پاسخ پایان‌ترم Take home — نیما سادینی 400101204

سوال ۱:

(I) $\mu_{emp} = \frac{1}{N}\sum_{i=1}^{N} x_i$, $x_i \overset{iid}{\sim} \mathcal{N}(\mu, \Sigma)$

اثبات

مدعا: $\mathbb{P}\left[\|\mu_{emp} - \mu\| \le C\sqrt{\frac{\mathrm{Tr}(\Sigma)}{N}} + C\sqrt{\frac{2\|\Sigma\|\log(1/\delta)}{N}}\right.$

We can claim that the distribution of $\mu_{emp}$ is a gaussian also. (linear combination of gaussians will be gaussian)

$\hookrightarrow \boxed{\mu_{emp} \sim \mathcal{N}(\mu, \frac{\Sigma}{n})}$ (1)

$\mathbb{E}\mu_{emp} = \frac{1}{N}\sum \mu = \vec{\mu}$.

$$\mathrm{Var}(\mu_{emp}) = \Sigma_{\mu_{emp}} = \mathbb{E}\left[(\vec{\mu}_{emp} - \vec{\mu})(\vec{\mu}_{emp} - \vec{\mu})^T\right] = \frac{1}{N^2}\mathbb{E}\left[\left(\sum_i \vec{X}_i - \vec{\mu}\right)\left(\sum_j \vec{X}_j - \vec{\mu}\right)^T\right]$$

$$= \sum_{i,j} \frac{1}{N^2}\Big(\mathbb{E}X_iX_j^T + \underbrace{\mu\mu^T - 2(\mathbb{E}X_i)\mu^T - \mu(\mathbb{E}X_j)^T}_{-\mu\mu^T}\Big) = \sum_{i,j} \mathbb{1}\{i=j\}\,\Sigma = \frac{1}{N}\Sigma$$

Now let $Z = \mu_{emp} - \mu \sim \mathcal{N}(0, \frac{\Sigma}{N})$:

$$\mathbb{E}\left[\left(\|\mu_{emp} - \mu\|_2^2\right)^{1/2}\right] \underset{\text{Jensen}}{\le} \sqrt{\mathbb{E}\|\vec{\mu}_{emp} - \vec{\mu}\|^2} = \sqrt{\mathbb{E}\left[(\vec{\mu}_{emp} - \vec{\mu})^T(\vec{\mu}_{emp} - \mu)\right]} = \sqrt{\mathbb{E}\left[\mathrm{Tr}\{(\vec{\mu}_{emp} - \vec{\mu})(\vec{\mu}_{emp} - \vec{\mu})\}\right]}$$

$X^TX = \mathrm{tr}\{XX^T\}$

$$= \sqrt{\mathrm{tr}\left\{\mathbb{E}\left[(\vec{\mu}_{emp} - \vec{\mu})(\vec{\mu}_{emp} - \vec{\mu})^T\right]\right\}} = \sqrt{\mathrm{tr}(\Sigma)\times\frac{1}{N}}$$

$\mathbb{E}\{\mathrm{Tr}\{\cdot\}\} = \mathrm{Tr}\{\mathbb{E}\{\cdot\}\}$ $\quad$ $\vec{\mu}_{emp} - \vec{\mu} \sim \mathcal{N}(0, \frac{\Sigma}{N})$

which suggests

$$\mathbb{E}\Big[\|\underbrace{\vec{\mu}_{emp} - \vec{\mu}}_{Z}\|_2\Big] \le \sqrt{\frac{1}{N}\mathrm{Tr}\{\Sigma\}} \quad (2)$$

Now we can whiten $Z$ by writing it as $Z = (\frac{1}{N}\Sigma)^{1/2}G$ where $G \sim \mathcal{N}(0, I)$ is standard gaussian. We want to extract the lipschitz constant!

$$\|Z\|_2 = \left\|\frac{1}{\sqrt{N}}\Sigma^{1/2}G\right\|_2 \rightarrow \nabla_G\{\|Z\|\} = \frac{1}{\sqrt{N}}\Sigma^{1/2}\cdot\frac{\Sigma^{1/2}G}{\|\Sigma^{1/2}G\|_2} \rightarrow \|\nabla_G\{\|Z\|\}\| = \frac{1}{\sqrt{N}}\cdot\frac{\|\Sigma G\|_2}{\|\Sigma^{1/2}G\|_2}$$

$$\|\nabla_G\{\|Z\|\}\|_2 \le \frac{1}{\sqrt{N}}\cdot\frac{\|\Sigma^{1/2}\|\cdot\|\Sigma^{1/2}G\|}{\|\Sigma^{1/2}G\|} = \frac{1}{\sqrt{N}}\cdot\sqrt{\|\Sigma\|} \longrightarrow L = \sqrt{\frac{\|\Sigma\|}{N}}$$

