Vapnik-Chervonenkis Classes (cont.)

• Why? $\mathcal{F}$: class of fcns $f : \mathcal{Z} \to \{0,1\}$

$$\mathcal{F}(z^n) = \{ (f(z_1), f(z_2), \ldots, f(z_n)) : f \in \mathcal{F} \} \subseteq \{0,1\}^n$$

How large is $\mathcal{F}(z^n)$ in the worst case?

$|\mathcal{F}(z^n)| \le 2^n$

$$R_n(\mathcal{F}(z^n)) \le 2\sqrt{\frac{\log |\mathcal{F}(z^n)|}{n}}$$

• $f : \mathcal{Z} \to \{0,1\}$ or $\{-1,+1\}$ $\longleftrightarrow$ $C_f := \{ z \in \mathcal{Z} : f(z) = 1 \}$

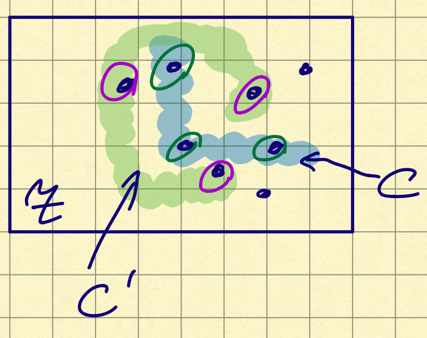
any $\mathcal{F}$ consisting of $\{0,1\}$ (or $\{-1,1\}$) valued fcns $\longleftrightarrow$ a family of subsets of $\mathcal{Z}$

$\mathcal{C}$: a class of subsets $C \subseteq \mathcal{Z}$

$$V(\mathcal{C}) := \sup \{ N \in \mathbb{N} : \text{a set of } N \text{ pts in } \mathcal{Z} \text{ can be shattered by } \mathcal{C} \}$$

$S = \{z_1, \ldots, z_n\} \subseteq \mathcal{Z}$ shattered by $\mathcal{C}$ if

$\forall S' \subseteq S \quad \exists C \in \mathcal{C} \quad$ s.t. $\quad S' = S \cap C$



$S' \subseteq S \longleftrightarrow b = (b_1, \ldots, b_n) \in \{0,1\}^n$

s.t. $b_i = 1_{\{z_i \in S'\}}$

VC dimension and linear independence (Dudley classes)

$g : \mathcal{Z} \to \mathbb{R} \longrightarrow \mathrm{pos}(g) := \{ z \in \mathcal{Z} : g(z) \geq 0 \}$

$\mathcal{G}$: a class of fcns $g : \mathcal{Z} \to \mathbb{R}$

$\longrightarrow \mathrm{pos}(\mathcal{G}) := \{ \mathrm{pos}(g) : g \in \mathcal{G} \}$

<u>Ex. 1:</u> $\mathcal{Z} = \mathbb{R}^d$

$\mathcal{G} := \{ z \mapsto \langle w, z \rangle + b : w \in \mathbb{R}^d \setminus \{0\}, b \in \mathbb{R} \}$

– nonconstant affine fcns on $\mathbb{R}^d$

$$\mathrm{pos}(\mathcal{G}) = \left\{ \{ z \in \mathbb{R}^d : \langle w, z \rangle + b \geq 0 \} : \begin{array}{l} w \in \mathbb{R}^d \setminus \{0\} \\ b \in \mathbb{R} \end{array} \right\}$$

– indicators of half-spaces

<u>Ex. 2:</u> $\mathcal{Z} = \mathbb{R}^d$

$\mathcal{G} := \{ z \mapsto r^2 - \|z - x\|^2 : x \in \mathbb{R}^d, r \in \mathbb{R} \}$

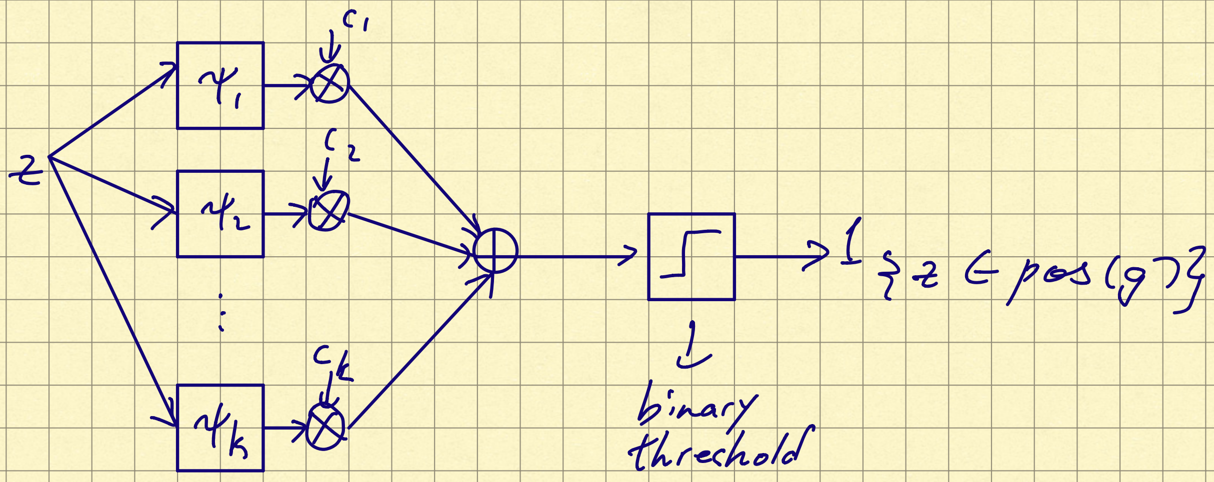
$\{ z \in \mathbb{R}^d : r^2 - \|z - x\|^2 \geq 0 \}$ = ball of rad. r centered at $x \in \mathbb{R}^d$

