Exam 1

xian chen

**Question1 (a)**

$x\_{i}=\left[\begin{matrix}cos\left(\frac{4i−3}{10}π\right)\\sin\left(\frac{4i−3}{10}π\right)\end{matrix}\right]$
$x\_{1}=(0.9510565,0.309017)^{T}$
$x\_{2}=(6.123233995736766e−17,1.0)^{T}$
$x\_{3}=(−0.9510565162951535,0.3090169943749475^{T}$
$x\_{4}=(−0.5877852522924732,−0.8090169943749473)^{T}$
$x\_{5}=(0.5877852522924729,−0.8090169943749476)^{T}$
$X\_{(5×2)}=\left[\begin{matrix}0.95&0.31\\0&1.0\\−0.95&0.31\\−0.59&−0.81\\0.59&−0.81\end{matrix}\right]$ $y\_{(5×1)}=\left[\begin{matrix}1\\1\\1\\−1\\−1\end{matrix}\right]$

Since $H^{(t)}=\left\{x|β\_{0}^{(t)}+x^{⊤}β\_{1}^{(t)}=0\right\}$, $β\_{0}^{(0)}=−0.5$ and $β\_{1}^{(0)}=\left[\begin{matrix}0\\1\end{matrix}\right]$, we have $0X1+X2=0.5\rightarrow X2=0.5$.



**(b)** We know that signed-distance of any point x to a hyperplane is $d(x\_{i},H)=\frac{x\_{i}^{T}β+β\_{0}}{∥β∥}$.

Hence, $d(x\_{i},H^{(0)})=\frac{x\_{i}^{T}β\_{1}^{(0)}+β\_{0}^{(0)}}{∥β\_{1}^{(0)}∥}$.

$d(x\_{1},H^{(0)})=−0.19$

$d(x\_{2},H^{(0)})=0.5$

$d(x\_{3},H^{(0)})=−0.19$

$d(x\_{4},H^{(0)})=−1.31$

$d(x\_{5},H^{(0)})=−1.31$

**(c)**

From part(b), we know $x\_{1}$ and $x\_{2}$ are misclassified because $d(x\_{1},H^{(0)})<0$ and $d(x\_{3},H^{(0)})<0$ but $y\_{1}=y\_{3}=+1$.

Thus, $M^{(0)}$ contains $x\_{1}$ and $x\_{3}$.

**(d)**

The sum of unsigned distances of misclassified data points is given by

$$D\left(β,β\_{0}\right)=−\sum\_{i\in M}^{}y\_{i}\left(x\_{i}^{⊤}β+β\_{0}\right)$$

Thus, for $β\_{0}=β\_{0}^{(0)},β\_{1}=β\_{1}^{(0)}$ and $M=M^{(0)}$,

$$\begin{matrix}D(β\_{1}^{(0)},β\_{0}^{(0)})&=−y\_{1}(x\_{1}β\_{1}^{(0)}+β\_{0}^{(0)})−y\_{3}(x\_{3}β\_{1}^{(0)}+β\_{0}^{(0)})\\&=0.38\end{matrix}$$

**(e)**

Given $\left(\begin{matrix}β\\β\_{0}\end{matrix}\right)\leftarrow \left(\begin{matrix}β\\β\_{0}\end{matrix}\right)+ρ\left(\begin{matrix}y\_{i}x\_{i}\\y\_{i}\end{matrix}\right)$ and $ρ=0.5$,

$$\begin{matrix}\left(\begin{matrix}β\_{1}^{(1)}\\β\_{0}^{(1)}\end{matrix}\right)&=\left(\begin{matrix}β\_{1}^{(0)}\\β\_{0}^{(0)}\end{matrix}\right)+0.5\left(\begin{matrix}y\_{1}x\_{1}\\y\_{1}\end{matrix}\right)\\&=\left(\begin{matrix}0\\1\\−0.5\end{matrix}\right)+0.5\left(\begin{matrix}0.95\\0.31\\1\end{matrix}\right)\\&=\left(\begin{matrix}0.475\\1.155\\0\end{matrix}\right)\end{matrix}$$

Hence, $β\_{0}^{(1)}=0$ and $β\_{1}^{(1)}=\left[\begin{matrix}0.475\\1.155\end{matrix}\right]$.

**(f)**

Since $H^{(1)}=\left\{x|β\_{0}^{(1)}+x^{⊤}β\_{1}^{(1)}=0\right\}$, where $β\_{0}^{(1)}=0$ and $β\_{1}^{(1)}=\left[\begin{matrix}0.475\\1.155\end{matrix}\right]$, we have $0.475\*X1+1.155\*X2=0\rightarrow X2=−0.475\*X1/1.155$.



**(g)**

The signed-distance of any point x to a hyperplane is $d(x\_{i},H)=\frac{x\_{i}^{T}β+β\_{0}}{∥β∥}$.

