VC dimension

František Hakl

ICS AS CR
hakl@cs.cas.cz

Mar 2013

Basic definitions

Let $\bar{X}$ be an universal set, $\bar{S} \subset \bar{X}$, $C \subset 2^{\bar{X}}$:

$$\Pi_C\left(\bar{S}\right) \stackrel{\text{def}}{=} \left\{\bar{S} \cap \bar{c} \,|\, \bar{c} \in C\right\}$$

$$\Pi_C\left(m\right) \stackrel{\text{def}}{=} \max\left\{\left|\Pi_C\left(\bar{S}\right)\right| \,\middle|\, \bar{S} \subset \bar{X}\,,\; \left|\bar{S}\right| = m\right\}$$

$$\mathrm{VC}_{dim}\left(C\right) \stackrel{\text{def}}{=} \sup\left\{m \,|\, \Pi_C\left(m\right) = 2^m\right\}$$

$\mathrm{VC}_{dim}\left(C\right)$ is maximal size of a set $\bar{S}$ shattered by system C

Examples:

VC_{dim} (intervals in $\Re^n$) $= 2n$.
VC_{dim} (union of n intervals in $\Re$) $= 2n$.
VC_{dim} (convex sets in $\Re^n$) $= +\infty$.

Let C be nonempty concept class over $\bar{X}$. Then

1. VC_{dim} (C) $= 0$ if and only if C contains exactly one set.
2. Let one of the following conditions is true:
   1. C is linearly ordered by inclusion, or
   2. any two sets in C are disjoint.

   Then VC_{dim} (C) $= 1$.

Example: $\mathrm{VC}_{dim}(\mathrm{C}) = +\infty$

Let $\mathrm{C} \stackrel{\text{def}}{=} \left\{ \bar{A}_\alpha \,\middle|\, (\exists \alpha \in \Re^n) \left(\bar{A}_\alpha = \left\{ x \in \Re \,\middle|\, \widetilde{\sin}(\alpha x) \geq 0 \right\} \right) \right\}$.
Then $\mathrm{VC}_{dim}(\mathrm{C}) = +\infty$.

Proof:
$z_i \stackrel{\text{def}}{=} \frac{1}{10^i}$, $\delta_1, \ldots, \delta_l$, $\delta_i \in \{0, 1\}$.
$\alpha \stackrel{\text{def}}{=} \pi \left(\sum_{i=1}^{l} (1 - \delta_i) 10^i + 1 \right)$.

$$\alpha z_j = \alpha \frac{1}{10^j} = \pi \left(\sum_{i=1}^{j-1} \frac{1 - \delta_i}{10^{j-i}} + \frac{1}{10^j} + (1 - \delta_j) + \sum_{i=j+1}^{l} (1 - \delta_i) 10^{i-j} \right) .$$

Sauer's lemma: (Norbert Sauer - 1972)

Let $\bar{X}$ be a finite set and $C \subset 2^{\bar{X}}$. Then

$$|C| \leq \sum_{i=0}^{\mathrm{VC}_{dim}(C)} \binom{|\bar{X}|}{i}.$$

Further, there exists $C \subset 2^{\bar{X}}$ such that equality holds.

Corollary: ($\mathrm{VC}_{dim}(C) = d$, $\Phi_{d,m} \stackrel{\text{def}}{=} \sum_{i=0}^{d} \binom{m}{i}$)

1. $\Pi_C(m) \leq \Phi_{d,m}$ for all $d, m \geq 0$.
2. $\Phi_{d,m} \leq m^d + 1$ for $d, m \geq 0$ and $\Phi_{d,m} \leq m^d$ for $d \geq 0$ and $m \geq 2$.
3. $\Phi_{d,m} \leq 2\frac{m^d}{d!} \leq \left(\frac{em}{d}\right)^d$ for $m \geq d \geq 1$.

