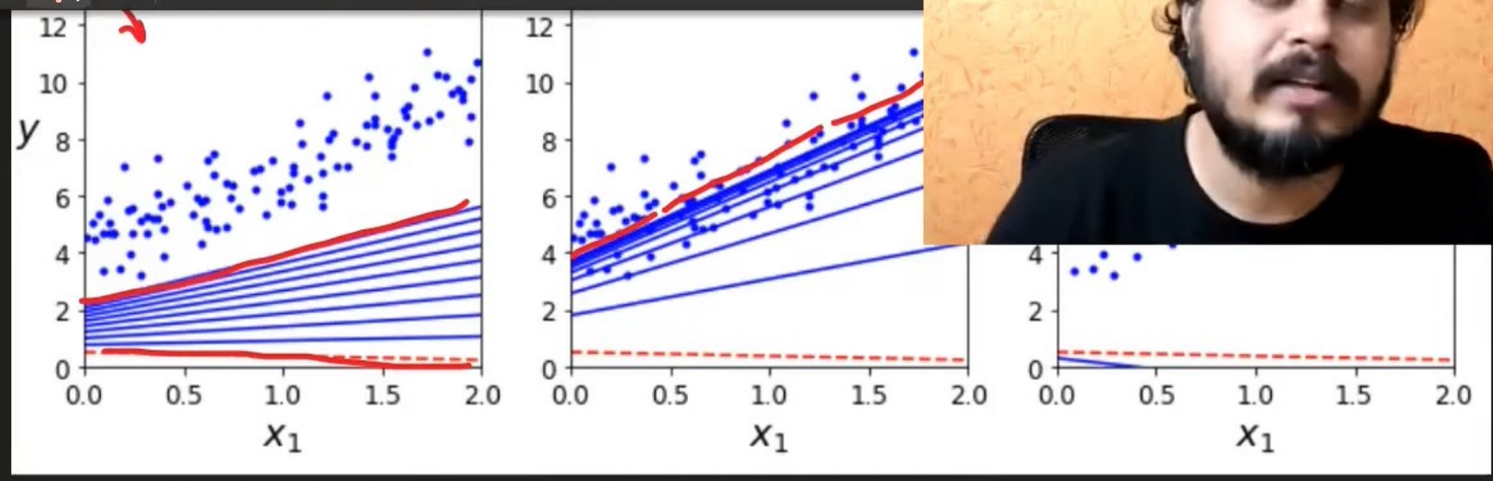
100 Days of ML

- Day 33 - Working-...
- Day 34 - Working...
- Day 35 - Complete...
- Day 36 - Handling...
- Day 37 - Handling...
- Day38-Missing Indi...
- Day39 - KNN Impu...
- Day40 - Iterative I...
- Day 41 - Outliers in...
- Day 42 - Outlier De...
- Day 43 - Outlier de...
- Day44 - Outlier De...
- Day 45 - Feature C...
- Day 46 - Curse of...
- Day 47 - PCA
- Day 48 - Simple Li...
- Day 49 - Regressio...
- Day 50 - Multiple L...
- Day 51 - Gradient...

Add section

- What is Gradient Descent?
- The Plan
- Intuition
- Mathematical Formulation
- Example
- Code from Scratch
- Visualization 1
- Few Discussions
- Adding m into the mix
- Code
- Visualization 2
- Effect of Learning Data
- Effect of Loss Function
- Effect of Data

Add page



$b = 0$

$b = b_{old} - \eta \cdot \text{slope}$

$\frac{dL}{db} = [\quad]$

$\sum (y_i - \hat{y}_i)^2$ (LR)

Deep →

LOR → [function

OneNote for Windows 10

Home Insert Draw View Help

Shapes Ink to Shape Ink to Text

100 Days of ML

Day 33 - Working-...
Day 34 - Working...
Day 35 - Complete...
Day 36 - Handling...
Day 37 - Handling...
Day38-Missing Indi...
Day39 - KNN Impu...
Day40 - Iterative I...
Day 41 - Outliers in...
Day 42 - Outlier De...
Day 43 - Outlier de...
Day44 - Outlier De...
Day 45 - Feature C...
Day 46 - Curse of...
Day 47 - PCA
Day 48 - Simple Li...
Day 49 - Regressio...
Day 50 - Multiple L...
Day 51 - Gradient...

