Intro
○○

Linear Regression
○○○○○

Question 1 (a → c)
○○○○○○○○○○

Multiple Linear Regression
○○○○○○

Question 2 (a → h)
○○○○○○○○○○○○

Regression I

COMP9417, 23T2

Intro
○○

Linear Regression
○○○○○

Question 1 (a → c)
○○○○○○○○○○

Multiple Linear Regression
○○○○○○

Question 2 (a → h)
○○○○○○○○○○○

Intro
●○

Linear Regression
○○○○○

Question 1 (a → c)
○○○○○○○○○○

Multiple Linear Regression
○○○○○○

Question 2 (a → h)
○○○○○○○○○○○○

Section 1

Intro

Intro

Who am I?

Intro

Linear Regression
00000

Question 1 (a → c)
0000000000

Multiple Linear Regression
000000

Question 2 (a → h)
00000000000

Intro

Who am I? Who are you?

Intro

Who am I? Who are you?

What you'll get from this course:

- Understand the basis of machine learning
- ML algorithms and the math behind them
- Ability to implement these ideas in Python

How to do well:

- Fully understand tut questions from week to week (they pile up)
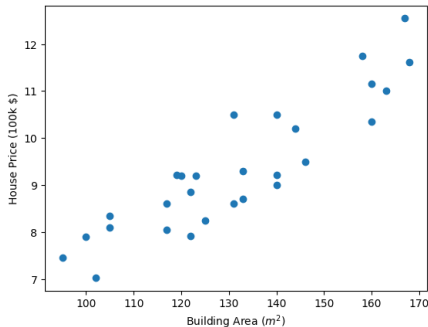- Don't be afraid of notation, break it all down
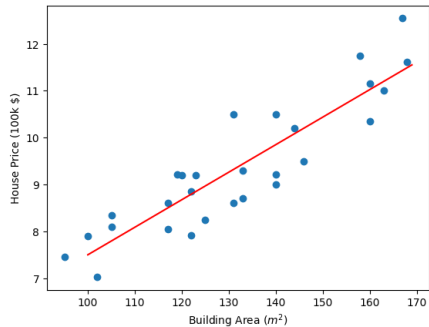- Keep researching

Intro
○○

Linear Regression
●○○○○

Question 1 (a → c)
○○○○○○○○○○

Multiple Linear Regression
○○○○○○

Question 2 (a → h)
○○○○○○○○○○○○

Section 2

Linear Regression

Intro
oo

Linear Regression
o●ooo

Question 1 (a → c)
oooooooooooo

Multiple Linear Regression
oooooo

Question 2 (a → h)
oooooooooooo

Linear Regression

Say we're given a task to explain the relationship of the prices of homes based on their size in square meters.

Intro
○○

Linear Regression
○●○○○

Question 1 (a → c)
○○○○○○○○○○

Multiple Linear Regression
○○○○○○

Question 2 (a → h)
○○○○○○○○○○○

Linear Regression

Say we're given a task to explain the relationship of the prices of homes based on their size in square meters.
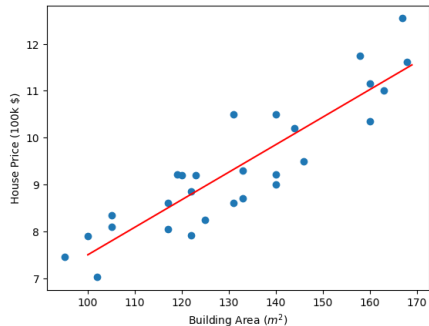
Intro
○○

Linear Regression
○○●○○

Question 1 (a → c)
○○○○○○○○○○

Multiple Linear Regression
○○○○○○

Question 2 (a → h)
○○○○○○○○○○○

Let's try fitting a line of best fit:

Intro
oo

Linear Regression
oo●oo

Question 1 (a → c)
ooooooooo

Multiple Linear Regression
oooooo

Question 2 (a → h)
ooooooooooo

Let's try fitting a line of best fit:

How do we know that this is the line of *best* fit?

