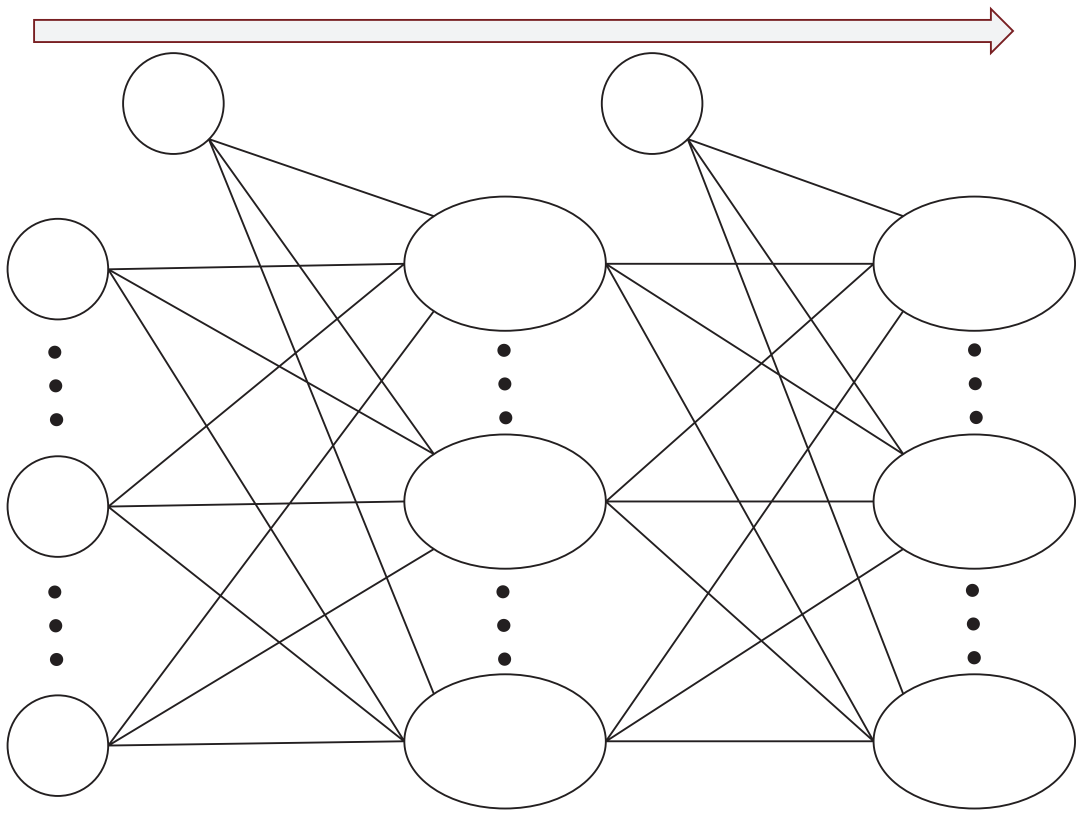
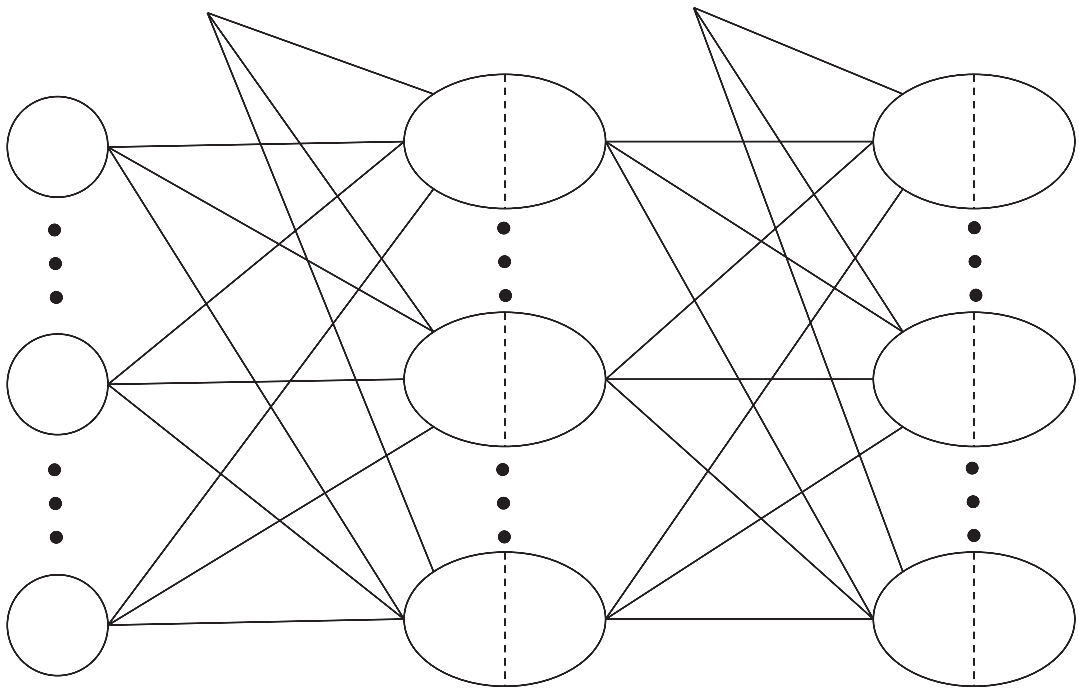
## 伝搬　　　の流れ：　　層→　　層

