Covering numbers: review + generalizations

Note Title 2015-11-13

Some notation:

Def's

Symmetric convex body: A compact set $K \subseteq \mathbb{R}^n$, whose interior contains 0, satisfying $K = -K$.

From now on, assume K is a symm. conv. body.
Ex) The unit ball of any norm is a symm. conv. body.

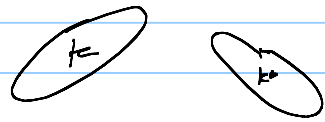
Minkowski functional For $x \in \mathbb{R}^n$, $\|x\|_K := \inf\{t \geq 0 : x \in tK\}$
(It is a norm if K is a symmetric convex body.)

- Convex body $K \to$ norm $\|\cdot\|_K$ w/ K as unit ball.

Ex) $\|x\|_{B_1} = \|x\|_1$.

Polar body $K^\circ := \{x \in \mathbb{R}^n : \langle x, y\rangle \leq 1 \quad \forall y \in K\}$

Ex) $K = B_2^n$, $K^\circ = B_2^n$



Ex) $K = B_1^n$, $K^\circ = B_\infty^n$

Thus, $\|x\|_{K^\circ} = \sup_{y \in K} \langle x, y\rangle$

Remark: $\|\cdot\|_K$ & $\|\cdot\|_{K^\circ}$ are dual norms.

Review

General simple covering arguments.

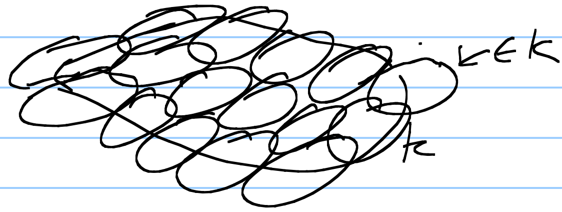
We controlled $\underset{\uparrow \text{ random matrix}}{\|A\|} = \sup_{x \in S^{n-1}} \|Ax\|_2$ by

discretizing S^{n-1} w/ a volumetric argument.

Same volumetric argument generalizes:

Thm Volumetric covering # — Let $K \subseteq \mathbb{R}^n$ (s.c.b.). Then

$$\boxed{N(K, \epsilon K) \leq \left(\frac{3}{\epsilon}\right)^n \qquad 0 < \epsilon}$$

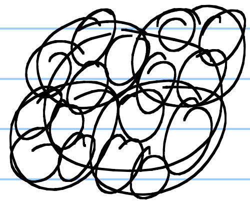


Application: More exotic random matrix norms.

Dudley's # corrected: Assume $X_t = 0$ a.s. for some $t \in K$.

$$\mathbb{E}\sup_{t \in K} |X_t| \leq C \int_0^{\operatorname{diam}(K)} \sqrt{\log N(K, d, \epsilon)}\, d\epsilon$$

sub-Gauss proc. w/ induced metric d.



- Gives proof of RIP for structured-random matrices.

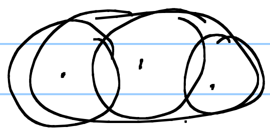
Generic chaining: $c\,\mathbb{E}\sup_{t \in T} X_t \leq \gamma_2(T, d) \leq C\,\mathbb{E}\sup_{t \in T} Y_t$

sub-Gauss. Gauss.

$\Rightarrow$ Master theorem

More Covering #'s

Def Packing: Given a set K and a metric d, an ϵ-packing is a subset $\mathcal{X} \subseteq K$ satisfying $d(x, y) > \epsilon$ for $x, y \in \mathcal{X}$, $x \neq y$.



- A maximal packing satisfies $d(\mathcal{X}, v) \leq \epsilon$ $\forall v \in K$ i.e., you can't add any more points to the packing.

- Packing number: $P(K, d, \epsilon)$ is cardinality of larges ϵ-packing.

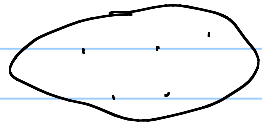
Lemma (Packing-covering equivalence)

$$\boxed{N(K, d, \epsilon) \leq P(K, d, \epsilon) \leq N\left(K, d, \frac{\epsilon}{2}\right)}$$

Proof: Exercize

Recall
Sudakov minoration:

$$w(K) \geq c\,\epsilon \sqrt{\log P(K, \epsilon B_2^n)}$$

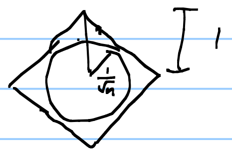


This gives a useful covering #:

$\Rightarrow$ **Lemma** (Sudakov minoration)

$$\boxed{\log(N(K, \epsilon B_2^n)) \leq \frac{c}{\epsilon^2} w(K)^2}$$

Ex) $K = B_1$



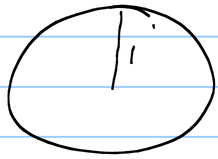
$w(K) = \mathbb{E}\|g\|_{K^\circ} = \mathbb{E}\|g\|_\infty = c\sqrt{\log n}$

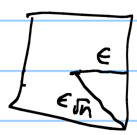
$$\log(N(B_1, \epsilon B_2^n)) \leq \frac{c}{\epsilon^2} \log(n)$$

"At coarse scales, B_1 behaves like a $\log(n)$-dim unit ball."

Lemma (Reverse Sudakov minoration)

$$\boxed{\log(N(B_2^n, \epsilon K)) \leq \frac{c}{\epsilon^2} w(K^\circ)^2}$$





Ex) $\log N(B_2^n, \epsilon B_\infty) \leq \frac{c}{\epsilon^2} w(B_1)^2 = \frac{c}{\epsilon^2}\sqrt{\log(n)}$

Application: Can be used to generalize Master theorem to a structured-random matrix, A, when $K = B_2^n$, i.e., control the sing. val's of A.