Proof of Sauer's lemma:

Define:

$$\begin{aligned}
\mathrm{C}(y) &\stackrel{\text{def}}{=} \left\{\bar{A}\dot{-}\{y\} \,\middle|\, \bar{A} \in \mathrm{C}\right\}\\
\mathrm{C}_y &\stackrel{\text{def}}{=} \left\{\bar{A} \in \mathrm{C} \,\middle|\, (\exists \bar{B} \in \mathrm{C})(\bar{A} \neq \bar{B} \text{ and } \bar{B} = \bar{A} \cup \{y\})\right\}\\
\mathrm{C}^y &\stackrel{\text{def}}{=} \left\{\bar{A} \in \mathrm{C} \,\middle|\, (\exists \bar{B} \in \mathrm{C}_y)(\bar{A} = \bar{B} \cup \{y\})\right\}.
\end{aligned}$$

Then:

$|\mathrm{C}| - |\mathrm{C}(y)| = |\mathrm{C}_y|$

$\mathrm{VC}_{dim}(\mathrm{C}_y) = n - 1 \Rightarrow \mathrm{VC}_{dim}(\mathrm{C}) \geq n$

$\left|\bar{X}\right| = k \Rightarrow \left|\bar{X}\right| = k + 1$

Radon's lemma: (Johann Radon - 1921), 1887-56

Let $\bar{S} \stackrel{\text{def}}{=} \{\vec{\boldsymbol{x}}_1, \ldots, \vec{\boldsymbol{x}}_k\} \subset \Re^n$, $k \geq n+2$, $\vec{\boldsymbol{x}}_i$ are mutually different. Then there exists sets $\bar{S}_1$ and $\bar{S}_2$ such that $\bar{S}_1 \cup \bar{S}_2 = \bar{S}$, $\bar{S}_1 \cap \bar{S}_2 = \emptyset$ and

$$[\bar{S}_1]_\kappa \cap [\bar{S}_2]_\kappa \neq \emptyset,$$

(symbol $[\bar{S}]_\kappa$ denotes a convex hull of the set $\bar{S}$).

Corollary:
$\mathrm{VC}_{dim}(\textbf{HALFSPACE}_n) = \mathrm{VC}_{dim}(\textbf{BALL}_n) = n + 1$

Cover's lemma (Thomas M. Cover - 1964), 1938-12

Let $\bar{X} \stackrel{\text{def}}{=} \{\vec{\boldsymbol{x}}_1, \ldots, \vec{\boldsymbol{x}}_d\} \subset \Re^N$ are linearly independent vectors. Than there exists

$$2 \sum_{k=0}^{d-1} \binom{N-1}{k}$$

mutually different disjoint splittings of the set $\bar{X}$ into sets $\bar{A}$ and $\bar{B}$ whose are homogeneously linearly separable (i.e. they can be separated via hyperplane which contains zero vector).

Corollary:
$\mathrm{VC}_{dim}\left(\mathbf{HALFSPACE}_{n,d}\right) = n - d$

Intersection and union

$$\mathrm{U}_{k,\mathrm{C}} \stackrel{\text{def}}{=} \left\{ \bigcup_{i=1}^{k} \bar{c}_i \,|\, (\forall i \in \{1,\ldots,k\})\,(\bar{c}_i \in \mathrm{C}) \right\}$$

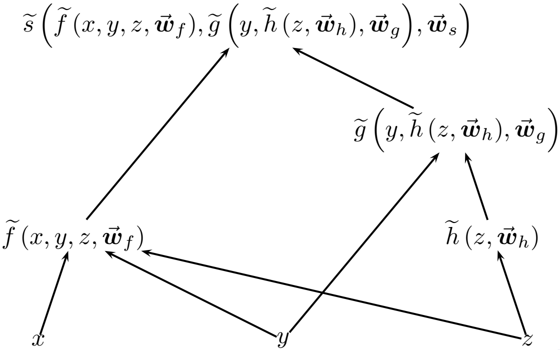
$$\mathrm{I}_{k,\mathrm{C}} \stackrel{\text{def}}{=} \left\{ \bigcap_{i=1}^{k} \bar{c}_i \,|\, (\forall i \in \{1,\ldots,k\})\,(\bar{c}_i \in \mathrm{C}) \right\}$$

Then

1. $(\forall \bar{a} \in \mathrm{C}) \left(\bar{X} \dot{-} \bar{a} \in \mathrm{C} \right) \Rightarrow \mathrm{VC}_{dim}\left(\mathrm{U}_{k,\mathrm{C}}\right) = \mathrm{VC}_{dim}\left(\mathrm{I}_{k,\mathrm{C}}\right)$
2. let $\mathrm{VC}_{dim}(\mathrm{C}) = d \geq 1$ be finite. Then $\mathrm{VC}_{dim}\left(\mathrm{U}_{k,\mathrm{C}}\right) \leq 2dk\log_2(3k)$ and $\mathrm{VC}_{dim}\left(\mathrm{I}_{k,\mathrm{C}}\right) \leq 2dk\log_2(3k)$.