Add section

What is Gradient Descent?
The Plan
Intuition
Mathematical Formulation
Example
Code from Scratch
Visualization 1
Few Discussions
Adding m into the mix
Code
Visualization 2
Effect of Learning Data
Effect of Loss Function
Effect of Data

Add page

Adding m into the mix

Thursday, May 20, 2021 1:48 PM

Steps

1) init random vals for m and b
$m = 1$ and $b = 0$

2) epochs = 100, lr = 0.01
for i in epochs:
$b = b - \eta \text{ slope}$
$m = m - \eta \text{ slope}$

OneNote for Windows 10

Home Insert Draw View Help

Shapes Ink to Shape Ink to Text

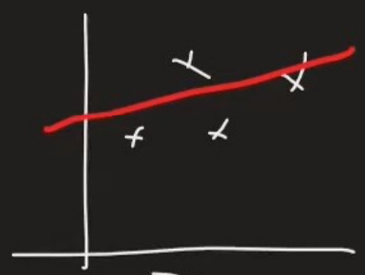
100 Days of ML

Day 33 - Working-...
Day 34 - Working...
Day 35 - Complete...
Day 36 - Handling...
Day 37 - Handling...
Day38-Missing Indi...
Day39 - KNN Impu...
Day40 - Iterative I...
Day 41 - Outliers in...
Day 42 - Outlier De...
Day 43 - Outlier de...
Day44 - Outlier De...
Day 45 - Feature C...
Day 46 - Curse of...
Day 47 - PCA
Day 48 - Simple Li...
Day 49 - Regressio...
Day 50 - Multiple L...
Day 51 - Gradient...

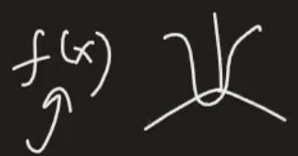
Add section

What is Gradient Descent?
The Plan
Intuition
Mathematical Formulation
Example
Code from Scratch
Visualization 1
Few Discussions
Adding m into the mix
Code
Visualization 2
Effect of Learning Data
Effect of Loss Function
Effect of Data

Add page



$z = f(x, y)$



$m = 1$ and b [illegible]

2) epochs = 100, lr = 0.01

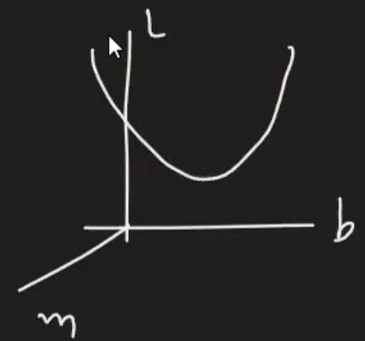
for i in epochs:

$b = b - \eta \cdot \text{slope}$

$m = m - \eta \cdot \text{slope}$

$L = \sum (y_i - \hat{y}_i)^2 = \sum (y_i - mx_i - b)^2$

$L(m, b)$



OneNote for Windows 10

Home Insert Draw View Help

Shapes Ink to Shape Ink to Text

100 Days of ML

- Day 33 - Working-...
- Day 34 - Working...
- Day 35 - Complete...
- Day 36 - Handling...
- Day 37 - Handling...
- Day38-Missing Indi...
- Day39 - KNN Impu...
- Day40 - Iterative I...
- Day 41 - Outliers in...
- Day 42 - Outlier De...
- Day 43 - Outlier de...
- Day44 - Outlier De...
- Day 45 - Feature C...
- Day 46 - Curse of...
- Day 47 - PCA
- Day 48 - Simple Li...
- Day 49 - Regressio...
- Day 50 - Multiple L...
- Day 51 - Gradient...

Add section

- What is Gradient Descent?
- The Plan
- Intuition
- Mathematical Formulation
- Example
- Code from Scratch
- Visualization 1
- Few Discussions
- Adding m into the mix
- Code
- Visualization 2
- Effect of Learning Data
- Effect of Loss Function
- Effect of Data

Add page

$m = 1$ and b ...

