第３回レポート課題

演習問題１～６から２題以上選んで解答する.

演習問題 1. $\mathbb{R} \times \{+1, -1\}$ に値をとる確率変数 (x, y) の分布が以下で与えられるとする.

$$\Pr(y = +1) = 0.4,\ \Pr(y = -1) = 0.6$$

$$p(x|y = +1):\ N(0, 1^2)\text{ の密度関数},\quad p(x|y = -1):\ N(3, 2^2)\text{ の密度関数}.$$

ただし $N(\mu, \sigma^2)$ は期待値 μ，分散 σ^2 の正規分布とする．このとき，x からラベル y を予測するベイズルールを $h_0(x)$ とする.

1. $h_0(1)$, $h_0(2)$ をそれぞれ求めよ.

2. $h_0(x) = +1$ となる x の範囲を求めよ．(数値的に適当な桁まで求めればよい)

演習問題 2. 独立に同一の分布にしたがうデータを $(X_1,Y_1),\dots,(X_n,Y_n)$ とする．仮説集合 $\mathcal{H}$ のなかで，学習誤差 $\widehat{e}(h)$ を最小にする仮説を $\widehat{h}$，また 予測誤差 $e(h)$ を最小にする仮説を $h_{\mathcal{H}}$ とする．$\mathbb{E}[\,\cdot\,]$ を観測データ $(X_1,Y_1),\dots,(X_n,Y_n)$ の分布に関する期待値とするとき，以下の不等式が成り立つことを示せ．

$$\mathbb{E}[\widehat{e}(\widehat{h})] \;\le\; e(h_{\mathcal{H}}) \;\le\; \mathbb{E}[e(\widehat{h})].$$

演習問題 3. ロジスティック損失を用いて２値判別を行う．データ (x,y) に対する判別関数 $f(x)$ のロジスティック損失は

$$\ell(yf(x)) = \log(1+e^{-yf(x)})$$

で与えられる (ただし対数の底は e)．期待損失

$$\mathbb{E}[\ell(yf(x))] = \int \sum_{y=\pm 1} \log(1+e^{-yf(x)})\Pr(Y=y|x)p(x)dx$$

を最小にする判別関数 $f(x)$ を求めよ．また $\mathrm{sign}(f(x))$ がベイズルールに一致するかどうか調べよ．

演習問題 4.

$\mathbb{R} \times \{+1, -1\}$ に値をとる確率変数 (x, y) の分布が以下で与えられるとする.

$$\Pr(y = +1) = q,\ \Pr(y = -1) = 1 - q, \quad q \in (0, 1)$$
$$p(x|y = +1): \text{区間 } [0, 4] \text{ 上の一様分布},$$
$$p(x|y = -1): \text{区間 } [2, 3] \text{ 上の一様分布}.$$

このとき x からラベル y を予測するベイズルールを $h_0(x)$ とする.

1. $h_0(x)$ を求めよ．$h_0(x) = +1$ となる x の範囲と $h_0(x) = -1$ となる x の範囲を, それぞれ明記すること.

2. 仮説集合 $\mathcal{H} = \{\mathrm{sign}(c - x) \,|\, c \in \mathbb{R}\}$ の中で最適な仮説 $h_{\mathcal{H}}$ を求めよ．また予測誤差 $e(h_{\mathcal{H}})$ を求めよ.

note: q の値で場合分けする.

演習問題 **5.1.** $Z_1, \ldots, Z_n \sim_{i.i.d.} P$, $\mathbb{E}[Z_i] = \mu$ とし，Z_i は 0 または 1 の値をとるとする.
チェビシェフの不等式を用いて

$$\Pr\left(\left|\frac{1}{n}\sum_{i=1}^{n} Z_i - \mu\right| \geq \varepsilon\right) \leq \frac{1}{4n\varepsilon^2} \tag{1}$$

が成り立つことを示せ.

2. 有限個の仮説からなる仮説集合 $\mathcal{H}$ に対して，学習誤差の最小化 **(ERM)** により仮説 $\widehat{h}$ を推定する．$\mathcal{H}$ の中で最適な仮説を $h_{\mathcal{H}}$ とする．$1-\delta$ 以上の確率で

$$e(\widehat{h}) < e(h_{\mathcal{H}}) + \varepsilon$$

が成り立つとする．**Hoeffding** 不等式を用いると，ε は $\sqrt{\frac{2}{n}\log\frac{|\mathcal{H}|+1}{\delta}}$ を上界にもつことを示すことができた **(講義参照)**．**Hoeffding** 不等式の代わりに (1) の不等式を用いると，ε の上界は $n, \delta, |\mathcal{H}|$ を用いてどのように表せるか？さらに，δ の値が小さいとき，**Hoeffding** 不等式から導出される上界との大小関係について考察せよ.

演習問題 6. **2**値ラベルデータ (x, y), $x \in \mathbb{R}^d$, $y \in \{+1, -1\}$ に対して，仮説 $h(x)$ の重み付き損失 $L(h(x), y)$ を

$$L(h(x), y) = w(y) \cdot \mathbf{1}[h(x) \neq y]$$

とする．ここで $w(y) > 0$ とする．データ (x, y) が $\mathrm{Pr}(Y = y|x)p(x)$ の分布にしたがうとき，仮説 $h_1(x)$ を

$$h_1(x) = \begin{cases} +1, & w(+1)\mathrm{Pr}(Y = +1|x) \geq w(-1)\mathrm{Pr}(Y = -1|x), \\ -1, & w(+1)\mathrm{Pr}(Y = +1|x) < w(-1)\mathrm{Pr}(Y = -1|x) \end{cases}$$

とする.

1. 次式が成り立つことを示せ.

$$\min_{y'} w(y')\mathrm{Pr}(Y = y'|x) = w(-h_1(x))\mathrm{Pr}(Y = -h_1(x)|x)$$

2. 期待損失

$$\mathbb{E}[L(h(x),y)] = \int \sum_{y=\pm 1} w(y)\mathbf{1}[h(x) \neq y]\Pr(Y=y|x)p(x)dx$$

を最小にする仮説は $h_1(x)$ であることを示せ.

3. 判別関数 $f(x)$ に対する重み付き指数損失を

$$\ell(f(x),y) = \begin{cases} e^{-f(x)}, & y=+1, \\ c\,e^{f(x)}, & y=-1 \end{cases}$$

とする．ここで $c>0$ とする．期待損失

$$\mathbb{E}[\ell(f(x),y)] = \int \sum_{y=\pm 1} \ell(f(x),y)\Pr(Y=y|x)p(x)dx$$

を最小にする判別関数 $f_1(x)$ を求めよ．さらに，$\mathrm{sign}(f_1(x))$ が $h_1(x)$ に一致するように c を定めよ.