1.5.4　推論と決定

茨城大学工学部情報工学科
05T4072R
相原　功昌

クラス分類問題

- 2つの段階

前半 : 推論段階

→訓練データからモデル$p(C_k|\mathbf{x})$を学習する

後半 : 決定段階

→事後確率を使って最適なクラスへ割り当てる

識別関数

- 推論段階と決定段階を同時に解く
  →入力**x**から直接、決定関数を学習する

識別関数

決定問題を解く

- ３つの異なるアプローチ
  →実際の応用では全て使用

アプローチ(a)

（１）クラスの条件付密度$p(\mathbf{x}|C_k)$をクラスC_kごとに決める推論問題を解く

（２）事前クラス確率$p(C_k)$をそれぞれ求める

（３）クラス事後確率$p(C_k|\mathbf{x})$を求める

（４）事後確率と決定理論から、新たな入力$\mathbf{x}$のクラス属性を決める

生成モデル

- 出力の分布だけでなく、入力の分布もモデル化するアプローチのこと

→モデルからのサンプリングによって入力空間で人工データ点を生成できるため

アプローチ(b)

（１）クラス事後確率$p(C_k|\mathbf{x})$を決める推論問題を解く

（２）事後確率と決定理論から、新たな入力$\mathbf{x}$のクラス属性を決める

識別モデル

- 事後確率を直接モデル化するアプローチのこと
- 判別モデルとも言う

アプローチ(c)

（1）識別関数f($\mathbf{x}$)を見つける
→確率は使わない

例　2クラスの場合
・f($\mathbf{x}$)は2値をとる
・f=0⇒C1　、　f=1⇒C2

アプローチ(a)の得失

- 最も手間がかかる
  →**x**とC$_k$の両方の同時分布を求めるから
- クラスの条件付き密度を求めるのに多くの訓練集合が必要
  →多くの応用では、**x**は高次元
- クラス事前分布p(C$_k$)は多くの場合、各クラスの訓練集合の比率で推定できる

アプローチ(a)の得失 （2）

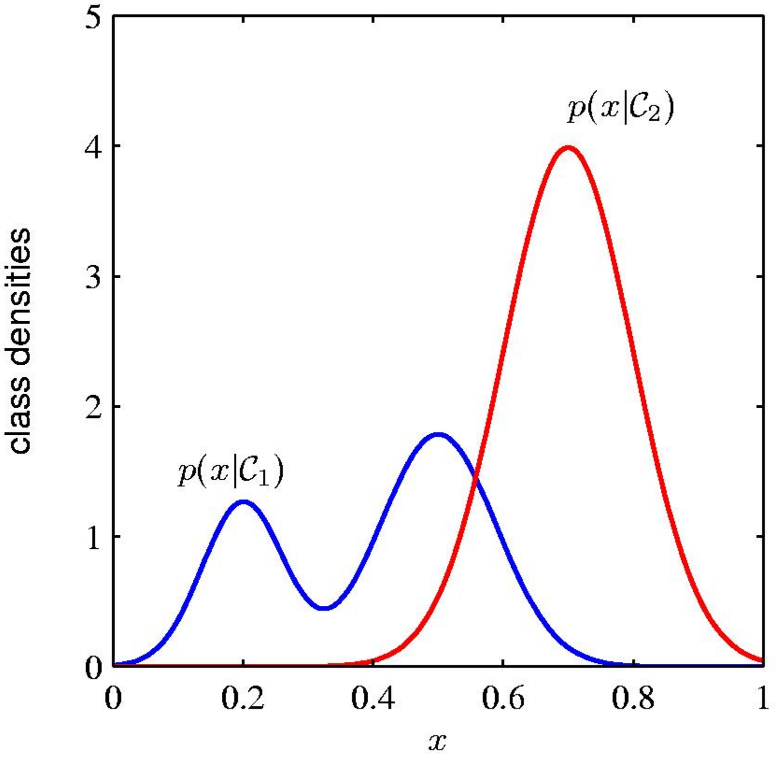
- データの周辺分布p(**x**)を決めることができる
- このモデル下で低い確率を取る新しいデータ点を見つけるのに役立つ

