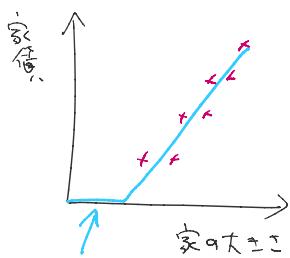


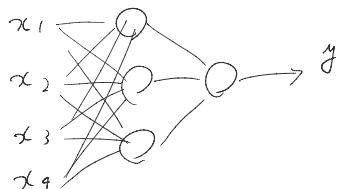
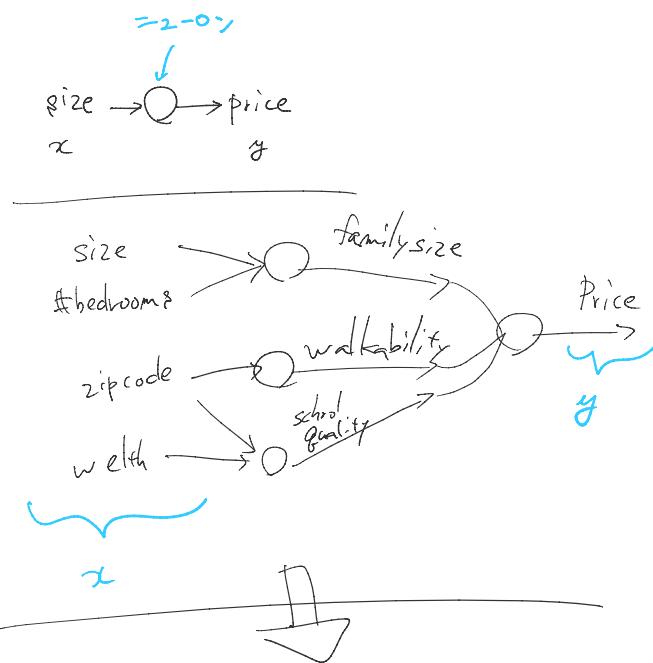
家賃予測



ReLU

Rectified Linear Unit

$$= \max(0, \text{val})$$



教師あり学習

CNN: 画像

RNN: 時系列、音声など

構造化データ: テキスト

非構造化データ: 音声、画像

Week1 ここまで

Computation Graph

$$J(a, b, c) = 3(a+bc)$$

$$u = bc$$

$$v = a + u$$

(2)

$$\begin{array}{c} Q = 5 \\ b = 3 \\ c = 2 \end{array} \xrightarrow{\quad} \boxed{u = bc} \xrightarrow{\quad} \boxed{v = a + u} \xrightarrow{\quad} \boxed{J = 3v} \xrightarrow{\quad} 33$$

(1)

$$\begin{array}{l} v = 11 \quad v = 11.001 \\ a = 5 \quad a = 5.001 \end{array} \quad \begin{array}{l} J = 33 \quad J = 33.003 \\ J = 33 \quad J = 33.003 \end{array}$$

$$\frac{dJ}{da} = 3 = \frac{dJ}{dv} \frac{dv}{da} \quad (1)$$

$$\frac{dv}{da} = 1 \quad (2)$$

$$\frac{du}{dc} = 1 \quad (3)$$

$$\begin{array}{ll} u = 6 & u = 6.002 \\ b = 3 & b = 3.001 \\ v = 11 & v = 11.002 \\ J = 33 & J = 33.006 \end{array}$$

(4)