2) epochs = 100, lr = 0.01

for i in epochs:

$b = b - \eta$ slope

$m = m - \eta$ slope

$z = f(x, y)$

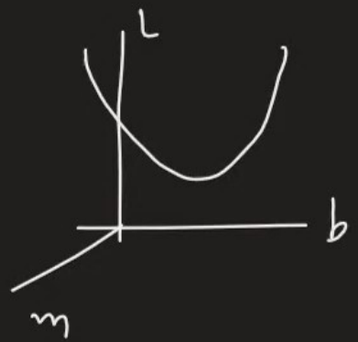
$L = \sum (y_i - \hat{y}_i)^2 = \sum (y_i - mx_i - b)^2$

$L(m, b)$

$\frac{dL}{db}$

b_slope = $\frac{\partial L}{\partial b}$

∂



m_slope = $\frac{\partial L}{\partial m}$

f(x) y

OneNote for Windows 10

Home Insert Draw View Help

Shapes Ink to Shape Ink to Text

100 Days of ML

- Day 33 - Working-...
- Day 34 - Working...
- Day 35 - Complete...
- Day 36 - Handling...
- Day 37 - Handling...
- Day38-Missing Indi...
- Day39 - KNN Impu...
- Day40 - Iterative I...
- Day 41 - Outliers in...
- Day 42 - Outlier De...
- Day 43 - Outlier de...
- Day44 - Outlier De...
- Day 45 - Feature C...
- Day 46 - Curse of...
- Day 47 - PCA
- Day 48 - Simple Li...
- Day 49 - Regressio...
- Day 50 - Multiple L...
- Day 51 - Gradient...

Add section

- What is Gradient Descent?
- The Plan
- Intuition
- Mathematical Formulation
- Example
- Code from Scratch
- Visualization 1
- Few Discussions
- Adding m into the mix
- Code
- Visualization 2
- Effect of Learning Data
- Effect of Loss Function
- Effect of Data

Add page

$z = f(x, y)$

$f(x)$

$b = 0$

$b\_slope = \frac{\partial L}{\partial b}$

$m\_slope = \frac{\partial L}{\partial m}$

$\sum (y_i - mx_i - b)^2$

$\frac{\partial L}{\partial b} = -2 \sum (y_i - mx_i - b)$

$= -2 \sum (y_i - mx_i - b)$

$= slope\_b \text{ at } b = 0$

$\frac{\partial L}{\partial m} = 2 \sum (y_i - mx_i - b)$

$= -2 \sum (y_i - mx_i - b) x_i$

$slope\_m \text{ at } m = 1$

OneNote for Windows 10

Home Insert Draw View Help

Shapes Ink to Shape Ink to Text

100 Days of ML

- Day 33 - Working-...
- Day 34 - Working...
- Day 35 - Complete...
- Day 36 - Handling...
- Day 37 - Handling...
- Day38-Missing Indi...
- Day39 - KNN Impu...
- Day40 - Iterative I...
- Day 41 - Outliers in...
- Day 42 - Outlier De...
- Day 43 - Outlier de...
- Day44 - Outlier De...
- Day 45 - Feature C...
- Day 46 - Curse of...
- Day 47 - PCA
- Day 48 - Simple Li...
- Day 49 - Regressio...
- Day 50 - Multiple L...
- Day 51 - Gradient...

Add section

- What is Gradient Descent?
- The Plan
- Intuition
- Mathematical Formulation
- Example
- Code from Scratch
- Visualization 1
- Few Discussions
- Adding m into the mix
- Code
- Visualization 2
- Effect of Learning Data
- Effect of Loss Function
- Effect of Data