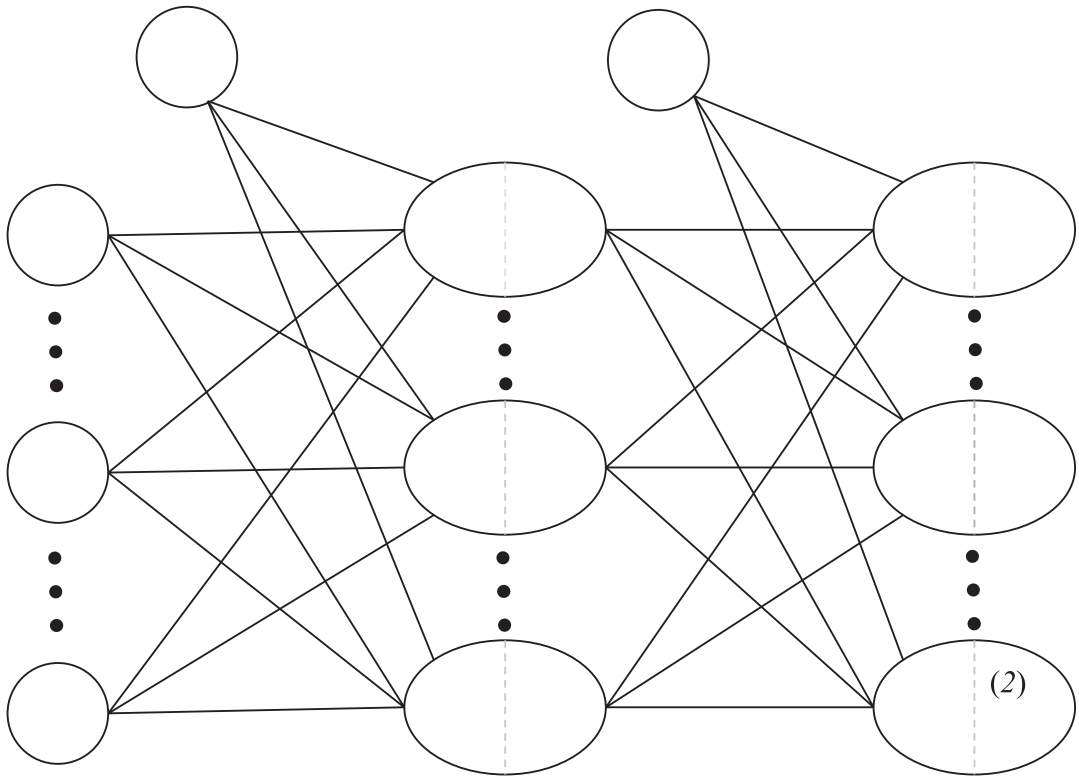





## 変数と定数

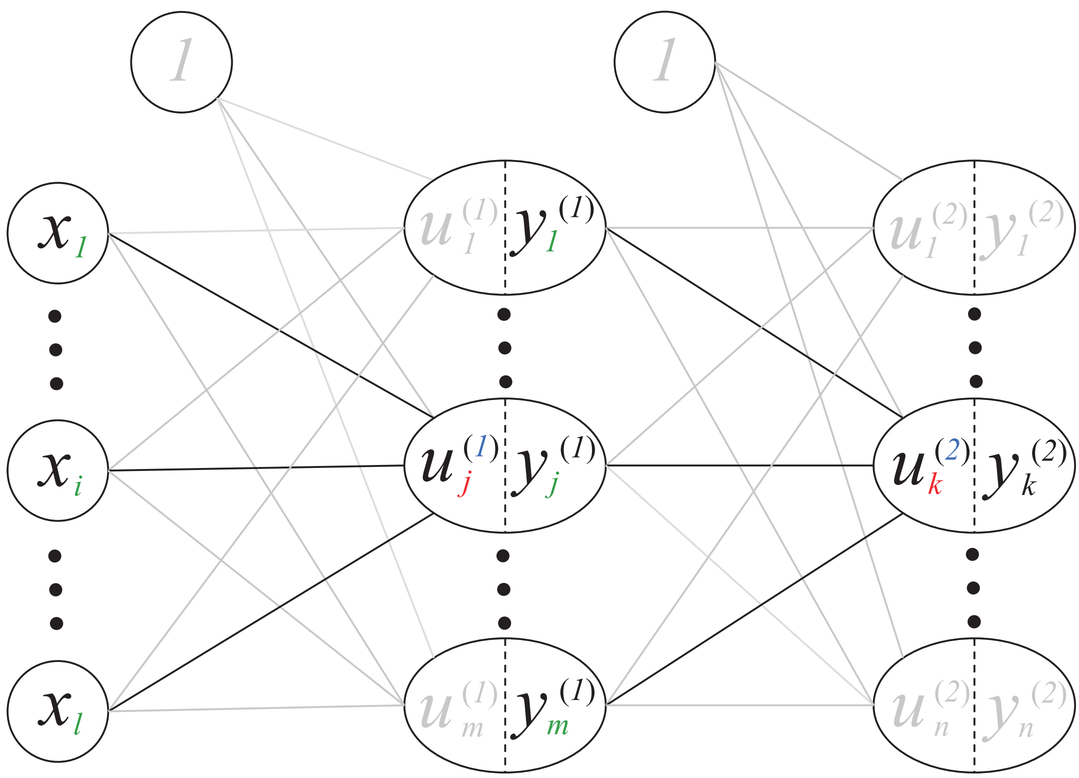


# 活性化関数

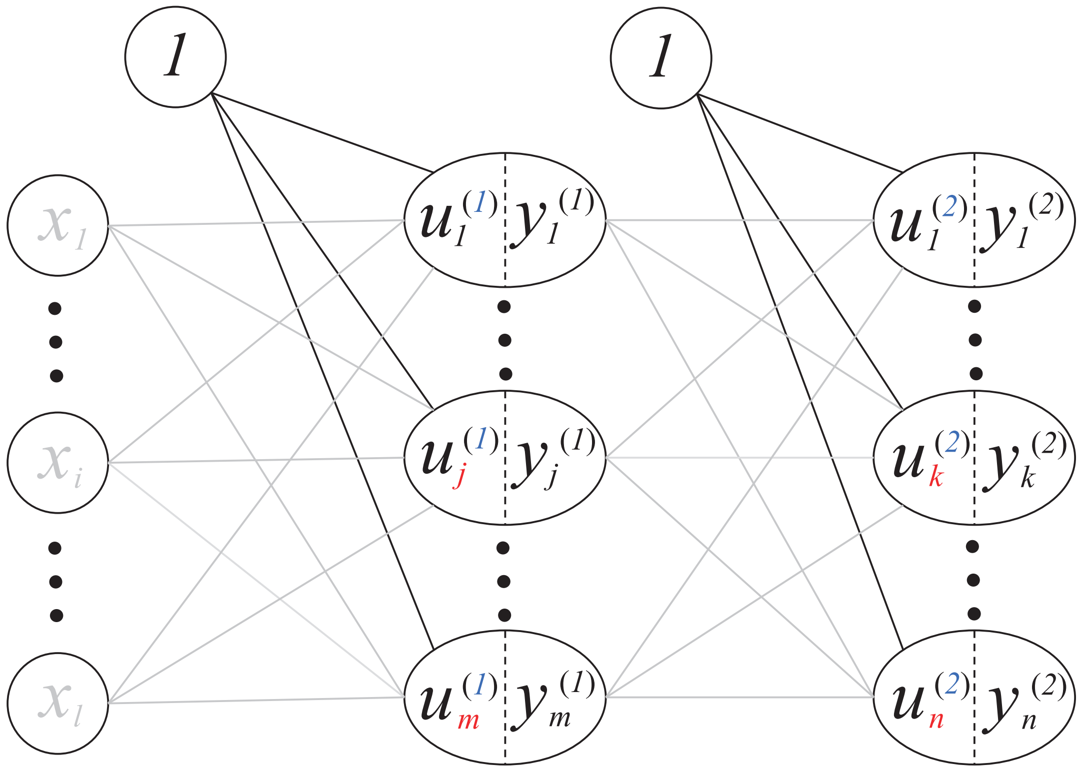





# ウェイトパラメータ



# バイアスパラメータ






# 活性化関数

■ の場合： 関数と 関数

$$h^{(2)}(u_k^{(2)}) =$$

$$h^{(1)}(u_j^{(1)}) = \underline{\qquad\qquad\qquad}$$

■ の場合： 関数と 関数

$$h^{(2)}(u_k^{(2)}) = \underbrace{\underline{\qquad\qquad\qquad}}$$

$$h^{(1)}(u_j^{(1)}) = \underline{\qquad\qquad\qquad}$$

