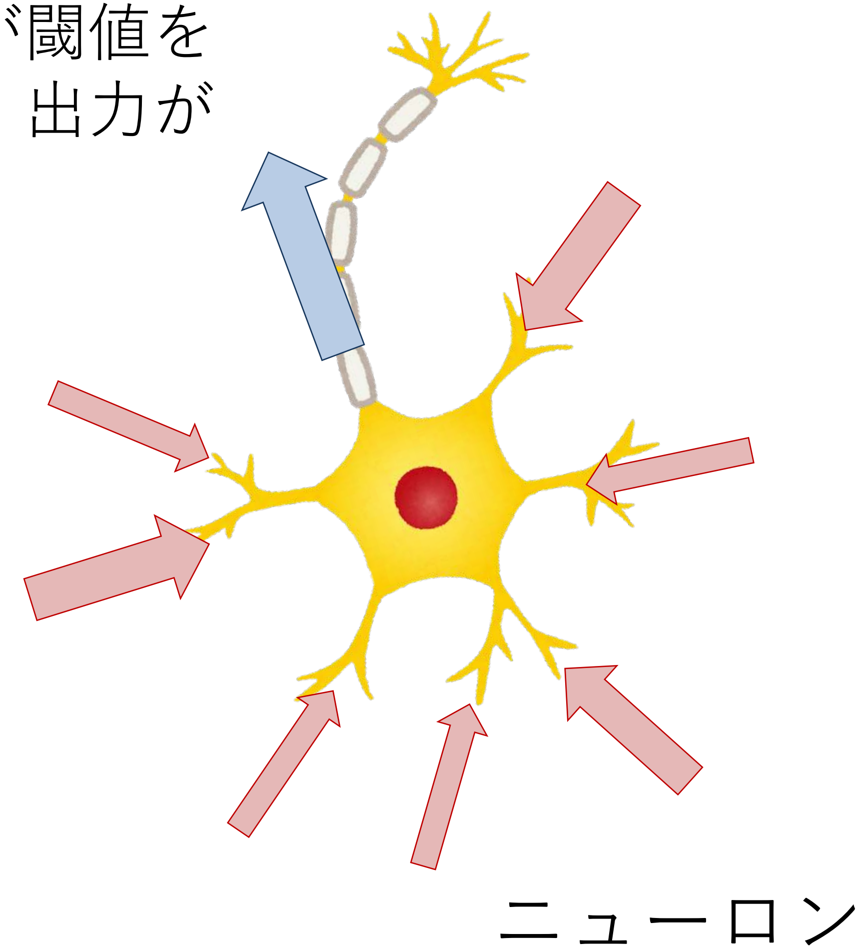
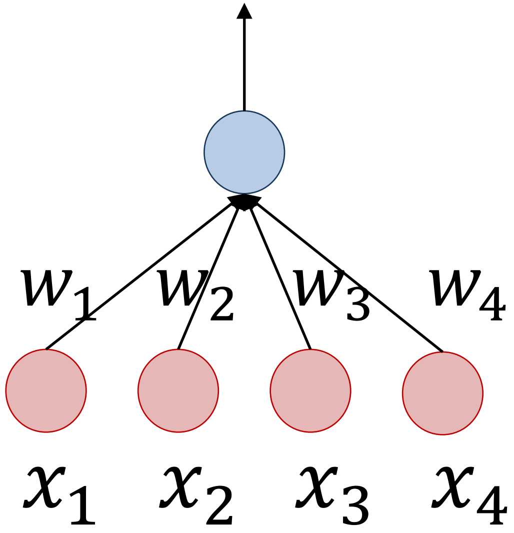
# 多層パーセプトロン
# MLP (Multilayer Perceptron)

# パーセプトロン

入力の和が閾値を
超えると，出力が
ある。

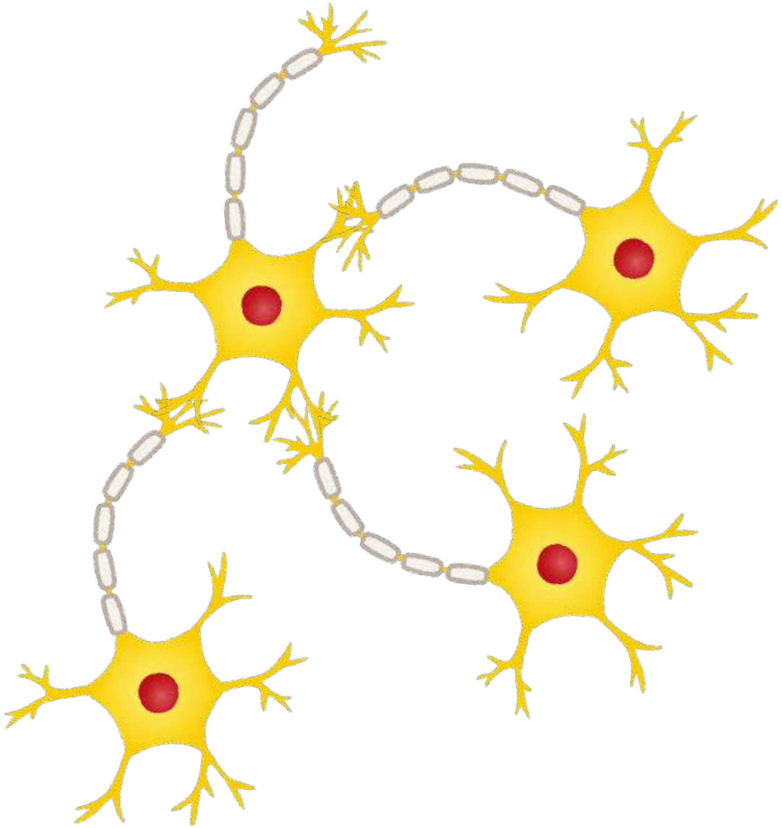


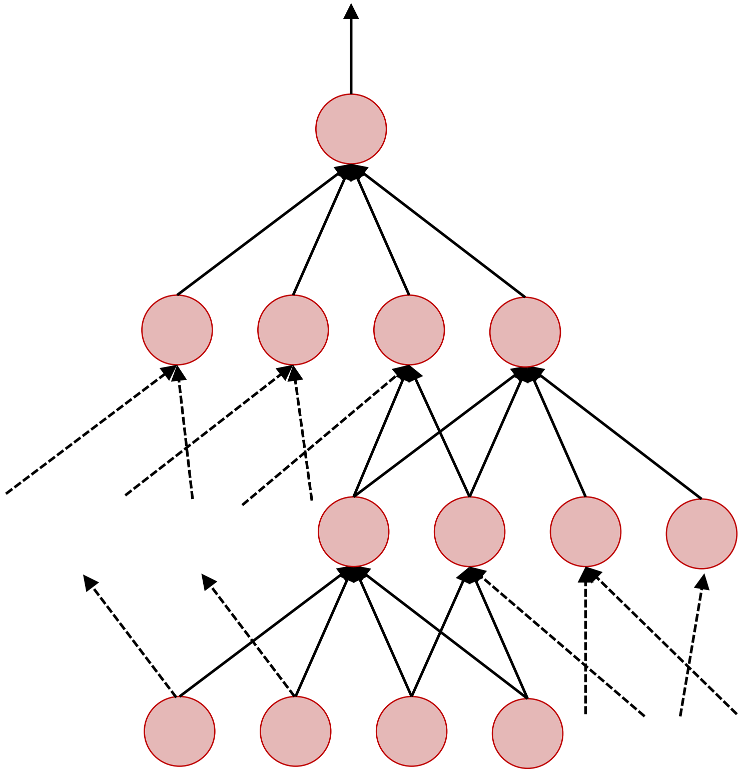

$$y = f\left(\sum_i w_i x_i\right)$$




パーセプトロンはニューロンに似せた，計算モデル

# ニューラルネットとは

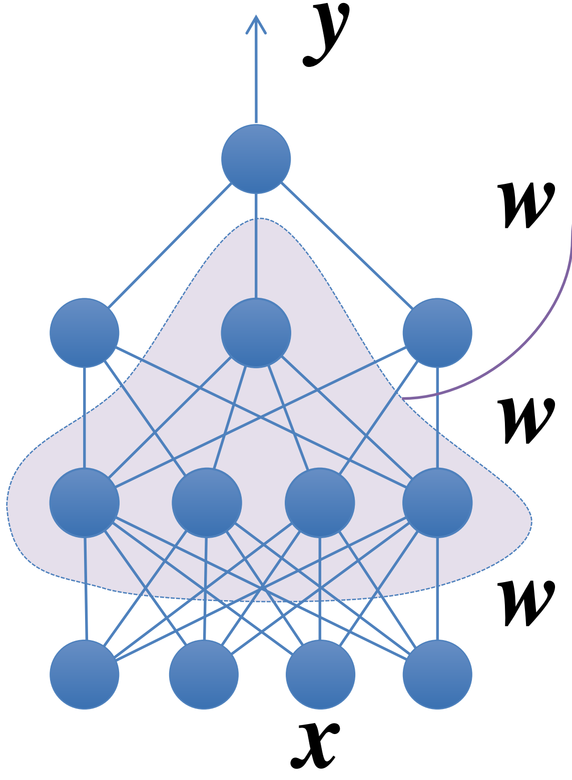


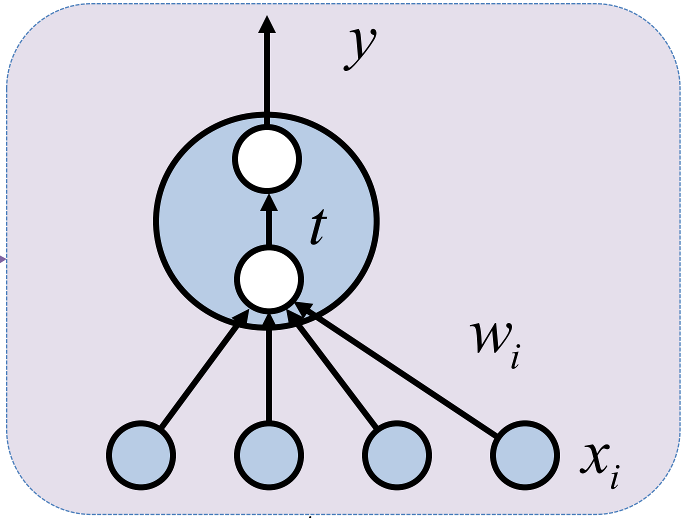


基本構造の組合せで
複雑なネットワークを作る

# 多層パーセプトロン：Multi-Layer Perceptron

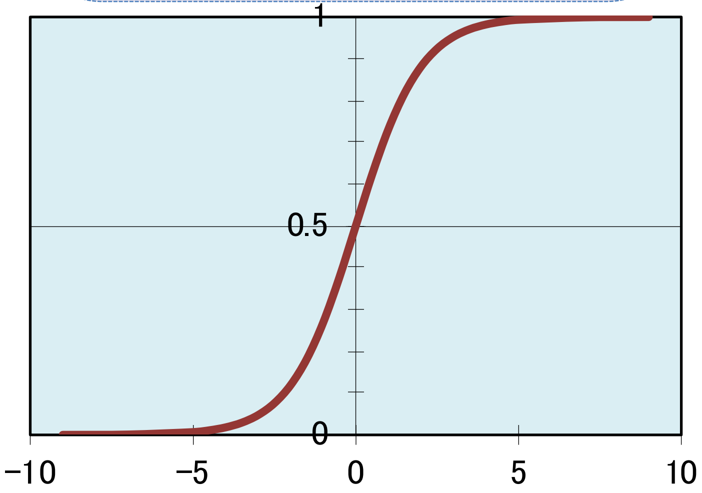
- MLPはパーセプトロンの組み合わせ




$$t = \sum_j w_j \cdot x_j$$

$$y = f(t)$$



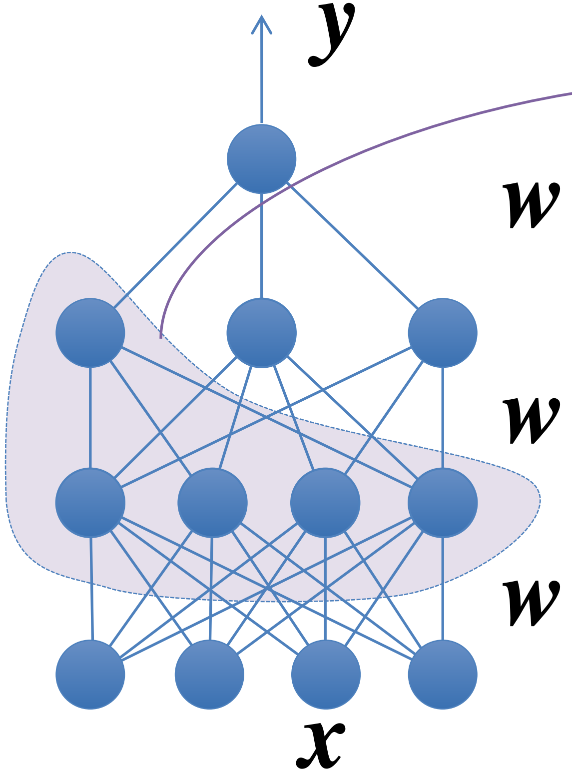

$f(t)$：非線形活性化関数

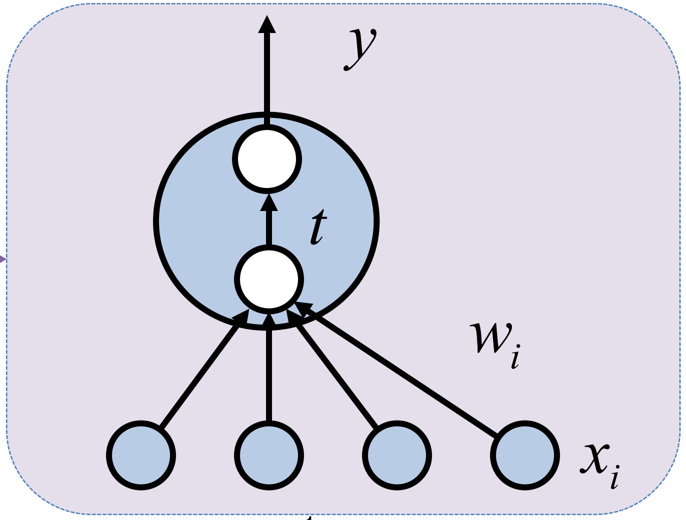
e.g.

$$f(t) = \frac{1}{1+e^{-t}}$$

# 多層パーセプトロン：Multi-Layer Perceptron

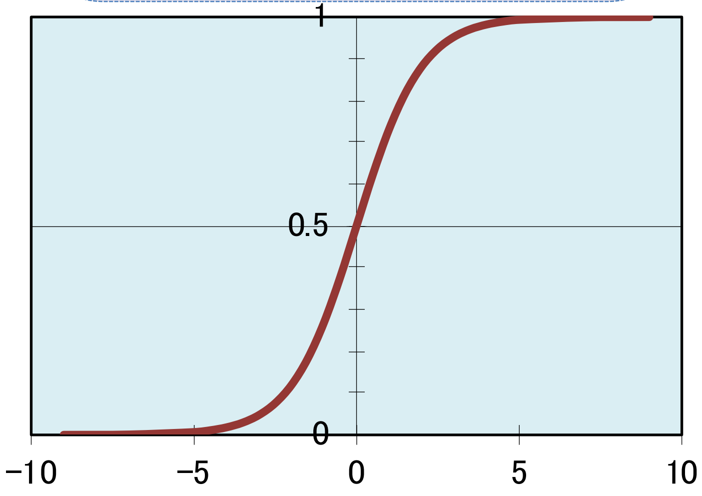
- MLPはパーセプトロンの組み合わせ




$$t = \sum_j w_j \cdot x_j$$

$$y = f(t)$$



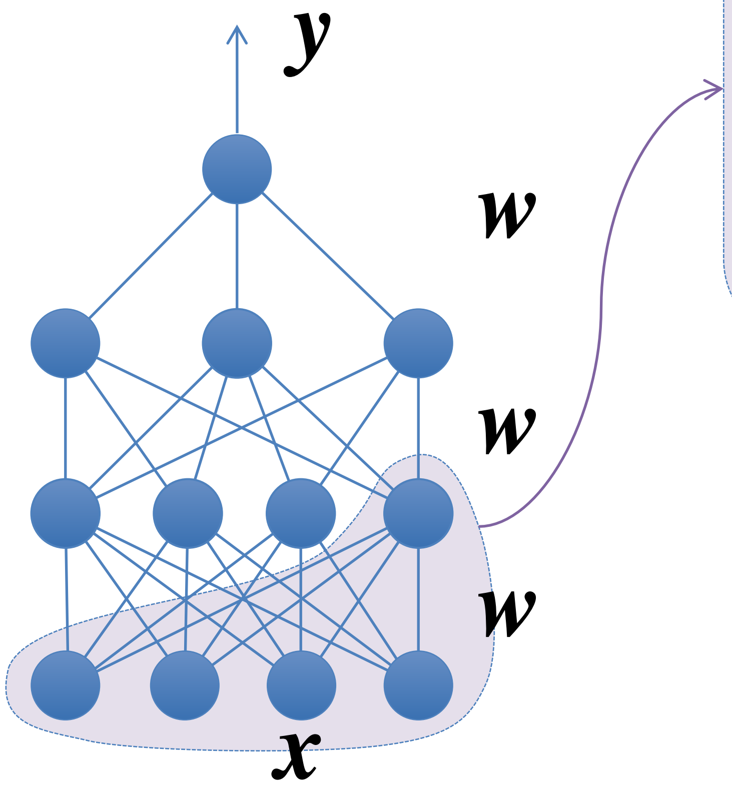

$f(t)$：非線形活性化関数

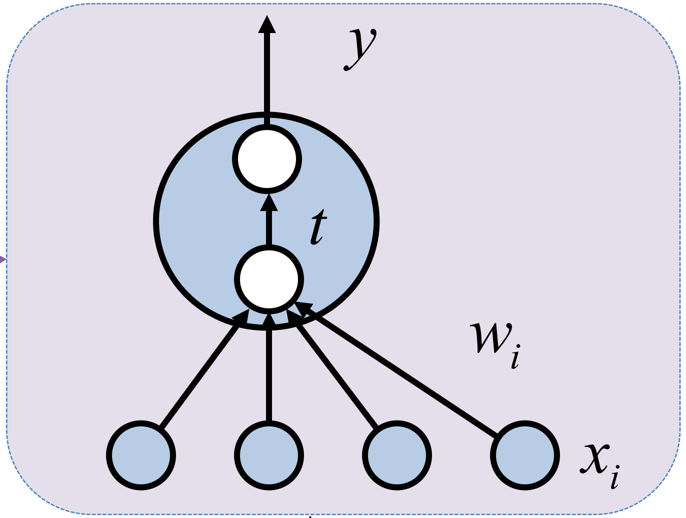
e.g.