Intro
○○

Linear Regression
○○○●○

Question 1 (a → c)
○○○○○○○○○○

Multiple Linear Regression
○○○○○○

Question 2 (a → h)
○○○○○○○○○○○

Let's define our error as

$$E = e_1 + e_2 + e_3 + \cdots + e_n$$
$$= \sum_{i=1}^{n} e_i$$

Intro
○○

Linear Regression
○○○●○

Question 1 (a → c)
○○○○○○○○○○

Multiple Linear Regression
○○○○○○

Question 2 (a → h)
○○○○○○○○○○○

Let's define our error as

$$E = e_1 + e_2 + e_3 + \cdots + e_n$$
$$= \sum_{i=1}^{n} e_i$$

We can generalise this to a function in nicer form:

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

Intro
oo

Linear Regression
ooo●o

Question 1 (a → c)
oooooooooo

Multiple Linear Regression
oooooo

Question 2 (a → h)
ooooooooooo

Let's define our error as

$$E = e_1 + e_2 + e_3 + \cdots + e_n$$
$$= \sum_{i=1}^{n} e_i$$

We can generalise this to a function in nicer form:

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

Something is wrong here.

Intro
○○

Linear Regression
○○○○●

Question 1 (a → c)
○○○○○○○○○○

Multiple Linear Regression
○○○○○○

Question 2 (a → h)
○○○○○○○○○○○

Formally, we define our error/loss function as:

$$L(\hat{y}) = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \qquad \text{a.k.a MSE}$$

$$L(w_0, w_1) = \frac{1}{n} \sum_{i=1}^{n} (y_i - w_0 - w_1 x_i)^2 \qquad \text{by definition}$$

The minimum of our loss function w.r.t w_0 and w_1 will be their optimal values respectively.

Intro
OO

Linear Regression
OOOOO

Question 1 (a → c)
●OOOOOOOOO

Multiple Linear Regression
OOOOOO

Question 2 (a → h)
OOOOOOOOOOOO

Section 3

Question 1 (a → c)

1a

Derive the least-squares estimates for the univariate linear regression model.

i.e Solve:

$$\underset{w_0, w_1}{\arg\min} \quad L(w_0, w_1)$$

$$\underset{w_0, w_1}{\arg\min} \quad \frac{1}{n}\sum_{i=1}^{n}(y_i - w_0 - w_1 x_i)^2$$

Intro
○○

Linear Regression
○○○○○

Question 1 (a → c)
○○●○○○○○○○

Multiple Linear Regression
○○○○○○

Question 2 (a → h)
○○○○○○○○○○○

First we differentiate $L(w_0, w_1)$ with respect to w_0,

Intro
○○

Linear Regression
○○○○○

Question 1 (a → c)
○○●○○○○○○○

Multiple Linear Regression
○○○○○○

Question 2 (a → h)
○○○○○○○○○○○

First we differentiate $L(w_0, w_1)$ with respect to w_0,

$$\frac{\partial L(w_0, w_1)}{\partial w_0} = -\frac{2}{n} \sum_{i=1}^{n} (y_i - w_0 - w_1 x_i)$$

$$= -\frac{2}{n} \left(\sum_{i=1}^{n} y_i - n w_0 - w_1 \sum_{i=1}^{n} x_i \right)$$

Intro
○○

Linear Regression
○○○○○

Question 1 (a → c)
○○●○○○○○○○

Multiple Linear Regression
○○○○○○

Question 2 (a → h)
○○○○○○○○○○○

First we differentiate $L(w_0, w_1)$ with respect to w_0,

$$\frac{\partial L(w_0, w_1)}{\partial w_0} = -\frac{2}{n} \sum_{i=1}^{n} (y_i - w_0 - w_1 x_i)$$

$$= -\frac{2}{n} \left(\sum_{i=1}^{n} y_i - n w_0 - w_1 \sum_{i=1}^{n} x_i \right)$$

For the minimum, $\frac{\partial L(w_0, w_1)}{\partial w_0} = 0$,

$$-\frac{2}{n} \left(\sum_{i=1}^{n} y_i - n w_0 - w_1 \sum_{i=1}^{n} x_i \right) = 0$$

Intro
○○

Linear Regression
○○○○○

Question 1 (a → c)
○○○●○○○○○○

Multiple Linear Regression
○○○○○○

Question 2 (a → h)
○○○○○○○○○○○

$$\frac{1}{n}\sum_{i=1}^{n} y_i - w_0 - w_1 \frac{1}{n}\sum_{i=1}^{n} x_i = 0$$
$$\bar{y} - w_0 - w_1 \bar{x} = 0 \tag{1}$$
$$w_0 = \bar{y} - w_1 \bar{x}$$

To find w_1, we follow a similar process and use simple simultaneous equations to solve for the final solution.

