L24: Linear Classifiers

Classification, Loss functions

Apr 8, 2026
FoDA
Jeff M. Phillips

Classification

Input $(X, y) \subset \mathbb{R}^d \times \{-1, +1\}$

(x_i, y_i)

$x_i \in \mathbb{R}^d$
stats

label
$y \in \{-1, +1\}$
win / lose
bug / not bug

Goal on new data $x \in \mathbb{R}^d$

predict $h(x) \to \{-1, +1\}$

$$h(x) = \text{sign}(g(x)) \qquad g(x) \in \mathbb{R}$$

-1 if $g(x) \leq 0$

$+1$ if $g(x) > 0$

Supervised Problem (data X, label y)

↳ cross-validation

Linear Classifier

$(x, y) \in \mathbb{R}^d \times \{-1, +1\}$

$g : \mathbb{R}^d \to \mathbb{R}$

$$g(x) \Longrightarrow \begin{cases} -1 & \text{if } g(x) < 0 \\ +1 & \text{if } g(x) > 0 \end{cases}$$

$\alpha = (\alpha_0, \alpha_1, \ldots, \alpha_d)$ parameter

$x = (x_1, x_2, \ldots, x_d)$

$$g_\alpha(x) = \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_d x_d + \alpha_0$$

$$= \langle \underbrace{(\alpha_1, \alpha_2, \ldots, \alpha_d)}_{w \in \mathbb{R}^d}, x \rangle + \underbrace{\alpha_0}_{b \in \mathbb{R}}$$

