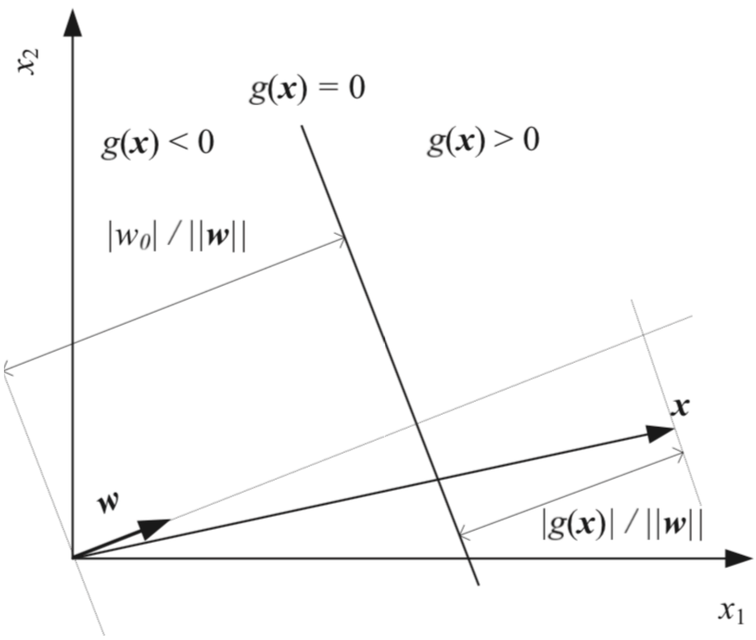
CSCI3320: Homework 7, Spring 2020 Deadline: April 13, 2020, 11:59 pm
Teacher: John C.S. Lui

1. *For the lecture of linear discriminant, we have:*

$$\boldsymbol{x} = \boldsymbol{x}_p + r\frac{\boldsymbol{w}}{\|\boldsymbol{w}\|}$$

and point x is in the corresponding figure.



Point ***x*** *and its projection distance on* ***w*** *is r, we have*

$$r = \frac{g(\boldsymbol{x})}{\|\boldsymbol{w}\|} \tag{10.4}$$

We see then that the distance to origin is

$$r_0 = \frac{w_0}{\|\boldsymbol{w}\|} \tag{10.5}$$

Prove the equalities of (10.4) and (10.5).

Answer:

Given figure 10.1, first let us take input $\boldsymbol{x}_0$ on the hyperplane. The angle between $\boldsymbol{x}_0$ and $\boldsymbol{w}$ is a_0 and because it is on the hyperplane

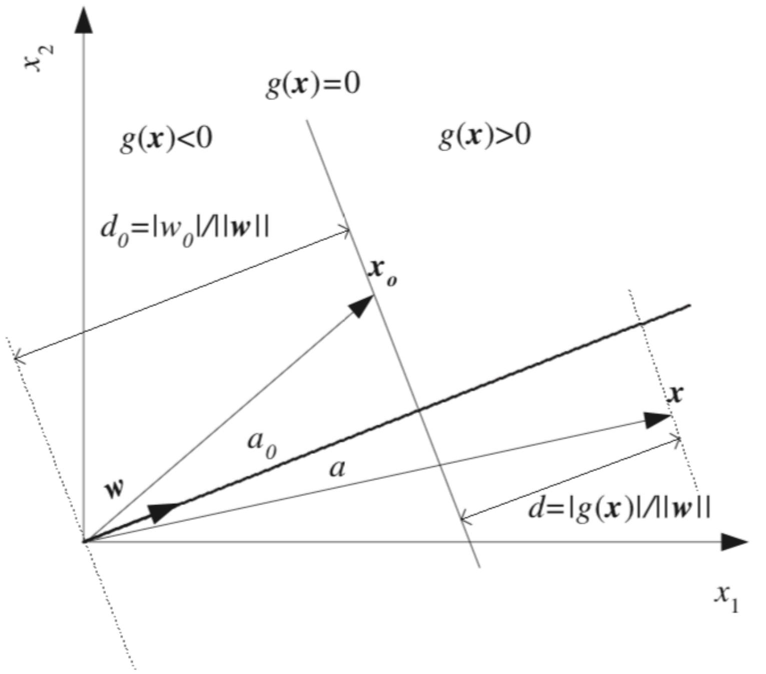


Figure 10.1 The geometric interpretation of the linear discriminant.

$g(\boldsymbol{x}_0) = 0$. Then

$$
\begin{aligned}
g(\boldsymbol{x}_0) &= \boldsymbol{w}^T\boldsymbol{x}_0 + w_0 = \|\boldsymbol{w}\|\|\boldsymbol{x}_0\|\cos a_0 + w_0 = 0 \\
d_0 &= \|\boldsymbol{x}_0\|\cos a_0 = \frac{|w_0|}{\|\boldsymbol{w}\|}
\end{aligned}
$$

For any $\boldsymbol{x}$ with angle a to $\boldsymbol{w}$, similarly we have

$$
\begin{aligned}
g(\boldsymbol{x}) &= \boldsymbol{w}^T\boldsymbol{x} + w_0 = \|\boldsymbol{w}\|\|\boldsymbol{x}\|\cos a + w_0 \\
d &= \|\boldsymbol{x}\|\cos a - \frac{w_0}{\|\boldsymbol{w}\|} = \frac{g(\boldsymbol{x}) - w_0 + w_0}{\|\boldsymbol{w}\|} = \frac{g(\boldsymbol{x})}{\|\boldsymbol{w}\|}
\end{aligned}
$$

2. *Show that the derivative of the softmax, $y_i = \exp(a_i)/\sum_j \exp(a_j)$, is $\partial y_i/\partial a_j = y_i(\delta_{ij} - y_j)$, where δ_{ij} is 1 if $i = j$ and 0 otherwise.*

Answer: Given that

$$y_i = \frac{\exp a_i}{\sum_j \exp a_j}$$

for $i = j$, we have

$$\begin{aligned}
\frac{\partial y_i}{\partial a_i} &= \frac{\exp a_i \left(\sum_j \exp a_j\right) - \exp a_i \exp a_j}{\left(\sum_j \exp a_j\right)^2} \\
&= \frac{\exp a_i}{\sum_j \exp a_j} \left(\frac{\sum_j \exp a_j - \exp a_i}{\sum_j \exp a_j}\right) \\
&= y_i(1 - y_i)
\end{aligned}$$

and for $i \neq j$, we have

$$\begin{aligned}
\frac{\partial y_i}{\partial a_j} &= \frac{-\exp a_i \exp a_j}{\left(\sum_j \exp a_j\right)^2} \\
&= -\left(\frac{\exp a_i}{\sum_j \exp a_j}\right)\left(\frac{\sum_j \exp a_j}{\sum_j \exp a_j}\right) \\
&= y_i(0 - y_j)
\end{aligned}$$

which we can combine in one equation as

$$\frac{\partial y_i}{\partial a_i} = y_i(\delta_{ij} - y_j)$$