

【事前に用意する情報】

入力 : $\mathbf{x}_n = [x_{n1} \dots x_{ni}]$

訓練データ : $\mathbf{d}_n = [d_{n1} \dots d_{nk}]$

【多層ネットワークのパラメータ】

$$\mathbf{w}^{(l)} = \begin{cases} \text{重み} & : \mathbf{W}^{(l)} = \begin{bmatrix} w_{11}^{(l)} & \dots & w_{1I}^{(l)} \\ \vdots & \ddots & \vdots \\ w_{J1}^{(l)} & \dots & w_{JI}^{(l)} \end{bmatrix} \\ \text{バイアス} & : \mathbf{b}^{(l)} = [b_1^{(l)} \dots b_J^{(l)}] \end{cases}$$

活性化関数 : $\mathbf{f}^{(l)}(\mathbf{u}^{(l)}) = [f^{(l)}(u_1^{(l)}) \dots f^{(l)}(u_J^{(l)})]$

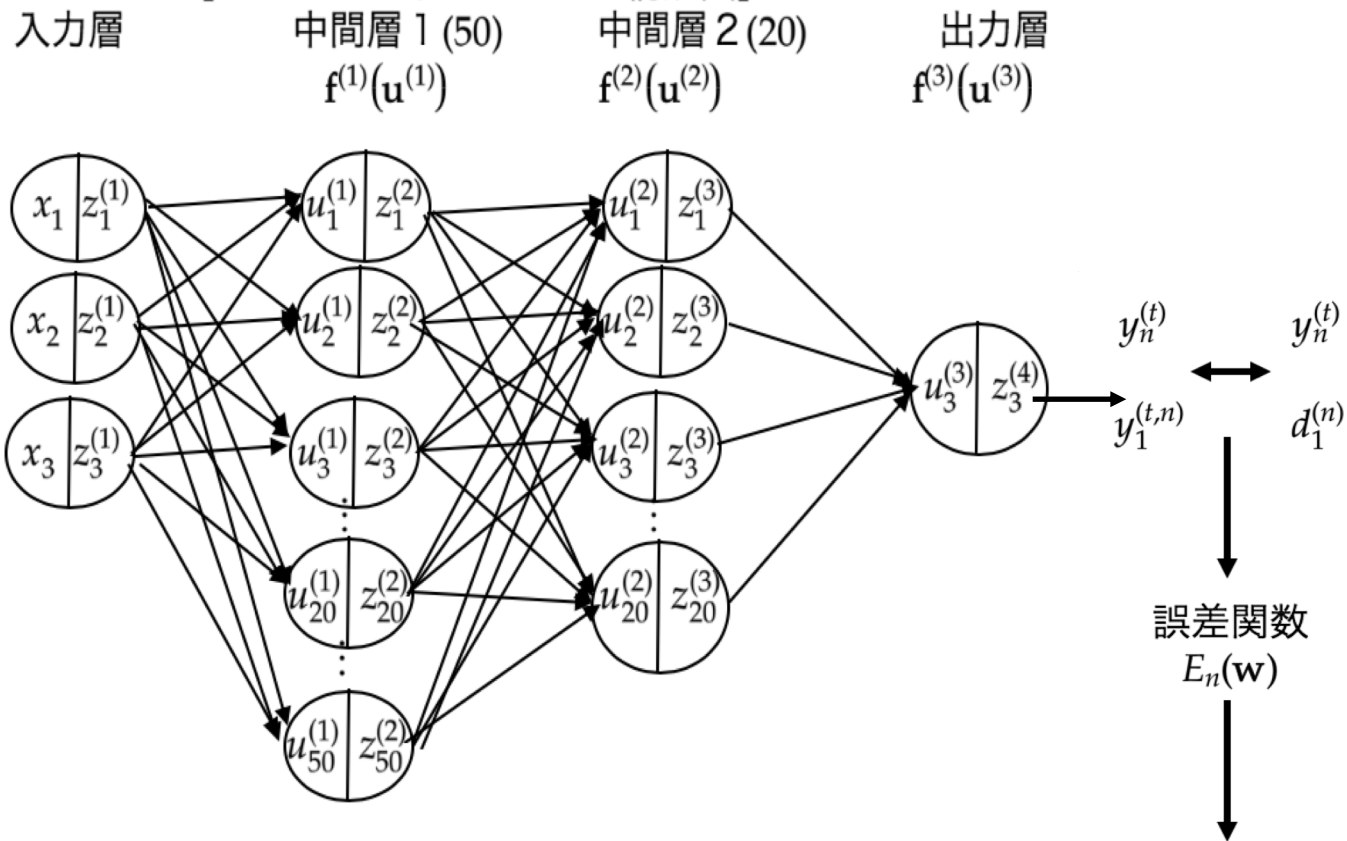
中間層出力 : $\mathbf{z}^{(l)} = [z_1^{(t)} \dots z_k^{(t)}] = \mathbf{f}^{(l)}(\mathbf{u}^{(l)})$

総入力 : $\mathbf{u}^{(l)} = \mathbf{W}^{(l)}\mathbf{z}^{(l-1)} + \mathbf{b}^{(l)}$

出力 : $\mathbf{y}_n^{(t)} = [y_{n1}^{(t)} \dots y_{nK}^{(t)}] = \mathbf{z}^{(L)}$

誤差関数 : $E_n(\mathbf{w})$

【ニューラルネットワーク構成図】



$$V_t = \mu V_{t-1} - \epsilon \nabla E_n(\mathbf{w}) \Leftarrow \nabla E_n(\mathbf{w}) = \frac{\partial E}{\partial \mathbf{w}}$$

$$\mathbf{w}^{(t+1)} = \mathbf{w}^{(t)} - V_t$$

構造図の必要事項

- 中間層は2層で1層目は50、2層目は20で構成
- 中間層の活性化関数は2層共にRelu関数を使用
- 出力層の活性化関数は恒等写像を使用
- 誤差関数は二乗誤差を使用
- Optimizerはモメンタムを使用