の関数であることに注意！

# 中間層の変数

■中間層での活性化関数への入力

$$u_j^{(1)} = \qquad + \cdots + \qquad + \cdots + \qquad +$$

■中間層からの出力

$$y_j^{(1)} =$$



# 出力層の変数

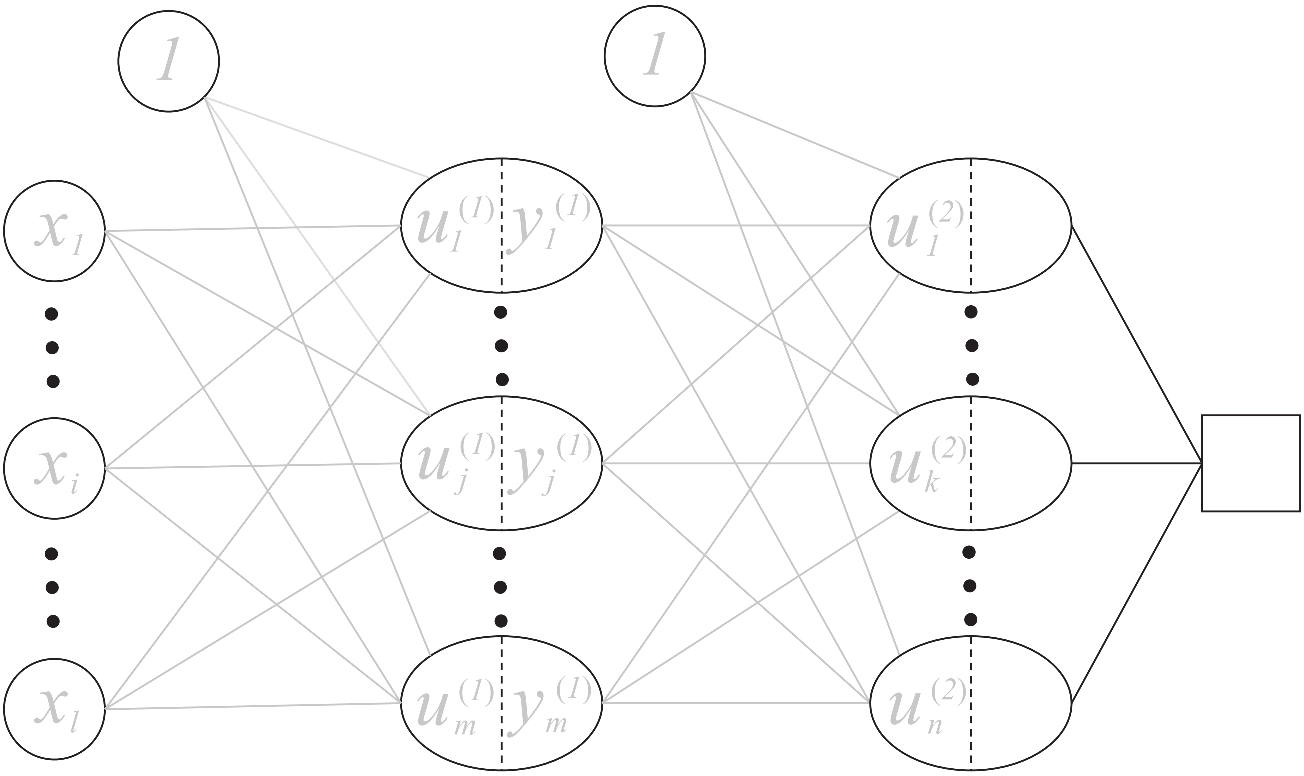
■出力層での活性化関数への入力

$$u_k^{(2)} = \qquad + \cdots + \qquad + \cdots + \qquad +$$

■出力層からの出力

$$y_k^{(2)} =$$

## 計算






## の 依存性

$$E = E( \quad , \quad , \quad , \quad ) = f( \quad , \quad )$$

$$W^{(\ )} = \begin{bmatrix} w^{(\ )} & \cdots & w^{(\ )} \\ \vdots & w^{(\ )} & \vdots \\ w^{(\ )} & \cdots & w^{(\ )} \end{bmatrix} \qquad B^{(\ )} = \begin{bmatrix} b^{(\ )} \\ \vdots \\ b^{(\ )} \end{bmatrix}$$

$$W^{(\ )} = \begin{bmatrix} w^{(\ )} & \cdots & w^{(\ )} \\ \vdots & w^{(\ )} & \vdots \\ w^{(\ )} & \cdots & w^{(\ )} \end{bmatrix} \qquad B^{(\ )} = \begin{bmatrix} b^{(\ )} \\ \vdots \\ b^{(\ )} \end{bmatrix}$$

