Duda & Hart 2.2Bayes Decision Theory-The Continuous Case

作成者: おおつかたく

November 14, 2025

目次

Purpose and Background

目的と背景

- 「ベイズ決定理論（BAYES DECISION THEORY）」は、与えられた入力データ（特徴ベクトル $\mathbf{x}$）を、複数のクラス（$\omega_1, \omega_2, \ldots$）のいずれかに分類する際の最適な意思決定を行うための枠組み
- 連続的な特徴量を持つ場合における各 action が生ずる cost を、損失関数をもちいて表し期待損失を最小化する

Bayes Decision Theory

ベイズの定理（復習）

観測ベクトル x が与えられたとき，事後確率は

$$P(\omega_j \mid x) = \frac{p(x \mid \omega_j)P(\omega_j)}{p(x)}, \tag{1}$$

$$p(x) = \sum_{i=1}^{S} p(x \mid \omega_i)P(\omega_i). \tag{2}$$

- $p(x \mid \omega_j)$: クラス条件付き確率密度（連続の場合）
- $P(\omega_j)$: 事前確率

Generalizations

一般化

1.1 の考えを形式化し 4 つの点で一般化する

1. 複数の特徴の使用を可能にする
2. 自然の状態（カテゴリ、クラス）が 2 つ以上あることを可能にする
3. 自然の状態を決定すること以外の行動（アクション）を可能にする
4. エラーの確率よりも一般的な損失関数（loss function）を導入する

Expected loss (risk) and Decision rule

期待損失（リスク）と決定則 (decision rule)

- 自然状態が ω_j という条件下で行動 α_i によって生じる損失を $\lambda(\alpha_i \mid \omega_j)$ とする。
- 特徴ベクトル $\vec{x}$ を d 個の成分を持つベクトル値確率変数とし、$p(\vec{x} \mid \omega_j)$ を $\vec{x}$ の状態条件付き確率密度関数、すなわち、自然の状態が ω_j であるという条件のもとでの $\vec{x}$ の確率密度関数とする
- 最後に、$P(\omega_j)$ を自然が状態 ω_j にある事前確率とする

期待損失（リスク）と決定則（続き 1）

この時、事後確率 $P(\omega_j \mid \vec{x})$ はベイズ定理によって $p(\vec{x} \mid \omega_j)$ から計算できる：

$$P(\omega_j \mid \vec{x}) = \frac{p(\vec{x} \mid \omega_j)P(\omega_j)}{p(\vec{x})}. \tag{3}$$

ただし、

$$p(\vec{x}) = \sum_{i=1}^{S} p(\vec{x} \mid \omega_i)P(\omega_i). \tag{4}$$

期待損失（リスク）と決定則（続き 2）

- 特定の $\vec{x}$ を観測し、行動 α_i をとることを考えていると仮定する。もし真の自然の状態が ω_j であれば、損失 $\lambda(\alpha_i \mid \omega_j)$ をこうむることになる。
- $P(\omega_j \mid \vec{x})$ は真の自然の状態が ω_j である確率であるため、行動 α_i をとることに関連する期待損失は単に以下のようになる。

$$R(\alpha_i \mid \vec{x}) = \sum_{j=1}^{S} \lambda(\alpha_i \mid \omega_j) P(\omega_j \mid \vec{x}). \tag{5}$$

期待損失（リスク）と決定則（続き 3）

- 期待損失はリスクと呼ばれ、 $R(\alpha_i \mid \vec{x})$ は条件付きリスクとして知られている。
- 特定の観測 $\vec{x}$ に遭遇した時は常に条件付きリスクを最小化する行動を選択することによって期待損失を最小化できる。

Optional Bayes Decision procedure

最適なベイズ決定手順であることを示す

- 全体的なリスク R を最小化し、事前確率 $P(\omega_j)$ に対する Bayes decision rule を見つける。
- Decision rule とはあらゆる観測 $\vec{x}$ に対してどの行動をとるべきかを示す関数 $\alpha(\vec{x})$ のこと
- 具体的にはすべての x に対して決定関数 (The decision function) $\alpha(x)$ は、$\alpha_1, \ldots, \alpha_a$ の a 個の値のいずれかをとる

最適なベイズ決定手順であることを示す（続き 1）

- 全体的なリスク R は与えられた decision rule に関連付けられた期待損失である
- 行動 α_i に関連する条件付きリスクは $R(\alpha_i \mid \vec{x})$ であり、decision rule がその行動を特定するため全体的なリスクは

$$R = \int R(\alpha(\vec{x})|\vec{x})p(\vec{x})\,d\vec{x} \tag{6}$$

この式はあらゆる観測 $\vec{x}$ について、特定の decision rule $\alpha(\vec{x})$ に従って行動した場合の平均期待損失（全体的なリスク）を計算するものであり、Bayes decision rule の核心。

最適なベイズ決定手順であることを示す（続き 2）

- ここで $d\vec{x}$ は d 次元空間の体積要素を示す我々の表記であり、積分は特徴空間全体にわたって行われる。
- 明らかにすべての $\vec{x}$ に対して $R(\alpha(\vec{x}) \mid \vec{x})$ が可能な限り小さくなるように $\alpha(\vec{x})$ を選ばれるならば、全体的なリスクは最小化される。

Bayes risk

ベイズリスク

このことから、Bayes decision rule は次のように正当化されている：

全体的なリスクを最小化するためには、 $\vec{x}$ を観測したときすべての $i = 1, \ldots, a$ について条件付きリスクの式 (5)

$$R(\alpha_i \mid \vec{x}) = \sum_{j=1}^{S} \lambda(\alpha_i \mid \omega_j) P(\omega_j \mid \vec{x}).$$

を計算し、そのうち $R(\alpha_i \mid \vec{x})$ が最小となるような行動 α_i を選択する。

ベイズリスク（続き 1）

- 複数の行動が $R(\alpha_i \mid \vec{x})$ を最小化する場合、これらのうちどの行動をとっても問題ない。
- 都合の良い任意のタイブレーク規則 (tie-breaking rule) を使用できることに注意しろ。
- 結果として得られる最小の全体的なリスクはベイズリスク (The Bayes risk) と呼ばれ、これは達成可能な上限である。

Conclusion

まとめ

- 期待損失 (リスク) $R(\alpha_i \mid \vec{x})$ は特定の観測 $\vec{x}$ に対して最小化する行動 α_i を選択することによって期待損失を最小化できる。
- decision rule に基づく全体的なリスクは式 (6)

$$R = \int R(\alpha(\vec{x})|\vec{x})p(\vec{x})\, d\vec{x}$$

であり、最小の全体的なリスクはベイズリスク (The Bayes risk) と呼ばれ、達成可能な最良な性能である。