

早稲田大学先進理工学部「パターン認識」

2009年度後期前半レポート課題

担当：赤穂昭太郎 s.akaho@aist.go.jp
http://www.neurosci.aist.go.jp/~akaho/waseda

問1 ある集団からランダムに選ばれた人が、確率 a で病気 X にかかっているとする。
 X を調べるための検査 A は、 X にかかっている人に対しては確率 $1-b$ で陽性、確率 b で陰性を示し、一方 X にかかっていない人に対しては確率 b で陽性、確率 $1-b$ で陰性を示す。

1. ランダムに選ばれた人が検査 A で陽性となる確率 p を求めよ。
2. 検査 A で陽性になった人が病気 X にかかっている確率 q を求めよ。
3. $a = 0.01, b = 0.1$ のとき、それぞれの確率 p, q がどのような値になるか計算せよ。

問2 $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ という n 組のデータに対し、

$$L = \sum_{i=1}^n (y_i - f(x_i))^2$$

が最小となるように関数をあてはめる手法を最小二乗法という。

1. パラメータ a, b をもつ直線 $f(x) = ax + b$ を最小二乗法によりあてはめるとき、 a, b を求める式を導け。
2. $y = f(x) + \epsilon$ というように、 $f(x)$ に ϵ というノイズがのって y が観測されるという確率モデルを考える。 ϵ が独立に同じ正規分布

$$p(\epsilon) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{\epsilon^2}{2\sigma^2}\right)$$

に従うと仮定すると、この確率モデルの最尤推定が最小二乗法と等価であることを示せ。
ここで、最尤推定とは、データが従う確率分布を最大化するようにパラメータの値を定めることをいう。

問3 ユニークな「パターン」の例を挙げ、それをパターン認識することによって得られるメリットをできるだけ具体的に詳しく説明せよ。

例：パターン：迷惑メール

迷惑メールを自動認識することにより迷惑メールを手で分類する手間が省ける

(ただし、この例はユニークではなく、説明が具体性にも欠け、詳しくもないので点数は低いことに注意)