情報認識 「識別関数のよさを測る規準」

■講師: 杉山 将(計算工学専攻)

■居室: W8E-505

■電子メール: <u>sugi@cs.titech.ac.jp</u>

識別関数のよさを測る規準

- ■よい識別関数を構成するためには、識別 関数の「よさ」を測る規準が必要.
 - 最大事後確率則
 - 最小誤識別率則
 - ベイズ決定則

最大事後確率則(1)

- ■最大事後確率則(maximum a posteriori probability rule): 入力パターン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.

$$x \in \omega_k : k = \underset{i}{\operatorname{arg\,max}} p(\omega_i \mid x)$$

■決定領域を次のように設定することと等価.

$$D_i = \{x \mid p(\omega_i \mid x) \ge p(\omega_j \mid x) \text{ for all } j \ne i\}$$

最大事後確率則(2)

■ベイズの定理を用いれば、事後確率は

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

p(x) はカテゴリによらないため,

$$p(\omega_i | x) \propto p(x | \omega_i) p(\omega_i)$$

条件付き確率 事前確率

■最大事後確率則は、条件付き確率と事前確率 の積を最大にするカテゴリに分類することと等価.

最小誤識別率則(1)

- ■最小誤識別率則(minimum misclassification rate rule):パターンが誤って分類される確率を最小にするように識別関数を決定
- $p_e(\omega_i \rightarrow \omega_j)$:カテゴリ ω_i に属するパターンが誤ってカテゴリ ω_j に分類される確率
- ■これはカテゴリ ω_i に属するパターンが決定領域 D_i に入る確率と等価.

$$p_e(\omega_i \to \omega_j) = \int_{x \in D_j} p(x \mid \omega_i) dx$$

最小誤識別率則(2)

 $p_e(\omega_i)$:カテゴリ ω_i に属するパターンが誤って他のカテゴリに分類される確率

$$p_{e}(\omega_{i}) = \sum_{j \neq i} p_{e}(\omega_{i} \to \omega_{j})$$

$$= \sum_{j \neq i} \int_{x \in D_{j}} p(x \mid \omega_{i}) dx$$

$$+ \int_{x \in D_{i}} p(x \mid \omega_{i}) dx - \int_{x \in D_{i}} p(x \mid \omega_{i}) dx$$

$$= 1 - \int_{x \in D_{i}} p(x \mid \omega_{i}) dx$$
正解率

最小誤識別率則(3)

■全体の誤識別率 P_e :

$$p_e(\omega_i)$$
 を全力テゴリーに対して平均したもの
$$p_e = \sum_{i=1}^m p_e(\omega_i) p(\omega_i)$$

■最小誤識別率則では、 p_e が最小になるように 識別関数を決定する.

ベイズ決定則(1)

- ■誤識別率最小則に従えば、降水確率40%の時は雨が降らないと識別.
- ■雨が降らないならば傘を持っていく必要はないが、 多くの人は降水確率40%ならば傘を持っていく.
- それは、傘を持っていかなくて雨が降ったときの損失(雨にぬれて風邪をひく)が、傘を持っていって雨が降らなかったときの損失(かばんが少し重くなる)よりもずっと大きいからである.
- ■ベイズ決定則(Bayes decision rule): 誤って識別した時の損失を最小にするように識別

ベイズ決定則(2)

 $l_{i,j}$:カテゴリ ω_i に属するパターンを誤ってカテゴリ ω_i に分類したときの損失(loss)

条件付き危険(conditional risk) $R(\omega_j | x)$: パターン x をカテゴリ ω_j に分類したときの損失の期待値

$$R(\omega_j \mid x) = \sum_{i=1}^m l_{i,j} p(\omega_i \mid x)$$

ベイズ決定則(3)

$$R(\omega_j \mid x) = \sum_{i=1}^m l_{i,j} p(\omega_i \mid x)$$

■条件付き危険が最小になるカテゴリに分類.

$$x \in \omega_k : k = \operatorname{arg\,min}_j R(\omega_j \mid x)$$

■決定領域を次のように設定することと等価.

$$D_i = \{x \mid R(\omega_i \mid x) \le R(\omega_j \mid x) \text{ for all } j \ne i\}$$

ベイズ決定則(4)

■全危険(total risk) *R*:条件付き危険の全ての *x* に関する期待値

$$R = \int_D R(\omega_j \mid x) p(x) dx$$

■ベイズ危険(Bayes risk):ベイズ決定則に 対する全危険の値

演習

- 1.最小後識別率則によって得られる識別規則は、 最大事後確率則によって得られるものと一致 することを示せ、
- 2.ベイズ決定則で、損失が $l_{i,j} = \begin{cases} 0 & (i=j) \\ l & (i \neq j) \end{cases}$

のとき, 即ち, 正しく識別した場合の損失が 0, 誤って識別した場合の損失がカテゴリによらず一定値 1 のとき, ベイズ決定則によって得られる識別規則は, 最大事後確率則によって得られるものと一致することを示せ.

演習1の答え

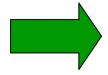
$$p_{e} = \sum_{i=1}^{m} \left(1 - \int_{x \in D_{i}} p(x \mid \omega_{i}) dx \right) p(\omega_{i})$$

$$= \sum_{i=1}^{m} p(\omega_{i}) - \sum_{i=1}^{m} \int_{x \in D_{i}} p(x \mid \omega_{i}) p(\omega_{i}) dx$$

$$= 1 - \sum_{i=1}^{m} \int_{x \in D_{i}} p(x, \omega_{i}) dx$$

$$= 1 - \sum_{i=1}^{m} \int_{x \in D_{i}} p(\omega_{i} \mid x) p(x) dx$$

- 第2項 $\sum_{i=1}^{m} \int_{x \in D_i}^{\infty} p(\omega_i | x) p(x) dx$ を最大にするように 決定領域を決めれば、誤識別率が最小.
- ■これは、全ての $j \neq i$ に対して $p(\omega_i | x) \geq p(\omega_i | x)$ となる x 全体を D_i とすることによって達成.



最小後識別率則=最大事後確率則

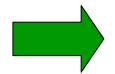
演習2の答え

$$R(\omega_{j} | x) = l \sum_{i \neq j} p(\omega_{i} | x)$$

$$= l \left(\sum_{i=1}^{m} p(\omega_{i} | x) - p(\omega_{j} | x) \right)$$

$$= l \left(1 - p(\omega_{j} | x) \right)$$

$$\propto -p(\omega_{j} | x)$$



ベイズ決定則=最大事後確率則

まとめ

- ■識別関数のよさを測る3つの規準: 最大事後確率則,最小誤識別率則,ベイズ決定則
- ■最大事後確率則と最小誤識別率則は等価.
- ■損失が一定のベイズ決定則は最大事後確率 則(及び最小誤識別率則)と等価.
- ■ベイズ決定則を用いるのが最も自然.
- ■しかし、現実には損失の値がはっきりしなかったり、計算が複雑になるといった理由から、最大事後確率則を用いることも多い。

小レポート(2)

■誤って識別した場合の損失がカテゴリによって異なるようなパターン認識の実例を3つ考えよ、また、それらの例では、損失の値は具体的にいくら位になるであろうか?