Intro
○○

Linear Regression
○○○○○

Question 1 (a → c)
○○○○●○○○○○

Multiple Linear Regression
○○○○○○

Question 2 (a → h)
○○○○○○○○○○○○

So,

Intro
oo

Linear Regression
ooooo

Question 1 (a → c)
ooooo●ooooo

Multiple Linear Regression
oooooo

Question 2 (a → h)
ooooooooooo

So,

$$\frac{\partial L(w_0, w_1)}{\partial w_1} = -\frac{2}{n} \sum_{i=1}^{n} x_i(y_i - w_0 - w_1 x_i)$$

$$= -\frac{2}{n} \left(\sum_{i=1}^{n} x_i y_i - w_0 \sum_{i=1}^{n} x_i - w_1 \sum_{i=1}^{n} x_i^2 \right)$$

So,

$$\frac{\partial L(w_0, w_1)}{\partial w_1} = -\frac{2}{n} \sum_{i=1}^{n} x_i(y_i - w_0 - w_1 x_i)$$

$$= -\frac{2}{n} \left(\sum_{i=1}^{n} x_i y_i - w_0 \sum_{i=1}^{n} x_i - w_1 \sum_{i=1}^{n} x_i^2 \right)$$

$\frac{\partial L(w_0, w_1)}{\partial w_1} = 0,$

$$\frac{1}{n} \left(\sum_{i=1}^{n} x_i y_i - w_0 \sum_{i=1}^{n} x_i - w_1 \sum_{i=1}^{n} x_i^2 \right) = 0$$

$$\overline{xy} - w_0 \bar{x} - w_1 \overline{x^2} = 0$$

Intro
○○

Linear Regression
○○○○○

Question 1 (a → c)
○○○○○●○○○○

Multiple Linear Regression
○○○○○○

Question 2 (a → h)
○○○○○○○○○○○

$$\overline{xy} - w_0\bar{x} - w_1\overline{x^2} = 0$$

$$w_1 = \frac{\overline{xy} - w_0\bar{x}}{\overline{x^2}} \tag{2}$$

Intro
○○

Linear Regression
○○○○○

Question 1 (a → c)
○○○○○●○○○○

Multiple Linear Regression
○○○○○○

Question 2 (a → h)
○○○○○○○○○○○

$$\overline{xy} - w_0\bar{x} - w_1\overline{x^2} = 0$$

$$w_1 = \frac{\overline{xy} - w_0\bar{x}}{\overline{x^2}} \qquad (2)$$

Sub (1) into (2):

Intro
○○

Linear Regression
○○○○○

Question 1 (a → c)
○○○○○●○○○○

Multiple Linear Regression
○○○○○○

Question 2 (a → h)
○○○○○○○○○○○

$$\overline{xy} - w_0\bar{x} - w_1\overline{x^2} = 0$$

$$w_1 = \frac{\overline{xy} - w_0\bar{x}}{\overline{x^2}} \tag{2}$$

Sub (1) into (2):

$$w_1 = \frac{\overline{xy} - (\bar{y} - w_1\bar{x})\bar{x}}{\overline{x^2}}$$

$$w_1 = \frac{\overline{xy} - \bar{x}\bar{y} + w_1\bar{x}^2}{\overline{x^2}}$$

$$w_1(\frac{\overline{x^2} - \bar{x}^2}{\bar{x}^2}) = \frac{\overline{xy} - \bar{x}\bar{y} + w_1\bar{x}^2}{\overline{x^2}}$$

$$w_1 = \frac{\overline{xy} - \bar{x}\bar{y}}{\overline{x^2} - \bar{x}^2}$$

Intro
○○

Linear Regression
○○○○○

Question 1 (a → c)
○○○○○○●○○○

Multiple Linear Regression
○○○○○○

Question 2 (a → h)
○○○○○○○○○○○

Finally, we have

$$w_1 = \frac{\overline{xy} - \bar{x}\bar{y}}{\overline{x^2} - \bar{x}^2} \text{ and } w_0 = \bar{y} - w_1\bar{x}$$

1b

Problem: Prove $(\bar{x}, \bar{y})$ is on the line.

