情報認識
「ガウスモデルに対する最大事後確率則」

- 講師:　杉山　将（計算工学専攻）
- 居室:　W8E-505
- 電子メール:　sugi@cs.titech.ac.jp

最大事後確率則

- 入力パターンが与えられてそれがどのカテゴリに属するかを決めるとき，それが属する可能性が最も高いカテゴリを選ぶ.
- これは，事後確率が最大になるカテゴリに分類することに対応.

$$x \in \omega_k : k = \arg\max_i p(\omega_i \mid x)$$

対数事後確率

- 対数をとっても大小関係は変わらないため，事後確率の対数をとったものを用いる.
- ベイズの定理より

$$p(\omega_i \mid x) = \frac{p(x \mid \omega_i)\, p(\omega_i)}{p(x)}$$

- 従って，

$$\log p(\omega_i \mid x) = \log p(x \mid \omega_i) + \log p(\omega_i) + C$$

$$C = -\log p(x) \quad : \text{定数}$$

推定方法

- $\{x_i\}_{i=1}^{n}$:訓練標本
- n_i :カテゴリ ω_i に属する訓練標本数
- 事前確率 $p(\omega_i)$:カテゴリ ω_i に含まれる標本の割合で推定

$$\hat{p}(\omega_i) = \frac{n_i}{n}$$

- 条件付き確率 $p(x \mid \omega_i)$:ガウスモデルに対する最尤推定

$$\hat{p}(x \mid \omega_i) = q(x; \hat{\mu}_i, \hat{\Sigma}_i)$$

ガウスモデル（正規分布）

- d 次元の確率ベクトル: $x = (x^{(1)}, x^{(2)}, \ldots, x^{(d)})^T$
- 2つのパラメータ:
  - 平均ベクトル　μ_i
  - 分散共分散行列　Σ_i

$$q(x; \mu_i, \Sigma_i) = \frac{1}{(2\pi)^{d/2} |\Sigma_i|^{1/2}} \exp\left(-\frac{1}{2} (x - \mu_i)^T \Sigma_i^{-1} (x - \mu_i) \right)$$

ガウスモデルにおける最尤推定

- ガウスモデルの最尤推定量

$$\hat{\mu}_i = \frac{1}{n_i} \sum_{j:x_j \in \omega_i} x_j$$

$$\hat{\Sigma}_i = \frac{1}{n_i} \sum_{j:x_j \in \omega_i} (x_j - \hat{\mu}_i)(x_j - \hat{\mu}_i)^T$$

対数事後確率の計算

- 分類したい入力パターンを x とすれば,

$$\log \hat{p}(\omega_i \mid x)$$

$$= -\frac{d}{2}\log 2\pi - \frac{1}{2}\log|\hat{\Sigma}_i| - \frac{1}{2}(x-\hat{\mu}_i)^T \hat{\Sigma}_i^{-1}(x-\hat{\mu}_i) + \log\frac{n_i}{n} + C$$

$$= -\frac{1}{2}\underbrace{(x-\hat{\mu}_i)^T \hat{\Sigma}_i^{-1}(x-\hat{\mu}_i)}_{\text{マハラノビス距離 (Mahalanobis distance)}} - \frac{d}{2}\log|\hat{\Sigma}_i| + \log\frac{n_i}{n} + C$$

- 対数事後確率は x の二次形式

共分散行列が共通のとき

- 各カテゴリの分散共分散行列が等しいという前提知識があるとき，共通の分散共分散行列の最尤推定量は

$$\hat{\Sigma} = \frac{1}{n}\sum_{i=1}^{m}\sum_{j:x_j \in \omega_i}(x_j - \hat{\mu}_i)(x_j - \hat{\mu}_i)^T$$

$$= \sum_{i=1}^{m}\frac{n_i}{n}\hat{\Sigma}_i$$

共分散行列が共通のとき（続き）

- 各カテゴリの分散共分散行列が等しい時，

$$\log \hat{p}(\omega_i \mid x)$$

$$= -\frac{1}{2} x^T \hat{\Sigma}^{-1} x + \hat{\mu}_i{}^T \hat{\Sigma}^{-1} x - \frac{1}{2} \hat{\mu}_i{}^T \hat{\Sigma}^{-1} \hat{\mu}_i - \frac{d}{2} \log |\hat{\Sigma}| + \log \frac{n_i}{n} + C$$

$$= \hat{\mu}_i{}^T \hat{\Sigma}^{-1} x - \frac{1}{2} \hat{\mu}_i{}^T \hat{\Sigma}^{-1} \hat{\mu}_i + \log \frac{n_i}{n} + C'$$

- 対数事後確率は x の一次形式

決定境界の計算

- カテゴリ数 $m = 2$ の場合，決定境界は

$$\hat{p}(\omega_1 \mid x) = \hat{p}(\omega_2 \mid x)$$

- ガウスモデルを用いたとき，決定境界は二次形式.
- 更に分散共分散行列が共通のとき，決定境界は一次形式，即ち，超平面(hyper-plane) .