$$f(t) = \frac{1}{1+e^{-t}}$$

# 多層パーセプトロン：Multi-Layer Perceptron

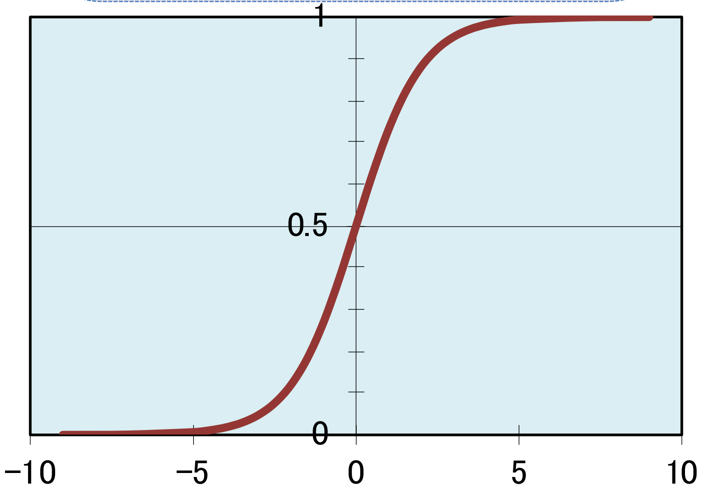
- MLPはパーセプトロンの組み合わせ




$$t = \sum_j w_j \cdot x_j$$

$$y = f(t)$$



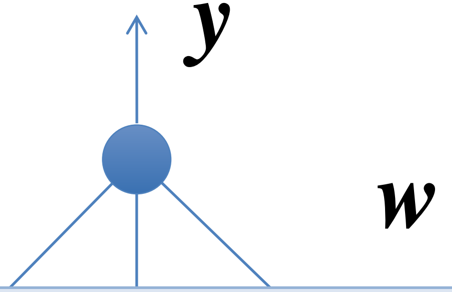

$f(t)$：非線形
活性化関数

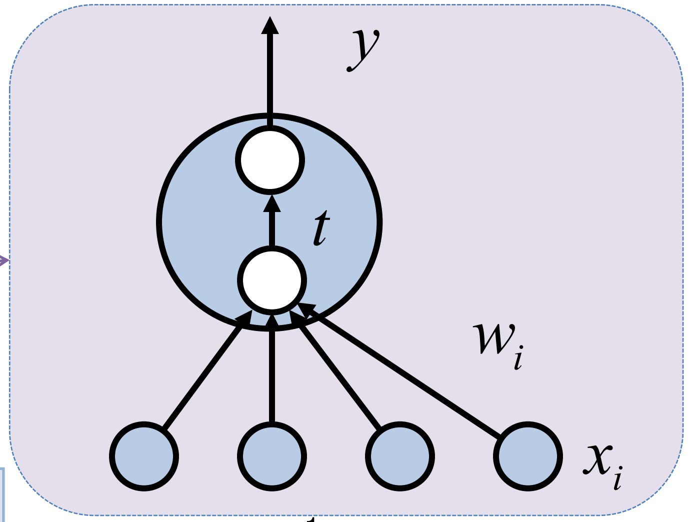
e.g.

$$f(t) = \frac{1}{1+e^{-t}}$$

# 多層パーセプトロン：Multi-Layer Perceptron

- MLPはパーセプトロンの組み合わせ



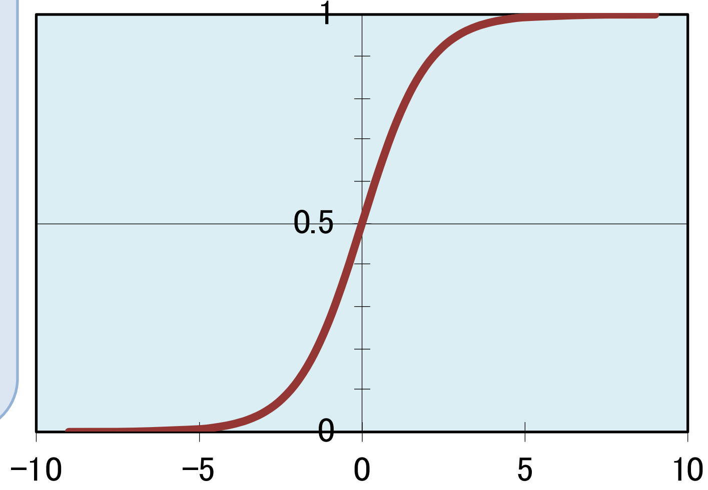




$$t = \sum_j w_j \cdot x_j$$

$$y = f(t)$$

MLPの機能

- 論理関数
- 座標変換
- 加法モデル

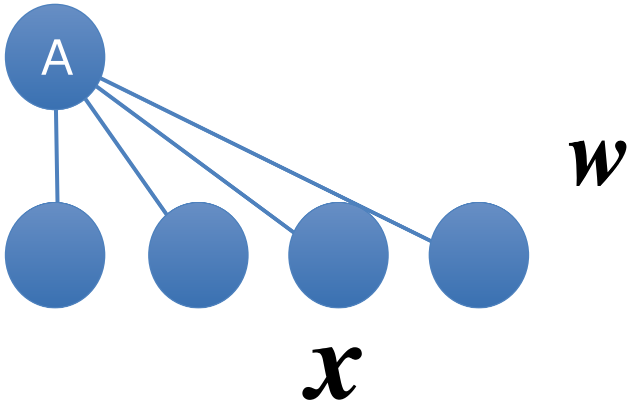


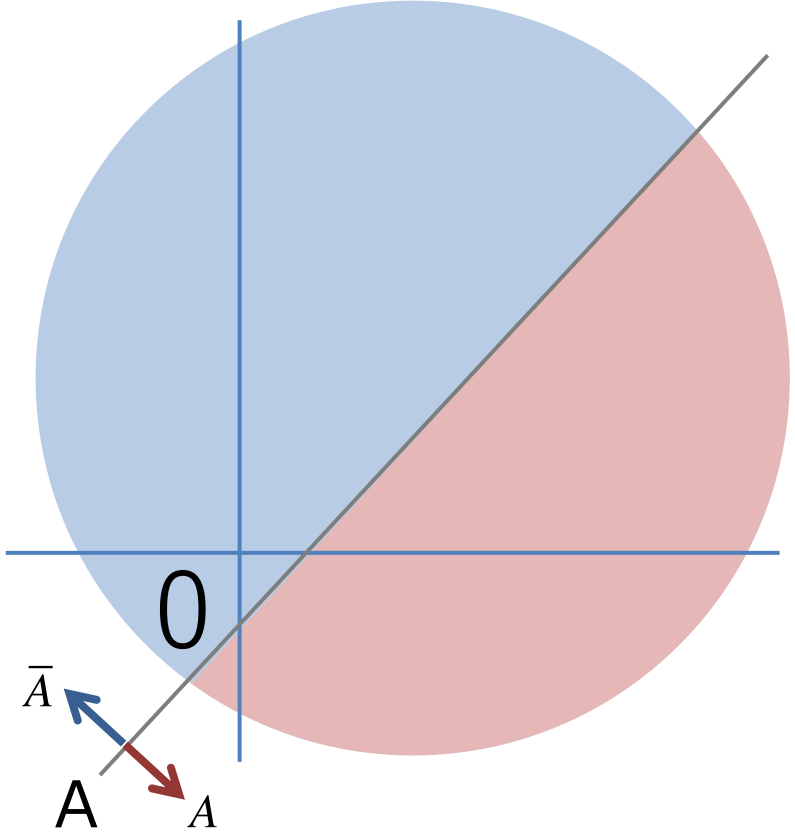

$f(t)$：非線形活性化関数

e.g.