From 1(a), the equation of our line ($\hat{y} = w_0 + w_1 x$) becomes:

$$\hat{y} = \bar{y} - \bar{x}\frac{\overline{xy} - \bar{x}\bar{y}}{\overline{x^2} - \bar{x}^2} + \frac{\overline{xy} - \bar{x}\bar{y}}{\overline{x^2} - \bar{x}^2}x$$

Sub $x = \bar{x}$,

$$\hat{y} = \bar{y} - \bar{x}\frac{\overline{xy} - \bar{x}\bar{y}}{\overline{x^2} - \bar{x}^2} + \frac{\overline{xy} - \bar{x}\bar{y}}{\overline{x^2} - \bar{x}^2}\bar{x}$$

$$\hat{y} = \bar{y} \qquad\qquad\qquad \therefore (\bar{x}, \bar{y}) \text{ is on the line}$$

Intro
○○

Linear Regression
○○○○○

Question 1 (a → c)
○○○○○○○○●○

Multiple Linear Regression
○○○○○○

Question 2 (a → h)
○○○○○○○○○○○

1c

Similar to 1a, though take care with the partial derivatives:

$$\frac{\partial L(w_0, w_1)}{\partial w_0} = -\frac{2}{n} \sum_{i=1}^{n} (y_i - w_0 - w_1 x_i)$$

$$\frac{\partial L(w_0, w_1)}{\partial w_1} = -\frac{2}{n} \sum_{i=1}^{n} x_i (y_i - w_0 - w_1 x_i) + 2\lambda w_1$$

Intro
○○

Linear Regression
○○○○○

Question 1 (a → c)
○○○○○○○○○●

Multiple Linear Regression
○○○○○○

Question 2 (a → h)
○○○○○○○○○○○

Final result is:

$$w_0 = \bar{y} - w_1 \bar{x}$$

$$w_1 = \frac{\overline{xy} - \bar{x}\bar{y}}{\overline{x^2} - \bar{x}^2 + \lambda}$$

Notice how the coefficients have an inverse relationship with λ.

Intro
○○

Linear Regression
○○○○○

Question 1 (a → c)
○○○○○○○○○○

Multiple Linear Regression
●○○○○○

Question 2 (a → h)
○○○○○○○○○○○○

Section 4

Multiple Linear Regression

Intro
oo

Linear Regression
ooooo

Question 1 (a → c)
ooooooooooo

Multiple Linear Regression
o●oooo

Question 2 (a → h)
ooooooooooo

Multiple Linear Regression

Recall the previous problem where we were tasked with finding price patterns of homes using the size of the home.

Intro
OO

Linear Regression
OOOOO

Question 1 (a → c)
OOOOOOOOOO

Multiple Linear Regression
O●OOOO

Question 2 (a → h)
OOOOOOOOOOO

Multiple Linear Regression

Recall the previous problem where we were tasked with finding price patterns of homes using the size of the home. Say we're now given the number of bedrooms in the house, how do we account for this in the model?

Intro
oo

Linear Regression
ooooo

Question 1 (a → c)
ooooooooooo

Multiple Linear Regression
o●oooo

Question 2 (a → h)
ooooooooooo

Multiple Linear Regression

Recall the previous problem where we were tasked with finding price patterns of homes using the size of the home. Say we're now given the number of bedrooms in the house, how do we account for this in the model?

Simple, just add another parameter:

$$\hat{y} = w_0 + w_1 x_1 + w_2 x_2$$

Intro
OO

Linear Regression
OOOOO

Question 1 (a → c)
OOOOOOOOOO

Multiple Linear Regression
O●OOOO

Question 2 (a → h)
OOOOOOOOOOOO

Multiple Linear Regression

Recall the previous problem where we were tasked with finding price patterns of homes using the size of the home. Say we're now given the number of bedrooms in the house, how do we account for this in the model?

Simple, just add another parameter:

$$\hat{y} = w_0 + w_1 x_1 + w_2 x_2$$

What if we're given the year the house was built and the coordinates? Let's say d more features?

