06T4027H　斎藤高行

4.3　カテゴリの事後確率の計算

カテゴリの事後確率の計算 (1)

各カテゴリ y の条件付き確率 $p(x \mid y)$ をガウスモデルで推定する場合を考える

$$\hat{p}(x \mid y) = \frac{1}{(2\pi)^{\frac{d}{2}} \det\left(\hat{\Sigma}_y\right)^{\frac{1}{2}}} \exp\left(-\frac{1}{2} \left(x - \hat{\mu}_y\right)^{\mathrm{T}} \hat{\Sigma}_y^{-1} \left(x - \hat{\mu}_y\right) \right)$$

- $\hat{\mu}_y$ ： 分布の期待値の推定量 (→P55 式4.2)
- $\hat{\Sigma}_y$ ： 分散共分散行列の推定量 (→P55 式4.2)

カテゴリの事後確率の計算 (2)

- 事後確立は対数をとっても、その大小関係に変化はない
- 事後確立の対数をとると、識別関数の計算が簡単になることが多い

対数事後確立 $\log p(y \mid x)$ を求めることが目標となる

カテゴリの事後確率の計算 (3)

- 条件付き確立の性質

$$p(y \mid x)p(x) = p(x \mid y)p(y)$$

$$\therefore \quad \log p(y \mid x) = \log p(x \mid y) + \log p(y) - \log p(x)$$

- 訓練標本の割合

$$\hat{p}(y) = \frac{n_y}{n}$$

n_y :カテゴリ y に属する訓練標本の数

n :全訓練標本の数

カテゴリの事後確率の計算 (4)

- 対数事後確立の推定量 $\log \hat{p}(y \mid x)$ を求める

$$
\begin{aligned}
\log \hat{p}(y \mid x) &= \log \hat{p}(x \mid y) + \log \hat{p}(y) - \log p(x) \\
&= -\frac{d}{2}\log(2\pi) - \frac{1}{2}\log\left(\det\left(\hat{\Sigma}_y\right)\right) \\
&\quad - \frac{1}{2}\left(x - \hat{\mu}_y\right)^T \hat{\Sigma}_y^{-1}\left(x - \hat{\mu}_y\right) + \log\frac{n_y}{n} - \log p(x) \\
&= -\frac{1}{2}\left(x - \hat{\mu}_y\right)^T \hat{\Sigma}_y^{-1}\left(x - \hat{\mu}_y\right) - \frac{1}{2}\log\left(\det\left(\hat{\Sigma}_y\right)\right) \\
&\quad + \log n_y + C
\end{aligned}
$$

(C :定数)

マハラノビス距離 (1)

- マハラノビス距離

$$\left(x-\hat{\mu}_y\right)^T \hat{\Sigma}_y^{-1}\left(x-\hat{\mu}_y\right)$$

- $\hat{\Sigma}$ によって定まる超楕円体上の点を等距離とみなす距離尺度

マハラノビス距離 (2)

- $\hat{\Sigma}_y$の固有値分解

$$\hat{\Sigma}\phi = \lambda\phi$$

$$\hat{\Sigma} = \sum_{j=1}^{d} \lambda_j \phi_j \phi_j^T \qquad \hat{\Sigma}^{-1} = \sum_{j=1}^{d} \frac{1}{\lambda_j} \phi_j \phi_j^T$$

$$\therefore \quad \left(x - \hat{\mu}_y\right)^T \hat{\Sigma}_y^{-1} \left(x - \hat{\mu}_y\right) = \sum_{j=1}^{d} \frac{\left(\phi_j^T \left(x - \hat{\mu}\right)\right)^2}{\lambda_j}$$

$\left\{\lambda_j\right\}_{j=1}^{d}$:固有値 $\qquad$ $\left\{\phi_j\right\}_{j=1}^{d}$:固有ベクトル

マハラノビス距離 (3)

マハラノビス距離が一定の場合、マハラノビス距離 (2) の式は超楕円体の方程式となる

(例：定数＝1の場合)