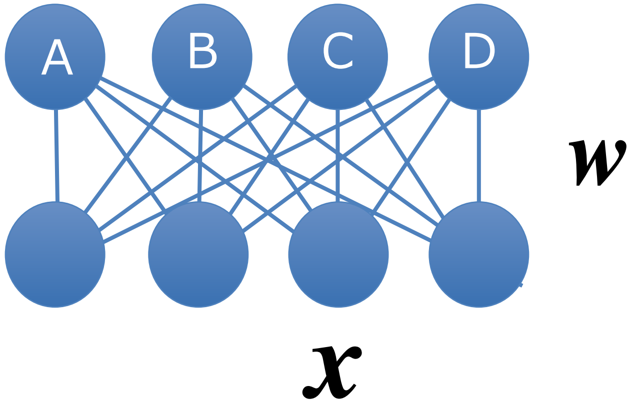
$$f(t) = \frac{1}{1+e^{-t}}$$

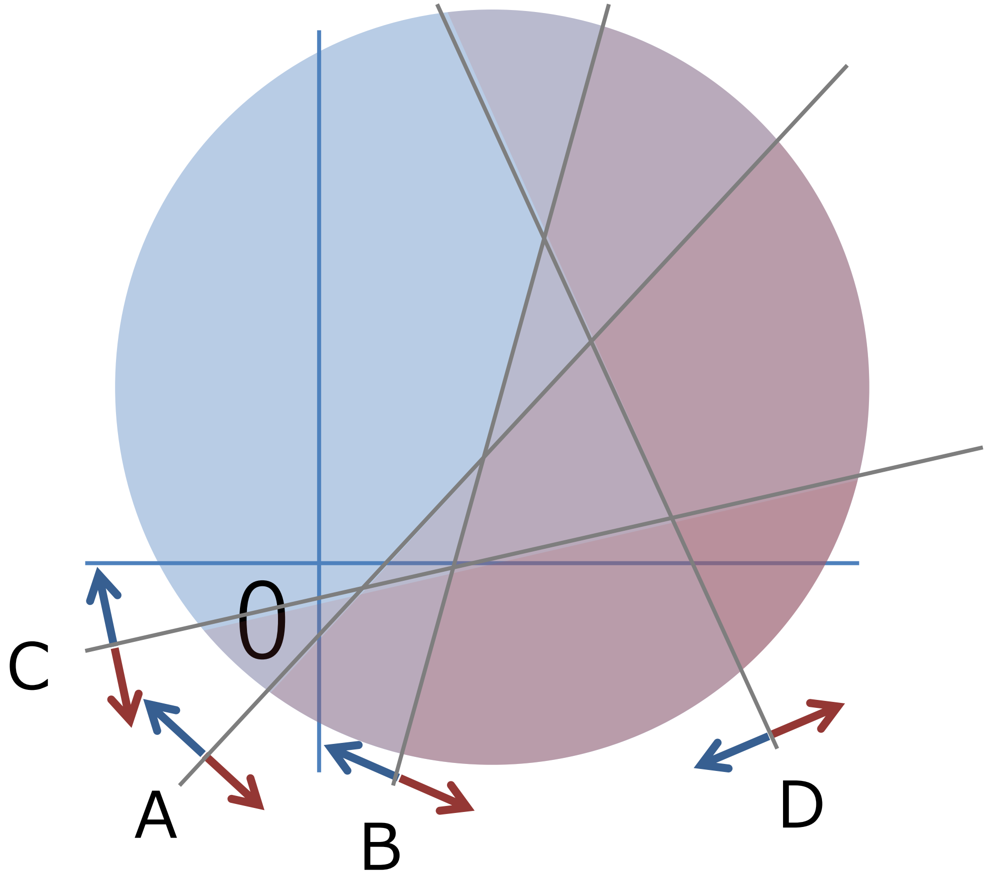
# MLPは何を実現するか:<br>論理関数としての解釈



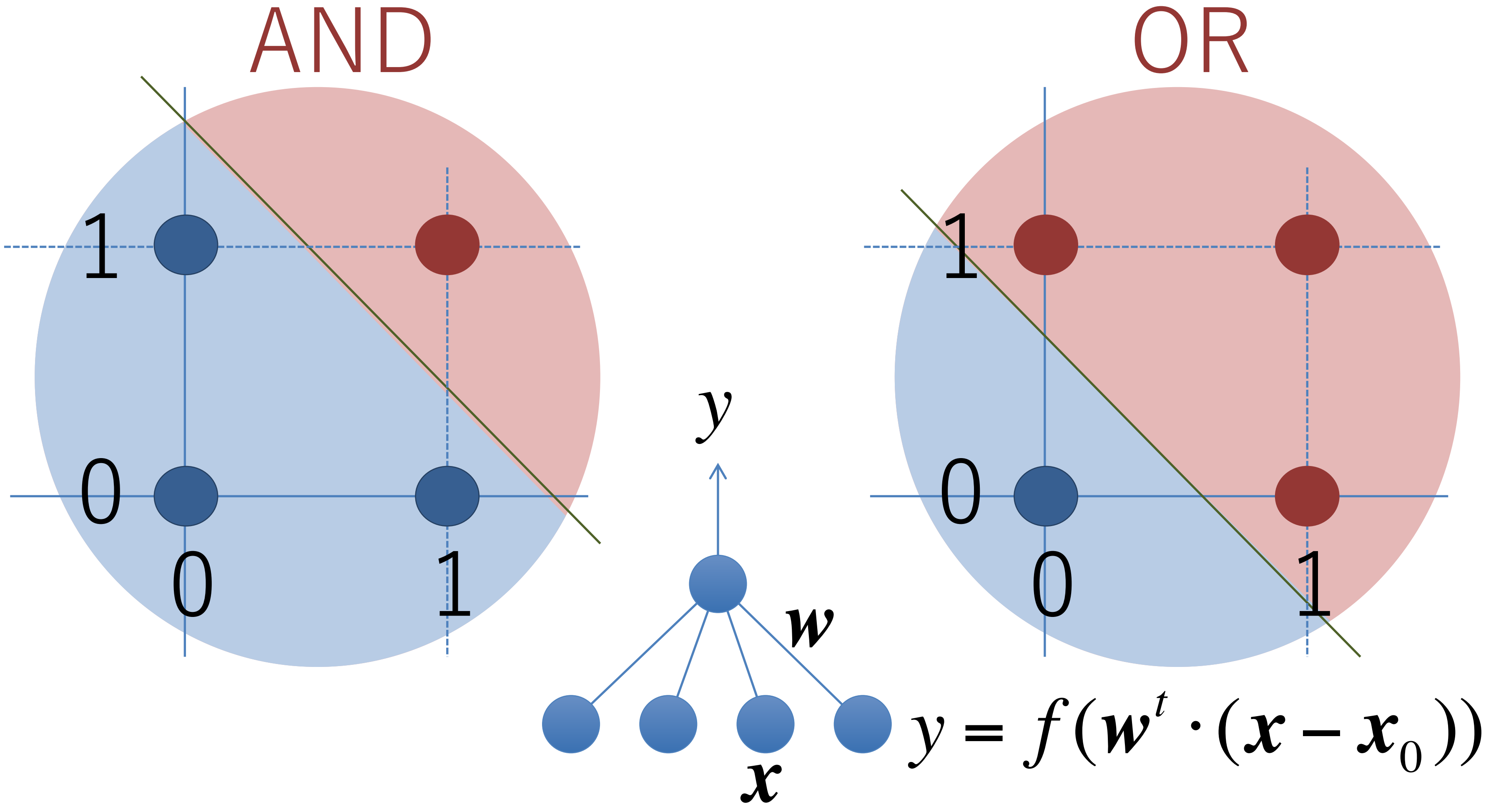



# MLPは何を実現するか:
# 論理関数としての解釈



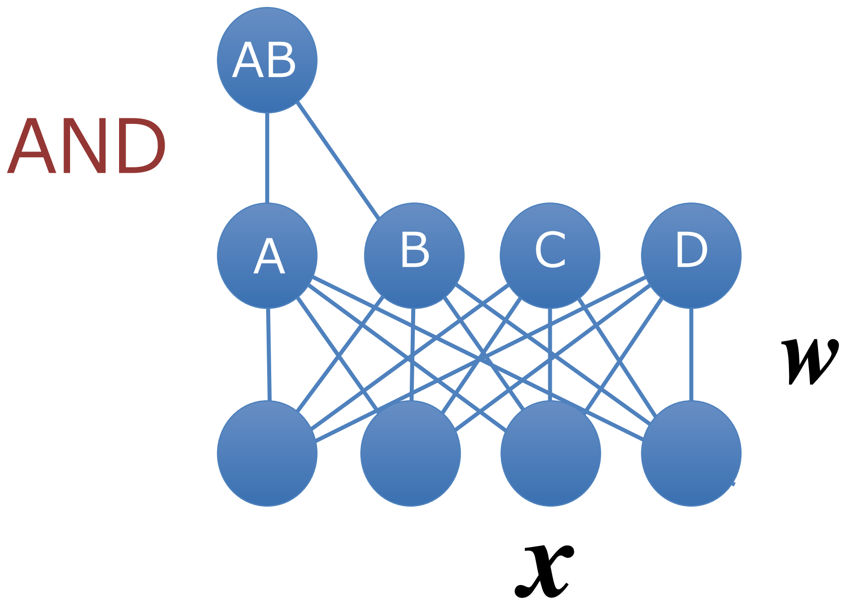



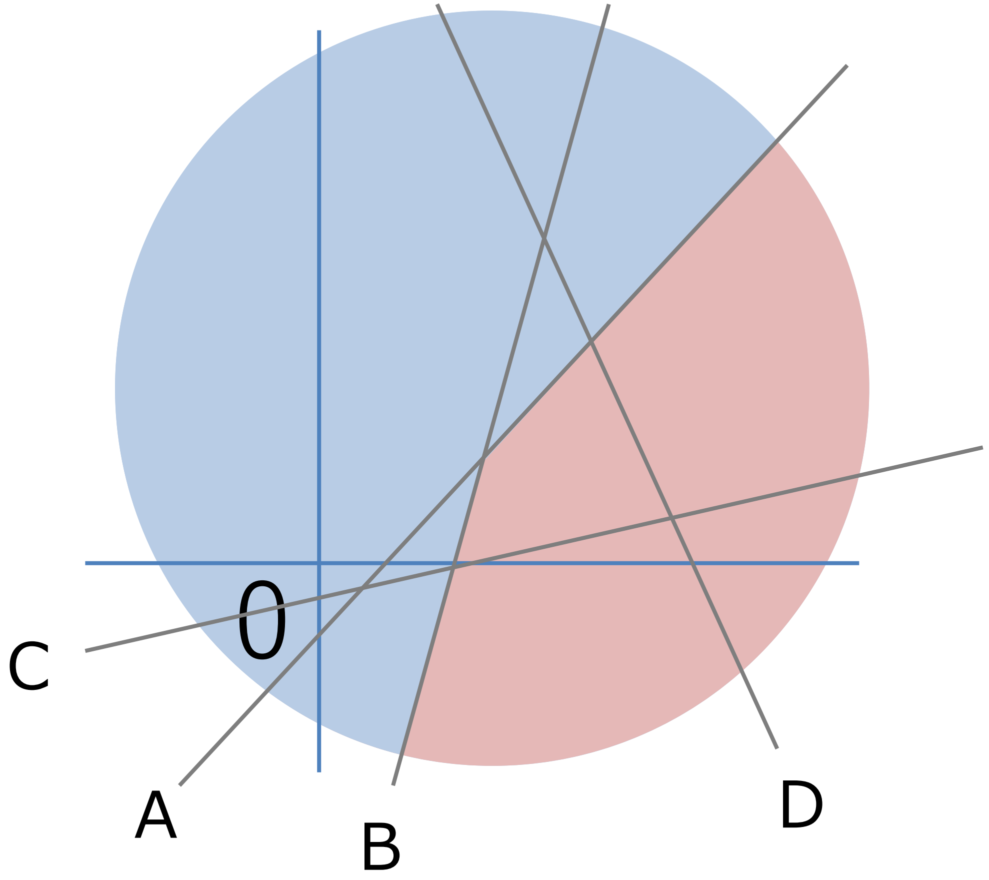
# MLPは何を実現するか: 論理関数としての解釈



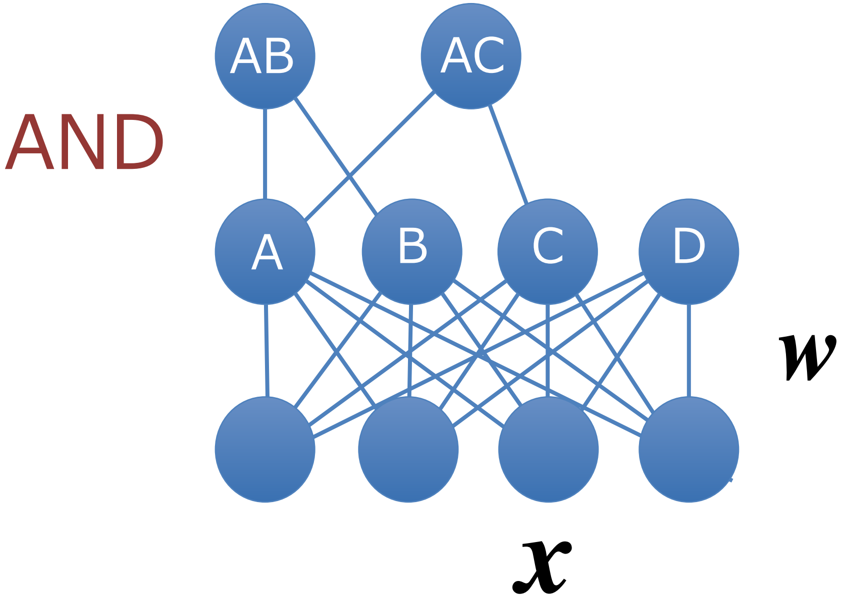

$$y = f(\boldsymbol{w}^t \cdot (\boldsymbol{x} - \boldsymbol{x}_0))$$

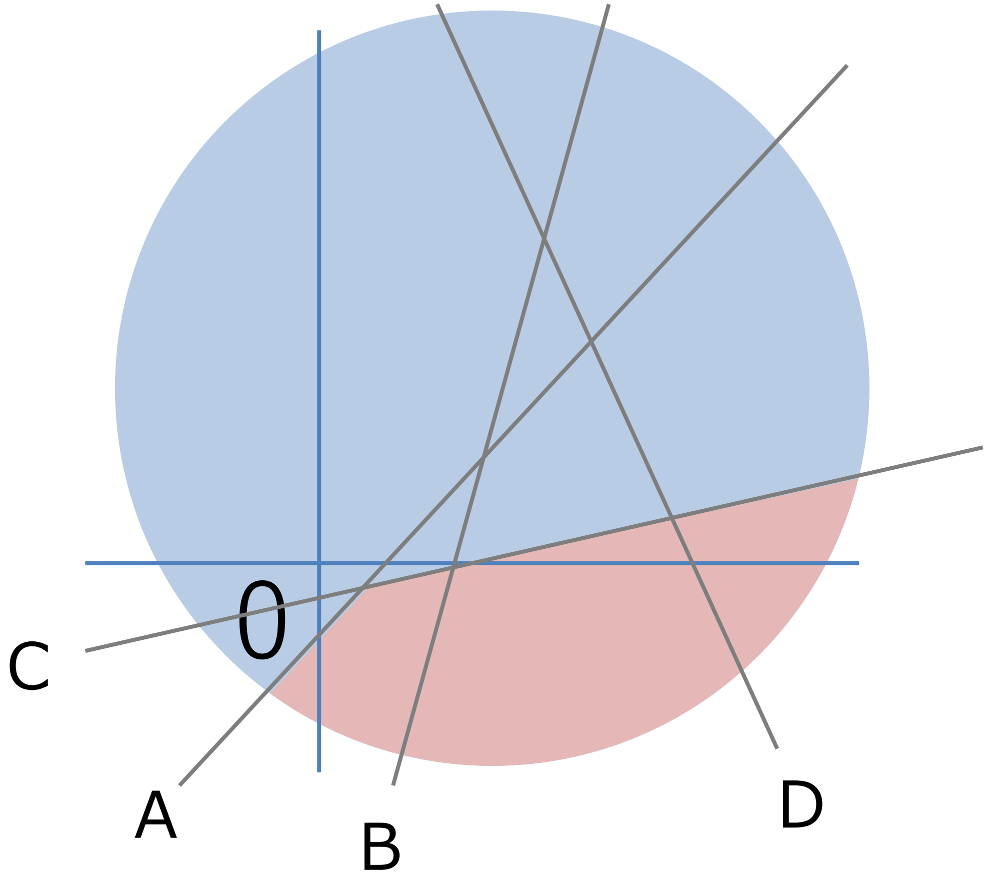
# MLPは何を実現するか:
# 論理関数としての解釈



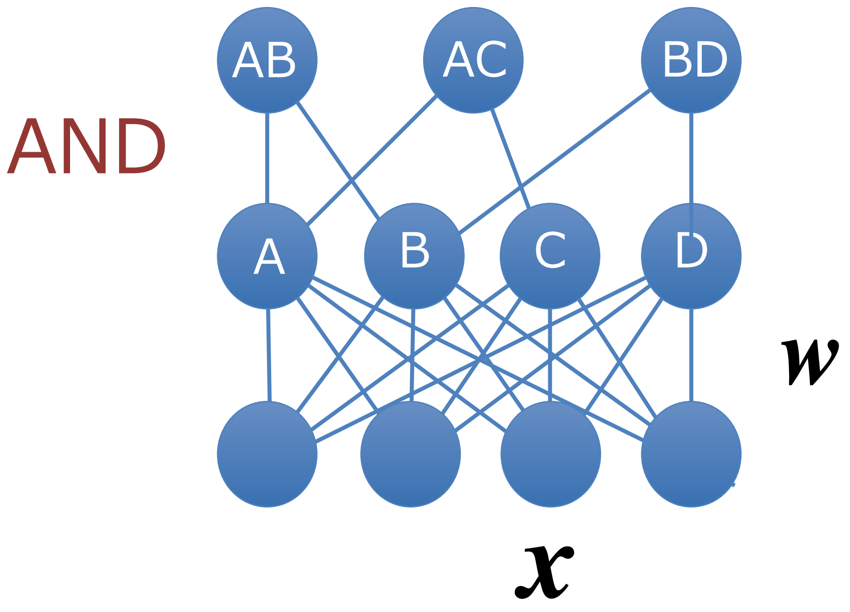



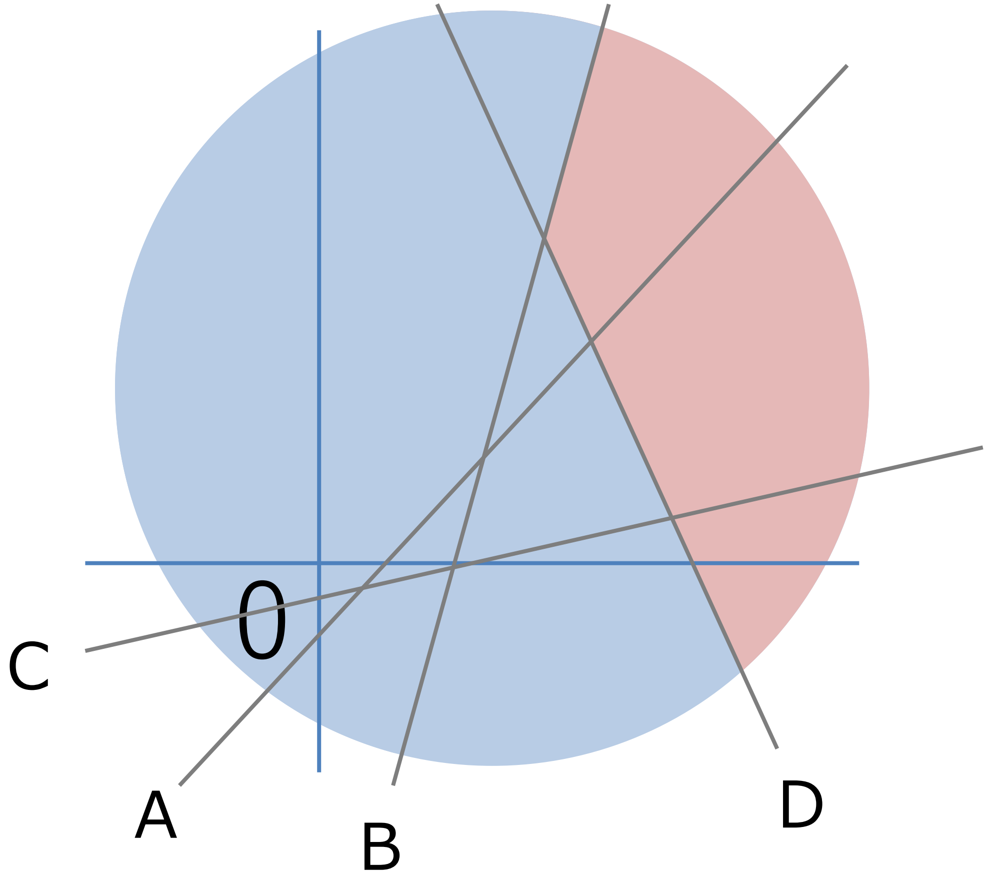
# MLPは何を実現するか:
# 論理関数としての解釈



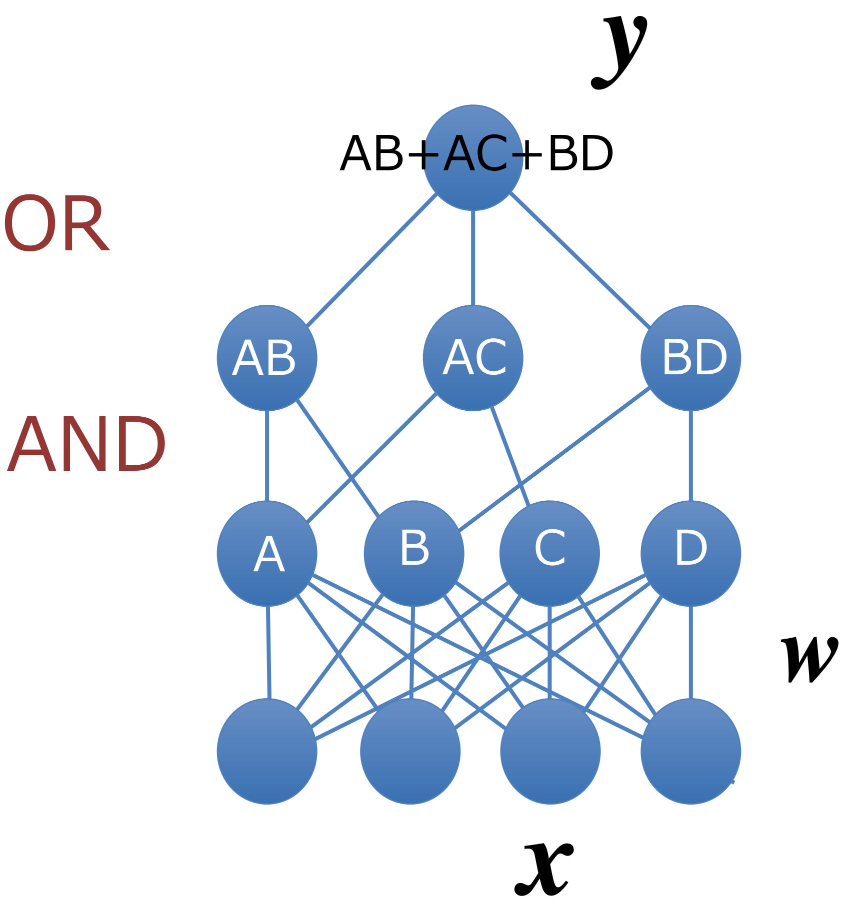



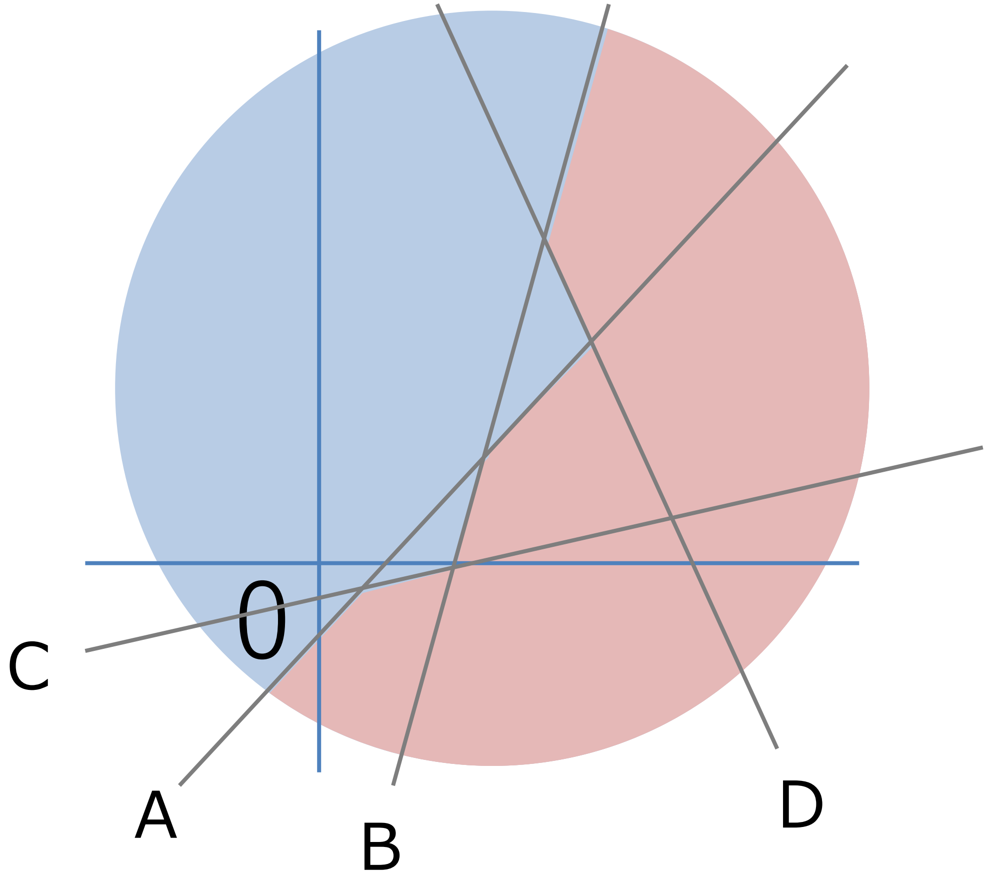
# MLPは何を実現するか:
# 論理関数としての解釈



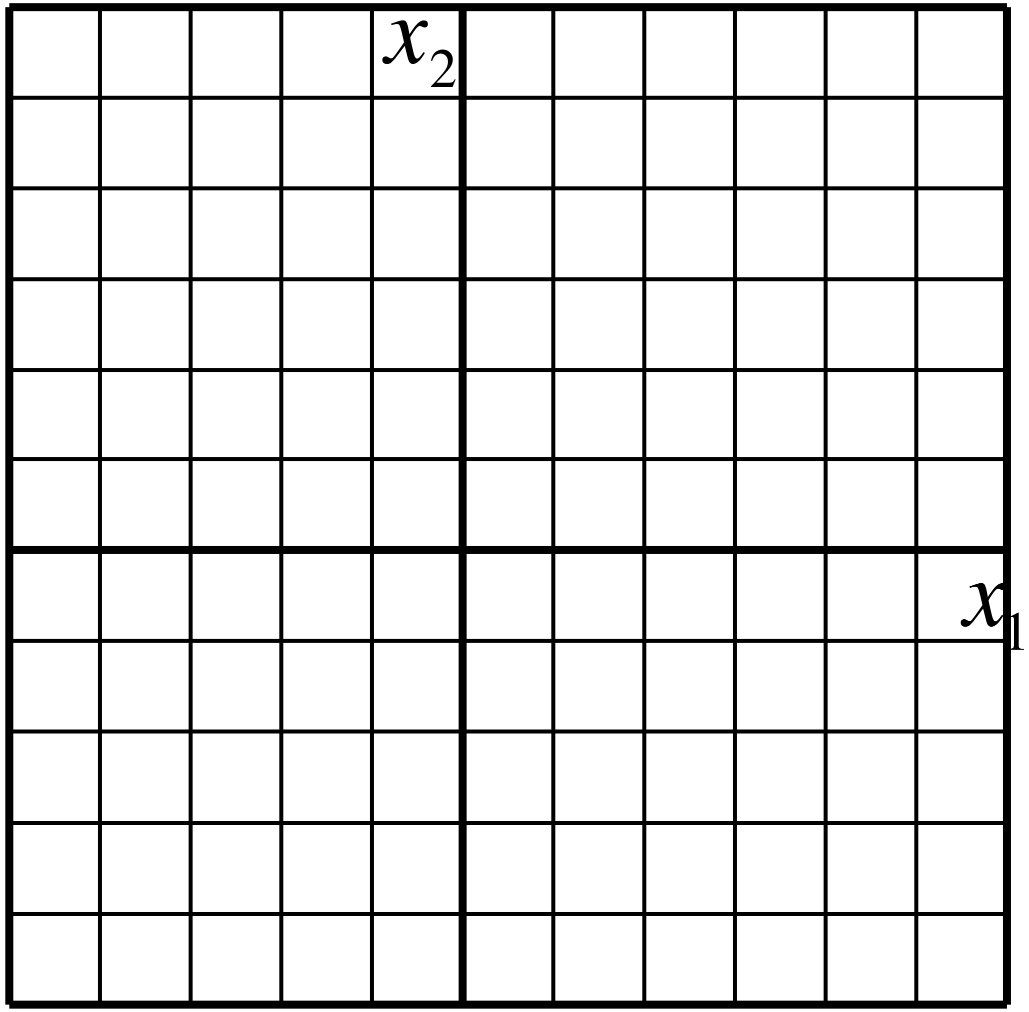



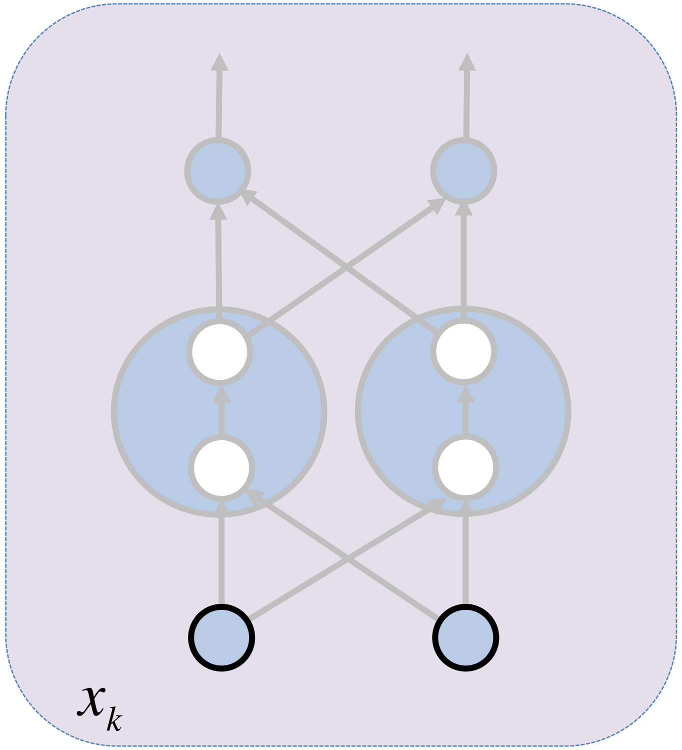