# 誤差関数

■ 問題の場合：

$E =$

$E$ は の関数

■ 問題の場合：

$E =$



# の 化 → パラメータの

法により誤差$E$を する

化にかかる 時間を大幅に短縮する目的で,
を に 必要がある！

$\overline{\qquad\qquad} = ?$ STEP 2 $\overline{\qquad\qquad} = ?$

STEP 1.5

$\overline{\qquad\qquad} = ?$ STEP 1 $\overline{\qquad\qquad} = ?$

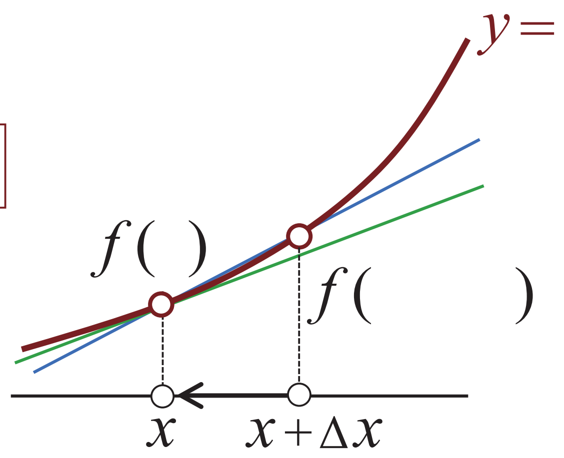
# 微分の基本　　微分（　変数関数の微分）

線分　　の傾き　$\dfrac{f(\qquad)-f(\;)}{\quad}$

線分　　の傾き　$\boxed{\dfrac{\quad}{\quad}} = \dfrac{f(\qquad)-f(\;)}{\quad}$

⇩

での　に対する　の変化率






# 微分の基本　　率（1変数の場合）

$$\begin{cases} y= \\ u= \end{cases}$$

に対する　の　　率

$\boxed{\dfrac{\quad}{\quad}} = \boxed{\dfrac{\quad}{\quad}}\boxed{\dfrac{\quad}{\quad}}$

に対する　の　　率

に対する　の　　率

## 率が成立する理由

$$\frac{}{} = \lim_{\Delta x \to 0} \frac{}{}$$

$$= \lim_{\Delta x \to 0} \left( \frac{}{} \quad \frac{}{} \right)$$

$$= \lim_{\Delta u \to 0} \frac{}{} \quad \lim_{\Delta x \to 0} \frac{}{}$$

$$= \frac{}{} \quad \frac{}{}$$



## 微分の基本　　微分（　変数関数の微分）

$$z = f(x, y)$$

$$\frac{\partial z}{\partial x} = \qquad \frac{}{}$$

$$\frac{\partial z}{\partial y} = \lim_{\Delta y \to 0} \frac{}{}$$

# 微分の基本　　微分

$z = f(x, y)$　変数　と　に対して　はどれだけ変化するか？

$\Delta z = f(\qquad , \qquad) - f(\ ,\ )$

$= \dfrac{f(\qquad , \qquad) - f(\ , \qquad)}{\quad}$

$+ \dfrac{f(\ ,\ + \quad) - f(\ ,\ )}{\quad}$

$\boxed{dz = \qquad + \qquad}$



# 全微分（　　　な場合）

$y = f(x_1, \cdots, x_n)$

$dy = \qquad + \cdots +$

$\boxed{dy = \qquad}$

## 微分の基本　　率（　変数の場合）

■　　　となる変数が　　の場合

$$y=f(\ ,\ )\quad u=g(\ )\quad v=h(\ )$$

$$dy= \qquad + \qquad \Rightarrow \boxed{\frac{dy}{dx}= \qquad + \qquad}$$

■　　　すると

$$y=f(\ ,\cdots,\ )\quad u=f(\ )\cdots u=f(\ )$$

$$\boxed{\frac{dy}{dx}= \qquad}$$



## 微分の基本　　鎖率（　変数の場合）

■　　　となる変数が　　の場合

$$z=f(\ ,\ )\quad u=g(\ ,\ )\quad v=h(\ ,\ )$$

$$dz= \qquad + \qquad \Rightarrow \boxed{\frac{\partial z}{\partial x}= \qquad + \qquad}\quad \boxed{\frac{\partial z}{\partial y}= \qquad + \qquad}$$

■一般化すると

$$z=f(\ ,\cdots,\ )\quad u=f(\ ,\ )\cdots f(\ ,\ )$$

$$\boxed{\frac{\partial z}{\partial x}= \qquad}\quad \boxed{\frac{\partial z}{\partial y}= \qquad}$$

## STEP 1 出力層での　　　　　　を求める

$$\frac{\partial E}{\partial w_{kj}^{(2)}} = \sum^{n} \frac{\quad}{\quad}\frac{\quad}{\quad} = \frac{\quad}{\quad}\frac{\quad}{\quad} + \sum^{n} \frac{\quad}{\quad}\boxed{\frac{\quad}{\quad}}$$