Hence, $d(x\_{i},H^{(1)})=\frac{x\_{i}^{T}β\_{1}^{(1)}+β\_{0}^{(1)}}{∥β\_{1}^{(1)}∥}$.

$d(x\_{1},H^{(1)})=0.65$

$d(x\_{2},H^{(1)})=0.92$

$d(x\_{3},H^{(1)})=−0.08$

$d(x\_{4},H^{(1)})=−0.97$

$d(x\_{5},H^{(1)})=−0.52$

**(h)**

Only $x\_{3}$ is misclassified because $d(x\_{3},H^{(1)})<0$ while $y\_{3}=+1$.

Thus, $M^{(1)}$ contains $x\_{3}$.

**(i)**

The sum of unsigned distances of misclassified data points is given by

$$D\left(β,β\_{0}\right)=−\sum\_{i\in M}^{}y\_{i}\left(x\_{i}^{⊤}β+β\_{0}\right)$$

Thus, for $β\_{0}=β\_{0}^{(1)},β\_{1}=β\_{1}^{(1)}$ and $M=M^{(1)}$,

$$\begin{matrix}D(β\_{1}^{(1)},β\_{0}^{(1)})&=−y\_{3}(x\_{3}β\_{1}^{(1)}+β\_{0}^{(1)})\\&=0.09\end{matrix}$$

We can see that sum of unsigned distances of misclassified data points is decreased.

**(j)**

Given $\left(\begin{matrix}β\\β\_{0}\end{matrix}\right)\leftarrow \left(\begin{matrix}β\\β\_{0}\end{matrix}\right)+ρ\left(\begin{matrix}y\_{i}x\_{i}\\y\_{i}\end{matrix}\right)$ and $ρ=0.5$,

$$\begin{matrix}\left(\begin{matrix}β\_{1}^{(2)}\\β\_{0}^{(2)}\end{matrix}\right)&=\left(\begin{matrix}β\_{1}^{(1)}\\β\_{0}^{(1)}\end{matrix}\right)+0.5\left(\begin{matrix}y\_{3}x\_{3}\\y\_{3}\end{matrix}\right)\\&=\left(\begin{matrix}0.475\\1.155\\0\end{matrix}\right)+0.5\left(\begin{matrix}−0.95\\0.31\\1\end{matrix}\right)\\&=\left(\begin{matrix}0.475\\1.155\\0\end{matrix}\right)\end{matrix}$$

Hence, $β\_{0}^{(2)}=0.5$ and $β\_{1}^{(2)}=\left[\begin{matrix}0\\1.31\end{matrix}\right]$.

**(k)**

Since $H^{(2)}=\left\{x|β\_{0}^{(2)}+x^{⊤}β\_{1}^{(2)}=0\right\}$, where $β\_{0}^{(2)}=−0.5$ and $β\_{1}^{(2)}=\left[\begin{matrix}0\\1.31\end{matrix}\right]$, we have $0\*X1+1.31\*X2+0.5=0\rightarrow X2=−0.5/1.31=−0.38$.



**(l)**

The signed-distance of any point x to a hyperplane is $d(x\_{i},H)=\frac{x\_{i}^{T}β+β\_{0}}{∥β∥}$.

Hence, $d(x\_{i},H^{(2)})=\frac{x\_{i}^{T}β\_{1}^{(2)}+β\_{0}^{(2)}}{∥β\_{1}^{(2)}∥}$.

$d(x\_{1},H^{(2)})=0.69$

$d(x\_{2},H^{(2)})=1.38$

$d(x\_{3},H^{(2)})=0.69$

$d(x\_{4},H^{(2)})=−0.43$

$d(x\_{5},H^{(2)})=−0.43$

**(m)**

Since the signed-distance of x1,x2 and x3 from $H^{2}$ are greater than 0 and the signed-distance of of x4 and x5 from $H^{2}$ are smaller than 0, all data points are correctly classified.

Thus, $M^{(2)}$ is empty.

**Question2 (a)**

$\begin{matrix}y&=Xβ+(c.v)\\&=X\_{β}+X⋅v⋅c\end{matrix}$

since $Xv=0$, we have $y=Xβ+(c.v)=X\_{β}$. Thus, $β+c.v$ also satisfies the linear equation $y=Xβ$, for any $c\in R$.

**(b)**

If the columns of $X$ are linearly independent, then $v=0$ satisfies $Xv=0$ $v\in R^{p}$.

**(c)**

If the columns of $X$ are not linearly independent($p>n$), then there are can be many solutions for $v$. $X\_{1}v=X\_{11}v\_{1}+X\_{12}v\_{2}+...+X\_{1p}v\_{p}=0$, not all $v\_{i}$ needs to be zero.