# MLPは何を実現するか:
# 論理関数としての解釈



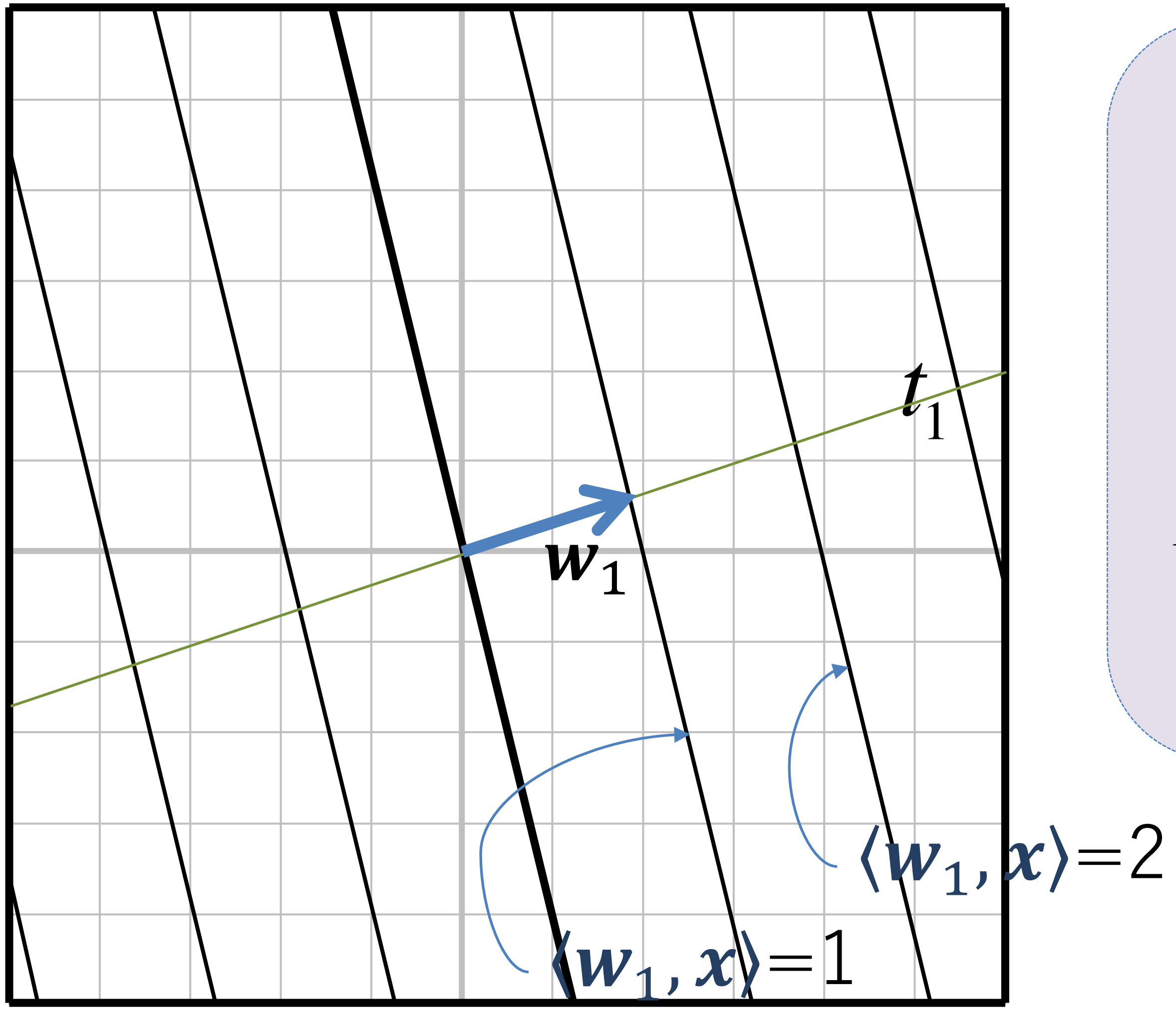
# MLPは何を実現するか: 座標変換器としての解釈

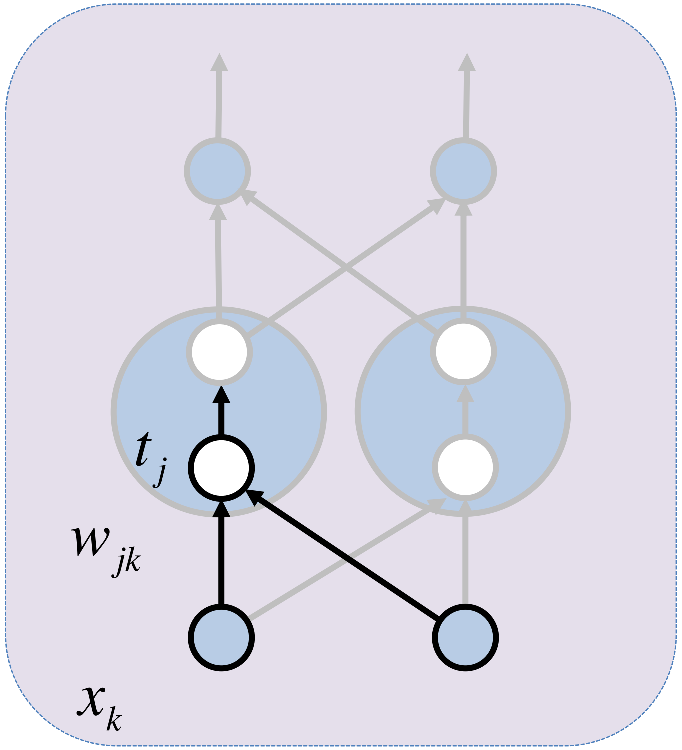



$$\boldsymbol{x} = (x_k)$$

# MLPは何を実現するか: 座標変換器としての解釈

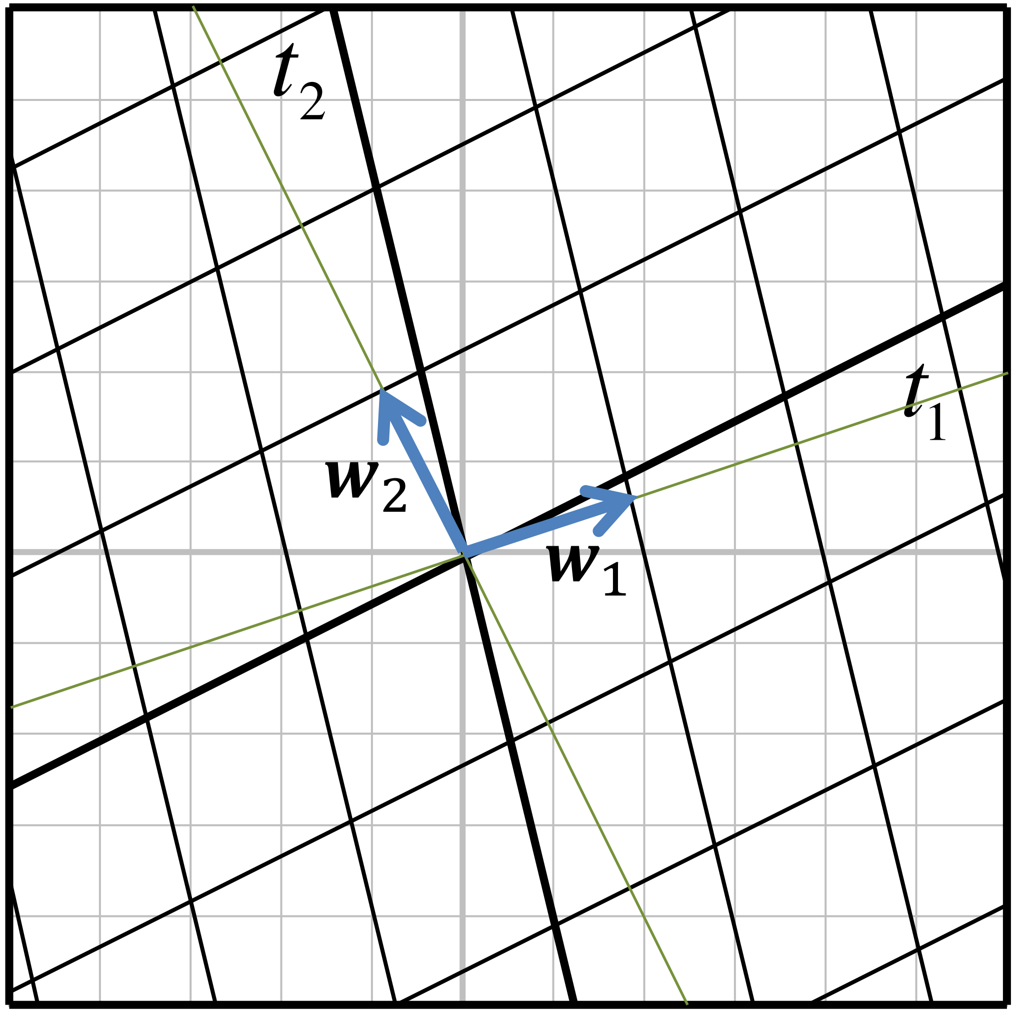


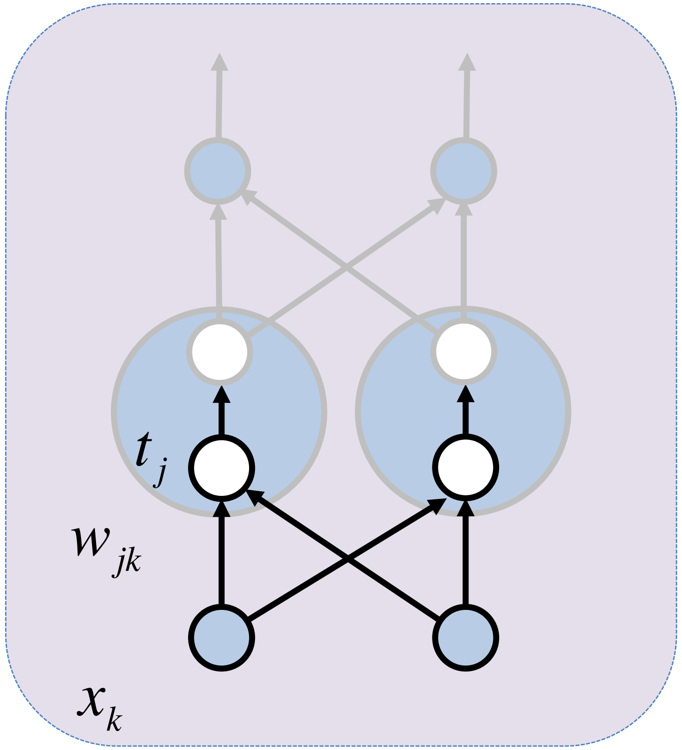




$$t_j = \sum_k w_{jk} x_k$$

$$\boldsymbol{t} = W\boldsymbol{x}$$

# MLPは何を実現するか:
# 座標変換器としての解釈

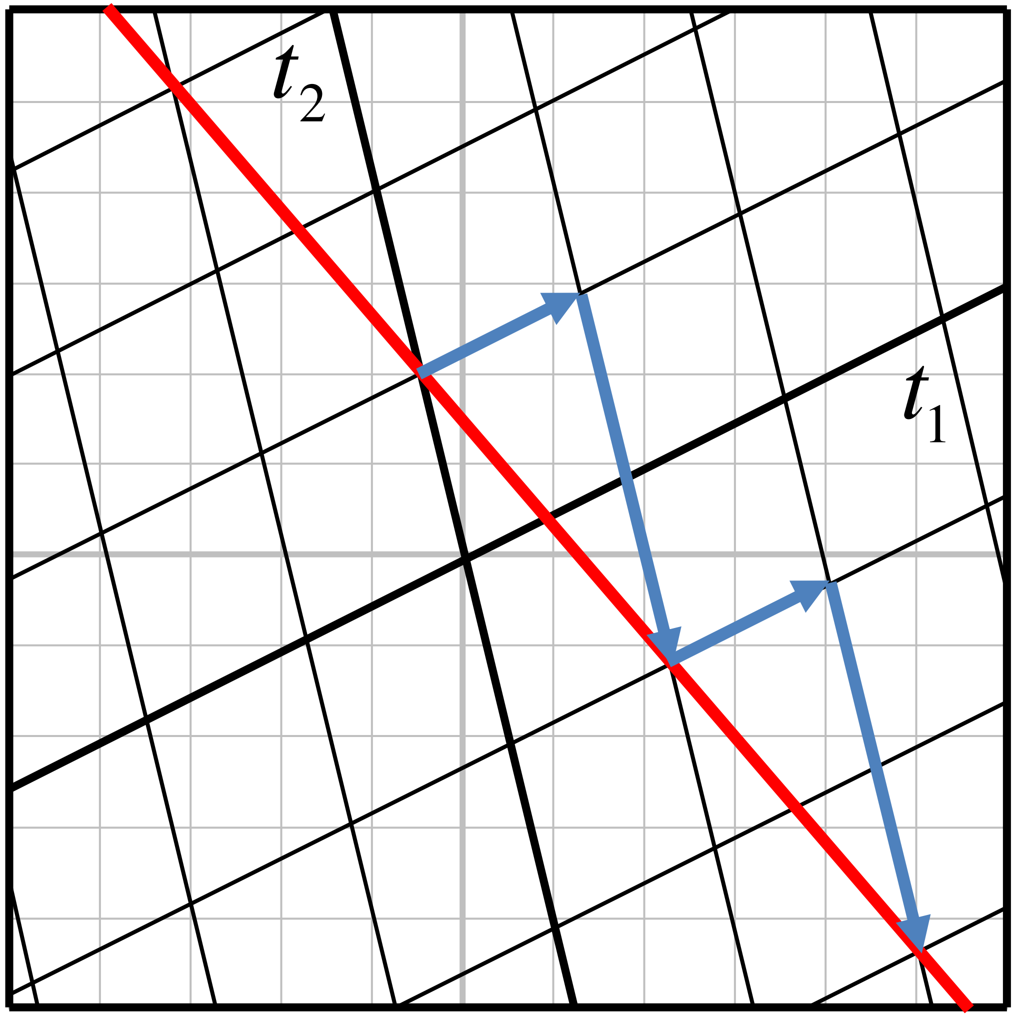


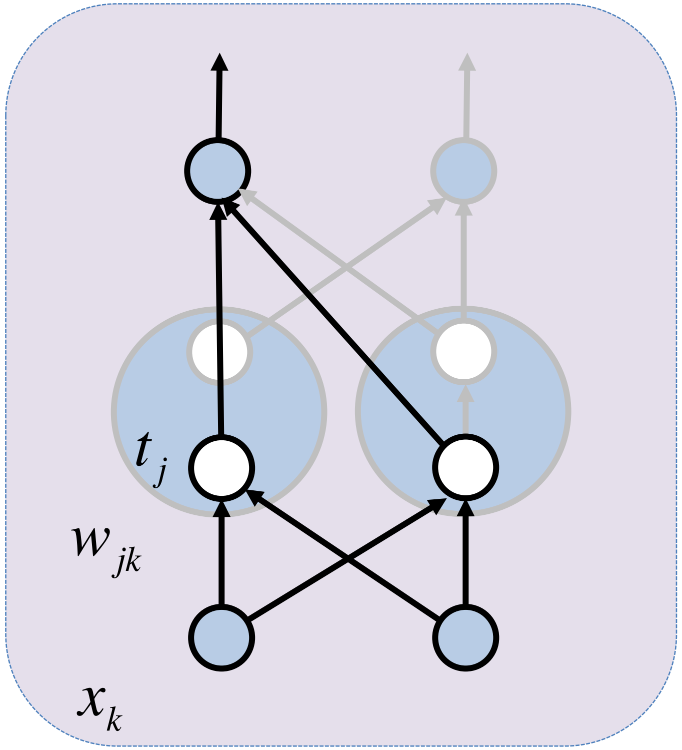




$$t_j = \sum_k w_{jk} x_k$$

$$\boldsymbol{t} = W\boldsymbol{x}$$

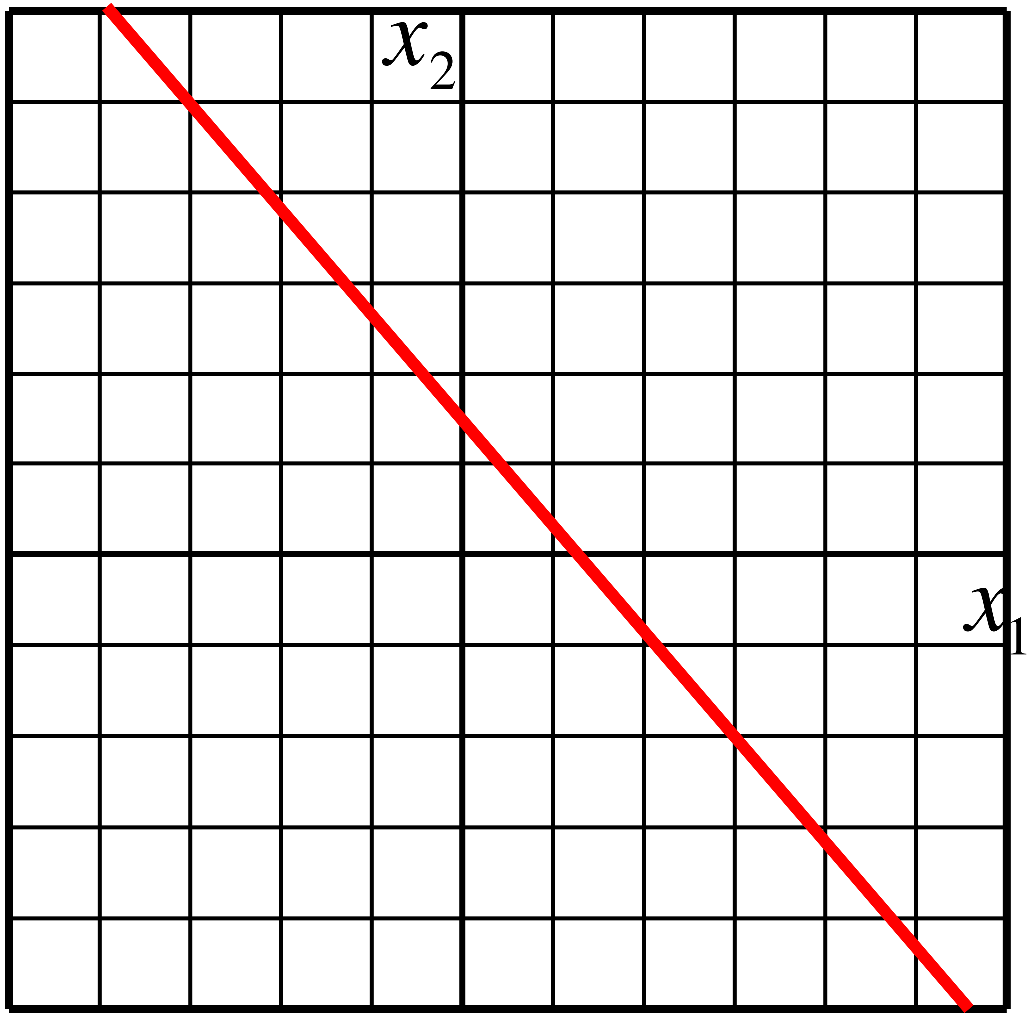
# MLPは何を実現するか:
# 座標変換器としての解釈

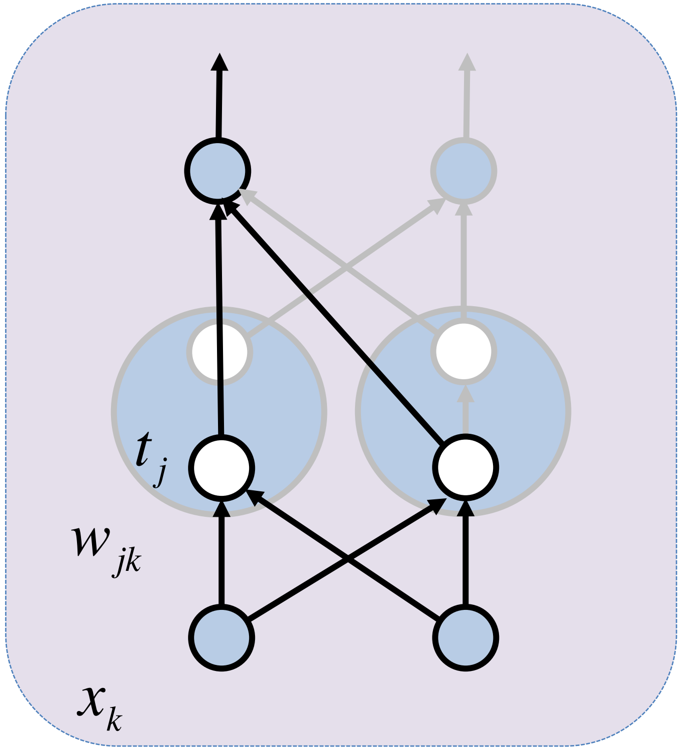



座標変換後の座標系における直線は，座標変換前の座標系においても直線

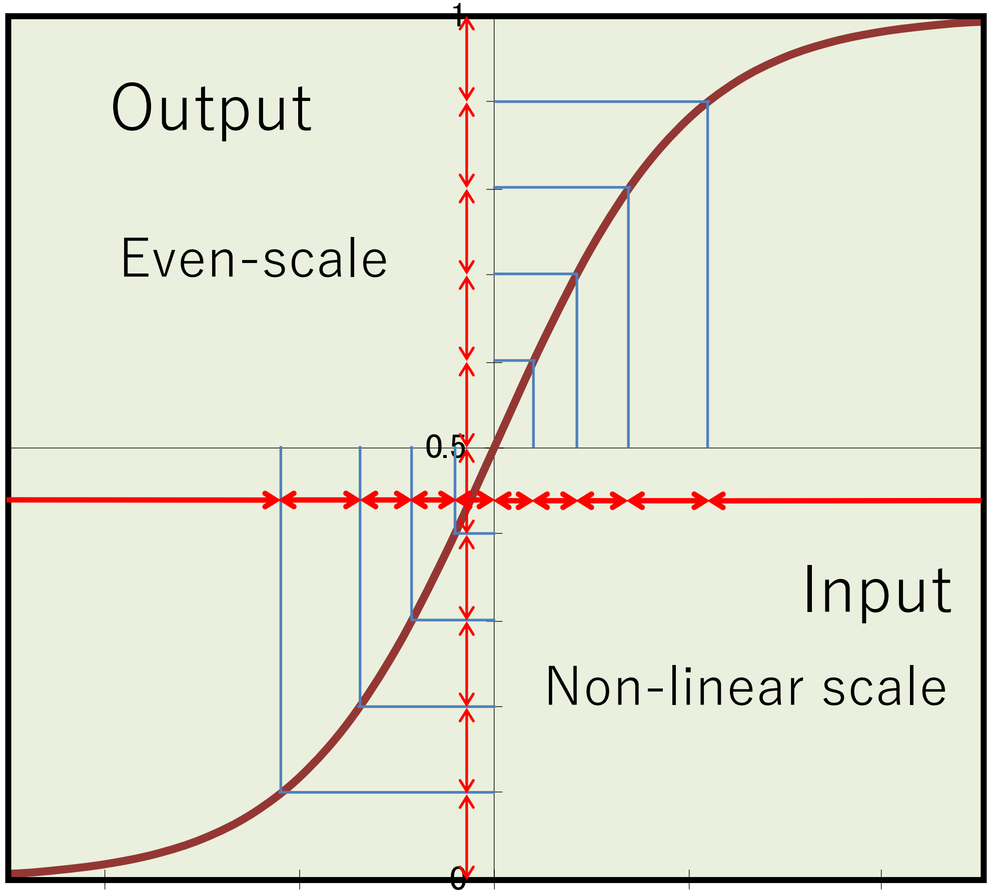
# MLPは何を実現するか: 座標変換器としての解釈




座標変換後の座標系における直線は，座標変換前の座標系においても直線

# MLPは何を実現するか: 座標変換器としての解釈

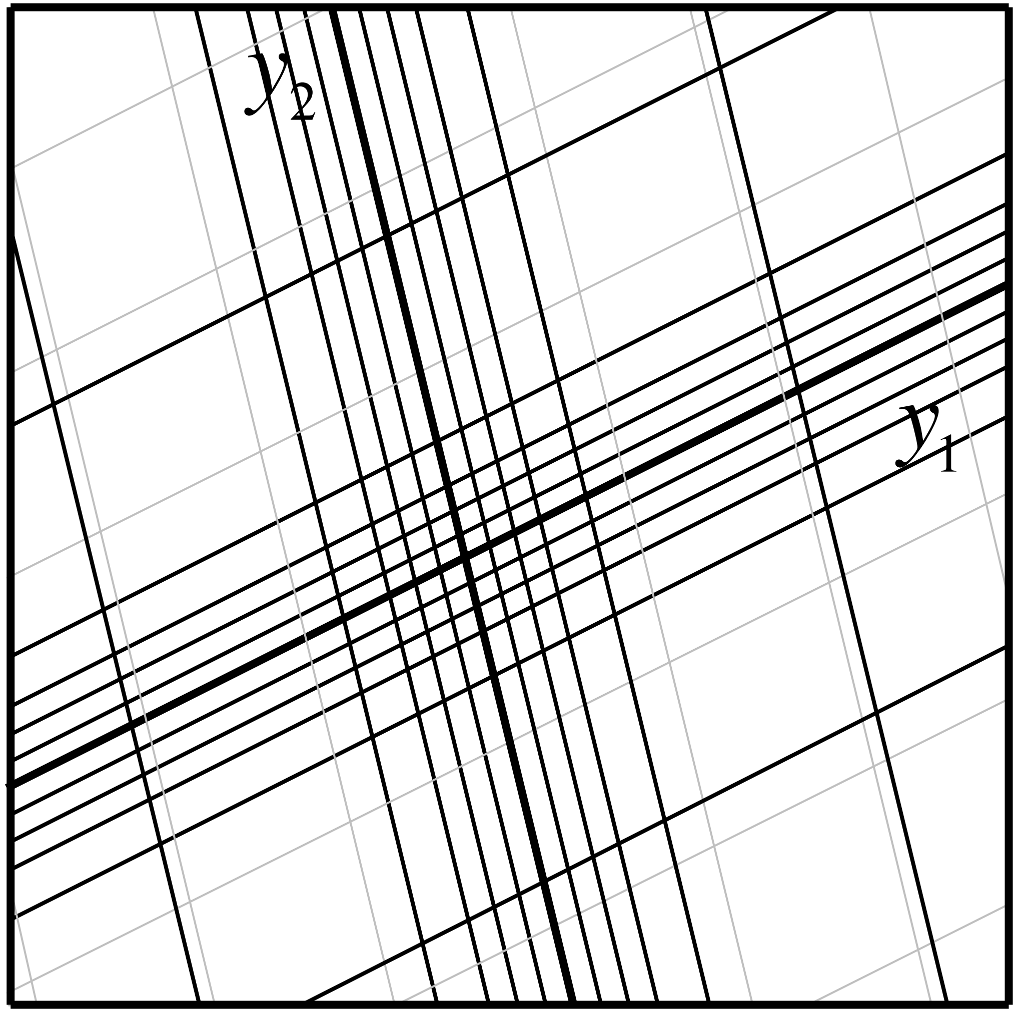