は　　の関数ではないため　　は

$$\frac{\partial E}{\partial w_{kj}^{(2)}} = \boxed{\frac{\quad}{\quad}}\frac{\quad}{\quad} = \boxed{\quad}\frac{\quad}{\quad}$$



## STEP 1 出力層での　　　　　　を求める

$$u_k^{(2)} = \quad + \cdots + \quad + \cdots + \quad +$$

$$\Downarrow$$

$$\frac{\quad}{\quad} =$$

$$\boxed{\quad = \quad}$$

## STEP 1 出力層での　　　　　　　を求める

$$\frac{\partial E}{\partial b_k^{(2)}} = \sum^{n} \frac{\quad}{\quad}\frac{\quad}{\quad} = \frac{\quad}{\quad}\frac{\quad}{\quad} + \sum^{n} \frac{\quad}{\quad}\boxed{\frac{\quad}{\quad}}$$

は　　の関数ではないため　　は

$$\frac{\partial E}{\partial b_k^{(2)}} = \boxed{\frac{\quad}{\quad}}\frac{\quad}{\quad} = \boxed{\quad}\frac{\quad}{\quad}$$



## STEP 1 出力層での　　　　　　　を求める

$$\frac{\partial E}{\partial b_k^{(2)}} = \boxed{\frac{\quad}{\quad}}\frac{\quad}{\quad} = \boxed{\quad}\frac{\quad}{\quad}$$

$$u_k^{(2)} = \quad + \cdots + \quad + \cdots + \quad +$$

$$\Downarrow$$

$$\frac{\quad}{\quad} =$$

$$\boxed{\frac{\quad}{\quad} =}$$

## STEP 1.5 出力層での　　　　　　　を求める

$$\frac{\partial E}{\partial y_j^{(1)}} = \sum_{r=1}^{n} \boxed{\frac{\quad}{\quad}} \frac{\quad}{\quad} = \sum_{r=1}^{n} \boxed{\quad} \frac{\quad}{\quad}$$

$$u_k^{(2)} = \quad + \cdots + \quad + \cdots + \quad + \quad$$

$k = 1,\cdots,n$

$$\frac{\quad}{\quad} =$$

$$\boxed{\frac{\quad}{\quad} = \sum_{r=1}^{n}}$$



## STEP 2 中間層での　　　　　　　を求める

$$\frac{\partial E}{\partial w_{ji}^{(1)}} = \sum^{m} \frac{\quad}{\quad}\frac{\quad}{\quad} = \frac{\quad}{\quad}\frac{\quad}{\quad} + \sum^{m} \frac{\quad}{\quad} \boxed{\frac{\quad}{\quad}}$$

は　　の関数ではないため　　は

$$\frac{\partial E}{\partial w_{ji}^{(1)}} = \boxed{\frac{\quad}{\quad}} \frac{\quad}{\quad} = \boxed{\quad} \frac{\quad}{\quad}$$

## STEP 2 中間層での を求める

$$u_j^{(1)} = \qquad + \cdots + \qquad + \cdots + \qquad +$$

$$\frac{\qquad}{\qquad} =$$

$$\boxed{\frac{\qquad}{\qquad} = \qquad\qquad}$$



## STEP 2 中間層での を求める

$$\frac{\partial E}{\partial b_j^{(1)}} = \sum^{m} \frac{\qquad}{\qquad}\frac{\qquad}{\qquad} = \frac{\qquad}{\qquad}\frac{\qquad}{\qquad} + \sum^{m} \frac{\qquad}{\qquad}\boxed{\frac{\qquad}{\qquad}}$$

は の関数ではないため は

$$\frac{\partial E}{\partial b_j^{(1)}} = \boxed{\frac{\qquad}{\qquad}}\frac{\qquad}{\qquad} = \boxed{\qquad}\frac{\qquad}{\qquad}$$

## STEP 2 中間層での対バイアス勾配を求める

$$u_j^{(1)} = w_{j1}^{(1)} x_1 + \cdots + w_{ji}^{(1)} x_i + \cdots + w_{jl}^{(1)} x_l + b_j^{(1)}$$

$$\Downarrow$$

$$\frac{\partial u_j^{(1)}}{\partial b_j^{(1)}} = 1$$

$$\boxed{\frac{\partial E}{\partial b_j^{(1)}} = \delta_j^{(1)}}$$



## 対パラメータ勾配の　　　表現

■ 　　層

$$\frac{\partial E}{\partial W^{(2)}} = \left( \qquad \right) \times \qquad \frac{\partial E}{\partial B^{(2)}} = \left( \quad \right) \times$$

■ 　　層

$$\frac{\partial E}{\partial W^{(1)}} = \left( \qquad \right) \times \qquad \frac{\partial E}{\partial B^{(1)}} = \left( \quad \right) \times$$

