

「パターン認識と機械学習」

1.5.2 期待損失の最小化

1.5.3 棄却オプション

小野寺喜行

例) 医療診断

□ 癌でない患者を癌と誤診断

□ 癌の患者を健康であると誤診断



こちらを出来るだけ減らした方が良い

損失関数の導入


- 損失関数(loss function)/コスト関数(cost function): 何らかの決定や行動を取ったときに生じる損失を表す尺度
- 全体の損失を最小化することが目標

損失行列

- 入力 \mathbf{X} の新たな値に対して

C_k : 真のクラス

C_j : \mathbf{X}

 このとき

損失の値を L_{kj} で表し k, j 成分とする行列を
損失行列 (loss matrix) と考える

損失行列の例

- 医療診断の場合の例

	癌	正常	←判断
癌	0	1000	
正常	1	0	
↑ 実際			

損失関数の最適解

- 損失関数を最小にするものが最適解
- 未知である真のクラスに依存するので、不確実性を表す同時確率分布 $p(\mathbf{x}, C_k)$ を用いて損失の平均を最小化する

損失の平均

- $$\mathbf{E}[L] = \sum_k \sum_j \int_{R_j} L_{kj} p(\mathbf{x}, C_k) d\mathbf{x}$$
で与えられる

- $p(\mathbf{x}, C_k)$: 真のクラスの不確実性を表す同時確率分布

損失の平均の最小化

- 各 \mathbf{x} は決定領域 R_j のどれかに独立に割り当てることができ、それを期待損失 $E[L]$ が最小になるように選ぶ



各 \mathbf{x} ごとに $\sum_k L_{kj} p(\mathbf{x}, C_k)$ を最小化することを意味する

最小化する決定規則

- 乗法定理 $p(\mathbf{x}, C_k) = p(C_k | \mathbf{x})p(\mathbf{x})$ より共通因子 $p(\mathbf{x})$ を除く
- 新たな \mathbf{X} を以下の量が最小になるようなクラス j に割り当てるのが、期待損失を最小化する決定規則となる

$$\sum_k L_{kj} p(C_k | \mathbf{x})$$

棄却オプション

- 同時確率 $p(x, C_k)$ の値が拮抗している場合、クラス分類の誤差が起きる
- 難しい場合には決定を避けるのが適切な場合もあり、そうすればクラス分類の決定に小さい誤差が見込める



棄却オプション (reject option)

棄却オプションの実現

- しきい値 (threshold) θ を導入し、事後確率 $p(C_k | \mathbf{x})$ の最大値が θ 以下となる入力 \mathbf{x} は棄却する
- θ の値によって棄却される事例の割合を制御
- $\theta = 1$ のとき: すべての事例が棄却
- $\theta < 1/K$ のとき: どの事例も棄却されない
(K はクラス数)