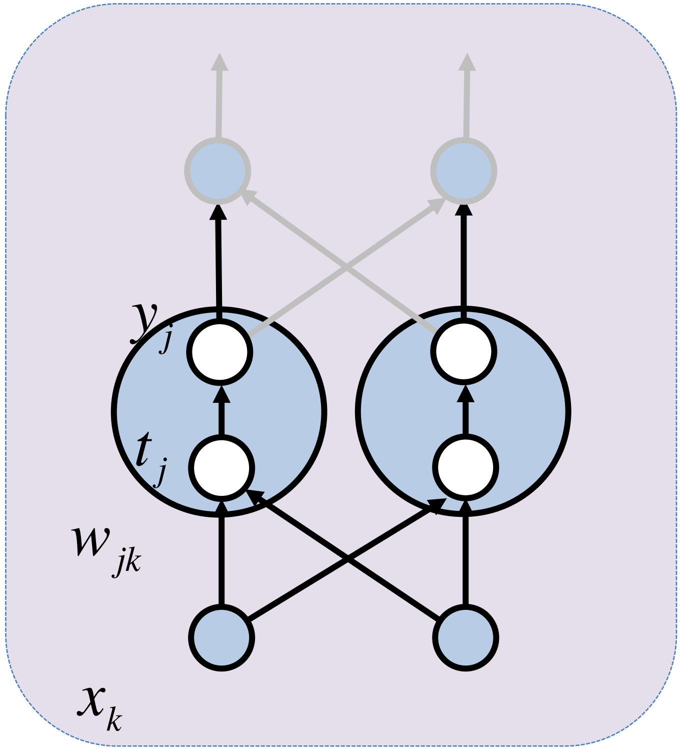

$f(t)$ :非線形関数

例.　シグモイド関数

$$f(t)=\frac{1}{1+e^{-t}}$$

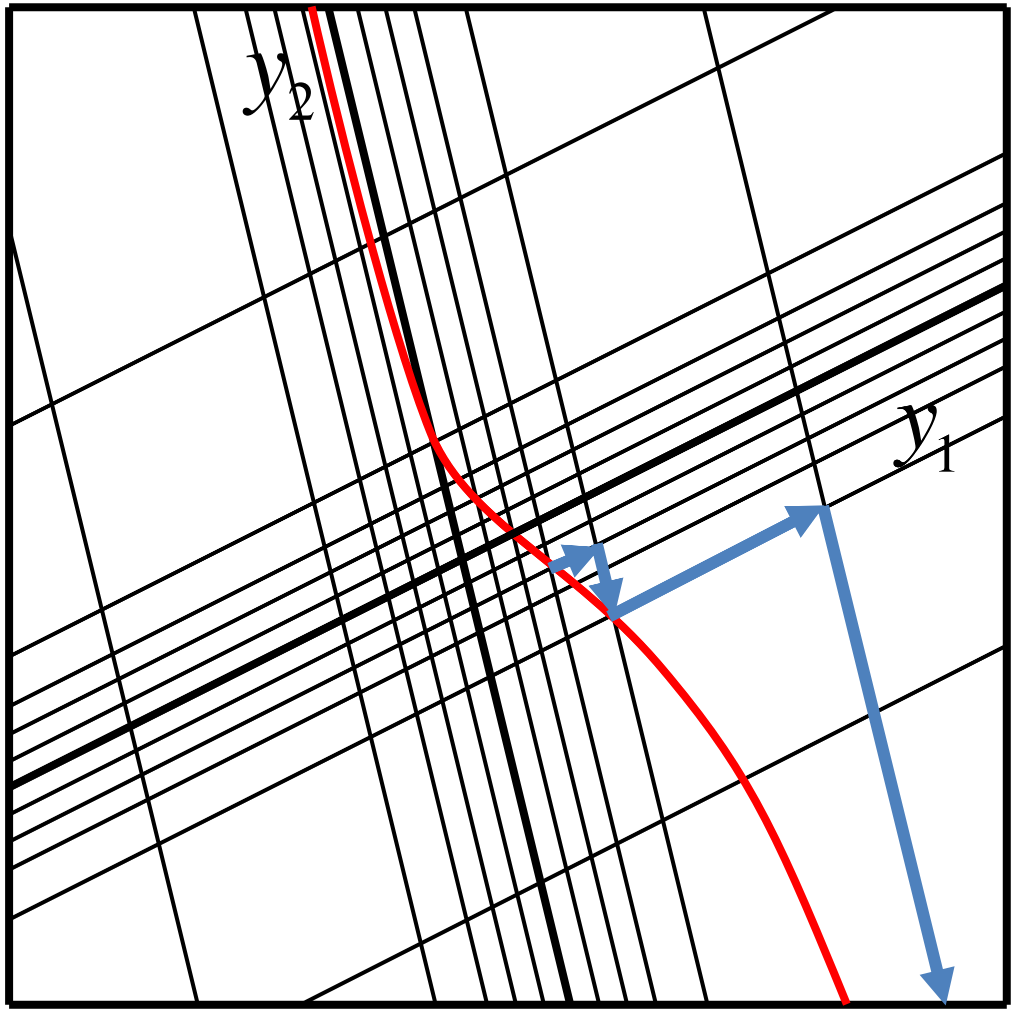
# MLPは何を実現するか:
# 座標変換器としての解釈

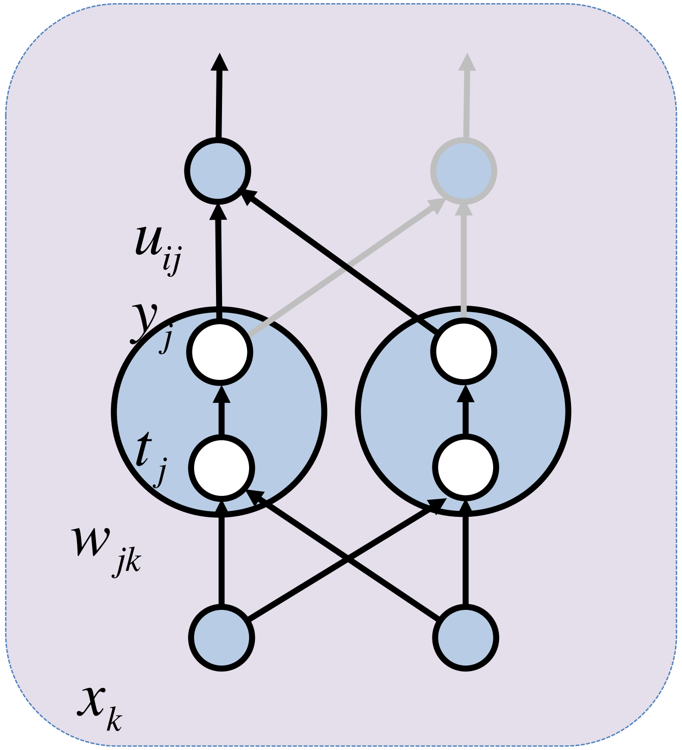


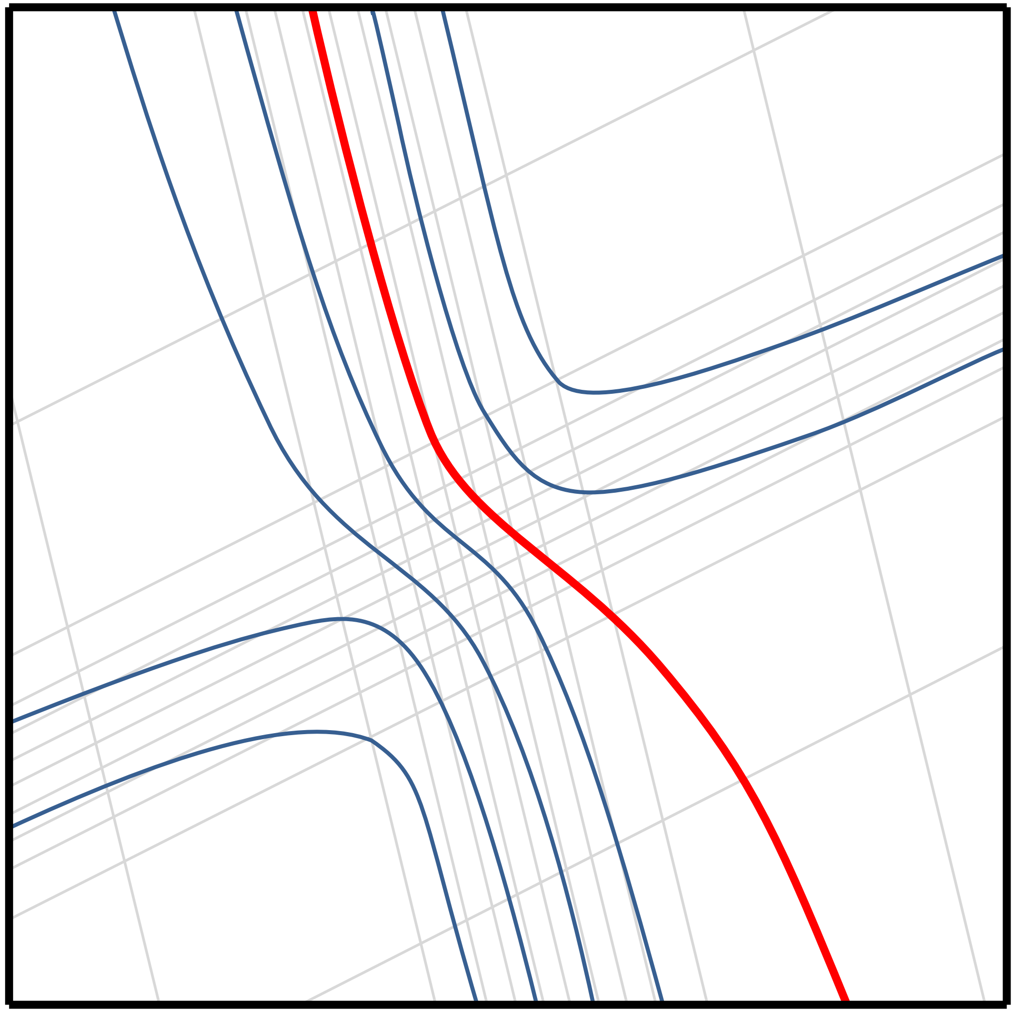

非線形活性化関数は，目盛りの振り直しをする。

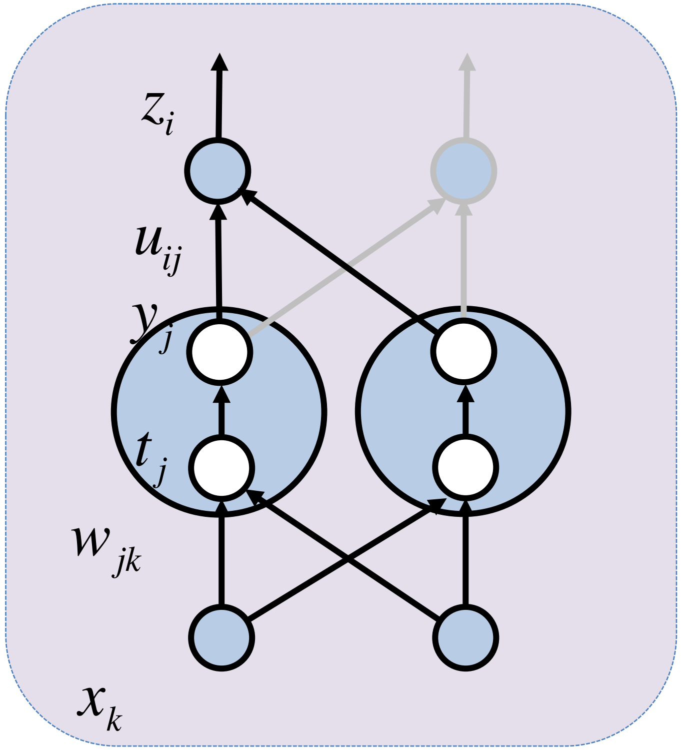
# MLPは何を実現するか:
# 座標変換器としての解釈



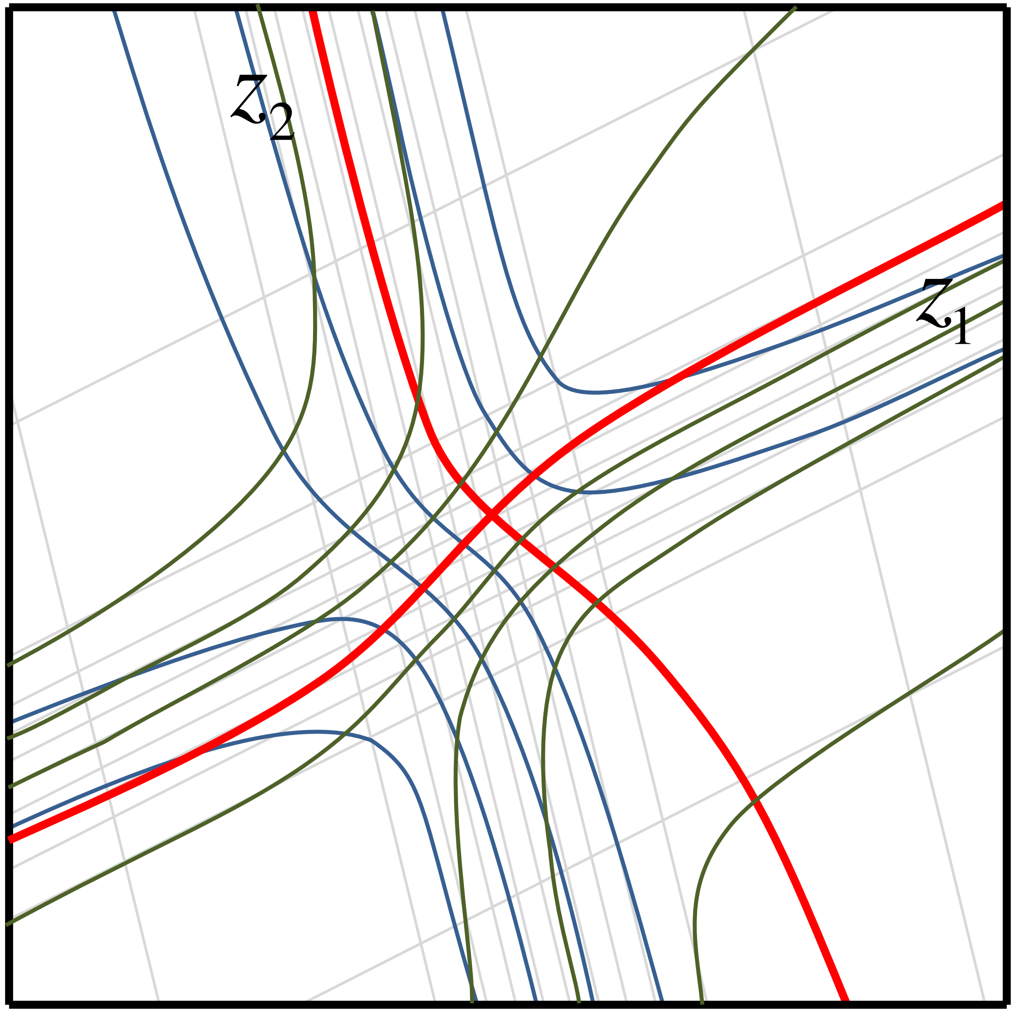
# MLPは何を実現するか:
# 座標変換器としての解釈

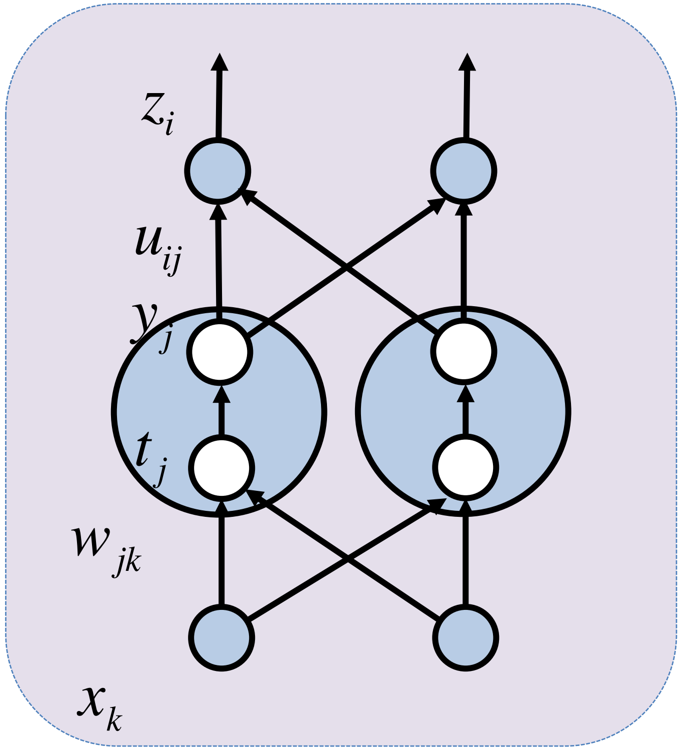





傾き -2 の直線の集合

# MLPは何を実現するか:
# 座標変換器としての解釈




傾き 1/2 の直線の集合

# MLPは何を実現するか: 加法モデルとしての解釈

## 加法モデル

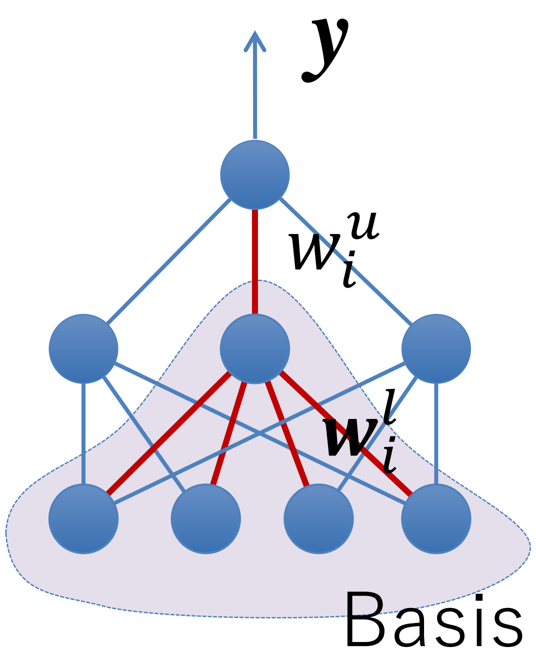
$$y(x) = \sum_{i=1}^{N} w_i \underline{\phi_i(\boldsymbol{x})}$$

Basis

## 多層パーセプトロン

$$y(x) = \sum_{i=1}^{N} w_i^u f\left(\boldsymbol{w}_i^{l\,T} \boldsymbol{x}\right),$$

$$\phi_i(\boldsymbol{x}) = f\left(\boldsymbol{w}_i^{l\,T} \boldsymbol{x}\right)$$




