B529: Homework 1

Nathan Byers

Tuesday, February 03, 2015

Question 1

Please use the perceptron algorithm to manually find a weight vector $\mathbf{w} = [\text{w0,w1,w2}]^{\mathrm{T}}$ *for linear classification of 3 data points* ($\mathbf{x}$i,yi) *satisfying sign* ($\mathbf{w}^{\mathrm{T}}\mathbf{x}$i) = yi. *The input data is*

$\mathbf{x}1=(1,2,2)^{\mathrm{T}}$ y1=+1

$\mathbf{x}2=(1,1,2)^{\mathrm{T}}$ y2=+1

$\mathbf{x}3=(1,2,0)^{\mathrm{T}}$ y3=-1

The first element in vector $\mathbf{x}$ is always x0=1 *(See slides in lecture 1).*

In the first step of the perceptron algorithm, the initial assignment of $\mathbf{w}$ is $[0,6,6]^{\mathrm{T}}$. *Please provide the intermediate steps of the computation. (30 points)*

Answer 1

- Step 1: $\mathbf{w} = [0,6,6]^{\mathrm{T}}$
- Step 2:

$$sign(\boldsymbol{w}^T\boldsymbol{x}_1) = sign\left(\begin{bmatrix}0 & 6 & 6\end{bmatrix} \cdot \begin{bmatrix}1\\2\\2\end{bmatrix}\right) = sign(24) = 1$$

$$sign(\boldsymbol{w}^T\boldsymbol{x}_1) = sign\left(\begin{bmatrix}0 & 6 & 6\end{bmatrix} \cdot \begin{bmatrix}1\\1\\2\end{bmatrix}\right) = sign(18) = 1$$

$$sign(\boldsymbol{w}^T\boldsymbol{x}_1) = sign\left(\begin{bmatrix}0 & 6 & 6\end{bmatrix} \cdot \begin{bmatrix}1\\2\\0\end{bmatrix}\right) = sign(12) \neq -1$$

$$\boldsymbol{w} = \boldsymbol{w} + y_3\boldsymbol{x}_3 = \begin{bmatrix}0\\6\\6\end{bmatrix} + -1 \cdot \begin{bmatrix}1\\2\\0\end{bmatrix} = \begin{bmatrix}-1\\4\\6\end{bmatrix}$$

- Step 2: iteration

$$sign(\boldsymbol{w}^T\boldsymbol{x}_1) = sign\left(\begin{bmatrix}-1 & 4 & 6\end{bmatrix} \cdot \begin{bmatrix}1\\2\\2\end{bmatrix}\right) = sign(19) = 1$$

$$sign(\boldsymbol{w}^T\boldsymbol{x}_1) = sign\left(\begin{bmatrix}-1 & 4 & 6\end{bmatrix} \cdot \begin{bmatrix}1\\1\\2\end{bmatrix}\right) = sign(15) = 1$$

$$sign(\boldsymbol{w}^T\boldsymbol{x}_1) = sign\left(\begin{bmatrix}-1 & 4 & 6\end{bmatrix} \cdot \begin{bmatrix}1\\2\\0\end{bmatrix}\right) = sign(7) \neq -1$$

$$\boldsymbol{w} = \boldsymbol{w} + y_3\boldsymbol{x}_3 = \begin{bmatrix}-1\\4\\6\end{bmatrix} + -1 \cdot \begin{bmatrix}1\\2\\0\end{bmatrix} = \begin{bmatrix}-2\\2\\6\end{bmatrix}$$

- Step 2: iteration

$$sign(\boldsymbol{w}^T\boldsymbol{x}_1) = sign\left(\begin{bmatrix}-2 & 2 & 6\end{bmatrix} \cdot \begin{bmatrix}1\\2\\2\end{bmatrix}\right) = sign(14) = 1$$

$$sign(\boldsymbol{w}^T\boldsymbol{x}_1) = sign\left(\begin{bmatrix}-2 & 2 & 6\end{bmatrix} \cdot \begin{bmatrix}1\\1\\2\end{bmatrix}\right) = sign(12) = 1$$

$$sign(\boldsymbol{w}^T\boldsymbol{x}_1) = sign\left(\begin{bmatrix}-2 & 2 & 6\end{bmatrix} \cdot \begin{bmatrix}1\\2\\0\end{bmatrix}\right) = sign(2) \neq -1$$

$$\boldsymbol{w} = \boldsymbol{w} + y_3\boldsymbol{x}_3 = \begin{bmatrix}-2\\2\\6\end{bmatrix} + -1 \cdot \begin{bmatrix}1\\2\\0\end{bmatrix} = \begin{bmatrix}-4\\0\\6\end{bmatrix}$$