$\|A\|\cdot\|B\| \ge \|AB\|$ $\quad$ $\|A^{1/2}\| = \|A\|^{1/2}$ $\quad$ $\nabla_G\{\|\mu - \mu_{emp}\|_2^2\} \le L$

From the course, we know that if $\nabla_G\{\cdot\} \le L$, then it will be $L^2$ subgaussian. thus we'd get

$t \overset{?}{=} O(L^2\log(\frac{1}{\delta}))$

$b \le c$, $\mathbb{P}(a > b) \le d \Rightarrow \mathbb{P}(a > c) \le d$

$$\mathbb{P}\Big[\|\vec{\mu} - \vec{\mu}_{emp}\|_2 \ge \underbrace{\mathbb{E}\big[\|\vec{\mu} - \vec{\mu}_{emp}\|_2\big]}_{\le \frac{1}{\sqrt{N}}\sqrt{\mathrm{Tr}\{\Sigma\}}} + t\Big] \le \exp\left(-\frac{t^2}{2L^2}\right) \le \delta$$

which leads to $\mathbb{P}\left[\|\vec{\mu} - \vec{\mu}_{emp}\|_2 \ge \sqrt{\frac{\mathrm{Tr}(\Sigma)}{N}} + \sqrt{\frac{2\|\Sigma\|\log(1/\delta)}{N}}\right] \le \delta$

$$\Longrightarrow \mathbb{P}\left[\|\vec{\mu} - \vec{\mu}_{emp}\|_2 \le \sqrt{\frac{\mathrm{Tr}(\Sigma)}{N}} + \sqrt{\frac{2\|\Sigma\|\log(1/\delta)}{N}}\right] \ge 1 - \delta \quad \text{Q.E.D.}$$

/1

$\text{III}$ $z_j = \frac{1}{m}\sum_{i\in B_j} X_i \sim N(\mu, \frac{\Sigma}{m})$ $\quad \|x\|_2 = x^T x = tr\{xx^T\}$ 

تمرین ۴ (II)

(i) $P[\|z_j-\mu\|_2 \geq s] \leq \frac{E[\|z_j-\mu\|_2^2]}{s^2} = \frac{E[tr\{(z_j-\mu)(z_j-\mu)^T\}]}{s^2}$

$tr\{E\} = E\{tr\}$

$= \frac{tr\{E[(z_j-\mu)(z_j-\mu)^T]\}}{s^2} = \frac{tr\{\Sigma\}}{ms^2} \Rightarrow \boxed{P[\|z_j-\mu\|^2 \geq s] \leq \frac{tr\{\Sigma\}}{ms^2}}$ (۵) P.L.D

(ii) $m = \frac{N}{k} = |B_j|$ size of each block. In order to get the ~~median of me~~ mean of medians, we must find at least $j_1, \dots, j_{\lfloor \frac{k}{2}\rfloor}$ such that $\|\mu - z_{j_i}\| \geq s$, thus:

all $\frac{k}{2}$ subsets

$$P[\exists j_1, \dots, j_{\lfloor \frac{k}{2}\rfloor}: \|\mu - z_{j_i}\| \geq s] \leq \sum_{j\in\{1,\dots,n\}} \sum_{|\pi| = \lfloor k/2\rfloor} P[\forall i\in\{1,\dots,\tfrac{k}{2}\}: |\mu - z_{\pi(i)}| \geq s]$$

$$\leq \binom{k}{\frac{k}{2}} P[\|\vec{z}_{\pi(i)} - \vec{\mu}\| \geq s] \leq \sum_{|\pi| = \lfloor k/2\rfloor} \left(P[\|\vec{z}_{\pi(i)} - \vec{\mu}\| \geq s]\right)^{\lfloor k/2\rfloor}$$

$$\leq 2^k \cdot \left(P(\|\vec{z}_{\pi(i)} - \vec{\mu}\| \geq s)\right)^{k/2} \overset{(۴)}{\leq} 2^k \cdot \left(\frac{tr\{\Sigma\}\, k}{N\delta^2}\right)^{\lfloor k/2\rfloor} \approx \left(\frac{4\, tr\{\Sigma\}\, k}{N\delta^2}\right)^{\lfloor \frac{k}{2}\rfloor}$$

now let $\frac{4\, tr\{\Sigma\} k}{N\delta^2} = O(1) = \Theta(\frac{1}{4}) \rightarrow \delta = O\left(\sqrt{\frac{Tr\{\Sigma\}\cdot k}{N}}\right)$