$\mathrm{pos}(\mathcal{G})$ – indicators of closed balls in $\mathbb{R}^d$

<u>Ex. 3:</u> $\mathcal{Z} = \mathbb{R}^d$

$g(z) = \sum_{i=1}^{k} c_i \psi_i(z)$ $\qquad c_i \in \mathbb{R}$, $\psi_i : \mathbb{R}^d \to \mathbb{R}$

$\mathrm{pos}(g) := \{ z \in \mathbb{R}^d : \sum_{i=1}^{k} c_i \psi_i(z) \geq 0 \}$



Dudley classes:

- $\mathcal{G}$ — an m-dim. linear space of fcns $g : \mathcal{Z} \to \mathbb{R}$
  $\exists\, \psi_1, \dots, \psi_m : \mathcal{Z} \to \mathbb{R}$ s.t. any $g \in \mathcal{G}$ can be uniquely represented as a lin. comb. of them:
  $$g(z) = \sum_{i=1}^{m} c_i \psi_i(z)$$
  for some $c_1, \dots, c_m \in \mathbb{R}$ (uniquely determined)
  $$g, g' \in \mathcal{G} \Rightarrow \alpha g + \beta g' \in \mathcal{G} \quad \forall \alpha, \beta \in \mathbb{R}$$
- $\forall h : \mathcal{Z} \to \mathbb{R}$ (not necessarily in $\mathcal{G}$):
  $$pos(g+h) := \{z \in \mathcal{Z} : g(z) + h(z) \geq 0\}$$
  $$pos(\mathcal{G}+h) := \{pos(g+h) : g \in \mathcal{G}\}$$
  fixed but arbitrary

Thm (Dudley) If $\mathcal{G}$ is a linear space of fcns $\mathcal{Z} \to \mathbb{R}$ of finite dim m, then
$$V(pos(\mathcal{G}+h)) = m \quad \forall h : \mathcal{Z} \to \mathbb{R}.$$

Proof (sketch)

① $\exists$ a set of m pts shattered by pos$(\mathcal{G}+h)$

$S = \{z_1, \dots, z_m\}$ is shattered by pos$(\mathcal{G}+h)$ if

$\forall\, b \in \{-1,+1\}^m \quad \exists\, g^b \in \mathcal{G}$ s.t.

$g^b(z_i) + h(z_i) \geq 0$ iff $b_i = 1$

$\downarrow$

find $c_1^b, \dots, c_m^b \in \mathbb{R}$ s.t. $g^b(z) = \sum_{j=1}^{m} c_j^b \psi_j(z) \; \forall z \in \mathcal{Z}$

$i=1$: $\quad g^b(z_1) + h(z_1) = \sum_{j=1}^{m} c_j^b \psi_j(z_1) + h(z_1)$

$i=2$: $\quad g^b(z_2) + h(z_2) = \sum_{j=1}^{m} c_j^b \psi_j(z_2) + h(z_2) \quad \dots$

$\vdots$

$i=m$: $\quad g^b(z_m) + h(z_m) = \sum_{j=1}^{m} c_j^b \psi_j(z_m) + h(z_m)$

$$M := \begin{pmatrix} \psi_1(z_1) & \psi_2(z_1) & \dots & \psi_m(z_1) \\ \psi_1(z_2) & \psi_2(z_2) & \dots & \psi_m(z_2) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_1(z_m) & \psi_2(z_m) & \dots & \psi_m(z_m) \end{pmatrix}$$

$$c^b := \begin{pmatrix} c_1^b \\ c_2^b \\ \vdots \\ c_m^b \end{pmatrix} \qquad \xi := \begin{pmatrix} h(z_1) \\ h(z_2) \\ \vdots \\ h(z_m) \end{pmatrix}$$

$\Rightarrow \quad g^b(z_i) + h(z_i) = (Mc^b)_i + \xi_i$

Given $b \in \{-1,+1\}^m \iff b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$

find $c^b \in \mathbb{R}^m$ s.t. $Mc^b + \xi = b$

$Mc^b = b - \xi$

$\Rightarrow \mathcal{S} = \{z_1, \dots, z_m\}$ will be shattered if $M = (M_{ij})$ with $M_{ij} = \psi_j(z_i)$ is invertible ($\det M \neq 0$).