$$\frac{du}{db} = 2$$

$$\frac{dv}{db} = \frac{dv}{du} \frac{du}{db} = 2$$

$$\frac{dJ}{db} = \frac{dJ}{dv} \frac{dv}{db} = 6$$

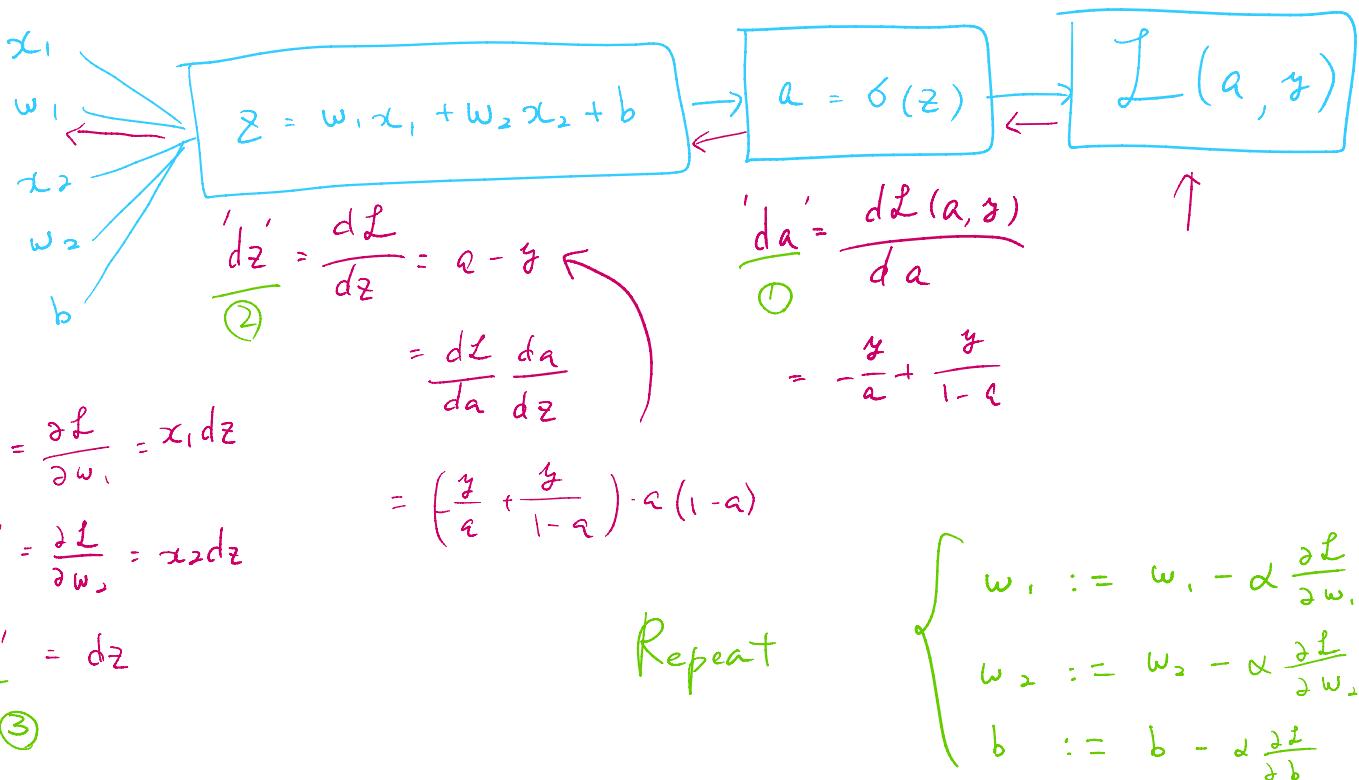
Logistic Regression

$$z = w^T x + b$$

$$z = w^T x + b$$

$$\hat{y} = \sigma(z) = a$$

$$L(a, y) = -(y \log(\hat{y}) + (1-y) \log(1-\hat{y}))$$

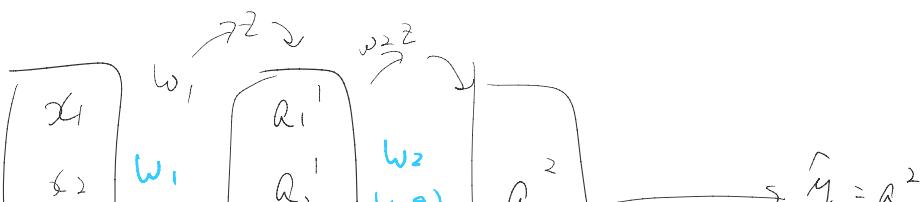


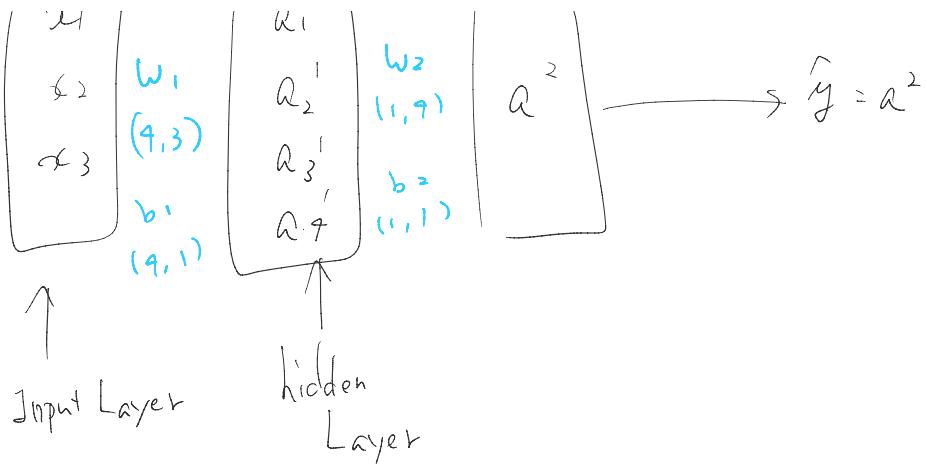
$$\bar{w}^T = \begin{bmatrix} m \times 1 \\ \vdots \end{bmatrix} \quad x = \begin{bmatrix} n \times m \\ \vdots \\ 0, \dots, 0 \end{bmatrix} \quad b = \begin{bmatrix} l \times m \\ b \ b \ \dots \ b \end{bmatrix}$$

$$w^T x = w^T x^{(1)} \ w^T x^{(2)} \dots w^T x^{(m)}$$

$$w^T x + b = w^T x^{(1)} + b \ \dots \ w^T x^{(m)} + b$$

Week 2 22.09.





$$x \in \mathbb{R}^{3 \times m}$$

$$w_1 \in \mathbb{R}^{3 \times 4}$$

$$b_1 \in \mathbb{R}^{4 \times 1}$$

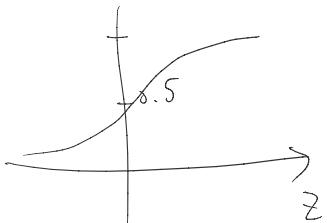
$$\omega_1 x^1 + b_1 = q' \in \mathbb{R}^{4 \times m}$$

$$w_2 \in \mathbb{R}^{4 \times 1}$$

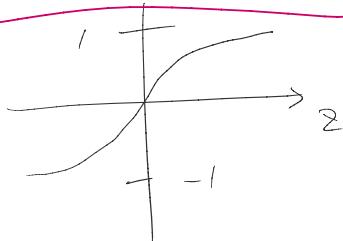
$$b_2 \in \mathbb{R}^{1 \times m}$$

$$\omega_2 q' + b_2 = a^2 \in \mathbb{R}^{1 \times m}$$

