HW # 3 (written)

Problem 5.2, 5.4

$\boxed{5.2}$

Consider a linear machine with discriminant functions
$g_i(x) = w^t x + w_{i0}$, $i = 1, \dots, c$
show that the decision regions are convex by showing that if $x_1 \in R_i$ and $x_2 \in R_i$, then
$\lambda x_1 + (1-\lambda) x_2 \in R_i$ if $0 \leq \lambda \leq 1$

g_i is an i-class discriminant

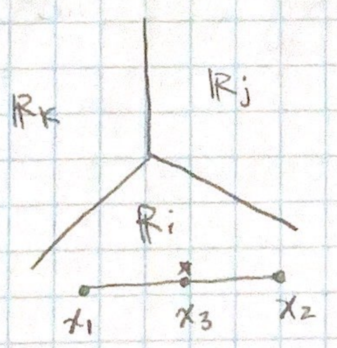
$$g_i(x) = w^t x + w_{i0} \quad , \quad i = 1, \dots, c$$

row vector ↑ w^t , column ↑ x

- The decison boundary between class C_j & C_k is given by $g_k(x) = g_j(x)$

- The hyperplane of this boundary is defined by
$(w_k - w_j)^T x + (w_{k0} - w_{j0}) = 0$

Proof: decision regions of these discrimants are convex



$x_1 \in R_i$, $x_2 \in R_i$

- any combination of these points will also be within R_i

$$\lambda x_1 + (1-\lambda) x_2 = \overset{*}{x_3} \in R_i$$

If we form the classifier w/ this point x_3
$g_i(x_3) = \lambda g_i(x_1) + (1-\lambda) g_i(x_2)$
- we have assumed x_1 & x_2 are in R_i

$\Rightarrow g_i(x_1) > g_{other}(x_1)$ and $g_i(x_2) > g_{other}(x_2)$
- multiplication by positive λ and $(1-\lambda)$ won't change the sign of this inequality

$\therefore$ $g_i(x_3) > g_{other}(x_3)$

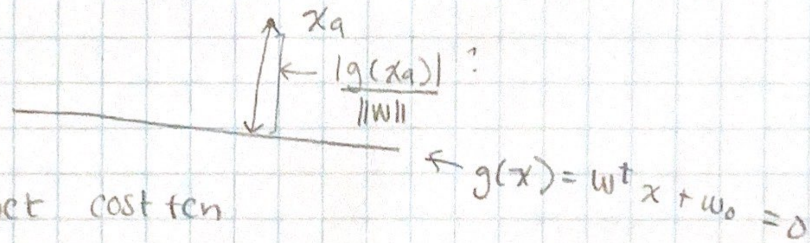
$\boxed{5.4}$

Consider the hyperplane used for discriminant fcns

a) Show that the distance from the hyperplane $g(x) = w^t x + w_0 = 0$ to the point x_a is $|g(x_a)|/\|w\|$ by minimizing $\|x - x_a\|^2$ subj to the constraint $g(x) = 0$

b) Show that the projection of x_a onto the hyperplane is given by $x_p = x_a - \dfrac{g(x_a)}{\|w\|^2} w$

$\boxed{\text{part a}}$



construct cost fcn

$$\operatorname{argmin} \; J(x) = \|x - x_a\|^2$$

$$\text{subj to } g(x) = 0$$

Use lagrangian multipliers

$$L(x, \lambda) = \|x - x_a\|^2 - \lambda(w^t x + w_0)$$

$$\frac{\delta L}{\delta x} = \frac{\delta J(x)}{\delta x} - \lambda \frac{\delta g(x)}{\delta x} = 0$$

$$(*) \quad = 2(x - x_a) - \lambda(w) = 0$$

transposed must be removed for math to work

$$\frac{\delta L}{\delta \lambda} = g(x) = w^t x + w_0 = 0$$

$$\lambda(w) = 2(x - x_a)$$

since $g(x) = 0$, $\lambda \neq 0$ from KKT conditions.

$$\lambda = \frac{2 w^T (x - x_a)}{w^T w}$$

multiply $(*)$ by $(x - x_a)^T$

$$2(x - x_a)^T (x - x_a) - \lambda (x - x_a)^T w = 0$$

over $\rightarrow$

$\boxed{\text{cont}}$

substitute expression for λ, move expression to one side

$$\cancel{2}(x - x_a)^T(x - x_a) = \frac{\cancel{2}w^T(x - x_a)}{w^T w}(x - x_a)^T w$$

$$(x - x_a)^2 = \frac{w^T(x - x_a)^2}{w^T w}$$

Square root of both sides

$$\|x - x_a\|_2 = \frac{w^T(x - x_a)}{\|w\|_2}$$

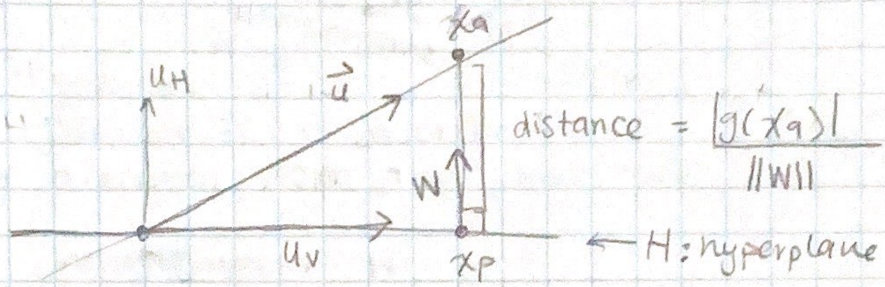
$$w^T(x) - w^T x_a = 0$$

$*(-1)$

$$w^T x_a + \underset{\substack{\uparrow \\ w_0}}{w^T(x)} = 0 \xleftarrow{??} g(x_a)$$

$$\|x - x_a\|_2 = \frac{w^T(x - x_a)}{\|w\|_2}$$

b.)



w is the normal vector that forms the hyperplane H

proj of a vector $\vec{u}$ onto another vector $\vec{v}$

$$\text{proj}_{\vec{v}}(\vec{u}) = \frac{\vec{u} \cdot \vec{v}}{||\vec{v}||^2} \vec{v}$$

proj of $\vec{u}$ onto a plane

- → subtract the component of $\vec{u}$ that is orthogonal to the plane from $\vec{u}$
- → $\vec{u}$ minus the vertical component of $\vec{u}$ ($\vec{u}_H$)
- → $\vec{u}_H$ can also be thought of as the projection of $\vec{u}$ onto the normal vector $\vec{w}$

$$u - \text{proj}_{\vec{n}}(\vec{u}) = \vec{u} - \frac{\vec{u} \cdot \vec{n}}{||\vec{n}||^2} \vec{n}$$

Projection of the vector x_a onto the hyperplane is:

$$x_p = x_a - \frac{w^T(x - x_a)}{||w||^2} w = x_a - \frac{g(x_a)}{||w||^2} w$$