✓ Linear Regr

— Computational Co

— Gradient De

ession

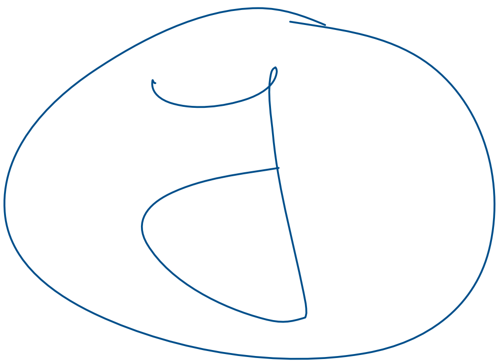
mplexity

escent

Age County Employ

Features

Income



model paran

$$\hat{y} = \theta_0 + \theta_1 x_1 + \ldots$$

pred

meter

$$\cdot + \underline{\theta_n} x_n$$

$\downarrow$

features

vectorised form

$\hat{y} = h_\theta(x) = \theta$

hypothesis func

feature
vector

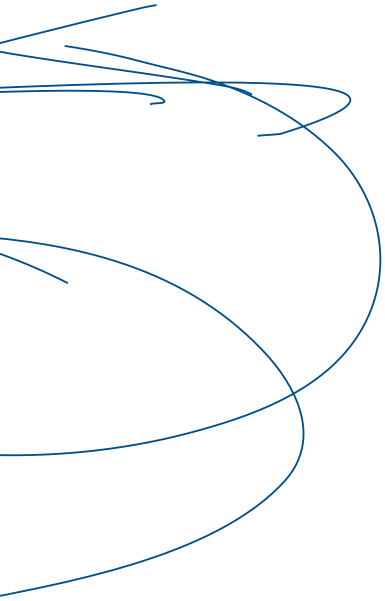
$\theta \cdot x$

tion

Friday, September 1, 2023 9:52 PM

$$y = x^3 + 4x^2 + \dots$$

