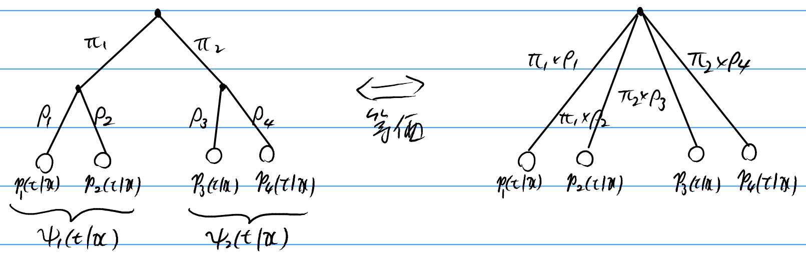
14.17

2レベルの階層的混合

(ゲート関数が定数の場合)



混合分布は

$$p(t|x) = \sum_{k=1}^{K} \pi_k \psi_k(t|x)$$

ここで

$$\psi_k(t|x) = \sum_{l \in \text{ゲート}k} \rho_l p_l(t|x)$$

なので

$$p(t|x) = \sum_{k=1}^{K} \pi_k \sum_{l \in \text{ゲート}k} \rho_l p_l(t|x) = \sum_{k=1}^{K} \sum_{l \in \text{ゲート}k} \pi_k \rho_l p_l(t|x)$$

$$= \pi_1 \rho_1 p_1(t|x) + \pi_1 \rho_2 p_2(t|x) + \cdots + \pi_K \rho_L p_L(t|x)$$

$$= \sum_{l=1}^{L} \pi_{l\text{の親}} \rho_l p_l(t|x)$$

$$= \sum_{l=1}^{L} \sigma_l p_l(t|x), \quad \sigma_l = \pi_{l\text{の親}} \times \rho_l$$

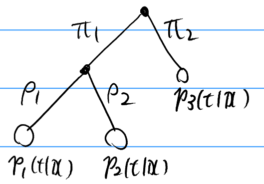
となる。これは1レベルの混合分布になっている

(ゲート関数がxの関数の場合)

ゲート関数が定数の場合と同様。

(ゲート関数が線形分類モデルの場合)

2レベルの線形ロジスティックモデルのゲート関数のHME



π_1, π_2は線形ロジスティックモデル、ρ_1, ρ_2も線形ロジスティックモデルとすると

$$\pi_1(x)=\sigma(w^T x),\quad \pi_2(x)=1-\pi_1(x)=1-\sigma(w^T x)$$

$$\rho_1(x)=\sigma(v^T x),\quad \rho_2(x)=1-\rho_1(x)=1-\sigma(v^T x)$$

である。混合分布は

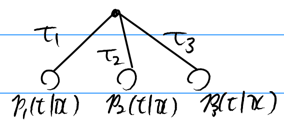
$$p(t|x)=\pi_1(x)\{\rho_1(x)p_1(t|x)+\rho_2(x)p_2(t|x)\}+\pi_2(x)p_3(t|x)$$

$$=\pi_1(x)\rho_1(x)p_1(t|x)+\pi_1(x)\rho_2(x)p_2(t|x)+\pi_2(x)p_3(t|x)$$

$$=\sigma(w^T x)\sigma(v^T x)p_1(t|x)+\sigma(w^T x)\{1-\sigma(v^T x)\}p_2(t|x)+\{1-\sigma(w^T x)\}p_3(t|x)\quad\cdots①$$

となる。

1レベルの線形ソフトマックスモデルのゲート関数のHME



$\tau_1(x)$, $\tau_2(x)$, $\tau_3(x)$は線形ソフトマックスモデルとすると

$$\tau_k(x)=\frac{\exp(u_k^T x)}{\sum_j \exp(u_j^T x)}$$

である。混合分布は

$$p(t|x)=\tau_1(x)p_1(t|x)+\tau_2(x)p_2(t|x)+\tau_3 p_3(t|x)$$

$$=\frac{\exp(u_1^T x)}{\sum_j \exp(u_j^T x)}p_1(t|x)+\frac{\exp(u_2^T x)}{\sum_j \exp(u_j^T x)}p_2(t|x)+\frac{\exp(u_3^T x)}{\sum_j \exp(u_j^T x)}p_3(t|x)\quad\cdots②$$

となる。

上記2つのHMEが等価であると仮定すると

① = ② となる。

p_1, p_2, p_3 は独立で0でないので

$$\sigma(w^T x)\sigma(v^T x) = \frac{\exp(u_1^T x)}{\sum_j \exp(u_j^T x)}$$

$$\sigma(w^T x)\{1-\sigma(v^T x)\} = \frac{\exp(u_2^T x)}{\sum_j \exp(u_j^T x)}$$

$$\{1-\sigma(w^T x)\} = \frac{\exp(u_3^T x)}{\sum_j \exp(u_j^T x)}$$

となる。ここで $\sigma(w^T x) = W$, $\sigma(v^T x) = V$, $\exp(u_j^T x) = U_j$ と書くと

$$WV = \frac{U_1}{U_1+U_2+U_3}$$

$$W(1-V) = \frac{U_2}{U_1+U_2+U_3}$$

$$1-W = \frac{U_3}{U_1+U_2+U_3}$$

となる。これより

$$(U_1+U_2+U_3)WV = U_1$$

$$(U_1+U_2+U_3)W(1-V) = U_2$$

$$(U_1+U_2+U_3)(1-W) = U_3$$

よって

$$\begin{pmatrix} WV-1 & WV & WV \\ W(1-V) & W(1-V)-1 & W(1-V) \\ 1-W & 1-W & -W \end{pmatrix} \begin{pmatrix} U_1 \\ U_2 \\ U_3 \end{pmatrix} = \mathbf{0}$$

となる。

左辺の係数行列の行列式は　　　　　　　　　　　　　　余因子展開

$$\{\underbrace{w(1-v)}_{\times}-1\}(1-w)\underbrace{-w(1-v)(1-w)}_{v}\underbrace{-w(1-v)(1-w)}_{\times}+\underbrace{w(1-v)(1-w)}_{v}+\underbrace{w(1-v)(1-w)}_{w}-\{\underbrace{w(1-v)}_{w}-1\}(1-w)$$

$$= w+(1-w) = 1 \neq 0$$

なので係数行列は逆行列を持つ。この逆行列を左から掛けて

$$\begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

を得る。

これは、$u_1+u_2+u_3=1$ と矛盾する。なので上記2つのHMEは等価ではない。

よって

2レベルの線形ロジスティックモデルのゲート関数のHMEに等価な

1レベルの線形ソフトマックスモデルのゲート関数のHMEが必ず存在するとは言えない。