Now putting rest will give us:

$$P\left[\exists j_1 - j_{\lceil \frac{k}{2}\rceil}: \|\mu - z_{j_i}\| \geq c\sqrt{\frac{k\cdot Tr\{\Sigma\}}{N}}\right] \leq 2^{-k} \leq \delta$$

$2^{-k} \leq \delta \rightarrow$ from rashyrin if $k = \Theta(\log(\frac{1}{\delta})) \rightarrow c \geq 4$ is ok!

(iii) u, v are both r-median. So from above we can say that u is r-median iff it is $\mu$-distant from its bound $c\sqrt{\frac{Tr\{\Sigma\}\log(1/\delta)}{N}}$. For the sake of rating, r-median is a point that has distance less than r, with at least $\lfloor \frac{k}{2}\rfloor$ points.

ادامه در پیش pdf (۲)

5 نمره (III)

(i) from definition $S_a = \{x \in \mathbb{R}^d, \exists J \subseteq \{1,\dots,k\}, |J| \geq \frac{k}{2}, \forall j \in J\ \|\vec{z}_j - \vec{x}\| \leq \|\vec{z}_j - \vec{a}\|\}$

let $\mu$ be a vector which doesn't lie on $S_a$ ($\vec{\mu} \in \bar{S}_a$), therefore,

$\exists J \in \{1,\dots,k\}, |J| \geq \frac{k}{2}$,

thus $\Rightarrow \forall J \in \{1,\dots,k\}, |J| \geq \frac{k}{2}, \exists j \in J : \|\vec{z}_j - \vec{\mu}\| \leq \|\vec{z}_j - \vec{a}\|$ (from $\vec{\mu} \in \bar{S}_a$)

from the idea of "chaining", we can write

$\|\vec{a} - \vec{\mu}\| \leq \|\vec{a} - \vec{z}_j + \vec{z}_j - \vec{\mu}\| \leq \|\vec{a} - \vec{z}_j\| + \|\vec{\mu} - \vec{z}_j\| = \sup \|\vec{x} - \vec{y}\|$

$x$ s.t. $\|x - z_j\| \leq \|\mu - z_j\|$

we assume that $x, y, z_j$ are colinear, in $\leftarrow$ $y$ s.t. $\|y - z_j\| \leq \|a - z_j\|$

a sense that $x$ replaces $\mu$, and $y$ replaces $a$. Note that not $x$, nor $y$ do not belong to $S_a$. thus:

$$\boxed{\|\vec{a} - \vec{\mu}\| \leq \sup_{\substack{x:\ - \\ y:\ -}} \{\|x - y\|\} \leq \underbrace{\sup_{x,y \in \bar{S}_a}}_{\text{bigger space}} \|x - y\| \underbrace{=}_{\text{by definition}} \operatorname{diam}(S_a^c)}$$

Q.E.D. ■

(ii) Now from definition, the optimum estimator is $\hat{\mu}_N = \underset{a \in \mathbb{R}^d}{\operatorname{argmin}} \operatorname{diam}(\bar{S}_a)$

we wish to prove $\|\hat{\mu}_N - \mu\| \leq \operatorname{diam}(\bar{S}_{\hat{\mu}_N})$.

let $\hat{\mu}_N = \underset{a \in \mathbb{R}^d}{\operatorname{argmin}} \{\operatorname{diam}(S_a^c)\}$ since, in the previous part we

proved that $\forall \vec{a} \in \mathbb{R}^d, \vec{\mu} \in \bar{S}_a$: $\|a - \mu\| \leq \operatorname{diam}(\bar{S}_a)$, so it also holds for $a = \hat{\mu}_N$

Q.E.D. ■

$\forall \mu \in \bar{S}_{\hat{\mu}_N} \Rightarrow \boxed{\|\hat{\mu}_N - \mu\| \leq \operatorname{diam}(\bar{S}_{\hat{\mu}_N})}$. However, it's doubtful to say that

$\mu \in \bar{S}_{\hat{\mu}_N}$ in general.

