Chapter10. Empirical vector quantization

July 5, 2021

Motivation

Given **continuous-valued sources** $Z \in \mathbb{R}^d$ (such as images), we want to map any source realization to a compact **binary string** b that would, upon **reconstruction** $\widehat{Z}$, differ from the original source as little as possible.

$z_1 \in \mathbb{R}^d \quad z_2 \in \mathbb{R}^d \quad \cdots \quad z_n \in \mathbb{R}^d$

The overall mapping $Z \mapsto b \mapsto \widehat{Z}$ is called vector quantizer.

1. **compactness** measured by expected rate: $\mathbf{E}\left[\mathrm{len}\left(b\right)\right]$
2. **accuracy** measured by expected distortion: $\mathbf{E}[d(z, \hat{z})]$, where $d(Z, \hat{Z}) : \mathbb{R}^d \times \mathbb{R}^d \to [0, \infty)$

$$d(z, \widehat{z}) = \|z - \widehat{z}\|^2 = \sum_{j=1}^{d} |z(j) - \widehat{z}(j)|^2$$

Definition 10.1.

Let Z $= \mathbb{R}^d, k \in \mathbb{N}$. A (d-dimensional) k-point vector quantizer is a (measurable) mapping $q :$ Z $\to \mathcal{C} = \{y_1, \ldots, y_k\} \subset$ Z, where the set $\mathcal{C}$ is called the codebook and its elements are called the codevectors.

Z $\tilde{z}$

- **source**: a random vector $Z \in \mathbb{R}^d$ with some probability distribution P_Z
- k-point quantizer q represents Z by the quantized **output** $\hat{Z} = q(Z)$
- **rate** of q (in bits):

$$R(q) := \lceil \log_2 k \rceil \quad = \log_2 k$$

- **expected distortion**:

$$D(P_Z, q) := \mathbf{E}\|Z - q(Z)\|^2 = \int_{\mathbb{R}^d} \|z - q(z)\|^2 P_Z(dz)$$

Goal: minimize the expected distortion $D(P_Z, q)$ subject to a fixed rate $R(q)$ $\quad$ k is fixed

Definition 10.2.

$|C| = k$

Let $\mathcal{Q}_k$ be the set of all k-point vector quantizers. Then the optimal performance on a given source distribution P_Z is defined by

$$D_k^*(P_Z) := \inf_{q \in \mathcal{Q}_k} D(P_Z, q) \equiv \inf_{q \in \mathcal{Q}_k} \mathbf{E}\|Z - q(Z)\|^2$$

$P_Z \in P(\mathbb{R})$

We say that a quantizer $q^* \in \mathcal{Q}_k$ is optimal for P_Z if

$$D(P_Z, q^*) = D_k^*(P_Z)$$

Definition 10.3.

A quantizer $q \in \mathcal{Q}_k$ with codebook $\mathcal{C} = \{y_1, \ldots, y_k\}$ is called nearestneighbor if, for all $z \in$ Z,

$$\|z - q(z)\|^2 = \min_{1 \leq j \leq k} \|z - y_j\|^2 .$$

$q(z) := \operatorname{argmin}_{y_j \in C} \|z - y_j\|$

Let $\mathcal{Q}_k^{\mathrm{NN}}$ denote the set of all k -point nearest-neighbor quantizers.

Lemma 10.1.

For any $q \in \mathcal{Q}_k$ we can always find some $q' \in \mathcal{Q}_k^{\mathrm{NN}}$, such that $D\left(P_Z, q'\right) \le D\left(P_Z, q\right)$

Proof.

Given a quantizer $q \in \mathcal{Q}_k$ with codebook $\mathcal{C} = \{y_1, \dots, y_k\}$, define q' by

$$q'(z) := \arg\min_{y_j \in \mathcal{C}} \|z - y_j\|^2$$

where ties are broken by going with the lowest index. Then q' is clearly a nearest-neighbor quantizer, and

$$\begin{aligned} D\left(P_Z, q'\right) &= \mathbf{E}\left\|Z - q'(Z)\right\|^2 \\ &= \mathbf{E}\left[\min_{1 \le j \le k} \|Z - y_j\|^2\right] \\ &\le \mathbf{E}\|Z - q(Z)\|^2 \\ &\equiv D\left(P_Z, q\right) \end{aligned}$$

$\square$

According to Lemma10.1, we have

$$D_k^*(P_Z) = \inf_{q \in \mathcal{Q}_k^{\mathrm{NN}}} \mathbf{E}\|Z - q(Z)\|^2 = \inf_{\mathcal{C}=\{y_1,\ldots,y_k\} \subset \mathsf{Z}} \mathbf{E}\left[\min_{1 \le j \le k} \|Z - y_j\|^2\right]$$

min

Theorem 10.1.