Intro
○○

Linear Regression
○○○○○

Question 1 (a → c)
○○○○○○○○○○

Multiple Linear Regression
○○●○○○

Question 2 (a → h)
○○○○○○○○○○○

Let's vectorise our model, say:

$x_i = \begin{bmatrix} 1 \\ x_{i1} \end{bmatrix}$ to represent our input & the bias (w_0)

$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$ to represent the target variable

$w = \begin{bmatrix} w_0 \\ w_1 \end{bmatrix}$ to represent the parameters

Intro
○○

Linear Regression
○○○○○

Question 1 (a → c)
○○○○○○○○○○

Multiple Linear Regression
○○○●○○

Question 2 (a → h)
○○○○○○○○○○○

Then, let's define our entire feature set X as:

$$X = \begin{bmatrix} 1 & x_{11} \\ 1 & x_{21} \\ \vdots & \vdots \\ 1 & x_{n1} \end{bmatrix}$$

Intro
○○

Linear Regression
○○○○○

Question 1 (a → c)
○○○○○○○○○○

Multiple Linear Regression
○○○●○○

Question 2 (a → h)
○○○○○○○○○○○

Then, let's define our entire feature set X as:

$$X = \begin{bmatrix} 1 & x_{11} \\ 1 & x_{21} \\ \vdots & \vdots \\ 1 & x_{n1} \end{bmatrix}$$

So,

$$Xw = \begin{bmatrix} w_0 + w_1 x_{11} \\ w_0 + w_1 x_{21} \\ \vdots \\ w_0 + w_1 x_{n1} \end{bmatrix}$$

$$\hat{y} = Xw$$

Intro
○○

Linear Regression
○○○○○

Question 1 (a → c)
○○○○○○○○○○

Multiple Linear Regression
○○○○●○

Question 2 (a → h)
○○○○○○○○○○○

Then, what does our error become?

Intro
OO

Linear Regression
OOOOO

Question 1 (a → c)
OOOOOOOOOO

Multiple Linear Regression
OOOO●O

Question 2 (a → h)
OOOOOOOOOOO

Then, what does our error become?

$$\mathcal{L}(w) = \frac{1}{n} \sum_{i=1}^{n} (y_i - [Xw]_i)^2$$

Intro
○○

Linear Regression
○○○○○

Question 1 (a → c)
○○○○○○○○○○

Multiple Linear Regression
○○○○●○

Question 2 (a → h)
○○○○○○○○○○○

Then, what does our error become?

$$\mathcal{L}(w) = \frac{1}{n} \sum_{i=1}^{n} (y_i - [Xw]_i)^2$$

Formally,

$$\mathcal{L}(w) = \frac{1}{n} \|y - Xw\|_2^2$$

Intro
○○

Linear Regression
○○○○○

Question 1 (a → c)
○○○○○○○○○○

Multiple Linear Regression
○○○○●○

Question 2 (a → h)
○○○○○○○○○○○

Then, what does our error become?

$$\mathcal{L}(w) = \frac{1}{n} \sum_{i=1}^{n} (y_i - [Xw]_i)^2$$

Formally,

$$\mathcal{L}(w) = \frac{1}{n} \|y - Xw\|_2^2$$

Squared 2-Norm Identity

For a vector v,

$$\|v\|_2^2 = v^T v$$

Intro
oo

Linear Regression
ooooo

Question 1 (a → c)
oooooooooo

Multiple Linear Regression
ooooo●

Question 2 (a → h)
ooooooooooo

Vector Calculus

If $u \in \mathbb{R}^n$ and $v \in \mathbb{R}^n$ is a fixed vector, then

$$\frac{\partial v^T u}{\partial u} = v$$

If $A \in \mathbb{R}^{n \times n}$ is a fixed matrix,

$$\frac{\partial (u^T A u)}{\partial u} = Au + A^T u$$

Intro
○○

Linear Regression
○○○○○

Question 1 (a → c)
○○○○○○○○○○

Multiple Linear Regression
○○○○○○

Question 2 (a → h)
●○○○○○○○○○○○

Section 5

Question 2 (a → h)

Intro
oo

Linear Regression
ooooo

Question 1 (a → c)
ooooooooo

Multiple Linear Regression
oooooo

Question 2 (a → h)
oooooooooooo

2a

Problem: Show that $\mathcal{L}(w) = \frac{1}{n}\|y - Xw\|_2^2$ has critical point $\hat{w} = (X^T X)^{-1} X^T y$.

To find optimal w, solve $\dfrac{\partial \mathcal{L}(w)}{\partial w} = 0$

Intro
oo

Linear Regression
ooooo

Question 1 (a → c)
ooooooooo

Multiple Linear Regression
oooooo

Question 2 (a → h)
o●oooooooooo

2a

Problem: Show that $\mathcal{L}(w) = \frac{1}{n}\|y - Xw\|_2^2$ has critical point $\hat{w} = (X^TX)^{-1}X^Ty$.