Add page

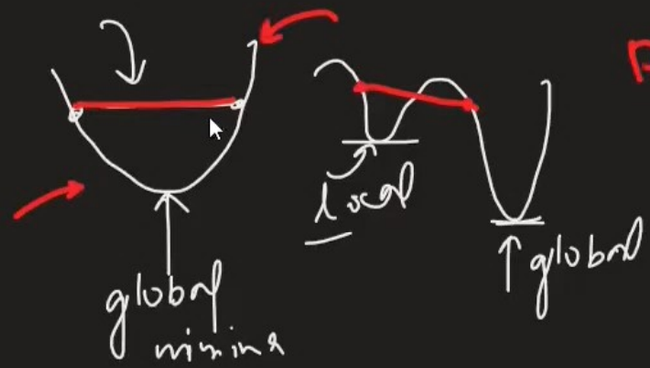
Effect of Loss Function

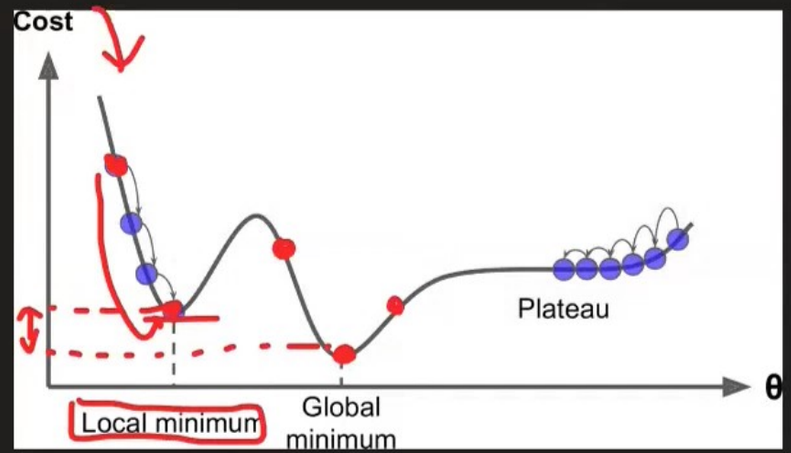
Thursday, May 20, 2021 1:53 PM

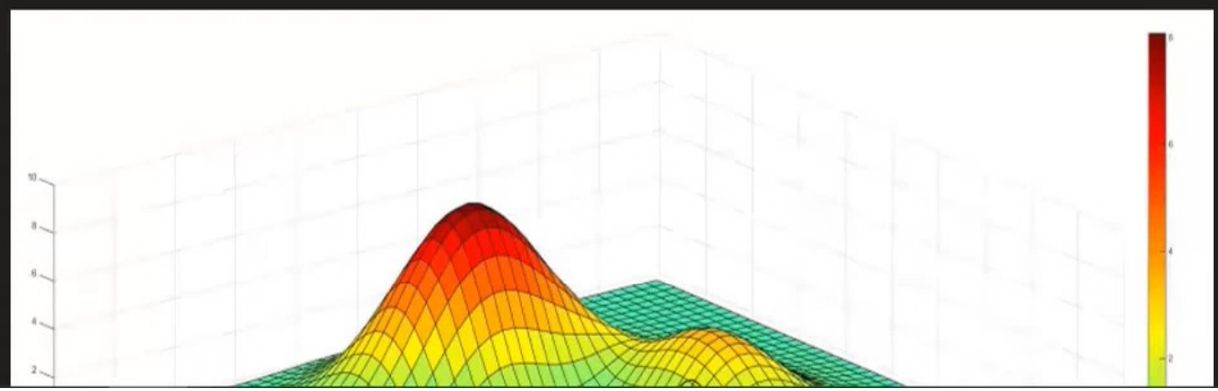
$L = \sum (y_i - \hat{y}_i)^2$ → convex function

10 cols

deep learn







OneNote for Windows 10

Home Insert Draw View Help

Shapes Ink to Shape Ink to Text

100 Days of ML

- Day 33 - Working-...
- Day 34 - Working...
- Day 35 - Complete...
- Day 36 - Handling...
- Day 37 - Handling...
- Day38-Missing Indi...
- Day39 - KNN Impu...
- Day40 - Iterative I...
- Day 41 - Outliers in...
- Day 42 - Outlier De...
- Day 43 - Outlier de...
- Day44 - Outlier De...
- Day 45 - Feature C...
- Day 46 - Curse of...
- Day 47 - PCA
- Day 48 - Simple Li...
- Day 49 - Regressio...
- Day 50 - Multiple L...
- Day 51 - Gradient...