→外れ値検出・新規性検出

アプローチ(b)の得失

- クラス分類を決定するだけが目的なら有効
- アプローチ(a)より計算資源を節約
- 事後確率$p(C_k|\mathbf{x})$を直接得られる

図



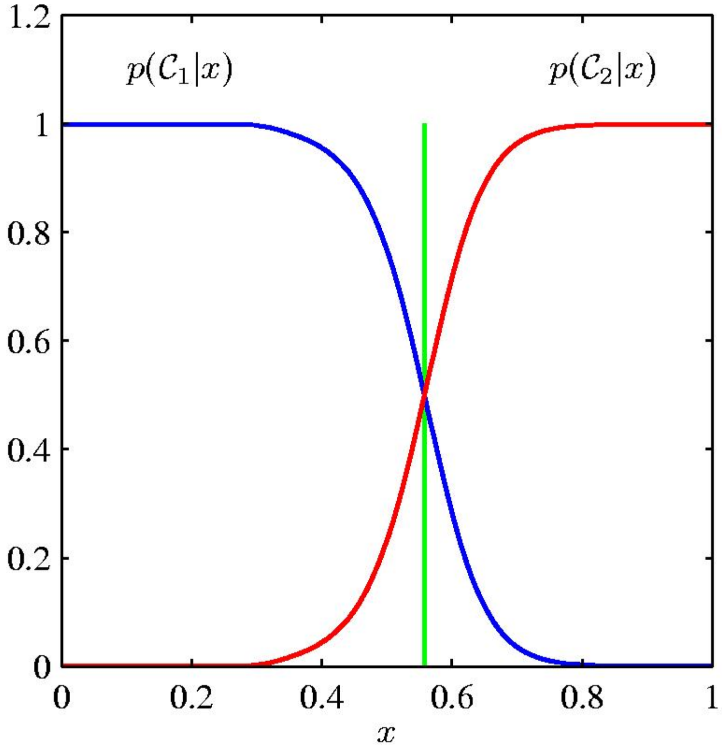


図1.27　2クラスで1入力変数xの場合の条件付き密度（左）と対応する事後確率（右）

アプローチ(c)の得失

- 推論と決定の段階は単一の学習問題に統合
- 誤識別率最小の決定境界を見つけることになる
- 事後確率$p(C_k|\mathbf{x})$に接近できなくなる

事後確率を計算したくなる理由

- リスク最小化
- 棄却オプション
- クラス事前確率の補正
- モデルの結合

リスク最小化

- 損失行列の要素が時間と共に変化するような場合

事後確率がわかっている

→式（1.81）を変更すればいい

$$\sum_k L_{kj}\, p(C_k|\mathrm{x}) \qquad \cdots(1.81)$$

事後確率がわからない

→訓練データに戻ってクラス分類問題を解きなおす

棄却オプション

- 事後確率を用いて棄却規準を定める

  →与えられた棄却データ点に対する誤識別率、あるいは期待損失を最小にすることができる

クラス事前確率の補正

- 各クラスに対して同じくらいの数になるようバランスされたデータ集合

→精緻なモデルを作成できる

- あるクラスに対するデータが少ない時

→事後確率を使用することで、少ないデータに重みをつける

モデルの結合(1)

- 複雑な問題を、多くの小さな部分問題に分割する

→別々のモジュールとして解く

- それぞれのモデルがそれぞれクラスの事後確率を持つ

→出力を確率の規則に則って、系統的に統合（式1.84）

$$p(x_I, x_B \mid C_k) = p(x_I \mid C_k)\, p(x_B \mid C_k) \qquad \cdots(1.84)$$

モデルの結合(2)

$$
\begin{aligned}
p(C_k \mid \mathrm{x}_I, \mathrm{x}_B) &\propto p(\mathrm{x}_I, \mathrm{x}_B \mid C_k)\, p(C_k) \\
&\propto p(x_I \mid C_k)\, p(x_B \mid C_k)\, p(C_k) \\
&\propto \frac{p(C_k \mid \mathrm{x}_I)\, p(C_k \mid \mathrm{x}_B)}{p(C_k)} \qquad \cdots (1.85)
\end{aligned}
$$

•式1.85を用いて、複雑な問題を解くことができる