Since $\psi_1, \dots, \psi_m$ are linearly indep., $\mathcal{S}$ can be constructed greedily:

given $z_1, \dots, z_i$ ($i < m$), can choose z_{i+1} s.t. the rows

$(\psi_1(z_1), \psi_2(z_1), \dots, \psi_m(z_1))$

$\vdots$

$(\psi_1(z_{i+1}), \psi_2(z_{i+1}), \dots, \psi_m(z_{i+1}))$

are linearly independent.

(<u>Hint:</u> use the fact that

$$\sum_{j=1}^{m} c_j \psi_j(z) = 0 \quad \forall z \in \mathcal{Z} \iff c_1 = \dots = c_m = 0.)$$

2 no set of $m+1$ pts in $\mathcal{Z}$ can be shattered

Suppose otherwise: $\mathcal{S} = \{z_1, \dots, z_{m+1}\}$ is shattered.

$$\mathcal{G}\big|_{\mathcal{S}} := \{(g(z_1), \dots, g(z_{m+1})) : g \in \mathcal{G}\} \subset \mathbb{R}^{m+1}$$

— linear subspace of dim. $\leq m$

$(g(z_1), \dots, g(z_{m+1})) \in \mathcal{G}|_{\mathcal{S}}$

$(g'(z_1), \dots, g'(z_{m+1})) \in \mathcal{G}|_{\mathcal{S}}$

$\Rightarrow (g(z_1) + g'(z_1), \dots, g(z_{m+1}) + g'(z_{m+1})) \in \mathcal{G}|_{\mathcal{S}}$

$\Rightarrow \exists\ v = (v_1, \dots, v_{m+1})$ (not all v_j's zero) orthogonal to $\mathcal{G}|_{\mathcal{S}}$:

$$v_1 g(z_1) + v_2 g(z_2) + \dots + v_{m+1} g(z_{m+1}) = 0 \quad \forall g \in \mathcal{G}$$

<u>Recall:</u> $\mathcal{S} = \{z_1, \dots, z_{m+1}\}$ is shattered by $\mathrm{pos}(\mathcal{G}+h)$

- $v_1 h(z_1) + v_2 h(z_2) + \dots + v_{m+1} h(z_{m+1}) = 0$

but at least one $v_i \neq 0 \Rightarrow$ w.l.o.g. $v_i < 0$

Let $b = (b_1, \dots, b_{m+1}) \in \{0,1\}^m$ s.t. $b_j = 1_{\{v_j \geq 0\}}$

By assumption, $\exists\ g^b \in \mathcal{G}$ s.t.

$$1_{\{g^b(z_1) + h(z_1)\}} = b_1, \dots, 1_{\{g^b(z_{m+1}) + h(z_{m+1})\}} = b_{m+1}$$

But $\displaystyle\sum_{j=1}^{m+1} v_j \left(g^b(z_j) + h(z_j)\right) = 0$

(b.c. $v \perp \mathcal{G}|_{\mathcal{S}}$ and $v_1 h(z_1) + \dots + v_{m+1} h(z_{m+1}) = 0$)

$$v_j \underbrace{\left(g^b(z_j) + h(z_j)\right)}_{\geq 0 \text{ iff } v_j \geq 0} \geq 0 \quad \forall j$$

$v_i < 0$ ($\exists$ at least one): $\underbrace{v_i}_{<0} \underbrace{\left(g^b(z_i) + h(z_i)\right)}_{<0} > 0$

— contradiction!

- $v_1 h(z_1) + \dots + v_{m+1} h(z_{m+1}) \neq 0$, w.l.o.g. < 0, at least one $v_i < 0$

$b_j = 1_{\{v_j \geq 0\}}$

By shattering, $\exists\ g^b \in \mathcal{G}$ s.t. $b_j = 1_{\{g^b(z_j) + h(z_j) \geq 0\}}$

$$\sum_{j=1}^{m+1} v_j \underbrace{\left(g^b(z_j) + h(z_j)\right)}_{\geqslant 0 \text{ iff } v_j \geqslant 0} = \sum_{j=1}^{m+1} v_j h(z_j) < 0$$

$$v_j \left(g^b(z_j) + h(z_j)\right) \geqslant 0 \qquad [v_i < 0 \iff g^b(z_i) + h(z_i) < 0]$$

— contradiction. $\blacksquare$