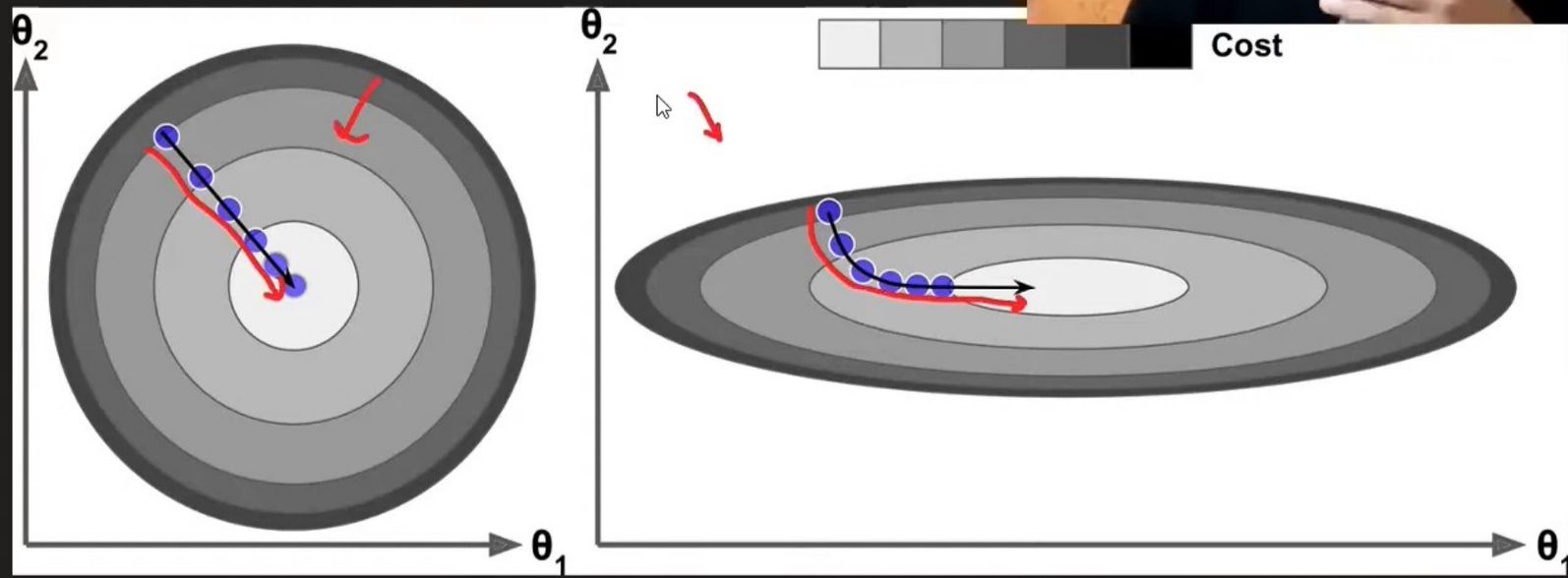
Add section

- What is Gradient Descent?
- The Plan
- Intuition
- Mathematical Formulation
- Example
- Code from Scratch
- Visualization 1
- Few Discussions
- Adding m into the mix
- Code
- Visualization 2
- Effect of Learning Data
- Effect of Loss Function
- Effect of Data

Add page

Effect of Data

Thursday, May 20, 2021 1:53 PM



OneNote for Windows 10

Home Insert Draw View Help

Shapes Ink to Shape Ink to Text

100 Days of ML

- Day 34 - Working...
- Day 35 - Complete...
- Day 36 - Handling...
- Day 37 - Handling...
- Day38-Missing Indi...
- Day39 - KNN Impu...
- Day40 - Iterative I...
- Day 41 - Outliers in...
- Day 42 - Outlier De...
- Day 43 - Outlier de...
- Day44 - Outlier De...
- Day 45 - Feature C...
- Day 46 - Curse of...
- Day 47 - PCA
- Day 48 - Simple Li...
- Day 49 - Regressio...
- Day 50 - Multiple L...
- Day 51 - Gradient...
- Day 52 - Types of...

Add section Add page

Types of Gradient Descent

Types of Gradient Descent

Saturday, May 22, 2021 1:30 PM