# 出力層デルタ

$$\delta_k^{(2)} = \underline{\qquad\qquad} = \sum^{n} \underline{\qquad\qquad}\ \underline{\qquad\qquad}$$

$$= \underline{\qquad\qquad}\ \underline{\qquad\qquad} + \sum^{n} \underline{\qquad\qquad}\ \underline{\qquad\qquad}$$



# 出力層デルタ：　　　問題の場合

$$\boxed{E = \sum_{r=1}^{n} \qquad\qquad\qquad}$$

$$\boxed{y_k^{(2)} = h^{(2)}(u_k^{(2)}) = \qquad}$$

$$\frac{\partial E}{\partial y_k^{(2)}} \frac{\partial y_k^{(2)}}{\partial u_k^{(2)}} = \qquad\qquad\qquad =$$

$$\sum_{r \neq k}^{n} \frac{\partial E}{\partial y_r^{(2)}} \frac{\partial y_r^{(2)}}{\partial u_k^{(2)}} = \sum_{r \neq k}^{n} \qquad\qquad\qquad =$$

$$\boxed{\delta_k^{(2)} = \qquad\qquad}$$

## 出力層デルタ：　　問題の場合

$$E = \sum_{r=1}^{n}$$

$$y_k^{(2)} = h^{(2)}(u_k^{(2)}) = \frac{\qquad}{\qquad}$$

$$\frac{\partial E}{\partial y_k^{(2)}}\frac{\partial y_k^{(2)}}{\partial u_k^{(2)}} = \qquad\qquad =$$

$$\sum_{r\neq k}^{n}\frac{\partial E}{\partial y_r^{(2)}}\frac{\partial y_r^{(2)}}{\partial u_k^{(2)}} = \sum_{r\neq k}^{n}\left(\qquad\qquad\qquad\right) =$$

$$\delta_k^{(2)} = \qquad\qquad =$$

問題と同じ



## 中間層デルタ

$$\delta_j^{(1)} = \frac{\partial E}{\partial u_j^{(1)}} = \sum^{n}\frac{\qquad}{\qquad}\frac{\qquad}{\qquad}$$

$$= \frac{\qquad}{\qquad}\frac{\qquad}{\qquad} + \sum^{m}\frac{\qquad}{\qquad}\boxed{\frac{\qquad}{\qquad}}$$

は　　の関数ではないため　　は

$$= \boxed{\frac{\qquad}{\qquad}}\frac{\qquad}{\qquad} = \boxed{\sum_{k=1}^{n}\qquad}\frac{\qquad}{\qquad}$$

STEP 1.5

を求めるには　　を先に算出しておく必要あり

## 中間層デルタ：　　問題と　　問題共通

■活性化関数が　　　　　関数の場合

$$\boxed{y_j^{(1)} = h^{(1)}(u_j^{(1)}) = \rule{3cm}{0.4pt}}$$

$$\frac{\partial y_j^{(1)}}{\partial u_j^{(1)}} =$$

$$\boxed{\delta_j^{(1)} = \left( \sum_{k=1}^{n} \qquad \right) \quad \left( \qquad \right)}$$



## 中間層デルタ：　　問題と　　問題共通

■活性化関数が　　　　関数の場合

$$\boxed{y_j^{(1)} = h(u_j^{(1)}) = \begin{cases} & \text{if } u_j^{(1)} \geq 0 \\ & \text{otherwise} \end{cases}}$$

$$\frac{\partial y_j^{(1)}}{\partial u_j^{(1)}} = \begin{cases} & \text{if } u_j^{(1)} \geq 0 \\ & \text{otherwise} \end{cases}$$