$\delta(z)$ あまり使われない



$$\tanh(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}}$$



多くの場合、 $\sigma(z)$ より

$\tanh(z)$ が用いられる。

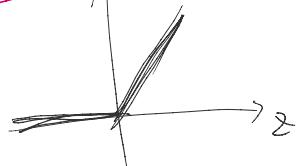
平均が 0 の近づいた時



よく使われる。

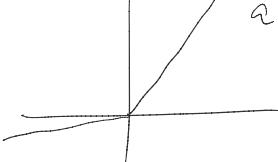
ReLU

$$a = \max(0, z)$$



Leaky ReLU

$$a = \max(0.01z, z)$$



Activation function.

$$g(z) = \sigma(z) \leftarrow この は 何故 えらぶる か?$$

$g(z) = z$ 恒等関数 \rightarrow 口ひきで回帰の性能が劣る。
 $(z=3)$

$tanh(z)$, 線形回帰の場合は出力層で
 出力層が実数。
 使う

導関数

$$z=0 \rightarrow g(z) = 0.5$$

$$\frac{dg(z)}{dz} = 0.5(1-0.5) \\ = 0.25$$

$$z=10 \rightarrow g(z) \approx 1$$

$$\frac{dg(z)}{dz} \approx 1(1-1) \\ \approx 0$$

$$z=-10 \rightarrow g(z) \approx 0$$

$$\frac{dg(z)}{dz} \approx 0(1-0) \\ \approx 0$$

$$g(z) = \tanh(z) = \frac{e^{-z} - e^z}{e^{-z} + e^z}$$

$$g'(z) = 1 - (\tanh(z))^2$$

$$z=10 \quad \tanh(z) \approx 1$$

$$g'(z) \approx 0$$

$$z=-10 \quad \tanh(z) \approx -1$$

$$g'(z) \approx 0$$

$$z=0 \quad \tanh(z) = 0 \\ g'(z) = 1$$

ReLU

$$g(z) = \max(0, z)$$

$$g'(z) = \begin{cases} 0 & \text{if } z < 0 \\ 1 & \text{if } z \geq 0 \end{cases}$$