MLPは加法モデル。
各々の基底はデータで学習可能なパラメタ $\boldsymbol{w}_i^l$ をもつ。
出力 $y$ は，基底の重み付き和。

# MLPは何を実現するか: 加法モデルとしての解釈

## 加法モデル

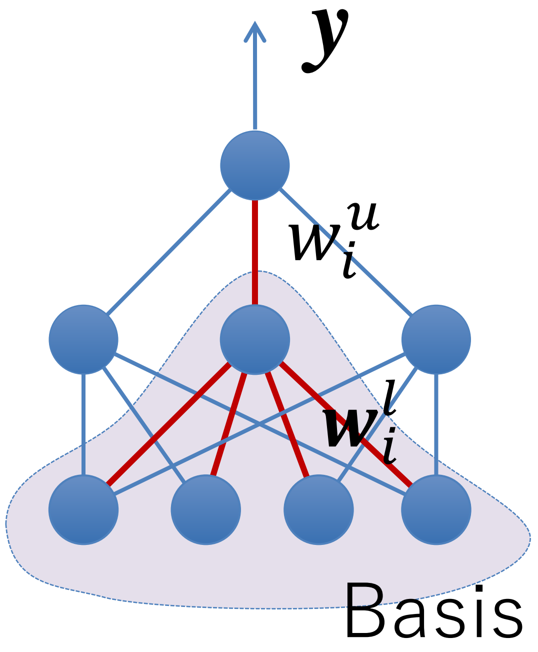
$$y(x) = \sum_{i=1}^{N} w_i \phi_i(\boldsymbol{x})$$

Basis

## 多層パーセプトロン

$$y(x) = \sum_{i=1}^{N} w_i^u f\left(\boldsymbol{w}_i^{l^T} \boldsymbol{x}\right),$$

$$\phi_i(\boldsymbol{x}) = f\left(\boldsymbol{w}_i^{l^T} \boldsymbol{x}\right)$$

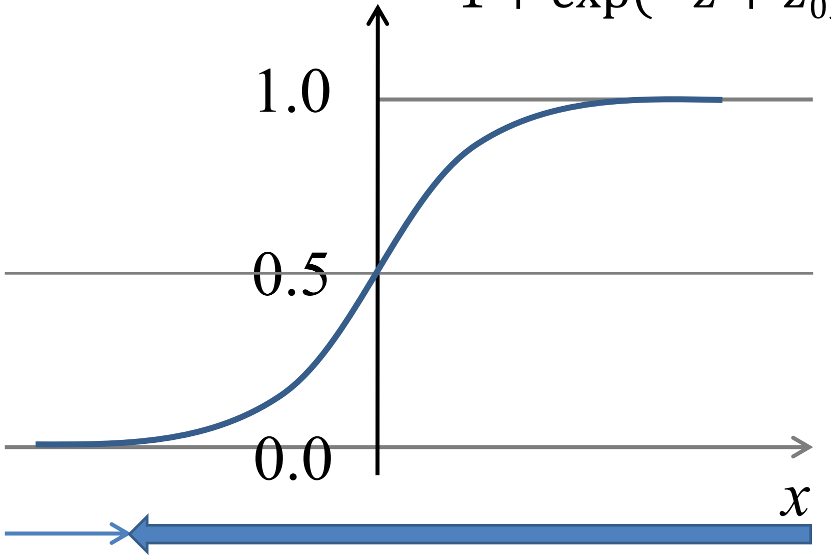



MLPは加法モデル。
各々の基底はデータで学習可能なパラメタ $\boldsymbol{w}_i^l$ をもつ。
出力 $y$ は，基底の重み付き和。

# MLPは何を実現するか:
# 加法モデルとしての解釈

## sigmoid

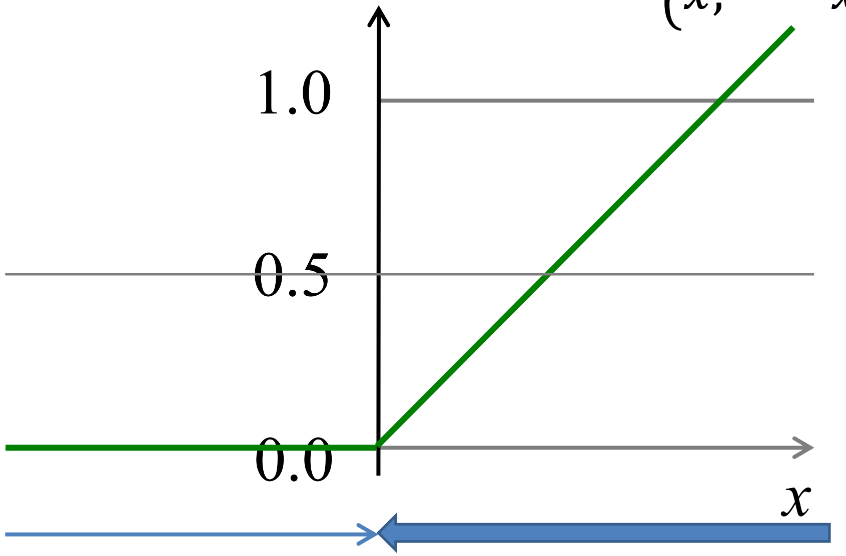
$$f(z) = \frac{1}{1 + \exp(-z + z_0)}$$




出力は 0

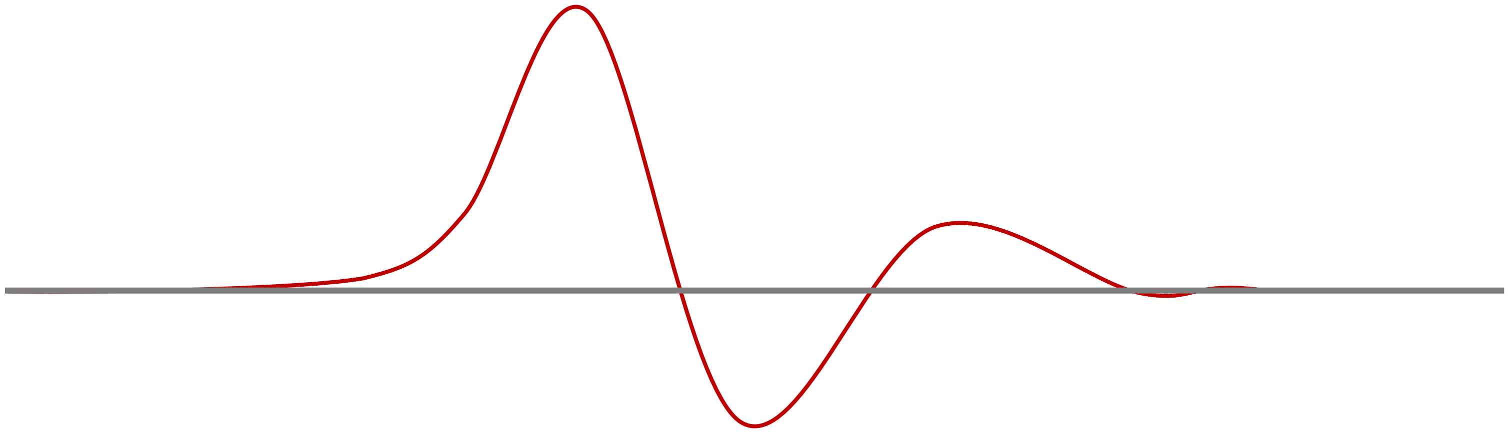
## Rectified Linear Unit

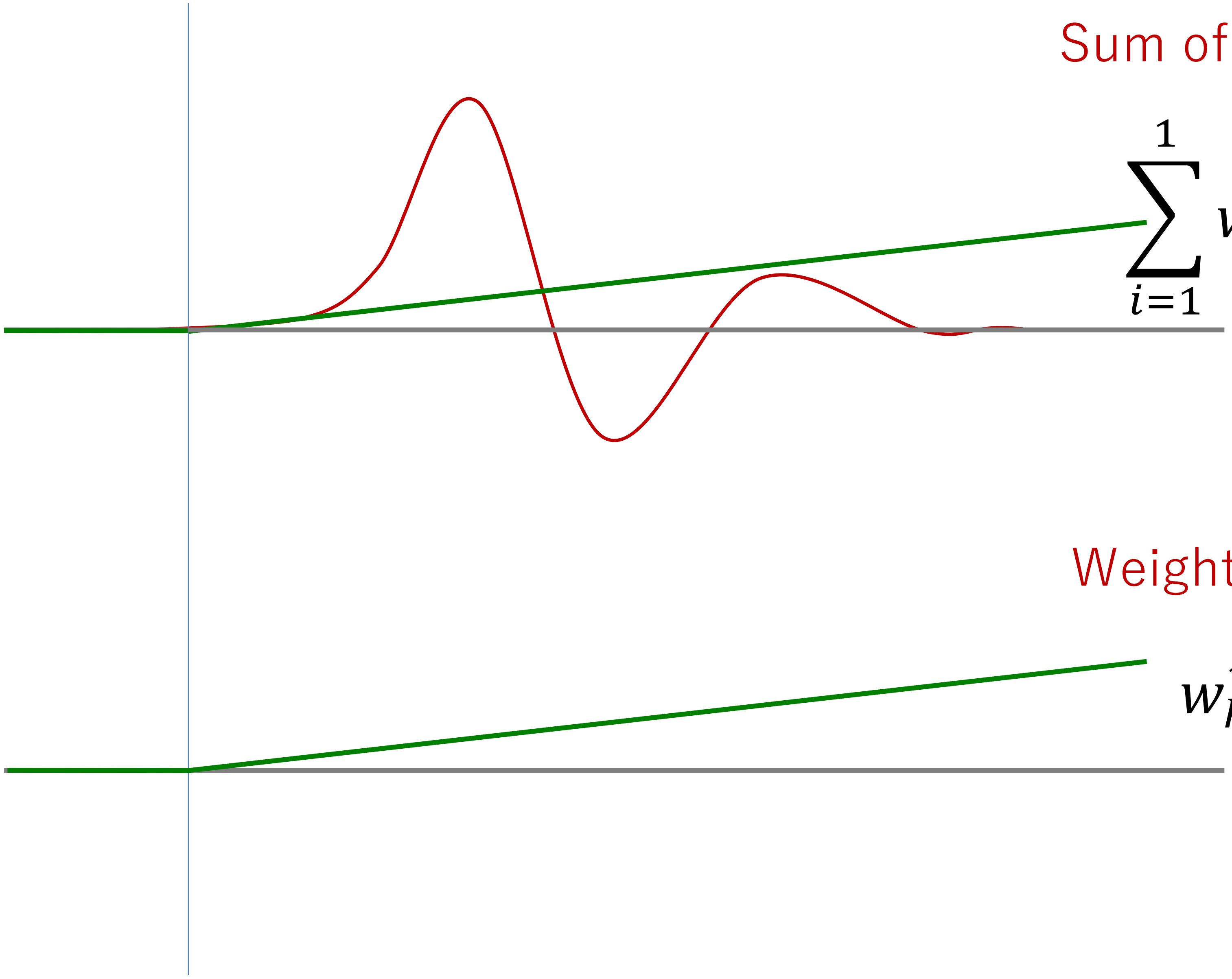
$$f(x) = \begin{cases} 0, & x < 0 \\ x, & x \geq 0 \end{cases}$$