$$\hat{y} = \theta_0 + \theta_1 x + \dots$$



RMSE $\longrightarrow$

$MSE(x, h_\theta) = -$

cost fun

Θ

ction

Friday, September 1, 2023 9:58 PM

Normal Equat

$$\hat{\theta} = (X^T X)$$

cost function optim

ion

$x^T y$

ize

Computational Comp

$$\underline{X^T \cdot X}$$

$$\left[x_1, x_2 \right.$$

$$n^2 \qquad \underline{\underline{(n+1)(}}$$

lexity

$$x_n)\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$n+1)$$

$O(n^2)$

$\boxed{\text{SVD}} \rightarrow 2$

$O(m)$

$O(n^3)$

OK

$4 \rightarrow O(n^2)$

8

$\cancel{16} \rightarrow O(n^{2.4}$

Gradient Desc

- minimize co

↓

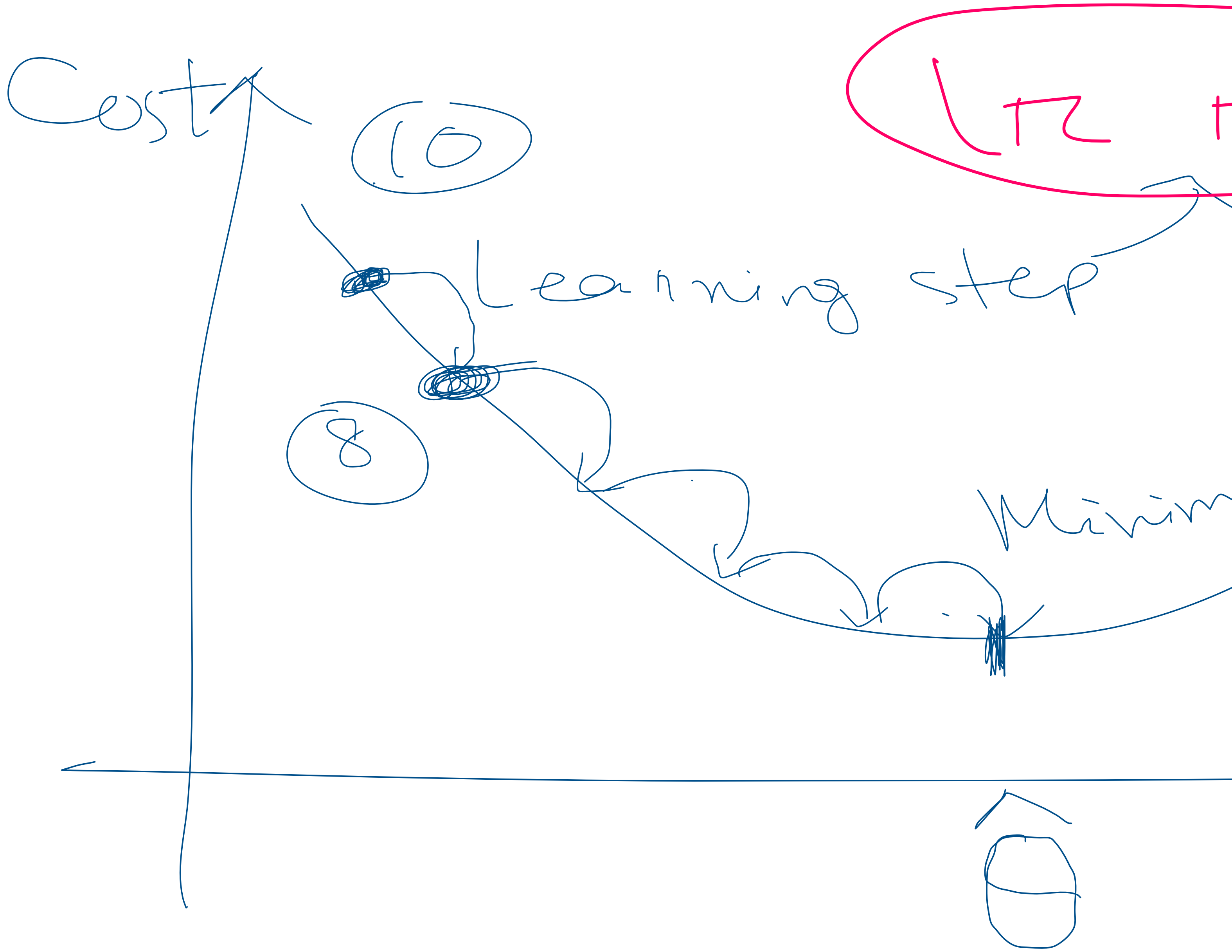
Parameter

ent

st func

S

Friday, September 1, 2023 10:30 PM



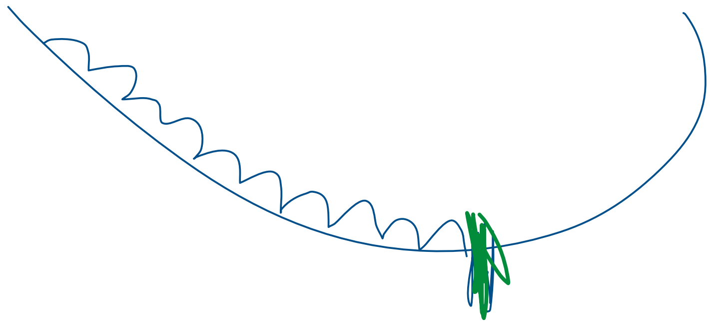
rate

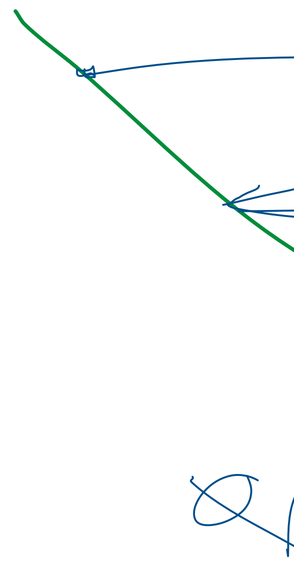
en

Friday, September 1, 2023 10:33 PM

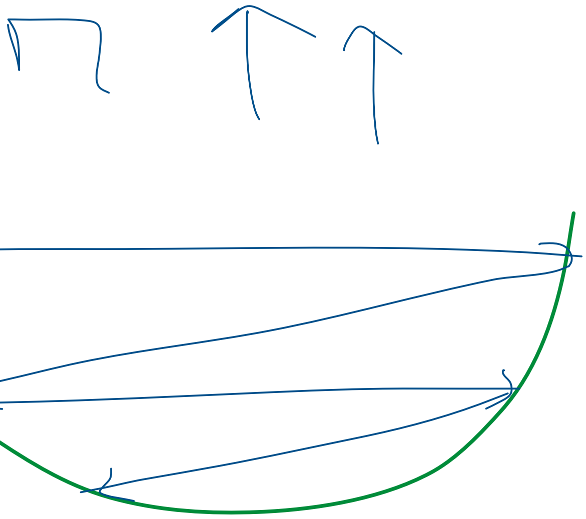
Lr ↓ ↓

L



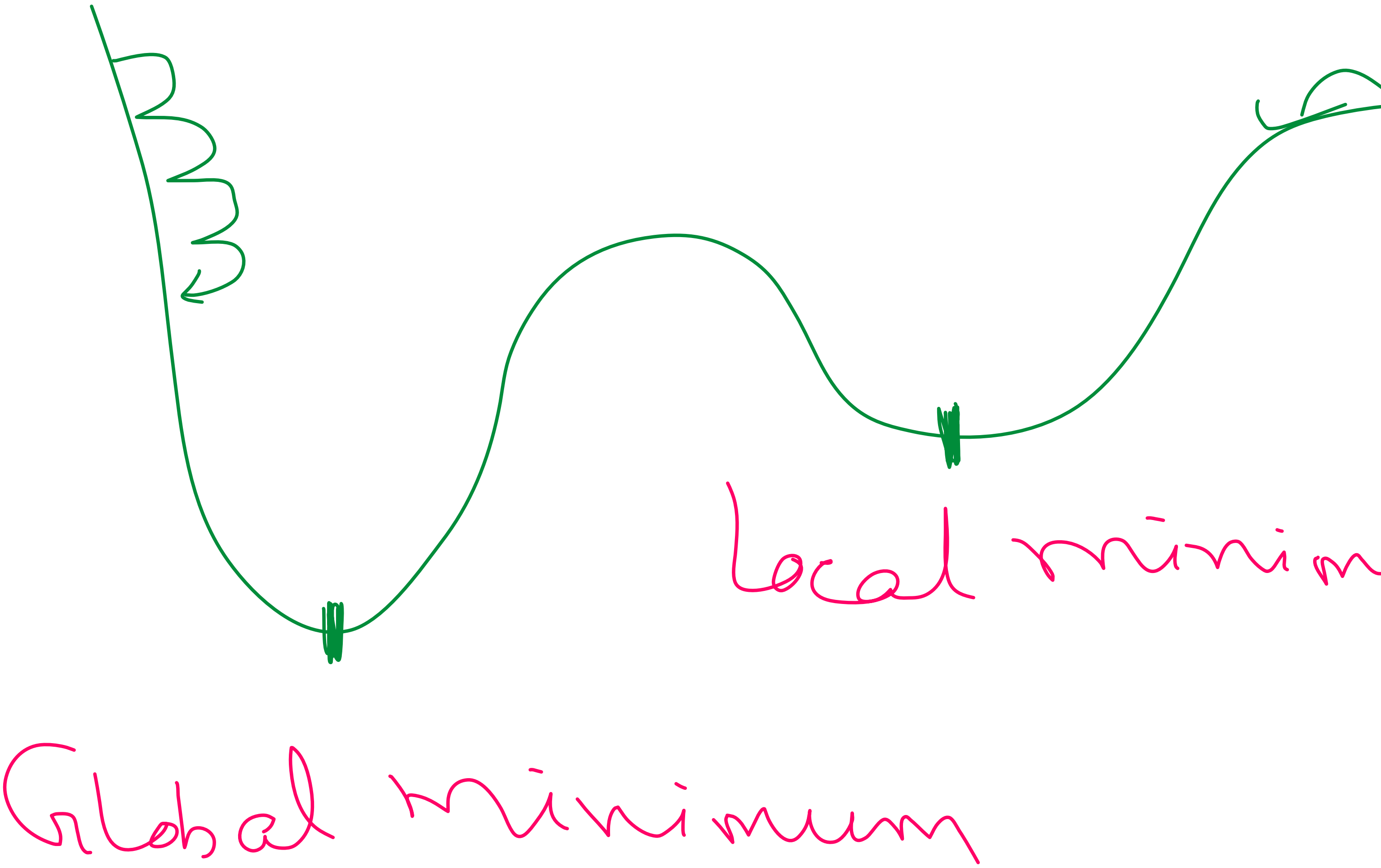


Computational time ↑↑



~~ethical~~

~~soln~~


Local minim
Global minimum

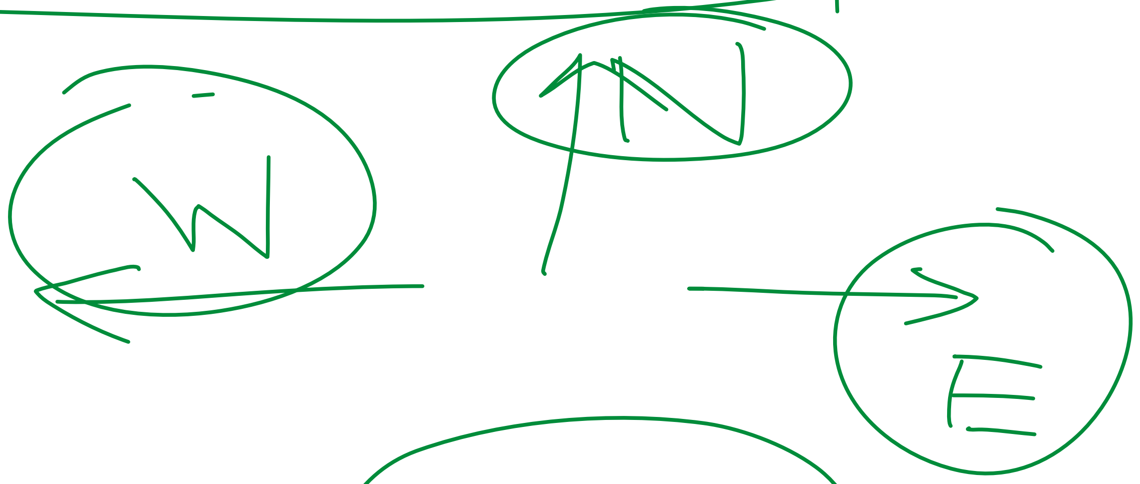
convex → globa

mini

el

inimum