Composed mapping - definition:

let $\widetilde{s} : \Re^3 \times \bar{W}_s \times \bar{W}_f \times \bar{W}_g \times \bar{W}_h \to \Re$, where $\bar{W}_s, \bar{W}_f, \bar{W}_g, \bar{W}_h$ are parameter spaces of mappings $\widetilde{s}$, $\widetilde{f}$, $\widetilde{g}$, $\widetilde{h}$, respectively, and $\widetilde{s}\left(\widetilde{f}\left(x, y, z, \vec{\boldsymbol{w}}_f\right), \widetilde{g}\left(y, \widetilde{h}\left(z, \vec{\boldsymbol{w}}_h\right), \vec{\boldsymbol{w}}_g\right), \vec{\boldsymbol{w}}_s\right)$.



Composed mapping - VC-dim:

$$\mathrm{C}_j^{loc} \stackrel{\text{def}}{=} \left\{ \bar{c} \,\middle|\, (\exists \vec{\boldsymbol{w}} \in W_j) \left(\bar{c} = \left\{ \vec{\boldsymbol{s}} \in \Re^{d_j^+} \,\middle|\, \widetilde{Z}_j\left(\vec{\boldsymbol{w}}, \vec{\boldsymbol{s}}\right) \le 0 \right\} \right) \right\}.$$

$$\mathrm{C}_j^{par} \stackrel{\text{def}}{=} \left\{ \bar{c} \,\middle|\, (\exists \omega \in \bar{W}_1 \times \cdots \times \bar{W}_j) \left(\bar{c} = \left\{ \vec{\boldsymbol{x}} \in \Re^{n} \,\middle|\, v_{j,\omega,\vec{\boldsymbol{x}}} \le 0 \right\} \right) \right\}.$$

Then

$$\Pi_{\mathrm{C}_\mathrm{k}^{\mathrm{par}}}(m) \le \Pi_{\mathrm{C}_1^{\mathrm{loc}}}(m) \cdot \Pi_{\mathrm{C}_2^{\mathrm{loc}}}(m) \cdot \cdots \cdot \Pi_{\mathrm{C}_\mathrm{k}^{\mathrm{loc}}}(m)$$

Linear composed mapping:

Let L be a composed linear mapping, w is the number of edges, z is the number of noninput vertices, and $q \stackrel{\text{def}}{=} w + z$. Then, for any $m > \max\left\{d_i^+ \,|\, i \in \{1, \ldots, q\}\right\}$ is

$$\Pi_{\mathrm{C_q^{par}}}(m) \leq \left(\frac{ezm}{q}\right)^q \qquad (1)$$

and further

$$\mathrm{VC}_{dim}\left(\mathrm{C}_q^{par}\right) < 2q\log_2(ez). \qquad (2)$$

Proof:

$\Pi_{\mathrm{C_i}}(m) \leq \left(\frac{em}{d_i^+ + 1}\right)^{d_i^+ + 1}$.

$\sum_{i=1}^{z} \alpha_i = 1 \Rightarrow -\sum_{i=1}^{z} \alpha_i \ln(\alpha_i) \leq \ln(z)$

Concept	params	VC-dim	method
$HS_{n,d}$	n-d	n -d +1	Perceptron
HS_n	n+1	n+1	Perceptron
$Ball_n$	n+1	n+1	NN
Int_n	2n	2n	cuts, DT
U_{k,HS_n}	$k(n+1)$	$\leq 2(n+1)k\log_2(3k)$	MLP(1hid.l.)
I_{k,HS_n}	dtto	dtto	dtto
$\mathrm{U}_{m,\mathrm{I}_{k,HS_n}}$	$mk(n+1)$	$\leq 4(n+1)km\cdot$ $\cdot\log_2(3k)\log_2(3m)+\cdots$	MLP ???
CLinM	$w+z$	$\leq 2(w+z)\log_2(e(w+z))$	MLP(2hid.l.)
CBallM	dtto	dtto	NN ???

??? VC-dim $\leq f$(params) ???