If Z has a finite second moment, $\mathbf{E}\|Z\|^2 < \infty$, then there exists a nearest-neighbor quantizer $q^* \in Q_k^{\mathrm{NN}}$ such that $D(P_Z, q^*) = D_k^*(P_Z)$.

10.3. Learning an (approximately) optimal quantizer

Finding an optimal q^* is a very difficult problem

- combinatorial search component: optimize over all k-point sets $\mathcal{C}$ in $\mathbb{R}^d$
- source distribution P_Z is often not known

Empirical method:
learn an (approximately) optimal quantizer for P_Z based on a sufficiently large training samples

- sample: $Z^n = (Z_1, \ldots, Z_n)$ be an i.i.d. from P_Z
- learning algorithm: take Z^n and output a $\hat{q}_n \in \mathcal{Q}_k$

DEFINITION 10.4.

We say that a quantizer $\widehat{q}_n \in Q_k^{\mathrm{NN}}$ is empirically optimal for Z^n if

$$\min_{q \in Q_k^{NN}} \cancel{\inf} D(P_n, q)$$

$$D(P_n, \widehat{q}_n) = D_k^*(P_n) = \min_{q \in \mathcal{Q}_k^{\mathrm{NN}}} D(P_n, q) = \min_{q \in \mathcal{Q}_k^{\mathrm{NN}}} \frac{1}{n} \sum_{i=1}^{n} \|Z_i - q(Z_i)\|^2$$

true risk : $D_k^*(P_Z) = \min_{q \in Q_k^{NN}} D(P_Z, q)$

By the nearest-neighbor property,

$$D_k^*(P_n) = \min_{\mathcal{C}=\{y_1,\dots,y_k\} \subset \mathsf{Z}} \frac{1}{n} \sum_{i=1}^{n} \min_{1 \le j \le k} \|Z_i - y_j\|^2$$

Expected distortion of $\widehat{q}_n$, given by

$$D(P_Z, \widehat{q}_n) = \mathbf{E}\left[\|Z - \widehat{q}_n(Z)\|^2 \mid Z^n\right] = \int_{\mathbf{Z}} \|z - \widehat{q}_n(z)\|^2 P_Z(dz)$$

r.v.

An abstract framework for ERM

$(\mathsf{Z}, \mathcal{P}, \mathcal{F})$

- Set Z $= \mathbb{R}^d$
- A class $\mathcal{P}$ of probability distributions on Z

  $P_Z \in \mathcal{P}(r)$
- A class $\mathcal{F}$ of functions $f : \mathsf{Z} \to \cancel{[0,1]}$ (induced losses)

  $\mathcal{F} = \{ d_q : q \in Q_k^{r,N} \}$

  $d_q(\vec{z}) = \| \vec{z} - q(\vec{z}) \|^2$

- The expected risk of any $f \in \mathcal{F}$:

$$P(f) := \mathbf{E}_P f(Z)$$

- The minimum risk :

$$L_P^*(\mathcal{F}) := \inf_{f \in \mathcal{F}} P(f)$$

- ERM algorithm:

$$\hat{f}_n = \arg\min_{f \in \mathcal{F}} P_n(f) = \arg\min_{f \in \mathcal{F}} \frac{1}{n} \sum_{i=1}^{n} f(Z_i) = D_k^*(P_n)$$

- Uniform deviation

$$\Delta_n(Z^n) := \|P_n - P\|_{\mathcal{F}} \equiv \sup_{f \in \mathcal{F}} |P_n(f) - P(f)|$$

Remarks about empirically optimal quantizers

$\mathbb{P}(| \ | > \delta) \leq \exp\left\{ \frac{-\delta^2}{r^2 n} \right\}$

1. Nearly optimal

$$\mathbf{E}\left[D\left(P_Z, \widehat{q}_n\right) - D_k^*\left(P_Z\right)\right] \leq \frac{C}{\sqrt{n}}$$

2. Strongly consistent

$$D\left(P_Z, \widehat{q}_n\right) - D_k^*\left(P_Z\right) \overset{n \to \infty}{\longrightarrow} 0 \quad \text{almost surely}$$

3. Exact solution is NP-complete, but there are various approximation techniques, e.g. K-means algorithm...

Assumption

For a given $r > 0$ and $z \in \mathbb{R}^d$, let $B_r(z)$ denote the ℓ_2 ball of radius r centered at z :

$$B_r(z) := \left\{ y \in \mathbb{R}^d : \|y - z\| \leq r \right\}$$

Let $\mathcal{P}(r)$ denote the set of all probability distributions P_Z on $\mathrm{Z} = \mathbb{R}^d$, such that

$$P_Z\left(B_r(0)\right) = 1$$

Theorem 10.2 .