出力は0

基底を加えることは，ある領域（⟸）を選んで，その領域の関数の値を変えることを意味する。



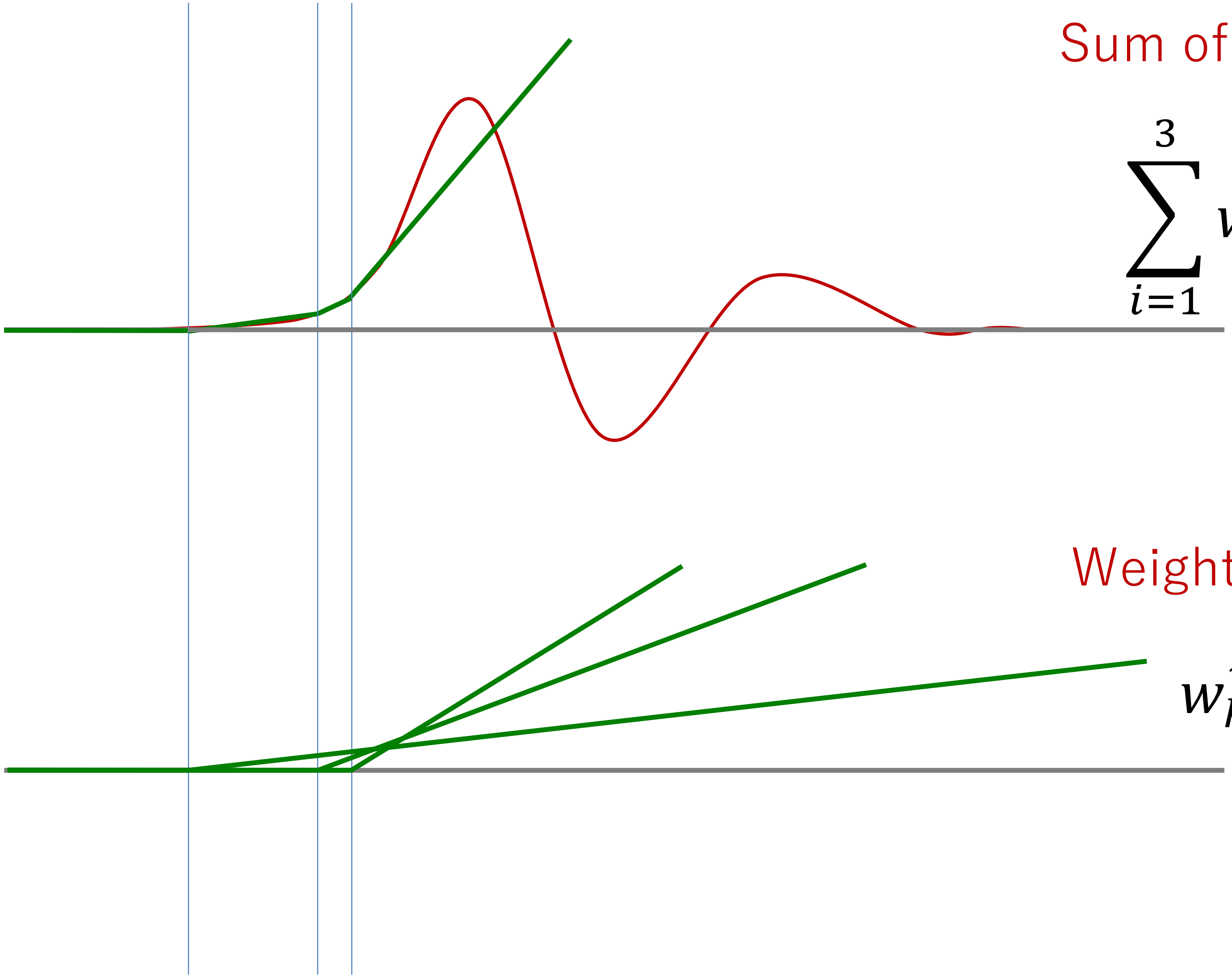


Sum of weighted basis

$$\sum_{i=1}^{1} w_i^u f\left({\boldsymbol{w}_i^l}^T \boldsymbol{x}\right)$$

Weighted basis

$$w_k^u f\left({\boldsymbol{w}_k^l}^T \boldsymbol{x}\right)$$

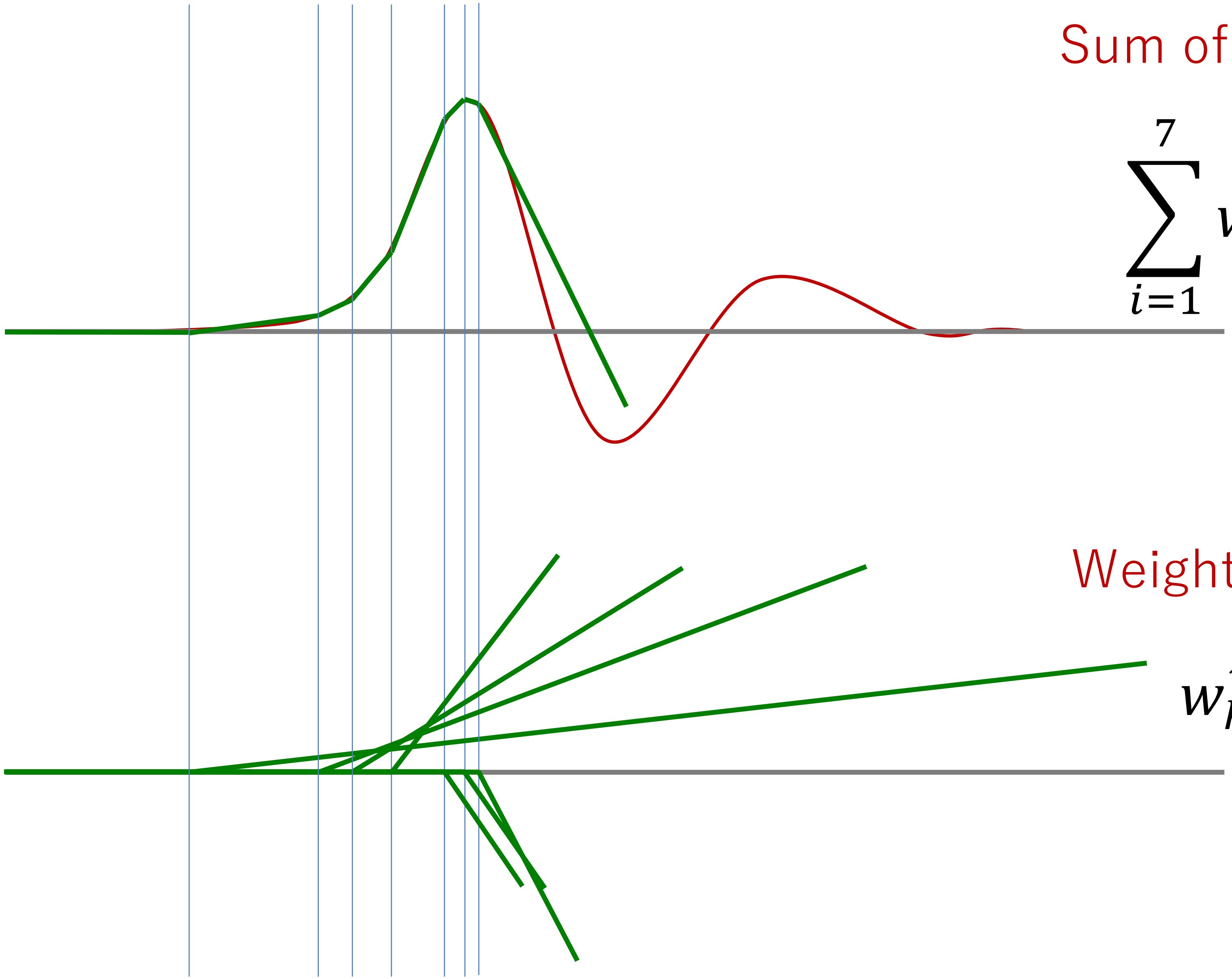


Sum of weighted basis

$$\sum_{i=1}^{3} w_i^u f\left(\boldsymbol{w}_i^{l\,T}\boldsymbol{x}\right)$$

Weighted basis

$$w_k^u f\left(\boldsymbol{w}_k^{l\,T}\boldsymbol{x}\right)$$



Sum of weighted basis

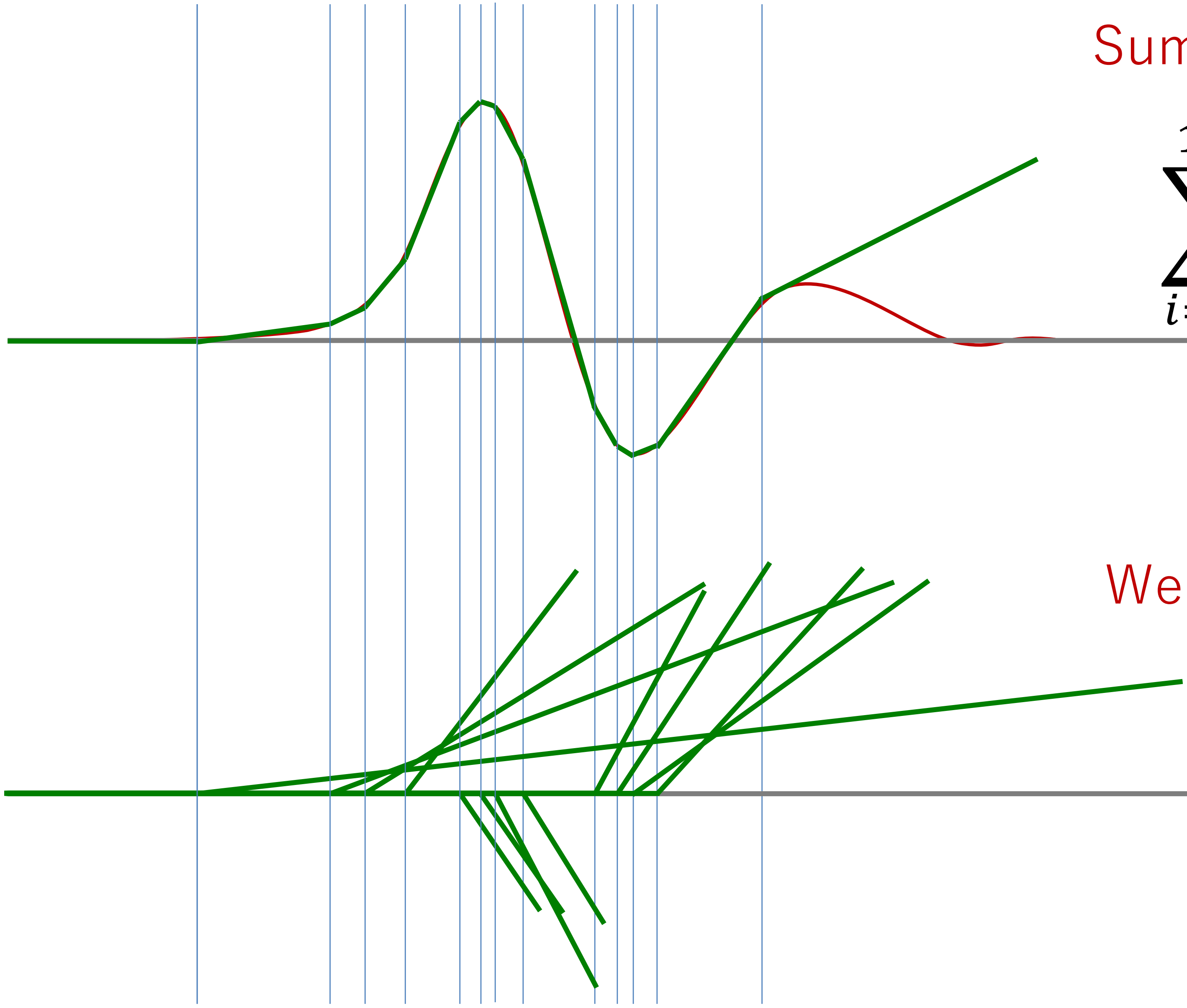
$$\sum_{i=1}^{7} w_i^u f\left({\boldsymbol{w}_i^l}^T \boldsymbol{x}\right)$$

Weighted basis

$$w_k^u f\left({\boldsymbol{w}_k^l}^T \boldsymbol{x}\right)$$

Sum of weighted basis

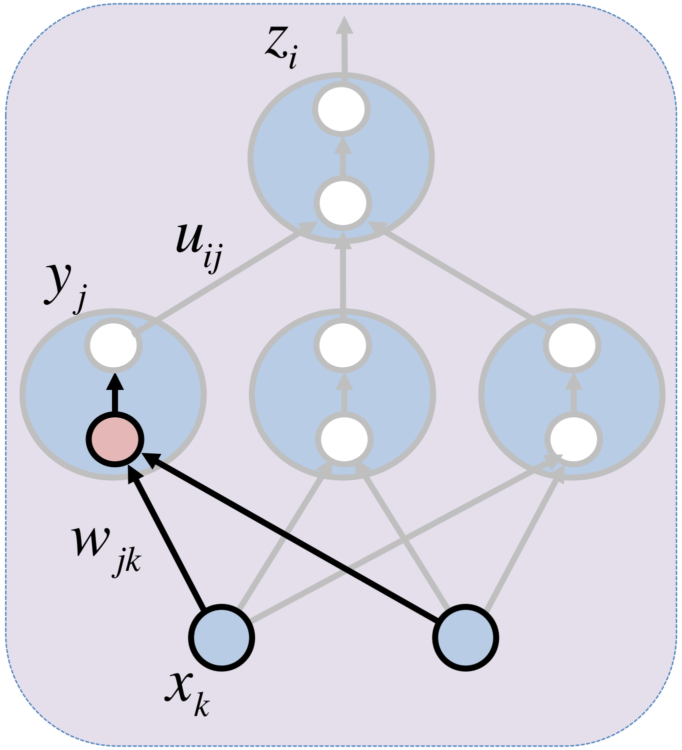
$$\sum_{i=1}^{13} w_i^u f\left(\boldsymbol{w}_i^{l^T}\boldsymbol{x}\right)$$



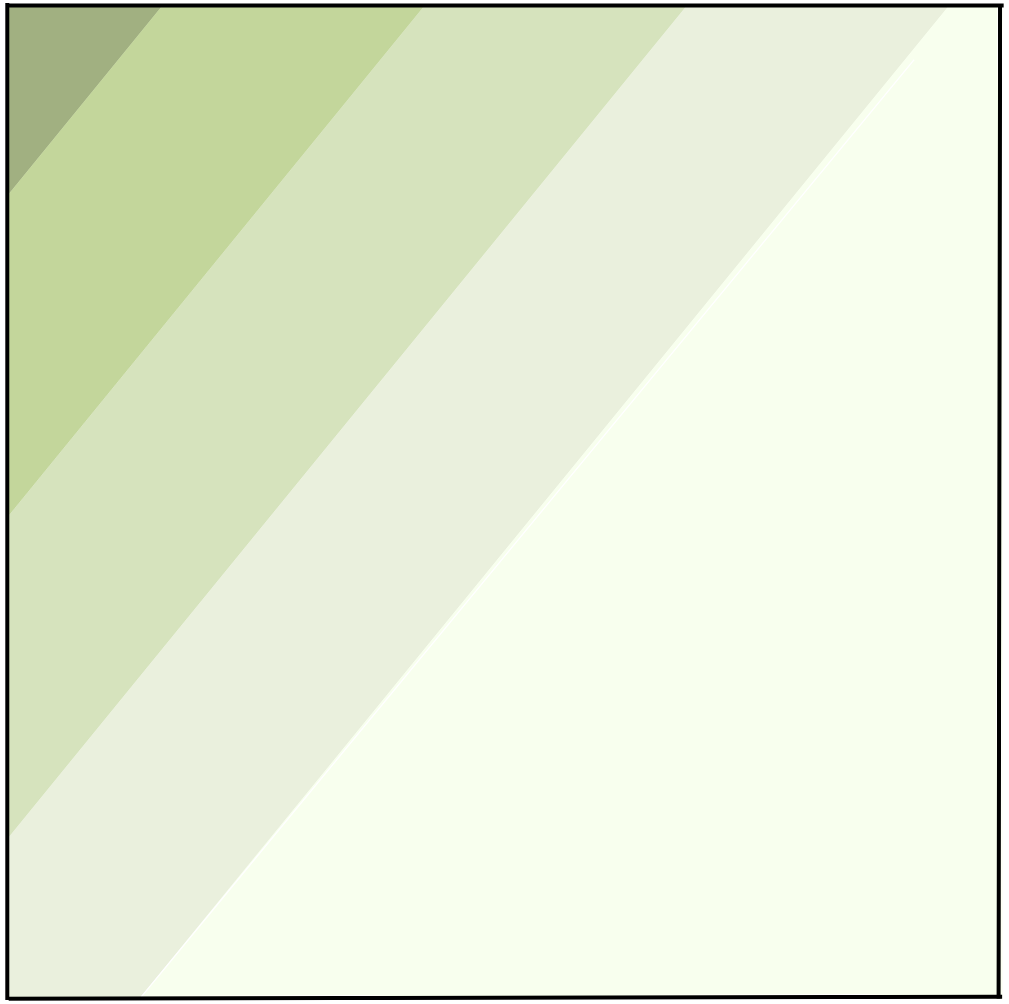
Weighted basis

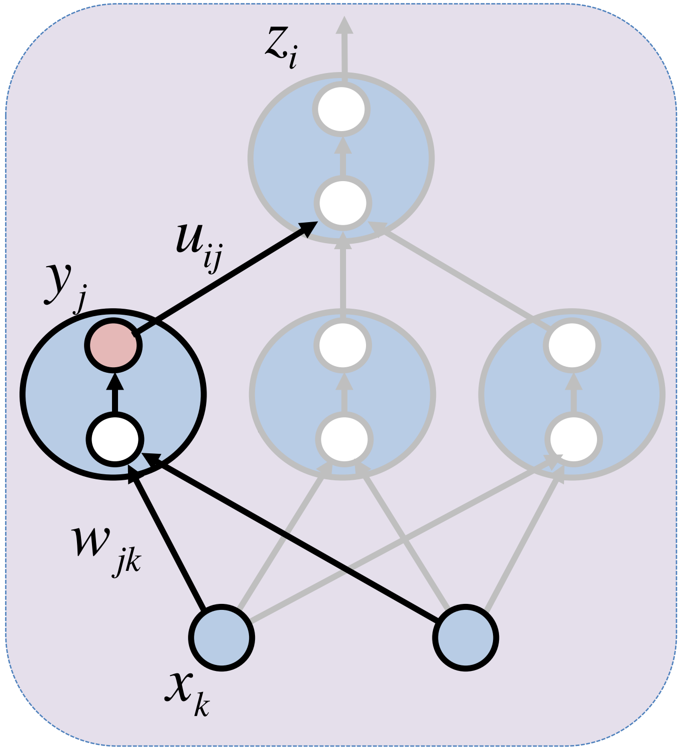
$$w_k^u f\left(\boldsymbol{w}_k^{l^T}\boldsymbol{x}\right)$$