Batch GD /



Partial der

Θ

ivative

$$MSE(\theta) = \begin{pmatrix} \frac{\partial}{\partial \ldots} \\ \frac{\partial}{\partial \theta} \\ \vdots \end{pmatrix}$$

- Smoothen win

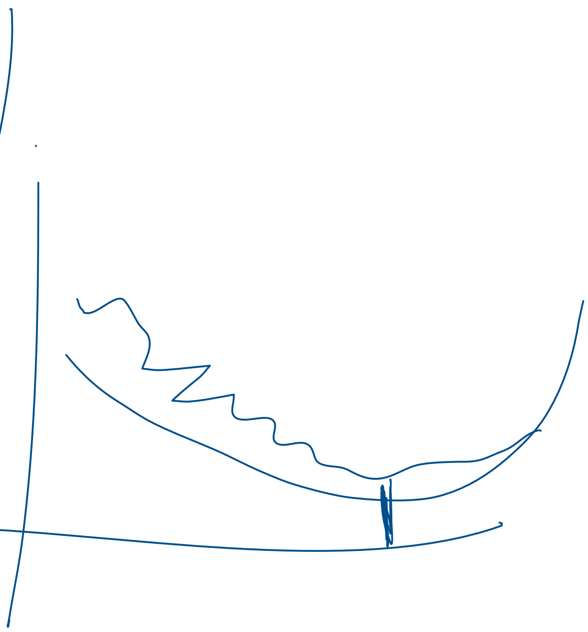
θ_0 MSE

T_1 MSF

ve

Stochastic GD

- an example
- feed it NN
- calc gradie
- update wei



ent

ghts

→ never reache[s]
minima.

→ vectorized impl[ementation]

→ larger data

→ faster converg[ence]

s

entation x

se t

ge

Mini_Batch GD

→ fixed

- a mini-batch p
- Feed
- Mean of gradien
- update