There exists some absolute constant $C > 0$, such that

$$\sup_{P_Z \in \mathcal{P}(r)} \mathbf{E}\left[D\left(P_Z, \widehat{q}_n\right) - D_k^*\left(P_Z\right)\right] \leq C r^2 \sqrt{\frac{k(d+1)\log(k(d+1))}{n}}$$

Here $\widehat{q}_n$ denotes an empirically optimal quantizer based on an i.i.d. sample Z^n.

Lemma 10.2.

Let $\mathcal{Q}_k^{\mathrm{NN}}(r)$ denote the set of all nearest-neighbor k -point quantizers whose codewords lie in $B_r(0)$. Then for any $P_Z \in \mathcal{P}(r)$,

$$D\left(P_Z, \widehat{q}_n\right) - D_k^*\left(P_Z\right) \leq 2 \sup_{q \in Q_k^{\mathrm{NN}}(r)} \left|D\left(P_n, q\right) - D\left(P_Z, q\right)\right|$$

6.1 : $\mathbb{E}_{P_Z}(\hat{f}_n) - \mathbb{E}_{P_Z}(f^*) \leq 2 \Delta_n(\mathcal{F})$

$= \sup_{f \in \mathcal{F}} |\mathbb{E}_{P_Z}(f) - \mathbb{E}_{P_n}(f)|$

$[0,1]$

Proof.

Fix P_Z and let $q^* \in \mathcal{Q}_k^{\mathrm{NN}}$ denote an optimal quantizer, i.e., $D(P_Z, q^*) = D_k^*(P_Z)$. Then, we can write

$$
\begin{aligned}
\Rightarrow D(P_Z, \widehat{q}_n) - D_k^*(P_Z) \\
&= D(P_Z, \widehat{q}_n) - D(P_n, \widehat{q}_n) + D(P_n, \widehat{q}_n) - D(P_n, q^*) + D(P_n, q^*) - D(P_Z, q^*) \\
&\leq D(P_Z, \widehat{q}_n) - D(P_n, \widehat{q}_n) + D(P_n, q^*) - D(P_Z, q^*) \\
&\leq 2 \sup_{q \in Q_k^{NN}(r)} \left| D(P_Z, q) - D(P_n, q) \right|
\end{aligned}
$$

Then we only need to show that both $\widehat{q}_n$ and q^* have all their codevectors in $B_r(0)$.

Since $B_r(0)$ is a convex set, for any point $y \notin B_r(0)$ we get $y' = ry/\|y\| \in B_r(0)$. Then

$$\|z - y'\| < \|z - y\|, \quad \forall z \in B_r(0).$$

Thus, if we take an arbitrary quantizer $q \in Q_k$ and replace all of its codevectors outside $B_r(0)$ by their projections, we will obtain another quantizer q', such that $\|z - q'(z)\| \leq \|z - q(z)\|$ for all $z \in B_r(0)$. □

$\theta^\circ > 90^\circ \Leftrightarrow \langle \vec{a}, \vec{b} \rangle < 0$

Proof for Theorem 10.2

Part.1:

Upper bound $\mathbf{E}|D(P_Z, \widehat{q}_n) - D_k^*(P_Z)|$ by $\mathbf{E}$[uniform deviation] (with $0-1$ loss)

Fact.1.1 Let $f_q(z) := \|z - q(z)\|^2$ Then, for any $z \in B_r(0)$ we have

$$0 \leq f_q(z) \leq 2\|z\|^2 + 2\|q(z)\|^2 \leq 4r^2$$

Fact.1.2 Using the integral identity, we have

r.v $X \geq 0$, $\mathbb{E}X = \int_0^\infty \mathbb{P}(X > t)\,dt$

$$D(P_Z, q) = \mathbf{E}_{P_Z}(f_q) = \int_0^{4r^2} \mathbb{P}_Z(f_q(Z) > u)\, du$$

$$D(P_n, q) = \mathbf{E}_{P_n}(f_q) = \int_0^{4r^2} \mathbb{P}_n(f_q(Z) > u)\, du \text{ a.s.}$$

