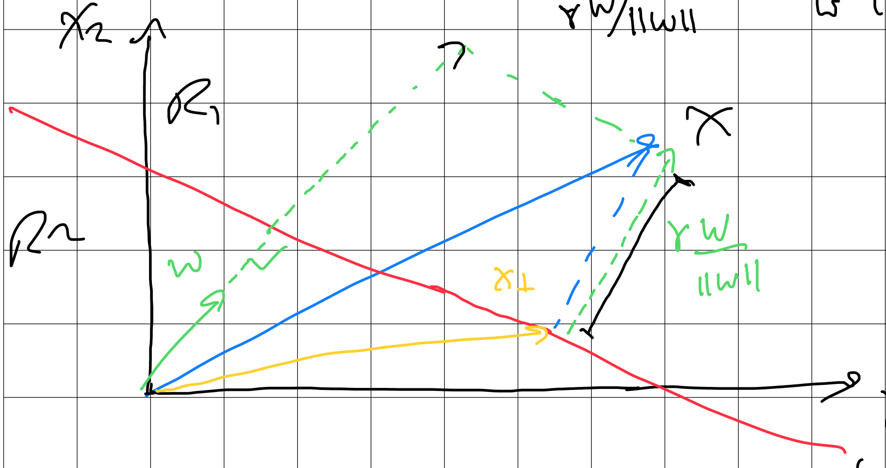
Función discriminante

$$g(x_A) = g(x_B) = 0$$
$$w^T x_A + b = w^T x_B + b = 0$$
$$w^T(x_A - x_B) = 0 \; ; \; w^T x_A = w^T x_B = 0$$



$$\|w\|^2 = w^T w$$

$$g(x) = w^T x + b$$

$$\|x - g(x)\| = ? \; ; \; x = x_\perp + \gamma \frac{w}{\|w\|}$$

$$w^T x + b = w^T x_\perp + b + \gamma \frac{w^T w}{\|w\|} = g(x) = \cancel{g(x_\perp)}^{0} + \gamma \frac{\|w\|^{\cancel{2}}}{\cancel{\|w\|}}$$

$$\boxed{\gamma = \frac{g(x)}{\|w\|}} \rightarrow \text{Distancia de } x \text{ a } g(x)$$

$\|x\| \to s(x)=0$?

$\min_x \|x\|; \ s.t. \ s(x)=0 \Rightarrow L(x,\lambda) = x^Tx - \lambda s(x)$

$= x^Tx - \lambda(w^Tx+b)$

$\frac{\partial L}{\partial x} = 2x - \lambda w = 0;$

$\frac{\partial L}{\partial \lambda} = w^Tx+b=0$

$ww^Tx = -wb$

$\boxed{x_{min} = -(ww^T)^{-1}wb}$

$\frac{\lambda}{2}w = x = -(ww^T)^{-1}wb$

$\frac{\lambda}{2}ww^T = -b\cancel{(ww^T)^{-1}}w\cancel{w^T}$

$\boxed{\frac{\lambda}{2} = -b(ww^T)^{-1}}$

$-b/\|w\|$

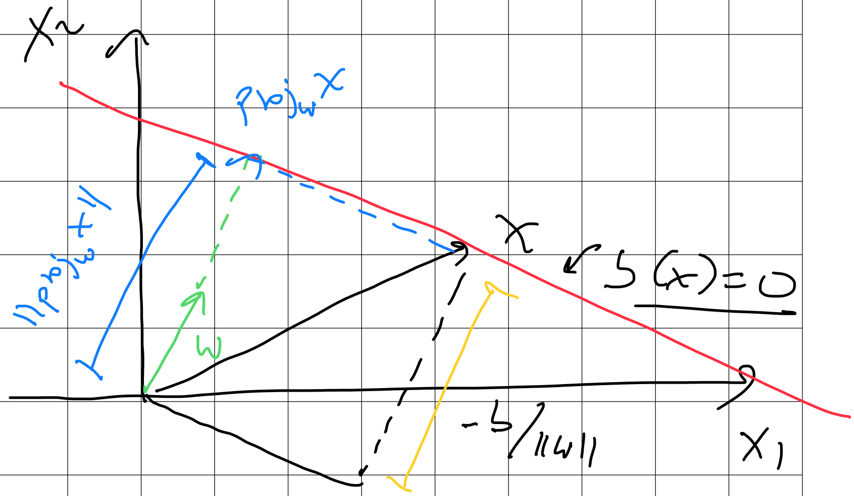
$\mathrm{proj}_w x = rw; \quad r = \frac{w^Tx}{\|w\|^2}$