$$g_{\alpha = (w, b)}(x) = \langle w, x \rangle + b$$

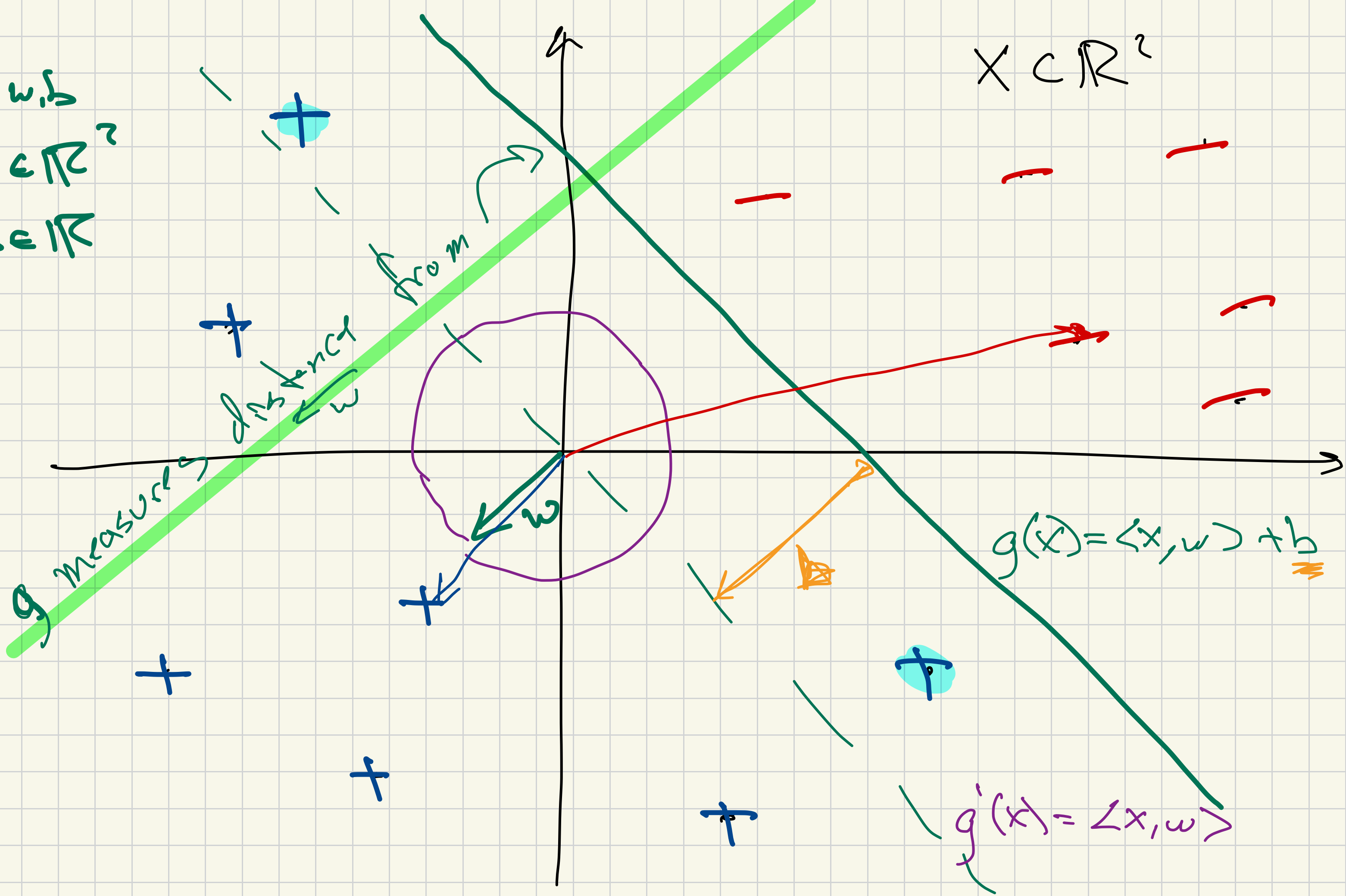
$\|w\| = 1 \qquad b \in \mathbb{R}$

$$= \langle \alpha, (1; x) \rangle \qquad \alpha, (1; x) \in \mathbb{R}^{d+1}$$

$g_{w,b}$

$w \in \mathbb{R}^2$

$b \in \mathbb{R}$



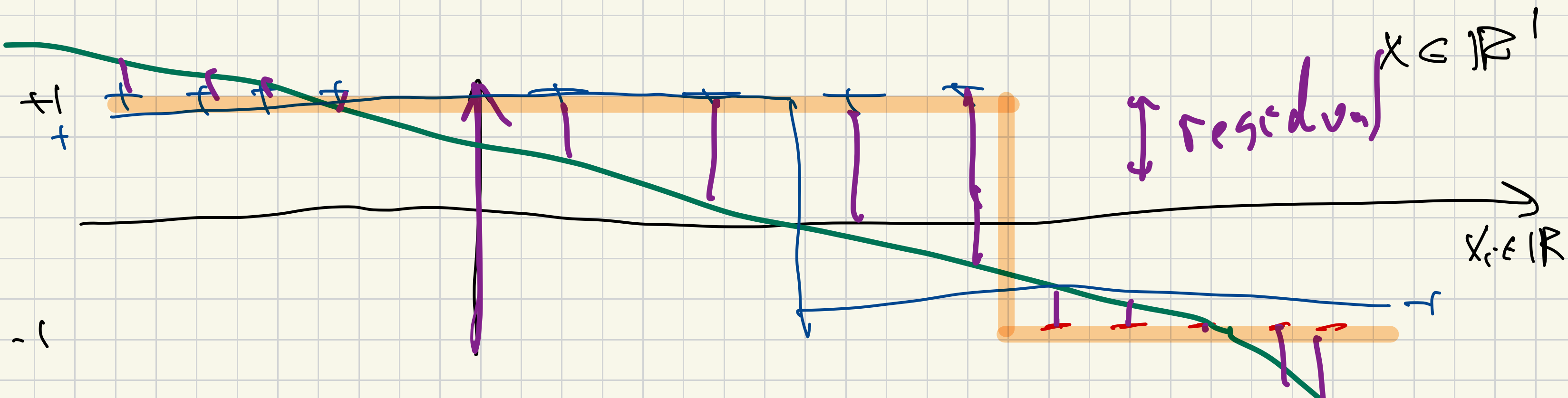
How to formulate "best" $g_{w,b}$

? sum of squared errors ?

$$f_{X,y}(\alpha) = \sum_{x_i \in X} \left(g_\alpha(x_i) - y_i \right)^2$$

not right option

$$h(x) = \mathrm{sign}(g(x))$$



Goal: Minimize # misclassified points.

$$\Delta(g_\alpha, (x, y)) = \sum_{i=1}^{n} \left(1 - \mathbb{1}\left(\text{sign}(y_i) = \text{sign}(g(x_i))\right)\right)$$

$\mathbb{1}$ indicator

$\mathbb{1}$: True $\Rightarrow$ 1

False $\Rightarrow$ 0

Run GD on $\Delta(\alpha)$

- compute gradient $\nabla \Delta(\alpha)$

  $\hookrightarrow$ always 0, undefined

- convex? No

Loss Function : convex approximation of Δ

$$f(\alpha) = \mathcal{L}\left(g_\alpha, (x, y)\right) = \sum_{i=1}^{n} \ell\left(\underbrace{g_\alpha, (x_i, y_i)}\right)$$

$$z_i = y_i\, g_\alpha(x_i) \in \mathbb{R} \qquad = \sum_{i=1}^{n} \ell_\alpha(z_i)$$

if $z_i > 0 \Rightarrow y_i = \mathrm{sign}(g_\alpha(x_i))$ $\qquad = \sum_{i=1}^{n} f_i(\alpha)$

$z_i < 0 \Rightarrow y_i \neq \mathrm{sign}(g_\alpha(x_i))$

Loss functions $\ell(z)$

$z = y \cdot g_a(x)$

hinge $\ell(z) = \max(0, 1-z)$

squared-hinge $\ell(z) = (\max(0, 1-z))^2$

ReLu $\ell(z) = \max(0, -z)$

logistic loss

$$\ell(z) = \ln(1 + \exp(-z))$$