# MLPは何を実現するか:
# 加法モデルとしての解釈

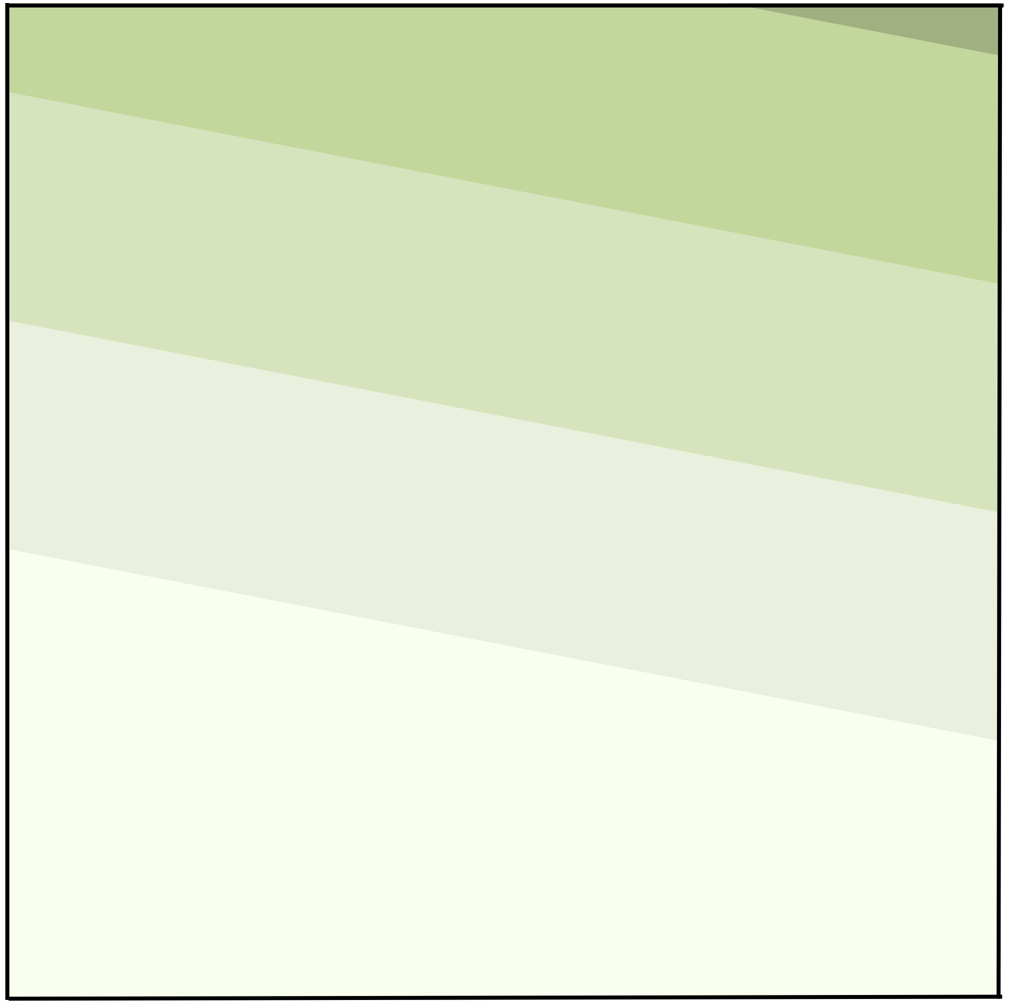


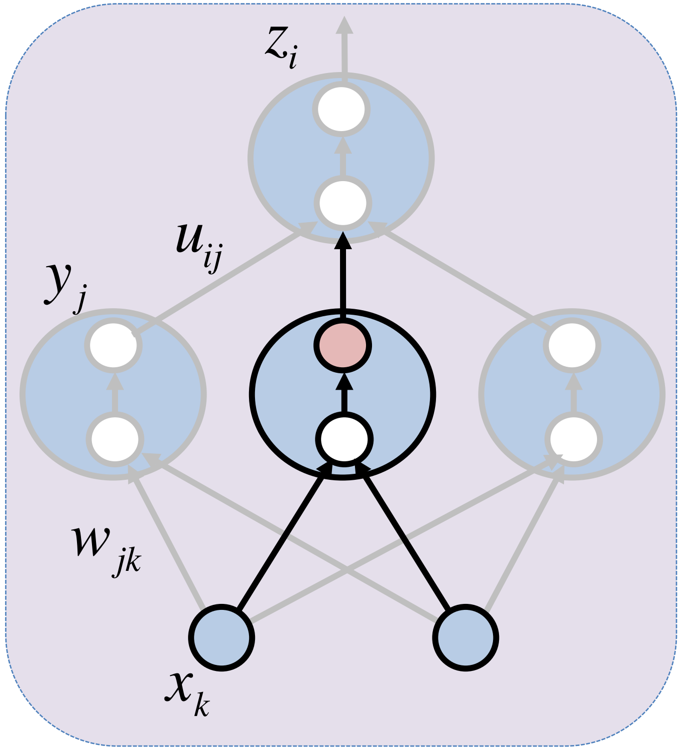
# MLPは何を実現するか:
# 加法モデルとしての解釈



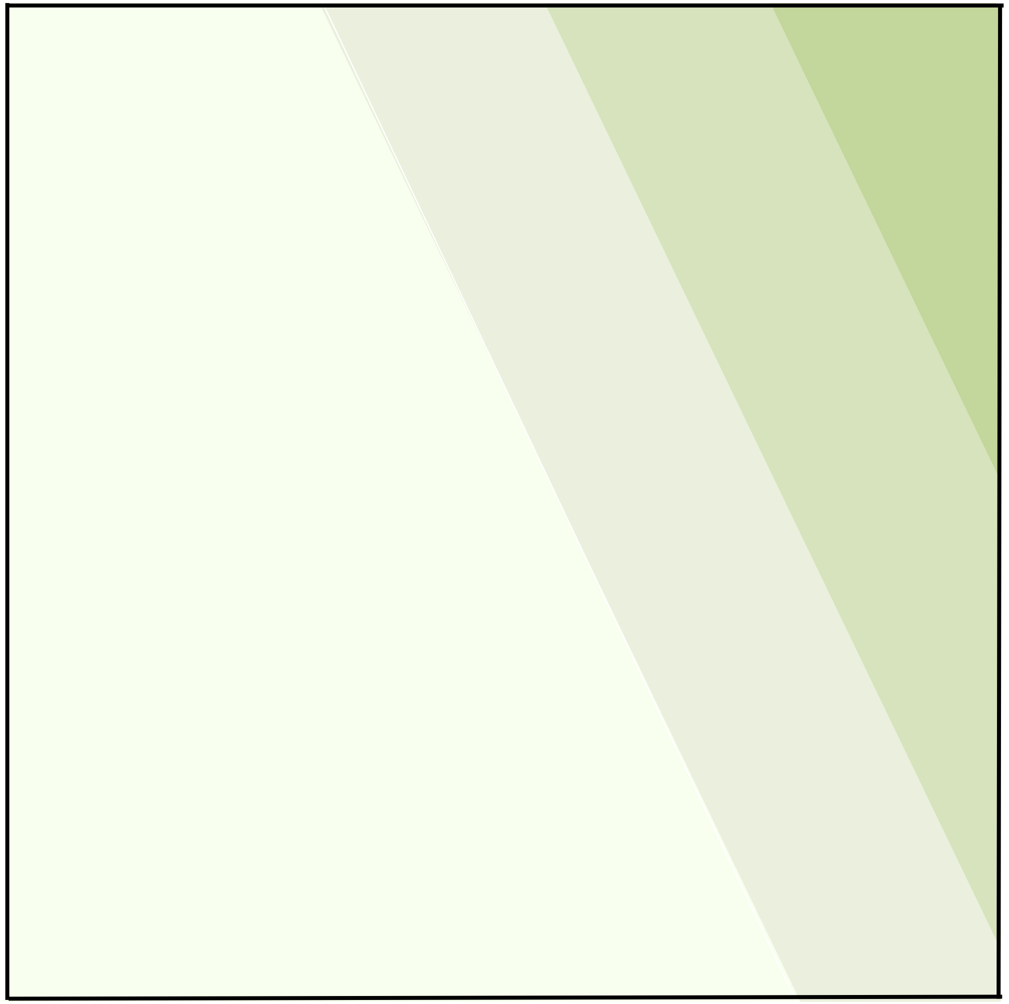


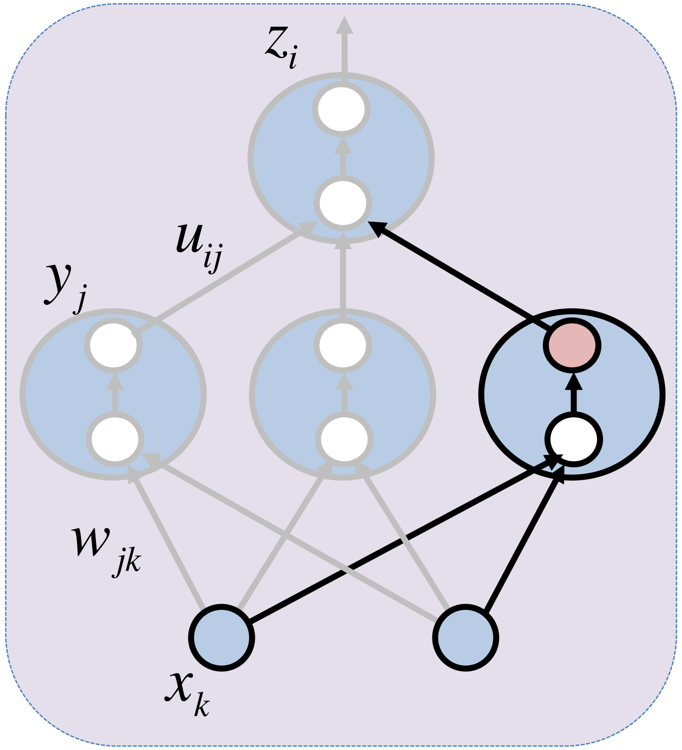
# MLPは何を実現するか:
# 加法モデルとしての解釈



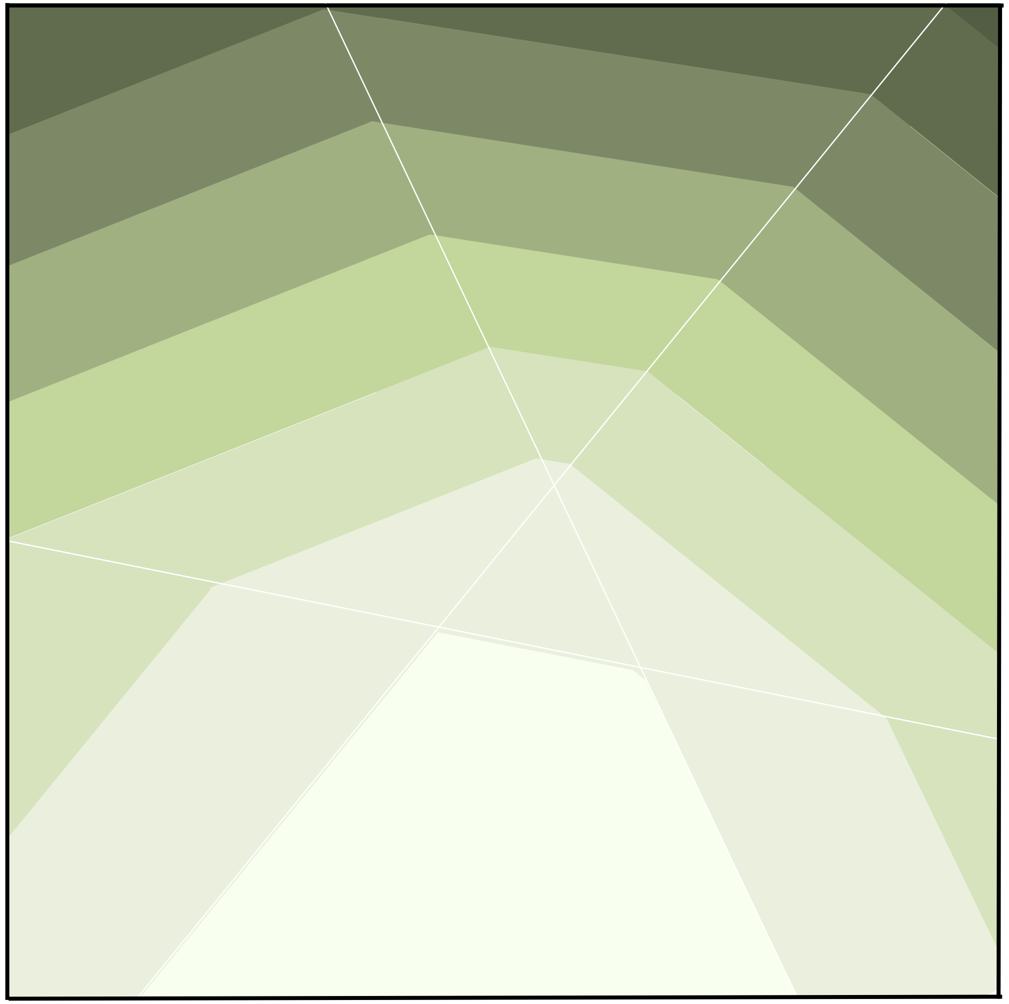


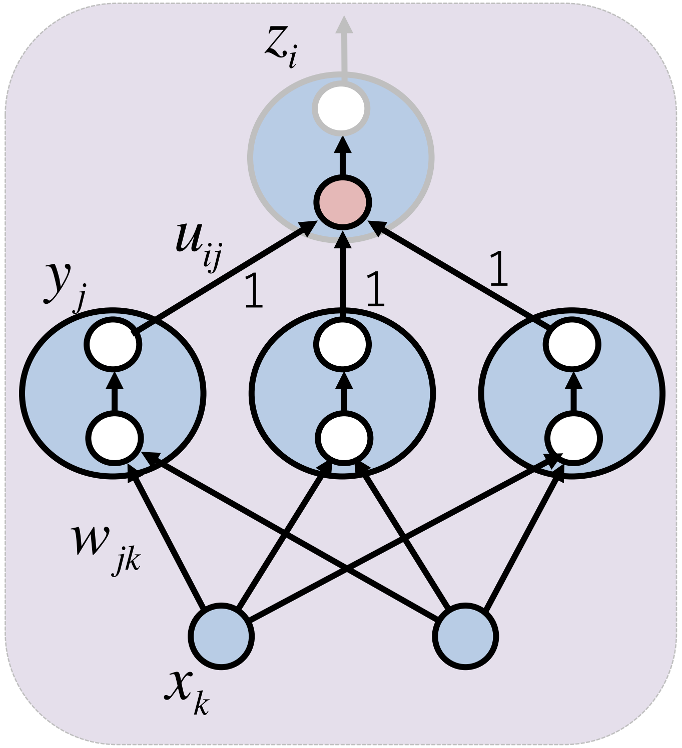
# MLPは何を実現するか:
# 加法モデルとしての解釈



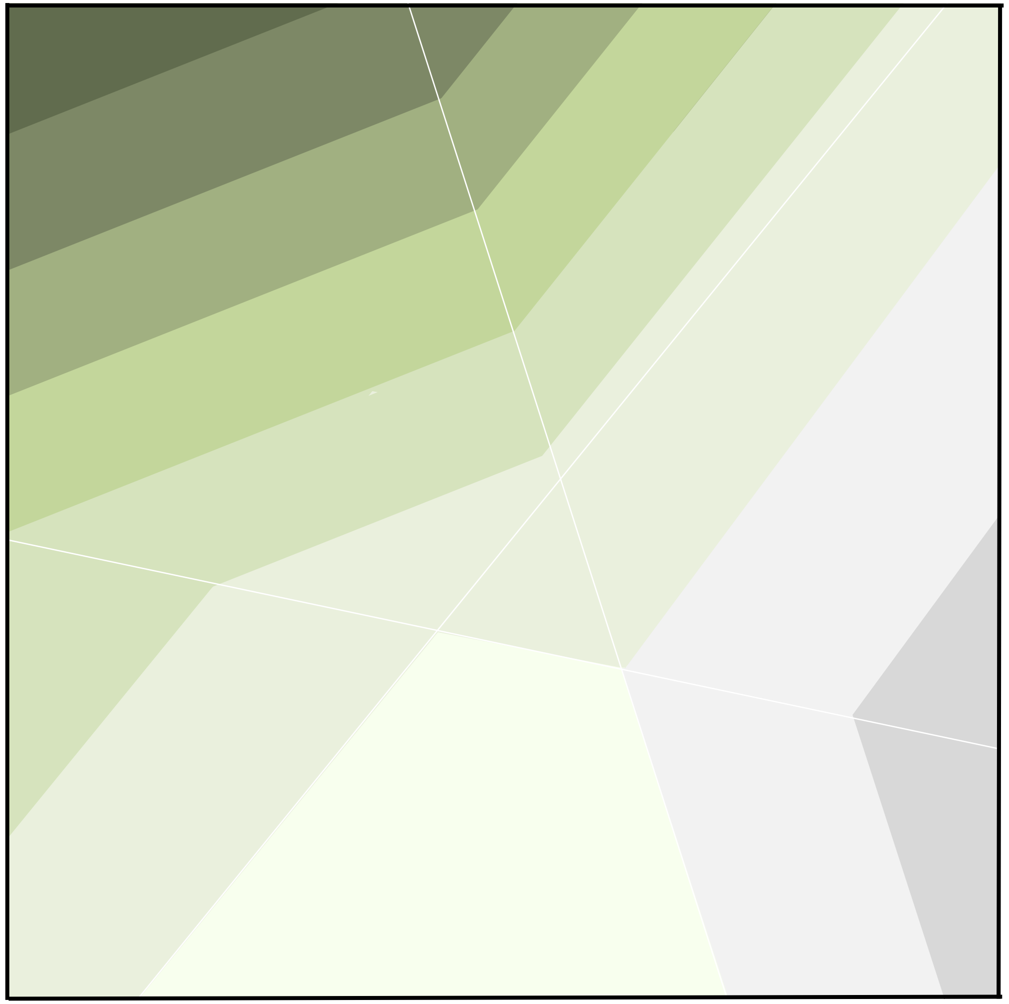


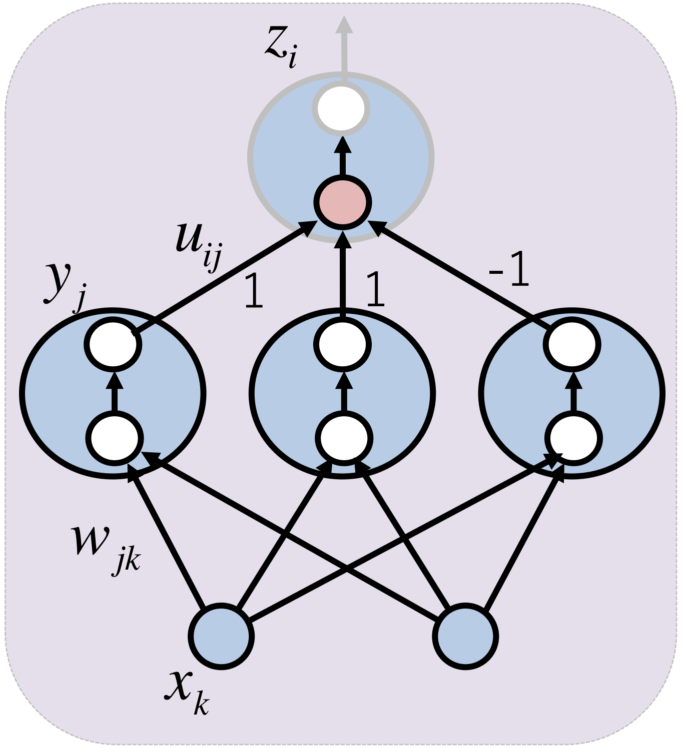
# MLPは何を実現するか:
# 加法モデルとしての解釈





# MLPは何を実現するか:
# 加法モデルとしての解釈





# まとめ

- MLP（多層パーセプトロン）は，パーセプトロンを層状に組み合わせたもの。
- MLPは任意の関数を構成できる。