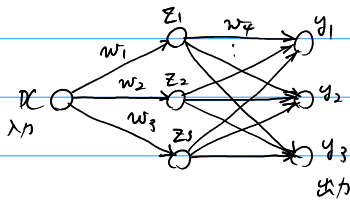


5.40

- ・ニューラルネットワークの出力ユニットの活性化関数をソフトマックス関数とする。
- ・目標変数 \mathcal{D} の条件付分布を多項分布とし、ニューラルネットの出力をこの多項分布の平均と解釈する。
 というモデルを考える



出力ユニット y_k の活性を a_k とすると

$$y_k(x, w) = \frac{\exp(a_k(x, w))}{\sum_j \exp(a_j(x, w))}$$

目標変数 \mathcal{D} (1 of K 表記) の条件付分布は

$$p(\mathcal{D} | x, w) = \prod_{k=1}^K y_k^{t_k}(x, w)$$

と仮定。このとき $D = \{\mathcal{D}_1, \mathcal{D}_2, \dots\}$, $X = \{x_1, x_2, \dots\}$, $y_{nk} = y_k(x_n, w)$ とおくと尤度は

$$p(D | w, X) = \prod_{n=1}^N p(\mathcal{D}_n | x_n, w) = \prod_{n=1}^N \prod_{k=1}^K y_{nk}^{t_{nk}}$$

対数尤度は

$$\ln p(D | w, X) = \sum_{n=1}^N \sum_{k=1}^K t_{nk} \ln y_{nk}$$

と仮定。

w の事前分布は (5.162) で与えられるとすると正規化誤差関数 $E(w)$ は (5.182) で与えられる。

$E(w)$ を最小化する w_{MAP} は $\partial = \frac{\partial E(w)}{\partial w}$ で与えられる。この微分は

ニューラルネットの誤差逆伝播で得られる。

ヘッセ行列 H は (5.91) 式、若しくは $(5.93), (5.94), (5.95)$ で得られる

w の事前分布のラプラス近似は (5.167) で得られる

超パラメータ α は (5.178) で得られる