Leaky ReLU

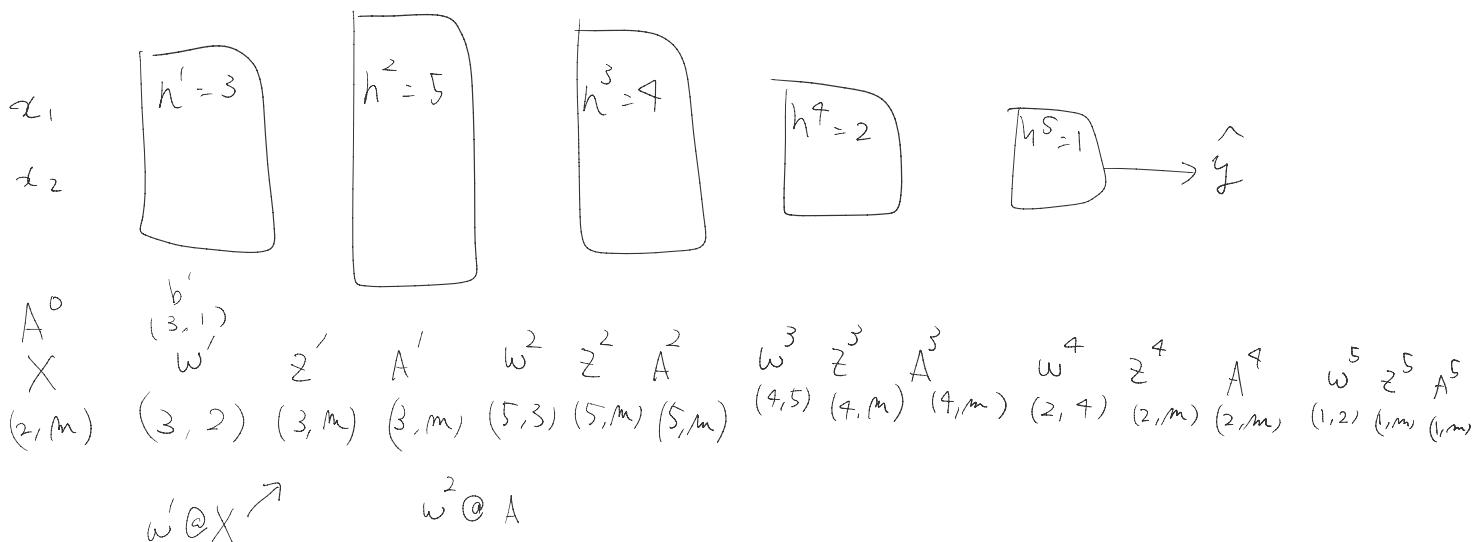
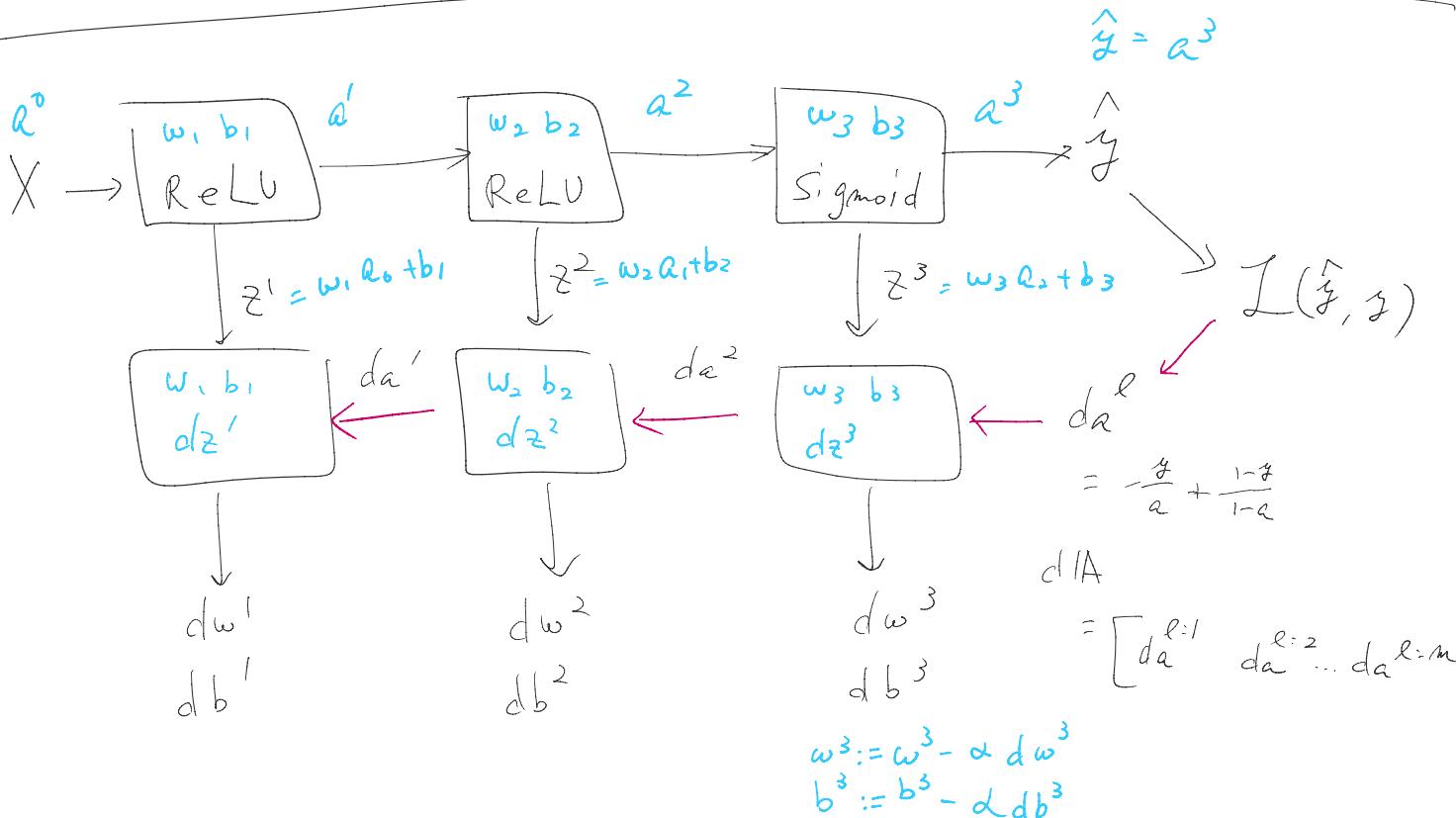
$$g(z) = \max(0.001z, z)$$