$\int_a^b f(x)dx \le (b-a) \sup_{x\in[a,b]} f(x)$

$$
\begin{aligned}
&\sup_{q\in\mathcal{Q}_k^{\mathrm{NN}}(r)} \left|D\left(P_n, q\right) - D\left(P_Z, q\right)\right| \\
&= \sup_{q\in\mathcal{Q}_k^{\mathrm{NN}}(r)} \left|\int_0^{4r^2} \left(\mathbb{P}_n\left(f_q(Z) > u\right) - \mathbb{P}_Z\left(f_q(Z) > u\right)\right) du\right| \\
&\le 4r^2 \sup_{q\in\mathcal{Q}_k^{\mathrm{NN}}(r)} \sup_{0\le u\le 4r^2} \left|\mathbb{P}_n\left(f_q(Z) > u\right) - \mathbb{P}_Z\left(f_q(Z) > u\right)\right| \\
&= 4r^2 \sup_{q\in\mathcal{Q}_k^{\mathrm{NN}}(r)} \sup_{0\le u\le 4r^2} \left|\mathbf{E}_{P_n}[\mathbf{1}_{f_q(Z)>u}] - \mathbf{E}_{P_Z}[\mathbf{1}_{f_q(Z)>u}]\right|
\end{aligned}
$$

For a given $q \in \mathcal{Q}_k^{\mathrm{NN}}(r)$ and a given $u > 0$ let us define the set

$$A_{u,q} := \left\{z \in \mathbb{R}^d : f_q(z) > u\right\}$$

and let $\mathcal{A}$ denote the class of all such sets: $\mathcal{A} := \left\{A_{u,q} : u > 0, q \in \mathcal{Q}_k^{\mathrm{NN}}(r)\right\}$.

Then we can write $\mathbf{1}_{\{f_q(z)>u\}} = \mathbf{1}_{\{z \in A_{u,q}\}}$

$$\sup_{q \in \mathcal{Q}_k^{\mathrm{NN}}(r)} \left| D\left(P_n, q\right) - D\left(P_Z, q\right) \right| \leq 4r^2 \sup_{A \in \mathcal{A}} \left| \mathbf{E}_{P_n}\left[\mathbf{1}_{z \in A}\right] - \mathbf{E}_{P_Z}\left[\mathbf{1}_{z \in A}\right] \right|$$

Therefore,

$$\begin{aligned} \mathbf{E}_{Z^n}\left[D\left(P_Z, \widehat{q}_n\right) - D_k^*\left(P_Z\right)\right] &\leq 2\mathbf{E}_{Z^n}\left[\sup_{q \in \mathcal{Q}_k^{\mathrm{NN}}(r)} \left| D\left(P_n, q\right) - D\left(P_Z, q\right) \right|\right] \\ &\leq 8r^2 \mathbf{E}_{Z^n}\left[\sup_{A \in \mathcal{A}} \left| \mathbf{E}_{P_n}\left[\mathbf{1}_{z \in A}\right] - \mathbf{E}_{P_Z}\left[\mathbf{1}_{z \in A}\right] \right|\right] \end{aligned}$$

$f(\vec{z}) = \mathbb{1}\{\vec{z} \in A\}$

$\mathcal{F} = \{ \mathbb{1}_{\vec{z} \in A} : A_{q,u} \in \mathcal{A} \}$

Part.2:

Upper bound the uniform deviation using Rademacher average and VC theory

$$\underbrace{\mathbf{E}\left[\sup_{A\in\mathcal{A}} \left|\mathbf{E}_{P_n}[\mathbf{1}_{z\in A}] - \mathbf{E}_{P_Z}[\mathbf{1}_{z\in A}]\right|\right]}_{\leq 2\mathbb{E} R_n(\mathcal{F}(Z^n)) \leq C\sqrt{\frac{V(\mathcal{F})}{n}} = C\sqrt{\frac{V(\mathcal{A})}{n}}} \leq C\sqrt{\frac{V(\mathcal{A})}{n}} \leq 2C\sqrt{\frac{k(d+1)\log(k(d+1))}{n}}$$

