

AI輪講

2021.09.01

陸研究室
土田真哉

出力の式

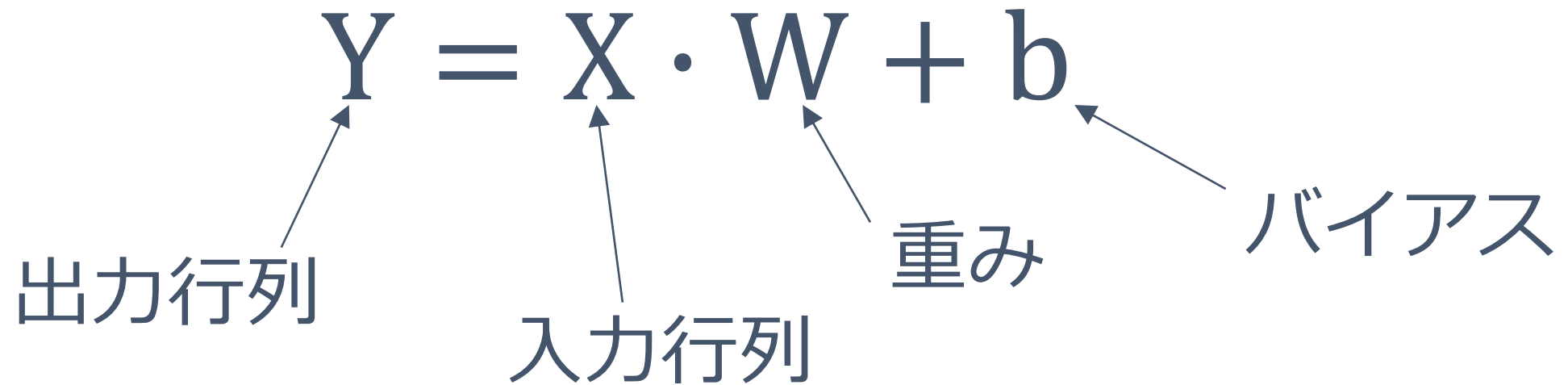
$$Y = X \cdot W + b$$

出力行列

入力行列

重み

バイアス



入力行列

X_{11}	X_{12}
X_{21}	X_{22}

重み

W

バイアス

b

出力行列

$X_{11} \cdot W + b$	$X_{12} \cdot W + b$
$X_{21} \cdot W + b$	$X_{22} \cdot W + b$

損失関数

平均二乗誤差(MSE Loss)

$$L = \frac{1}{N} \sum_{i=1}^N (y_i - t_i)^2$$

Loss

出力行列の要素の数

出力

教師データ

出力行列

Y_{11}	Y_{12}
Y_{21}	Y_{22}

教師行列

T_{11}	T_{12}
T_{21}	T_{22}

損失関数(MSELoss)

$$L = \frac{1}{4} \left((Y_{11} - T_{11})^2 + (Y_{12} - T_{12})^2 + (Y_{21} - T_{21})^2 + (Y_{22} - T_{22})^2 \right)$$

重みの勾配

損失関数の重みの偏微分

$$\frac{\partial L}{\partial W} = \frac{\partial}{\partial W} \left(\frac{1}{N} \sum_{i=1}^N (y_i - t_i)^2 \right) = \frac{\partial}{\partial W} \left(\frac{1}{N} \sum_{i=1}^N (x_i \cdot W + b - t_i)^2 \right)$$

$$\frac{\partial L}{\partial W} = \frac{1}{N} \sum_{i=1}^N 2 \cdot (y_i - t_i) \cdot x_i$$

入力行列

X_{11}	X_{12}
X_{21}	X_{22}

出力行列

Y_{11}	Y_{12}
Y_{21}	Y_{22}

教師行列

T_{11}	T_{12}
T_{21}	T_{22}

重みの勾配

$$\frac{\partial L}{\partial W_1} = \frac{1}{4} (2 \cdot (Y_{11} - T_{11}) \cdot X_{11} + 2 \cdot (Y_{12} - T_{12}) \cdot X_{12} + 2 \cdot (Y_{21} - T_{21}) \cdot X_{21} + 2 \cdot (Y_{22} - T_{22}) \cdot X_{22})$$

バイアスの勾配

損失関数の重みの偏微分

$$\frac{\partial L}{\partial b} = \frac{\partial}{\partial b} \left(\frac{1}{N} \sum_{i=1}^N (y_i - t_i)^2 \right) = \frac{\partial}{\partial b} \left(\frac{1}{N} \sum_{i=1}^N (x_i \cdot W + b - t_i)^2 \right)$$

$$\frac{\partial L}{\partial b} = \frac{1}{N} \sum_{i=1}^N 2 \cdot (y_i - t_i)$$

入力行列

X_{11}	X_{12}
X_{21}	X_{22}

出力行列

Y_{11}	Y_{12}
Y_{21}	Y_{22}

教師行列

T_{11}	T_{12}
T_{21}	T_{22}

バイアスの勾配

$$\frac{\partial L}{\partial b} = \frac{1}{4} (2 \cdot (Y_{11} - T_{11}) + 2 \cdot (Y_{12} - T_{12}) + 2 \cdot (Y_{21} - T_{21}) + 2 \cdot (Y_{22} - T_{22}))$$

重みの更新

確率的勾配降下法(SGD)

$$W \leftarrow W - \eta \frac{\partial L}{\partial W}$$

学習率(今回は0.01)

重みの勾配

バイアスの更新

確率的勾配降下法(SGD)

$$b \leftarrow b - \eta \frac{\partial L}{\partial b}$$

学習率(今回は0.01)

バイアスの
勾配