例33の補足

新納浩幸

2003/5/21

カテゴリ k の生起確率を $p(k)$、カテゴリ k のときに入力が x である条件付き確率を $p(x|k)$ で表すと、(x,y) の同時確率は以下で示せる。

$$p(x,y) = \sum_{k=1}^{N} p(k)p(x|k)\delta(y-\eta_k) \tag{1}$$

これはよく考えればわかる。実質 y が k_0 だとすれば、$p(y) = p(k_0)$ である。また $\delta(y-\eta_{k_0}) = 1$。それ以外の k では。$\delta(y-\eta_k) = 0$。なので、式 1 のように k ですべて足しこんでもよい。

回帰関数の定義から、

$$r_k(x) = \frac{\int y_k p(x,y)dy}{\int p(x,y)dy} \tag{2}$$

式 2 の分母は式 1 を積分することで求まる。厳密には面倒だけど、

$$\int \delta(y-\eta_k)dy = 1$$

は言えるので、式 2 の分母は

$$\sum_{j=1}^{N} p(j)p(x|j)$$

となる。次に式 2 の分子を考える。

$$\sum_{j=1}^{N} y_k p(j)p(x|j)$$

に注目すると、これは $j = k$ の時のみ、$y_k = 1$ でそれ以外は $y_k = 0$ なので、結局、

$$\sum_{j=1}^{N} y_k p(j)p(x|j) = p(k)p(x|k)$$

となる。再び、

$$\int \delta(y-\eta_k)dy = 1$$

を使えば、式 2 の分母は、$p(k)p(x|k)$ となる。