

下記問題の解答をレポートにまとめ、LETUS からレポートを提出すること。提出期限 8 月 7 日 (火) 23:59 まで。提出するレポートは pdf ファイルにすること。

1 (制約付き最適化問題) R パッケージ `Rsolnp` では、等式・不等式制約をもつ最適化問題を逐次 2 次計画法を用いて解く `solnp` が用意されている。例えば

$$\begin{aligned} &\text{minimize } f(\mathbf{x}) = (x_1 - 1)^2 + x_2 - 2 \\ &\text{subject to } h(\mathbf{x}) = x_2 - x_1 - 1 = 0 \\ &\quad g(\mathbf{x}) = x_1 + x_2 - 2 \leq 0 \end{aligned}$$

は次のようにして解くことができる(かもしれない)。

```
install.packages("Rsolnp") # パッケージのインストール
library(Rsolnp) # パッケージの読み込み
obj <- function(x) (x[1]-1)^2+x[2] - 2 # 最小化問題の目的関数
eqfun <- function(x) x[2]-x[1]-1
eqB <- 0
ineqfun <- function(x) x[1]+x[2] -2
UB =0
result <- solnp(c(0,0), fun= obj, eqfun=eqfun, eqB=eqB,
               ineqfun = ineqfun, ineqUB =UB,ineqLB=-Inf)
result # 出力を確認すること。
```

また,

$$\begin{aligned} &\text{maximize } f(x_1, x_2) = x_1^2 + (x_2 - 5)^2 \\ &\text{subject to } g_1(x_1, x_2) = x_1 \geq 0 \\ &\quad g_2(x_1, x_2) = x_2 \geq 0 \\ &\quad g_3(x_1, x_2) = 2x_1 + x_2 \leq 4 \end{aligned}$$

に対しては,

```
obj <- function(x) - (x[1]^2+(x[2] - 5)^2) # 最小化問題の目的関数
ineqfun <- function(x) c( x[1], x[2], -2*x[1]-x[2]+4)
ineqLB <- c(0,0,0)
result <- solnp(c(0,.2), fun= obj, eqfun=eqfun, eqB=eqB,
               ineqfun = ineqfun, ineqLB =ineqLB,ineqUB=rep(Inf,3))
result # 出力を確認すること。
```

これを参考に、2018 年期末試験3のサポートベクトルマシンのデータセットを用いて、制約付き最適化問題を定式化して解きなさい。(制約付きの最適化問題のソルバーは `Rsolnp` 以外にもたくさんあるので何を用いてもよいが、関数 `svm` 等を直接用いてはならない。)

2 (パーセプトロン型学習)

1. 講義資料「パターン認識第 10 回: サポートベクトルマシン」から, 完全分離できるデータとして LDA を用いた 2 次元に縮約したワインデータの任意の 2 つのクラスラベルのデータセットを用いてパーセプトロン型学習を行いたい. ただし, 学習規則は次のようであった. 学習データの系列を $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_i, \dots$ とする. パーセプトロンの学習規則は $i+1$ 番目の係数ベクトルを \mathbf{w}_{i+1} を, i 番目の学習データ \mathbf{x}_i を入力したときの出力 $f(\mathbf{x}_i)$ に応じて,

$$\begin{cases} f(\mathbf{x}_i) \geq 0 \text{ であれば} & \mathbf{w}_{i+1} = \mathbf{w}_i \\ f(\mathbf{x}_i) < 0 \text{ であれば} & \mathbf{w}_{i+1} = \mathbf{w}_i + \eta \mathbf{x}_i \end{cases}$$

とする.

2. パーセプトロン型学習規則のアルゴリズムのコーディングを行い, 上記データに適用したときの, 最適な線形識別規則のベクトル \mathbf{w}_0 とバイアス項 b_0 を求めよ. ただし, 初期値は η は必要に応じて各自調整すること.
3. 平面上にデータ点と識別境界の直線をプロットせよ.

3 (ニューラル・ネットワーク) R パッケージ `gclus` には `wine` データが用意されている. 3 つのぶどう品種の 178 ワインについて 13 変数の化学成分を記録したデータである. このデータを用いて誤差逆伝搬法の学習規則 (ニューラル・ネットワークモデル) を用いた分類を行え. このとき, 隠れ素子数の数 $M(\text{size})$ や `decay` パラメータの最適化は, CV 法やホールドアウト法を用いた誤り率の推定から行え.

4 (多クラスのサポートベクトルマシン) サポートベクトルマシンを用いた多クラス分類には, 一対一法を使う. 例えば, ワインデータの分類は次のように行う.

#事前に `wine` データの読み込みを済ましていることが前提

```
svmfit=svm( as.factor(wine$Class)~., data=wine[, -13], kernel="linear", cost=10, gamma=1)
table(wine$Class, predict(svmfit)) ## 誤ることなくすべてのデータを分類できた.
```

R パッケージ `ISLR` のデータ `Khan` は, 異なる 4 類の線維形成性小細胞腫瘍 (`small round blue cell tumors`) の遺伝子発現データである.

```
library(ISLR) # install.packages("ISLR") が済んでいる場合.
```

```
data(Khan)
```

```
?Khan
```

```
table(Khan$ytrain)
```

```
table(Khan$ytest)
```

SVM を用いた他クラス分類を行い, テストデータによる予測の精度を評価せよ. また, このようなデータの分類には SVM が適している理由を述べよ.

5 (サポートベクトルマシン) 講義資料「パターン認識 第8回: パーセプトロン型学習」で紹介した縮小した手書き数字認識のデータセット (MNIST) データを用いて, SVMを用いた他クラス分類を行え. 得られた結果はニューラルネットワークを用いた結果と比較し, その違いや長短について述べよ.