$\|\mathrm{proj}_w x\| = \|rw\| = (r^2w^Tw)^{1/2} = r(\|w\|^2)^{1/2} = r\|w\|$

$= \frac{w^Tx}{\|w\|^{\cancel{2}}}\cancel{\|w\|} = \frac{w^Tx}{\|w\|}$

$s(x) = w^Tx+b=0 \quad w^Tx = -b.$

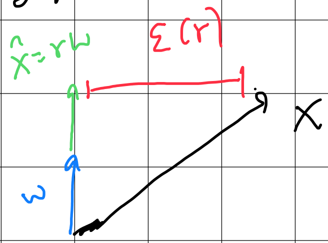
$$\boxed{\|\mathrm{proj}_w x\| = \frac{-b}{\|w\|}}$$



$\hat{x} = r w$ $\qquad \varepsilon(r) = \| x - \hat{x} \| = \| x - r w \|$

$\varepsilon(r) = \langle x - r w, x - r w \rangle = x^T x - 2 r w^T x + r^2 w^T w.$

$$\frac{\partial}{\partial r} \varepsilon(r) = -2 w^T x + 2 r w^T w = 0 \Rightarrow \boxed{r = \frac{w^T x}{\|w\|^2}}$$



Repaso proyección ortogonal

Función discriminante de Fisher

$$m_1 = \frac{1}{N_1}\sum_{n\in C_1} x_n \;;\quad m_2 = \frac{1}{N_2}\sum_{n\in C_2} x_n \;;$$

$$m_{1w} = w^T m_1$$

$$m_{2w} = w^T m_2$$

$$\boxed{m_{1w} - m_{2w} = w^T(m_1 - m_2)}$$

$$S_1^2 = \sum_{n\in C_1} (w^T x_n - m_{1w})^2 .$$

$$J(w) = \frac{(m_{2w} - m_{1w})^2}{S_1^2 + S_2^2}$$

$$(m_{2w} - m_{1w})^2 = (w^T m_2 - w^T m_1)^2 = (w^T(m_2 - m_1))^2$$

$$= w^T (m_2 - m_1)(m_2 - m_1)^T w$$

$$= w^T S_B w$$

$$S_k^2 = \sum_{n\in C_k} (w^T x_n - m_{kw})^2 = \sum (w^T x_n - w^T m_k)^2 = \sum (w^T (x_n - m_k))^2$$

$$= \sum w^T (x_n - m_k)(x_n - m_k)^T w = w^T \sum (x_n - m_k)(x_n - m_k)^T w$$

$$= w^T S_k w$$

$$\boxed{J(w) = \frac{w^T S_B w}{w^T S_w w}} \;;\quad S_w = S_1 + S_2 .$$

$$w^* = \underset{w}{\operatorname{argmax}}\, J(w).$$

$$\frac{\partial J(w)}{\partial w} = \frac{\frac{\partial}{\partial w}\{w^T S_B w\}\, w^T S_w w - \frac{\partial}{\partial w}(w^T S_w w)\, w^T S_B w}{(w^T S_w w)^2} = 0$$

$$\frac{\partial J(w)}{\partial w} = 2 S_B w \underbrace{(w^T S_W w)}_{\text{escalar}} - 2 S_W w \underbrace{(w^T S_B w)}_{\text{escalar}} = 0$$

$$\frac{w^T S_W w \, S_B w}{w^T S_W w} - \frac{w^T S_B w \, S_W w}{w^T S_W w} = 0$$

$$S_B w - \underbrace{\frac{w^T S_B w}{w^T S_W w}}_{\lambda \text{ escalar}} S_W w = 0$$

$$\boxed{S_B w = \lambda S_W w} \rightarrow \text{eig generalizado.}$$

Si S_W^{-1} existe: $\boxed{S_W^{-1} S_B w = \lambda w}$ eig estándar

$$S_B w = (m_1 - m_2)\underbrace{(m_1 - m_2)^T w}_{\alpha \text{ escalar}} = \alpha (m_1 - m_2)$$

$$S_W^{-1} \alpha (m_1 - m_2) = \lambda w \Rightarrow \boxed{w \propto S_W^{-1} (m_1 - m_2)}$$

Análisis para múltiples clases:

$$\mu_k = \frac{1}{N_k} \sum_{n \in C_k} x_n \qquad \mu = \frac{1}{N} \sum_n x_n \ ; \quad s_n = w^T x_n$$

$$J(w) = \det(w^T S_B w) / \det(w^T S_W w) \ ; \quad S_W = \sum S_k$$

$$S_B = \sum_{k=1}^{K} N_k (\mu_k - \mu)(\mu_k - \mu)^T \qquad S_W = \sum_{k=1}^{K} \sum_{n \in C_k} (x_n - \mu_k)(x_n - \mu_k)^T$$

$$S_B w = \lambda S_W w \rightarrow \text{eig generalizado} \rightarrow K-1 \text{ eig distintos}$$