$$g'(z) = \begin{cases} 0.001 & \text{if } z < 0 \\ 1 & \text{if } z \geq 0 \end{cases}$$

$$f^{(z)} = \begin{cases} 1 & \text{if } z \geq 0 \\ \text{undefined} & \text{if } z = 0 \end{cases}$$

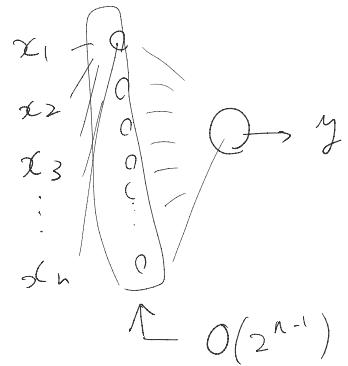
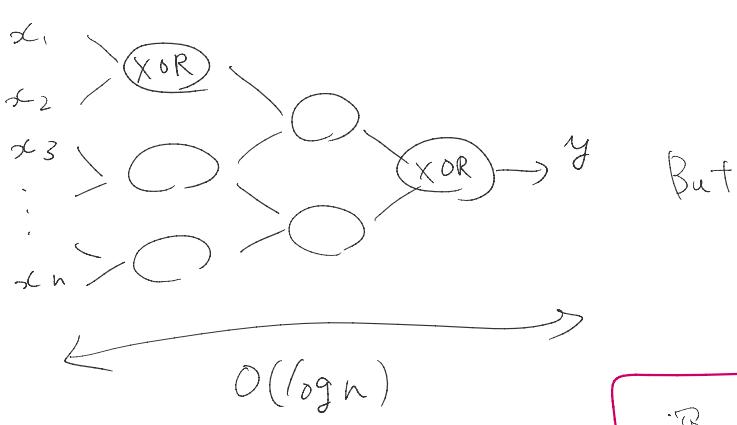
$$f(z) = 1 \quad \text{if } z \geq 0$$

Week 3 二进制



NN 从 $\mathbb{R}^d \rightarrow \mathbb{R}^n$

$$y = x_1 \text{ XOR } x_2 \text{ XOR } x_3 \dots \text{ XOR } x_n$$



深め方が良い（重みが少ない）例。

誤差関数

学習率 α

max iter

batch size