- Step 2: iteration

$$sign(\boldsymbol{w}^T\boldsymbol{x}_1) = sign\left(\begin{bmatrix}-4 & 0 & 6\end{bmatrix} \cdot \begin{bmatrix}1\\2\\2\end{bmatrix}\right) = sign(8) = 1$$

$$sign(\boldsymbol{w}^T\boldsymbol{x}_1) = sign\left(\begin{bmatrix}-4 & 0 & 6\end{bmatrix} \cdot \begin{bmatrix}1\\1\\2\end{bmatrix}\right) = sign(8) = 1$$

$$sign(\boldsymbol{w}^T\boldsymbol{x}_1) = sign\left(\begin{bmatrix}-4 & 0 & 6\end{bmatrix} \cdot \begin{bmatrix}1\\2\\0\end{bmatrix}\right) = sign(-4) = -1$$

Final $\mathbf{w} = (-4, 0, 6)^{\mathrm{T}}$

Question 2

Please use R to manually compute the weight vector $\mathbf{w} = [w0, w1, w2]^{\mathrm{T}}$ *for linear regression of N input data points* (xi,yi) *that minimizes the sum of squared error* $(\mathbf{w}^{\mathrm{T}}\mathbf{x}i - yi)^2$. *The input data is*

$\mathbf{x}1=(1,2,3)^{\mathrm{T}}$ y1 = 2

$\mathbf{x}2=(1,1,2)^{\mathrm{T}}$ y2 = 1

$\mathbf{x}3=(1,1,0)^{\mathrm{T}}$ y3 = -1

$\mathbf{x}4=(1,-1,-2)^{\mathrm{T}}$ y4= -2

The first element in vector $\mathbf{x}$ *is always* x0=1 *(See slides in lecture 1)*

Please provide the intermediate steps of the computation (20 points).

Answer 2

```
X <- matrix(c(1, 2, 3, 1, 1, 2, 1, 1, 0, 1, -1, -2), nrow = 4, byrow = TRUE)
X
```

```
##      [,1] [,2] [,3]
## [1,]    1    2    3
## [2,]    1    1    2
## [3,]    1    1    0
## [4,]    1   -1   -2
```

```
y <- matrix(c(2, 1, -1, -2), ncol = 1)
y
```

```
##      [,1]
## [1,]    2
## [2,]    1
## [3,]   -1
## [4,]   -2
```

```
Xt <- t(X)
Xt
```

```
##      [,1] [,2] [,3] [,4]
## [1,]    1    1    1    1
## [2,]    2    1    1   -1
## [3,]    3    2    0   -2
```

```
w <- solve(Xt %*% X) %*% Xt %*% y
w
```

```
##       [,1]
## [1,] -0.45
## [2,] -0.45
## [3,]  1.05
```

Question 3

In Fisher's linear discriminant analysis, we search for a vector **w** *such that all data points are well separately after they are projected to the direction defined by* **w**. *When the input data points are*

x1=$(3,4)^T$ y1 = 1

x2=$(1,2)^T$ y2 = 1

x3=$(1,0)^T$ y3 = -1

x4=$(1,-2)^T$ y4= -1

(Note that the first element in vector **x** *is NOT x0=1 in this problem), please use R to manually find a vector* **w**=[w1,w2] *with Fisher's LDA and provide the intermediate steps (20 points).*

Answer 3

Here I create a matrix with each column containing an **x** vector as well as the y scalar.

```
Xy <- matrix(c(3, 4, 1, 1, 2, 1, 1, 0, -1, 1, -2, -1), ncol = 4)
colnames(Xy) <- c("x1hat", "x2hat", "x3hat", "x4hat")
rownames(Xy) <- c("x1", "x2", "y")
Xy
```

```
##    x1hat x2hat x3hat x4hat
## x1     3     1     1     1
## x2     4     2     0    -2
## y      1     1    -1    -1
```

Now I calculate the **m** vectors.

```
m1hat <- matrix(apply(Xy[, Xy[3, ] == 1], 1, mean)[-3], ncol = 1)
rownames(m1hat) <- c("m1", "m2")
m1hat
```

```
##    [,1]
## m1    2
## m2    3
```

```
m2hat <- matrix(apply(Xy[, Xy[3, ] == -1], 1, mean)[-3], ncol = 1)
rownames(m2hat) <- c("m1", "m2")
m2hat
```

```
##    [,1]
## m1    1
## m2   -1
```