To find optimal w, solve $\dfrac{\partial \mathcal{L}(w)}{\partial w} = 0$

$$
\begin{aligned}
\mathcal{L}(w) &= \frac{1}{n}(y - Xw)^T(y - Xw) \\
&= \frac{1}{n}\left(y^Ty - y^TXw - w^TX^Ty + w^TX^TXw\right) \\
&= \frac{1}{n}\left(y^Ty - 2y^TXw + w^TX^TXw\right)
\end{aligned}
$$

Intro
○○

Linear Regression
○○○○○

Question 1 (a → c)
○○○○○○○○○○

Multiple Linear Regression
○○○○○○

Question 2 (a → h)
○○●○○○○○○○○○

Let's find the derivative w.r.t w,

Intro
○○

Linear Regression
○○○○○

Question 1 (a → c)
○○○○○○○○○○

Multiple Linear Regression
○○○○○○

Question 2 (a → h)
○○●○○○○○○○○○○

Let's find the derivative w.r.t w,

$$\frac{\partial \mathcal{L}(w)}{\partial w} = -\frac{1}{n}(-2X^T y + 2X^T X w)$$

Intro
oo

Linear Regression
ooooo

Question 1 (a → c)
oooooooooo

Multiple Linear Regression
oooooo

Question 2 (a → h)
oo●ooooooooo

Let's find the derivative w.r.t w,

$$\frac{\partial \mathcal{L}(w)}{\partial w} = -\frac{1}{n}(-2X^T y + 2X^T X w)$$

To solve for $\hat{w}$,

$$-2X^T y + 2X^T X \hat{w} = 0$$
$$\hat{w} = (X^T X)^{-1} X^T y$$

Intro
○○

Linear Regression
○○○○○

Question 1 (a → c)
○○○○○○○○○○

Multiple Linear Regression
○○○○○○

Question 2 (a → h)
○○○●○○○○○○○○

2b

Problem: Prove $\hat{w} = (X^T X)^{-1} X^T y$ is a global minimum.

Intro
OO

Linear Regression
OOOOO

Question 1 (a → c)
OOOOOOOOOO

Multiple Linear Regression
OOOOOO

Question 2 (a → h)
OOO●OOOOOOO

2b

Problem: Prove $\hat{w} = (X^T X)^{-1} X^T y$ is a global minimum.

$$\begin{aligned}
\nabla_w^2 \mathcal{L}(w) &= \nabla_w(\nabla_w \mathcal{L}(w)) \\
&= \nabla_w(-2X^T y + 2X^T X w) \\
&= 2X^T X
\end{aligned}$$

2b

Problem: Prove $\hat{w} = (X^T X)^{-1} X^T y$ is a global minimum.

$$\nabla_w^2 \mathcal{L}(w) = \nabla_w(\nabla_w \mathcal{L}(w))$$
$$= \nabla_w(-2X^T y + 2X^T X w)$$
$$= 2X^T X$$

So, for a vector $u \in \mathbb{R}^p$,

$$u^T(2X^T X)u = 2(u^T X^T)(Xu)$$
$$= 2(Xu)^T(Xu)$$
$$= 2\|Xu\|_2^2 \geq 0$$

Intro
○○

Linear Regression
○○○○○

Question 1 (a → c)
○○○○○○○○○○

Multiple Linear Regression
○○○○○○

Question 2 (a → h)
○○○○●○○○○○○

Therefore, $\mathcal{L}$ is convex and $\hat{w}$ is the unique global minimum.

Intro
oo

Linear Regression
ooooo

Question 1 (a → c)
oooooooooo

Multiple Linear Regression
oooooo

Question 2 (a → h)
ooooo●ooooo

2c

$$x_i = \begin{bmatrix} 1 \\ x_{i1} \end{bmatrix}$$ to represent our input & the bias (w_0)

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$ to represent the target variable

$$w = \begin{bmatrix} w_0 \\ w_1 \end{bmatrix}$$ to represent the parameters

Intro
○○

Linear Regression
○○○○○

Question 1 (a → c)
○○○○○○○○○○

Multiple Linear Regression
○○○○○○

Question 2 (a → h)
○○○○○○●○○○○

$$X = \begin{bmatrix} 1 & x_{11} \\ 1 & x_{21} \\ \vdots & \vdots \\ 1 & x_{n1} \end{bmatrix}$$

$$X^T y = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_{11} & x_{21} & \cdots & x_{n1} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