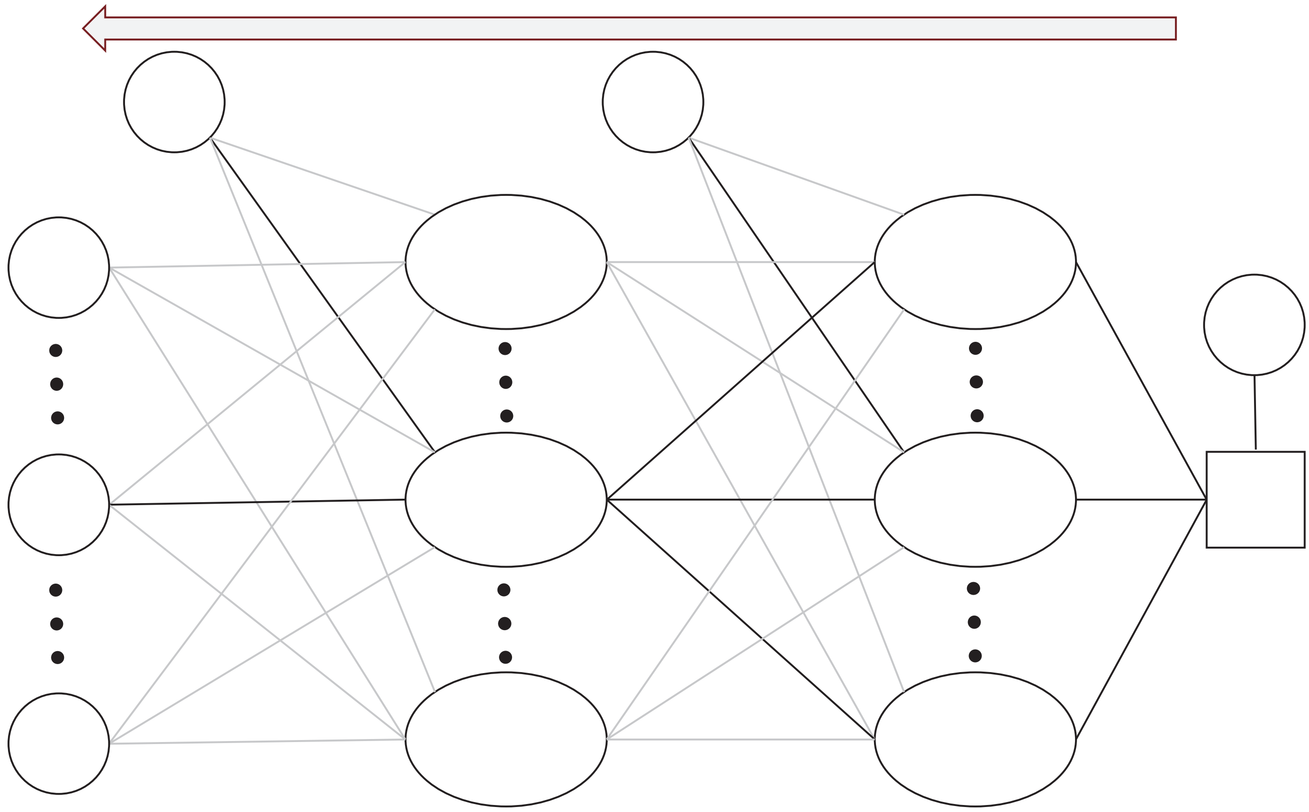
$u_j^{(1)}$　　の場合

$$\boxed{\delta_j^{(1)} =}$$

$u_j^{(1)}$　　の場合

$$\boxed{\delta_j^{(1)} =}$$

## 伝搬　　の流れ：　　層→　　層



教師信号 $D = \left(d_1 \cdots d_n\right)^T$



## まとめ　対パラメータ勾配の計算：出力層

$$\frac{\partial E}{\partial W^{(2)}} = \left( \qquad \right) \times$$

$$\boxed{\frac{\partial E}{\partial W^{(2)}} = \left( \left( \quad - \quad \right) \quad \right) \times}$$

$$\frac{\partial E}{\partial B^{(2)}} = \left( \qquad \right) \times$$

$$\boxed{\frac{\partial E}{\partial B^{(2)}} = \left( \quad - \quad \right) \times}$$

## 中間層（　　　　　　関数の場合）

$$\frac{\partial E}{\partial W^{(1)}} = \left(\qquad\right) \times = \left(\left(\sum_{r=1}^{n} \qquad\right) \left(\quad - \quad\right)\right) \times$$

$$\Downarrow$$

$$\boxed{\frac{\partial E}{\partial W^{(1)}} = \left(\left(\sum_{r=1}^{n} \left(\quad - \quad\right) \qquad\right) \left(\quad - \quad\right)\right) \times}$$

$$\frac{\partial E}{\partial B^{(1)}} = \left(\qquad\right) \times$$

$$\Downarrow$$

$$\boxed{\frac{\partial E}{\partial B^{(1)}} = \left(\left(\sum_{r=1}^{n} \left(\quad - \quad\right) \qquad\right) \left(\quad - \quad\right)\right) \times}$$



## 中間層（　　関数の場合）　$u_j^{(1)} \geq 0$ の場合

$$\frac{\partial E}{\partial W^{(1)}} = \left(\qquad\right) \times = \left(\left(\sum_{r=1}^{n} \qquad\right) \quad\right) \times$$

$$\Downarrow$$

$$\boxed{\frac{\partial E}{\partial W^{(1)}} = \left(\left(\sum_{r=1}^{n} \left(\quad - \quad\right) \qquad\right) \quad\right) \times}$$

$$\frac{\partial E}{\partial B^{(1)}} = \left(\qquad\right) \times$$

$$\Downarrow$$

$$\boxed{\frac{\partial E}{\partial B^{(1)}} = \left(\left(\sum_{r=1}^{n} \left(\quad - \quad\right) \qquad\right)\right) \times}$$