THEOREM 6.1. *Fix a space* Z *and let* $\mathcal{F}$ *be a class of functions* $f : \mathsf{Z} \to [0,1]$*. Then for any* $P \in \mathcal{P}(\mathsf{Z})$

$$\mathbf{E}\Delta_n(Z^n) \leq 2\mathbf{E}R_n(\mathcal{F}(Z^n)). \tag{6.22}$$

$$\Delta(Z^n) = \sup_{f\in\mathcal{F}} \left| \mathbb{E}_{P_n}(f) - \mathbb{E}_{P_Z}(f) \right|$$

THEOREM 7.1. *Let* Z *be an arbitrary set and let* $\mathcal{F}$ *be a class of binary-valued functions* $f : \mathsf{Z} \to \{0,1\}$*, or a class of functions* $f : \mathsf{Z} \to \{-1,1\}$*. Let* Z^n *be an i.i.d. sample of size* n *drawn according to an arbitrary probability distribution* $P \in \mathcal{P}(\mathsf{Z})$*. Then, with probability one,* THEOREM 7.2. *There exists an absolute constant $C > 0$, such that under the conditions of the preceding theorem, with probability one,*

$$R_n(\mathcal{F}(Z^n)) \le C\sqrt{\frac{V(\mathcal{F})}{n}}.$$

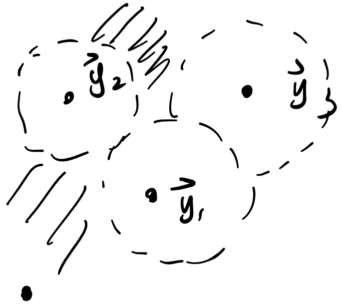
DEFINITION 7.3. *Let $\mathcal{F}$ be a class of functions $f : \mathsf{Z} \to \{0,1\}$, or let $\mathcal{F}$ be a class of functions $f : \mathsf{Z} \to \{-1,1\}$. We say that a finite set $S = \{z_1, \ldots, z_n\} \subset \mathsf{Z}$ is shattered by $\mathcal{F}$ if it is shattered by the class*

$$\mathcal{C}_{\mathcal{F}} := \{C_f : f \in \mathcal{F}\},$$

where $C_f := \{z \in \mathsf{Z} : f(z) = 1\}$. The nth shatter coefficient of $\mathcal{F}$ is $\mathbb{S}_n(\mathcal{F}) = \mathbb{S}_n(\mathcal{C}_{\mathcal{F}})$, and the VC dimension of $\mathcal{F}$ is defined as $V(\mathcal{F}) = V(\mathcal{C}_{\mathcal{F}})$.

$A_{u,q} \in \mathcal{A} \quad : \quad \vec{z} \in A_{u,q} \iff f_q(\vec{z}) = \min_{\vec{y}_j \in C} \|\vec{z} - \vec{y}_j\|^2 > u$

$$= \min_{\vec{y}_j \in C} \|\vec{z} - \vec{y}_j\| > \sqrt{u}$$



$$B_r(\vec{z}) = \{\vec{y} \in \mathbb{R}^d : \|\vec{y} - \vec{z}\| \le r\}$$

$$\|\vec{z} - \vec{y}_i\| > \sqrt{u} \quad = \quad B_{\sqrt{u}}(\vec{y}_i)^c$$

$$A_{u,q} = \underbrace{\left(\bigcap_{j=1}^{k} B_{\sqrt{u}}(\vec{y}_i)^c\right)} \cap B_r(0)$$

$$\mathcal{A} = \{A_{u,q} : q \in Q_k^{NN}(r),\ u \in [0, 4r^2]\}$$

$$\mathcal{A} \subset \tilde{\mathcal{A}} := \left\{\bigcap_{j=1}^{k} B_j^c : B_j \in \mathcal{B},\ \forall j\right\}$$

$$V(\mathcal{A}) \le V(\tilde{\mathcal{A}}) \in$$

Fact.2.1 For any class of sets $\mathcal{M}$, let $\overline{\mathcal{M}}$ denote the class $\{M^c : M \in \mathcal{M}\}$ formed by taking the complements of all sets in $\mathcal{M}$. Then for any n

$$\boxed{\mathbb{S}_n(\overline{\mathcal{M}}) = \mathbb{S}_n(\mathcal{M})}$$

① $S_n(\mathcal{M}) \leq S_n(\overline{\mathcal{M}})$

② $S_n(\overline{\overline{\mathcal{M}}}) \geq S_n(\overline{\mathcal{M}})$

$$S_n(\mathcal{M}) = \sup_{\substack{S \subset Z \\ |S|=n}} \left| \{ S \cap M : M \in \mathcal{M} \} \right|$$

$$S_n(\overline{\mathcal{M}}) = \cdots \left| \{ S \cap M^c : M \in \mathcal{M} \} \right|$$

