Exercise 5.7

Proof of lemma 5.5.1 part 2

Jules Coret

May 25, 2017

Theorem

Given an $r \in (0,1)$, there exist constants $\epsilon, c > 0$. Such that for all n, s with $s < rn$ the set of s-sparse vectors in $\mathbb{R}^n$ has a cardinality of at least $n \vee \exp(cs\log(n/s))$

Proof

Let S be the set of s-sparse vectors whose non-zero coordiantes are equal to $1/\sqrt{s}$. Let X, Y be independ random variables uniformly distributed on S. Since $\|X - Y\|^2 = \|X - Y\|_0/s$ we have

$$\mathbb{P}(\|X - Y\| \leq \epsilon) = \mathbb{P}(\|X - Y\|_0 \leq \epsilon^2 s)$$

$$\begin{aligned}
\mathbb{P}(\|X - Y\|_0 \leq \epsilon^2 s | X) &= \frac{|\{y \in S : \|y - X\|_0 \leq \epsilon^2 s\}|}{|S|} \\
&= \binom{n}{s}^{-1} \binom{s}{(1-\epsilon^2)s} \binom{n - (1-\epsilon^2)s}{\epsilon^2 s} \\
&= \binom{n}{s}^{-1} \binom{s}{E} \binom{n-E}{s-E} \\
&= \frac{s!(n-s)!s!(n-E)!}{n!E!(s-E)!(s-E)!(n-s)!}
\end{aligned}$$

Proof

$$
\begin{aligned}
&= \frac{s!s!(n-E)!E!}{n!E!E!(s-E)!(s-E)!} \\
&= \binom{n}{E}^{-1}\binom{s}{s-E}^{2} \\
&= \binom{n}{(1-\epsilon^2)s}^{-1}\binom{s}{\epsilon^2 s}^{2} \\
&\leq (n/(1-\epsilon^2)s)^{(-1+\epsilon^2)s}(e/\epsilon^2)^{2\epsilon^2 s} \\
&\leq ((1-\epsilon^2)r)^{(1-\epsilon^2)rn} e^{2\epsilon^2 rn} \epsilon^{-4\epsilon^2 rn} \\
&= \exp\{(1-\epsilon^2)rn\left(\log(r)+\log(1-\epsilon^2)\right) + 2\epsilon^2 rn - \log(\epsilon)4\epsilon^2 rn\} \\
&\leq \exp\{rn((\log(r)) + 2\epsilon^2 - \log(\epsilon)4\epsilon^2)\}
\end{aligned}
$$

Proof

Let $X_1, X_2, .., X_M$ be M independent vectors uniform on S then

$$\mathbb{P}(\|X_i - X_j\| \leq \epsilon \text{ for some } i \neq j) \leq M^2 \exp\{rn((\log(r)) + 2\epsilon^2 - \log(\epsilon)4\epsilon^2)\}$$

for $M = \exp(cs\log(n/s))$ we have

$$\begin{aligned}\mathbb{P} &\leq \exp\{2cs\log(n/s) + rn(\log(r) + 2\epsilon^2(1 - 2\log(\epsilon)))\} \\ &\leq \exp\{-2crn\log(r) + rn(\log(r) + 2\epsilon^2(1 - 2\log(\epsilon)))\} \\ &= \exp\{rn((1-2c)\log(r) + 2\epsilon^2(1 - 2\log(\epsilon)))\}\end{aligned}$$

Proof

$$\begin{aligned} \mathbb{P} \leq \exp\{rn((1-2c)\log(r) + 2\epsilon^2(1-2\log(\epsilon)))\} &< 1 \\ (1-2c)\log(r) + 2\epsilon^2(1-2\log(\epsilon)) &< 0 \\ c &< 1/2 \\ \epsilon \text{ such that } 2\epsilon^2(1-2\log(\epsilon)) &< -\log(r)(1-2c) \end{aligned}$$

We also have that

$$\mathbb{P}(\|X_i - X_j\| \geq \epsilon \text{ for all } i \neq j) = 1 - \mathbb{P}(\|X_i - X_j\| \leq \epsilon \text{ for some } i \neq j)$$

Thus we have for these parameters that $\mathbb{P}(\|X_i - X_j\| \geq \epsilon \text{ for all } i \neq j) > 0$
Thus there has to be a set S of s-sparse vectors separated by ϵ with cardinality greater or equal to $\exp(cs\log(n/s))$