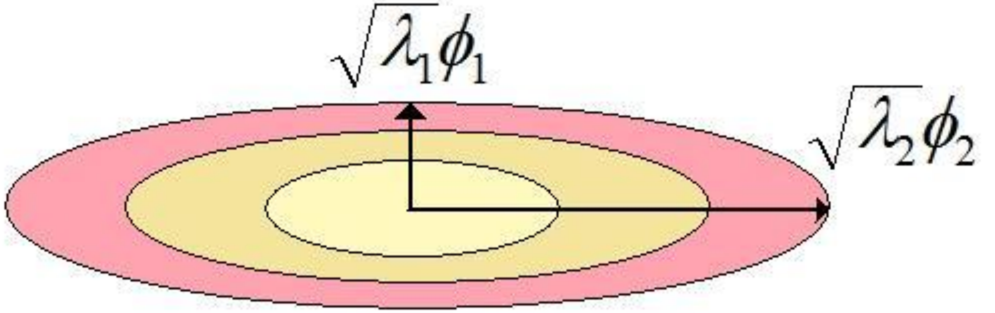
$$\sum_{j=1}^{d} \frac{\left(\phi_j^T \left(x - \hat{\mu}\right)\right)^2}{\lambda_j} = 1$$

マハラノビス距離 (4)

- 具体的には以下の通りである

各主軸の向き: $\left\{\sqrt{\phi_j}\right\}_{j=1}^{d}$　　各主軸の長さ: $\left\{\sqrt{\lambda_j}\right\}_{j=1}^{d}$

- 尚、次元 $d=2$ のとき、下図のような楕円となる



マハラノビス距離 (5)

- カテゴリ数 $c=2$ のとき、決定境界は事後確立が等しくなる $1/2$ の点の集合となる

$$p(y=1|x)=p(y=2|x)$$

決定境界

条件付き確立 $p(x \mid y)$ をガウスモデルで近似する場合、決定境界は x の二次超曲面となる

06T4027H 斎藤高行

4.4 線形判別分析

線形判別分析 (1)

- 各カテゴリの分散共分散行列 Σ_y が等しいと仮定して、それを Σ と表現する

$$\Sigma_1 = \cdots = \Sigma_C = \Sigma$$

- このとき、Σ の最尤推定量 $\hat{\Sigma}$ は

$$\hat{\Sigma} = \frac{1}{n}\sum_{y=1}^{c}\sum_{i:y_i=1}\left(x_i - \hat{\mu}_i\right)^T\left(x_i - \hat{\mu}_i\right) = \sum_{y=1}^{c}\frac{n_y}{n}\hat{\Sigma}_y$$

線形判別分析 (2)

- カテゴリの事後確率の計算 (4)の式に $\hat{\Sigma}$ を用いる

$$\log \hat{p}(y \mid x) = -\frac{1}{2}(x - \hat{\mu}_y)^T \hat{\Sigma}_y^{-1}(x - \hat{\mu}_y) - \frac{1}{2}\log(\det(\hat{\Sigma}_y))$$

$$+ \log n_y + C$$

$$= x^T \hat{\Sigma}^{-1} \hat{\mu}_y - \frac{1}{2}\hat{\mu}_y^T \hat{\Sigma}^{-1} \hat{\mu}_y + \log n_y + C'$$

(C' : 定数)

線形判別分析 (3)

- 分散共分散行列が共通のとき、対数事後確立は x の一次形式となる

カテゴリ数 $c = 2$ の場合、決定境界は

$$\hat{a}^T x + \hat{b} = 0$$

ただし

$$\hat{a} = \hat{\Sigma}^{-1}\left(\hat{\mu}_1 - \hat{\mu}_2\right)$$

$$\hat{b} = -\frac{1}{2}\left(\hat{\mu}_1^T \hat{\Sigma}^{-1} \hat{\mu}_1 - \hat{\mu}_2^T \hat{\Sigma}^{-1} \hat{\mu}_2\right) + \log \frac{n_1}{n_2}$$

決定境界

- 決定境界は x の超平面になる
- この決め方を**フィッシャーの線形判別分析**という