Intro
○○

Linear Regression
○○○○○

Question 1 (a → c)
○○○○○○○○○○

Multiple Linear Regression
○○○○○○

Question 2 (a → h)
○○○○○○●○○○○

$$X = \begin{bmatrix} 1 & x_{11} \\ 1 & x_{21} \\ \vdots & \vdots \\ 1 & x_{n1} \end{bmatrix}$$

$$X^T y = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_{11} & x_{21} & \cdots & x_{n1} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

$$= \begin{bmatrix} n\bar{y} \\ n\overline{xy} \end{bmatrix}$$

Intro
○○

Linear Regression
○○○○○

Question 1 (a → c)
○○○○○○○○○○

Multiple Linear Regression
○○○○○○

Question 2 (a → h)
○○○○○○●○○○○

$$X = \begin{bmatrix} 1 & x_{11} \\ 1 & x_{21} \\ \vdots & \vdots \\ 1 & x_{n1} \end{bmatrix}$$

$$X^T X = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_{11} & x_{21} & \cdots & x_{n1} \end{bmatrix} \begin{bmatrix} 1 & x_{11} \\ 1 & x_{11} \\ \vdots & \vdots \\ 1 & x_{n1} \end{bmatrix}$$

$$X^T y = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_{11} & x_{21} & \cdots & x_{n1} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

$$= \begin{bmatrix} n\bar{y} \\ n\overline{xy} \end{bmatrix}$$

Intro
○○

Linear Regression
○○○○○

Question 1 (a → c)
○○○○○○○○○○

Multiple Linear Regression
○○○○○○

Question 2 (a → h)
○○○○○○●○○○○

$$X = \begin{bmatrix} 1 & x_{11} \\ 1 & x_{21} \\ \vdots & \vdots \\ 1 & x_{n1} \end{bmatrix}$$

$$X^T y = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_{11} & x_{21} & \cdots & x_{n1} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

$$= \begin{bmatrix} n\bar{y} \\ n\overline{xy} \end{bmatrix}$$

$$X^T X = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_{11} & x_{21} & \cdots & x_{n1} \end{bmatrix} \begin{bmatrix} 1 & x_{11} \\ 1 & x_{11} \\ \vdots & \vdots \\ 1 & x_{n1} \end{bmatrix}$$

$$= \begin{bmatrix} n & \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} x_i^2 \end{bmatrix}$$

Intro
○○

Linear Regression
○○○○○

Question 1 (a → c)
○○○○○○○○○○

Multiple Linear Regression
○○○○○○

Question 2 (a → h)
○○○○○○●○○○○

$$X = \begin{bmatrix} 1 & x_{11} \\ 1 & x_{21} \\ \vdots & \vdots \\ 1 & x_{n1} \end{bmatrix}$$

$$X^T X = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_{11} & x_{21} & \cdots & x_{n1} \end{bmatrix} \begin{bmatrix} 1 & x_{11} \\ 1 & x_{11} \\ \vdots & \vdots \\ 1 & x_{n1} \end{bmatrix}$$

$$X^T y = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_{11} & x_{21} & \cdots & x_{n1} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

$$= \begin{bmatrix} n & \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} x_i^2 \end{bmatrix}$$

$$= \begin{bmatrix} n\bar{y} \\ n\overline{xy} \end{bmatrix}$$

$$= \begin{bmatrix} n & n\bar{x} \\ n\bar{x} & n\overline{x^2} \end{bmatrix}$$

Intro
○○

Linear Regression
○○○○○

Question 1 (a → c)
○○○○○○○○○○

Multiple Linear Regression
○○○○○○

Question 2 (a → h)
○○○○○○○○●○○○

We have:

$$X^T X = \begin{bmatrix} n & n\bar{x} \\ n\bar{x} & n\overline{x^2} \end{bmatrix}$$

Intro
○○

Linear Regression
○○○○○

Question 1 (a → c)
○○○○○○○○○○

Multiple Linear Regression
○○○○○○

Question 2 (a → h)
○○○○○○○●○○○

We have:

$$X^T X = \begin{bmatrix} n & n\bar{x} \\ n\bar{x} & n\overline{x^2} \end{bmatrix}$$

Recall the inverse of a matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is $A^{-1} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$.

