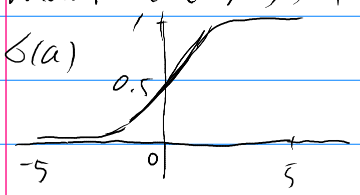
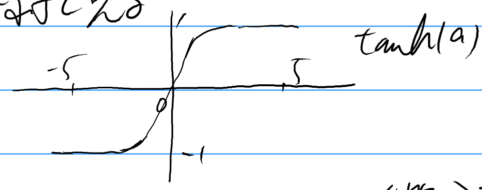
演習5.1

$$\sigma(a)=\frac{1}{1+e^{-a}},\quad \tanh(a)=\frac{e^{a}-e^{-a}}{e^{a}+e^{-a}}$$

maximaでグラフを描いて様子を見る





$\sigma'(0)=\frac{1}{4}$ $\qquad$ $\tanh'(0)=1$ ← 微分もmaximaで計算

これより視察により

$$\sigma(a)=\frac{1}{2}\left(\tanh\left(\tfrac{1}{2}a\right)+1\right)\quad\cdots ①$$

を得る。

計算して確認する

$$\frac{1}{2}\left(\tanh\left(\tfrac{1}{2}a\right)+1\right)=\frac{1}{2}\left(\frac{e^{\frac{a}{2}}-e^{-\frac{a}{2}}}{e^{\frac{a}{2}}+e^{-\frac{a}{2}}}+1\right)=\frac{1}{2}\,\frac{e^{\frac{a}{2}}-e^{-\frac{a}{2}}+e^{\frac{a}{2}}+e^{-\frac{a}{2}}}{e^{\frac{a}{2}}+e^{-\frac{a}{2}}}$$

$$=\frac{e^{\frac{a}{2}}}{e^{\frac{a}{2}}+e^{-\frac{a}{2}}}=\frac{1}{1+e^{-a}}=\sigma(a)$$

(5.7) $\displaystyle y_k(x,w)=\sigma\left(\sum_{j=1}^{M}w_{kj}^{(2)}h\left(\sum_{i=1}^{D}w_{ji}^{(1)}x_i+w_{j0}^{(1)}\right)+w_{k0}^{(2)}\right)$

(5.7)のhにσを使うと

$$y_k(x,w)=\sigma\left(\sum_j w_{kj}^{(2)}\sigma\left(\sum_i w_{ji}^{(1)}x_i+w_{j0}^{(1)}\right)+w_{k0}^{(2)}\right)$$

①より

$$y_k(x,w)=\sigma\left(\sum_j w_{kj}^{(2)}\frac{1}{2}\left(\tanh\left(\frac{1}{2}\left(\sum w_{ji}^{(1)}x_i+w_{j0}^{(1)}\right)\right)+1\right)+w_{k0}^{(2)}\right)$$

$$=\sigma\left(\sum_j \frac{1}{2}w_{kj}^{(2)}\tanh\left(\sum_i \frac{1}{2}w_{ji}^{(1)}x_i+\frac{1}{2}w_{j0}^{(1)}\right)+\sum_j \frac{1}{2}w_{kj}^{(2)}+w_{k0}^{(2)}\right)$$

$$=\sigma\left(\sum_j v_{kj}^{(2)}\tanh\left(\sum_i v_{ji}^{(1)}x_i+v_{j0}^{(1)}\right)+v_{k0}^{(2)}\right)=z_k(x,v)$$

但し

$$v_{kj}^{(2)}=\frac{1}{2}w_{kj}^{(2)},\quad v_{ji}^{(1)}=\frac{1}{2}w_{ji}^{(1)},\quad v_{j0}^{(1)}=\frac{1}{2}w_{j0}^{(1)},\quad v_{k0}^{(2)}=\sum_j\frac{1}{2}w_{kj}^{(2)}+w_{k0}^{(2)}$$

を得る。つまり、あるxに対して$y_k(x,w)$と同じ値を返す、隠れ活性化関数が$\tanh$の$z_k(x,v)$が存在する。