For $S \subset Z$, $|S| = n$: $\quad S \cap M = S/T \qquad S \cap M_1 \neq S \cap M_2 \Leftrightarrow T_1 \neq T_2$

$S \cap M^c = T \qquad S \cap M_1^c \neq S \cap M_2^c \Leftrightarrow T_1 \neq T_2$

Fact.2.2 For any class of sets $\mathcal{N}$, let $\mathcal{N}_k$ denote the class $\{N_1 \cap N_2 \cap \ldots \cap N_k : N_j \in \mathcal{N}, 1 \leq j \leq k\}$, formed by taking intersections of all possible choices of k sets from $\mathcal{N}$. Then

$$\mathbb{S}_n(\mathcal{N}_k) \leq \mathbb{S}_n^k(\mathcal{N})$$

Lemma: $\mathcal{N}_1, \mathcal{N}_2, \quad \mathcal{N} = \{ N_1 \cap N_2 : N_1 \in \mathcal{N}_1, N_2 \in \mathcal{N}_2 \}$

$$S_n(\mathcal{N}) \leqslant S_n(\mathcal{N}_1) \cdot S_n(\mathcal{N}_2)$$

$$S_n(\mathcal{N}) = \sup_{\substack{S \subset Z \\ |S| = n}} \left| \{ S \cap N_1 \cap N_2 : N_1 \in \mathcal{N}_1, N_2 \in \mathcal{N}_2 \} \right|$$

- For S : $\left| \{ S \cap N_1 : N_1 \subset \mathcal{N}_1 \} \right| = \left| \{ B_1, B_2 \cdots B_l \} \right| \leq \underline{S_n(\mathcal{N}_1)}$

  - For $i = 1 : l$: $\quad \underbrace{S \cap N_1} \cap N_2 = \underline{B_i} \cap N_2$

$$\{ B_i \cap N_2 : N_2 \in \mathcal{N}_2 \} \leq S_{|B_i|}(\mathcal{N}_2) \leq S_n(\mathcal{N}_2)$$

$|B_i| \leq n$

$$|\{ S \cap N_1 \cap N_2 \}| \leqslant \sum_{i=1}^{\overset{\#\text{ of dinstinct subsets } \{S \cap N_1\}}{\overset{\|}{l}}} \#\text{ of distinct subsets } \{ B_i \cap N_2 \}$$

$$\leq l \cdot \max \{ \#\text{ of distinct subsets } \{ B_i \cap N_2 \} \}$$

$$\leq S_n(\mathcal{N}_1) \cdot S_n(\mathcal{N}_2)$$

$\tilde{\mathcal{A}} = \{ B_1^c \cap \cdots \cap B_k^c : B_j^c \in \overline{\mathcal{B}} \}$

$S_n(\tilde{\mathcal{A}}) \leq S_n^k(\overline{\mathcal{B}}) = S_n^k(\mathcal{B})$

$V(\mathcal{B}) = d+1$

LEMMA 7.2 (Sauer–Shelah lemma). *Let $\mathcal{C}$ be a class of subsets of some space Z with $V(\mathcal{C}) = d < \infty$. Then for all n,*

$$\mathbb{S}_n(\mathcal{C}) \leq \binom{n}{\leq d}. \tag{7.5}$$

Also, $\binom{n}{\leq d} \leq (n+1)^d$ and, for $n \geq d$, $\binom{n}{\leq d} \leq \left(\frac{ne}{d}\right)^d$.

$$S_n^k(\mathcal{B}) \leq \left(\frac{ne}{d+1}\right)^{(d+1)k}$$

$$S_n(\tilde{\mathcal{A}}) \leq \left(\frac{ne}{d+1}\right)^{(d+1)k}$$

$$V(\tilde{\mathcal{A}}) := \inf_n \{ n \in \mathbb{N} : 2^n > S_n(\tilde{\mathcal{A}}) \}$$

$$\Leftarrow \quad \text{for } n > V(\tilde{\mathcal{A}}) \quad 2^n > S_n(\tilde{\mathcal{A}})$$

$$2^n > \left(\frac{ne}{d+1}\right)^{(d+1)k}$$

$$\hookrightarrow \quad n = 4(k+1)d \log((d+1)k) \geq V(\tilde{\mathcal{A}})$$