Intro
○○

Linear Regression
○○○○○

Question 1 (a → c)
○○○○○○○○○

Multiple Linear Regression
○○○○○○

Question 2 (a → h)
○○○○○○○●○○○

We have:

$$X^T X = \begin{bmatrix} n & n\bar{x} \\ n\bar{x} & n\overline{x^2} \end{bmatrix}$$

Recall the inverse of a matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is $A^{-1} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$.

$$(X^T X)^{-1} = \frac{1}{n^2\overline{x^2} - n^2\bar{x}^2} \begin{bmatrix} n\overline{x^2} & -n\bar{x} \\ -n\bar{x} & n \end{bmatrix}$$

$$= \frac{1}{n(\overline{x^2} - \bar{x}^2)} \begin{bmatrix} \overline{x^2} & -\bar{x} \\ -\bar{x} & 1 \end{bmatrix}$$

Intro
oo

Linear Regression
ooooo

Question 1 (a → c)
ooooooooo

Multiple Linear Regression
oooooo

Question 2 (a → h)
ooooooooo●oo

2d

$$(X^T X)^{-1} X^T y = \frac{1}{n(\overline{x^2} - \bar{x}^2)} \begin{bmatrix} \overline{x^2} & -\bar{x} \\ -\bar{x} & 1 \end{bmatrix} \begin{bmatrix} n\bar{y} \\ n\overline{xy} \end{bmatrix}$$

Intro
○○

Linear Regression
○○○○○

Question 1 (a → c)
○○○○○○○○○○

Multiple Linear Regression
○○○○○○

Question 2 (a → h)
○○○○○○○○○●○○

2d

$$(X^T X)^{-1} X^T y = \frac{1}{n(\overline{x^2} - \bar{x}^2)} \begin{bmatrix} \overline{x^2} & -\bar{x} \\ -\bar{x} & 1 \end{bmatrix} \begin{bmatrix} n\bar{y} \\ n\overline{xy} \end{bmatrix}$$

$$= \frac{1}{\overline{x^2} - \bar{x}^2} \begin{bmatrix} \overline{x^2}\bar{y} - \bar{x}\overline{xy} \\ \overline{xy} - \bar{x}\bar{y} \end{bmatrix}$$

Intro
○○

Linear Regression
○○○○○

Question 1 (a → c)
○○○○○○○○○○

Multiple Linear Regression
○○○○○○

Question 2 (a → h)
○○○○○○○○○●○○

2d

$$(X^T X)^{-1} X^T y = \frac{1}{n(\overline{x^2} - \bar{x}^2)} \begin{bmatrix} \overline{x^2} & -\bar{x} \\ -\bar{x} & 1 \end{bmatrix} \begin{bmatrix} n\bar{y} \\ n\overline{xy} \end{bmatrix}$$

$$= \frac{1}{\overline{x^2} - \bar{x}^2} \begin{bmatrix} \overline{x^2}\bar{y} - \bar{x}\overline{xy} \\ \overline{xy} - \bar{x}\bar{y} \end{bmatrix}$$

$$= \begin{bmatrix} \bar{y} - \hat{w}_1 \bar{x} \\ \frac{\overline{xy} - \bar{x}\bar{y}}{\overline{x^2} - \bar{x}^2} \end{bmatrix}$$

Intro
○○

Linear Regression
○○○○○

Question 1 (a → c)
○○○○○○○○○○

Multiple Linear Regression
○○○○○○

Question 2 (a → h)
○○○○○○○○○○●○

2e - Lab

Given $x_1, \cdots, x_5 = 3, 6, 7, 8, 11$ and $y_1, \cdots, y_5 = 13, 8, 11, 2, 6$ compute the least squares solution by hand and using Python. Check your results with the sklearn implementation.

Intro
○○

Linear Regression
○○○○○

Question 1 (a → c)
○○○○○○○○○○

Multiple Linear Regression
○○○○○○

Question 2 (a → h)
○○○○○○○○○○●

2g

$$\mathsf{MSE}(w) = \arg\min_{w} \frac{1}{n}\|y - Xw\|_2^2 \text{ and } \mathsf{SSE}(w) = \arg\min_{w} \|y - Xw\|_2^2$$

i) Is the minimiser of MSE and SSE the same?

ii) Is the minimum value of MSE and SSE the same?