---

6 نمره (IV)

let $k = \lceil 360 \log \frac{2}{\delta} \rceil$, $r = \max \left\{ 4e\eta \sqrt{\frac{Tr(\Sigma)}{N}},\ 240 \sqrt{\frac{\|\Sigma\| \log \frac{2}{\delta}}{N}} \right\}$

we wish to bound the concentration for $\sum_i \langle X_i - \mu, v \rangle$.

(1)

$= \langle v, \Sigma v \rangle$

$v \in \mathbb{R}^d$, $\|v\|_2 = r$

$\Rightarrow \mathbb{E}[\langle X - \mu, v \rangle^2] = \mathbb{E}[\{(X - \mu)^T v\}^2] = \mathbb{E}[v^T (X - \mu)(X - \mu)^T v] = v^T \Sigma v$

$\mathbb{E}[(X - \mu)(X - \mu)^T] = \Sigma$

$\Rightarrow \mathbb{P}\left[\frac{1}{m} \sum_{i \in B_j} \langle X_i - \mu, v \rangle \geq \frac{r^2}{4}\right] = \mathbb{P}\left[\left|\frac{1}{m} \sum_{i \in B_j} \langle X_i - \mu, v \rangle\right|^2 \geq \frac{r^4}{16}\right] \underset{\text{chebyshev}}{\leq} \frac{16 \mathbb{E}\left[\left|\frac{1}{m} \sum_{i \in B_j} \langle X_i - \mu, v \rangle\right|^2\right]}{r^4}$

$\leq \frac{16}{r^4} \mathbb{E}\left[\frac{1}{m^2} \sum_{i \in B_j} \sum_{k \in B_j} \langle X_i - \mu, v \rangle \langle X_j - \mu, v \rangle\right] = \frac{16}{r^4} \sum_{i \in B_j} \sum_{k \in B_j} \mathbb{1}\{i = k\} \mathbb{E}[\langle X_i - \mu, v \rangle^2]$

(1) $= \frac{16}{r^4} \times m \left(\frac{1}{m^2} v^T \Sigma v\right) = \frac{16 \cdot v^T \Sigma v}{m r^4} = \longrightarrow$ ادامه در صفحه بعد

$$\mathbb{P}\left[\frac{1}{m}\sum_{i\in B_j}\langle X_i-\mu, v\rangle^2 \geq \frac{r^2}{4}\right] \leq \frac{16}{m r^4} v^T\Sigma v = \frac{16k}{N r^4} v^T \Sigma v \leq \frac{16k}{Nr^4}\|\Sigma\|\cdot\underbrace{\|v\|_2^2}_{1}$$ (IV)

$m=\frac{N}{k}$ from blocks

$$\leq \frac{16k\|\Sigma\|}{Nr^2} = \frac{16\|\Sigma\|}{N}\left\lceil 360\log\frac{2}{\delta}\right\rceil \times \left\{\frac{1}{\max\left\{400\sqrt{\frac{Tr(\Sigma)}{N}},\ 240\sqrt{\frac{\|\Sigma\|\log(2/\delta)}{N}}\right\}}\right\}^2$$

opening the definition of $k, r$

$$= \frac{16\|\Sigma\|}{N}\left\lceil 360\log\frac{2}{\delta}\right\rceil\cdot \min\left(\frac{400^2 Tr\{\Sigma\}}{N},\ \frac{240^2\|\Sigma\|\log\frac{2}{\delta}}{N}\right)^{-1}$$

$$= 16\|\Sigma\|\cdot\left\lceil 360\log\frac{2}{\delta}\right\rceil\cdot\left\{\min\left(\cancel{160000\, Tr\{\Sigma\}},\ 240^2\|\Sigma\|\log\frac{2}{\delta}\right)\right\}^{-1}$$

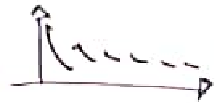
$$\leq \frac{\cancel{\|\Sigma\|}\left\lceil 360\log\frac{2}{\delta}\right\rceil\cdot\cancel{16}}{\cancel{16}\cdot 60^2\ \cancel{\|\Sigma\|}\log\frac{2}{\delta}} = \frac{1}{10}\ \frac{\left\lceil 360\log\frac{2}{\delta}\right\rceil}{360\log\frac{2}{\delta}}$$

for certain amount of $\delta$, we can make it such that the ratio above will be $\frac{1}{10}$.

since $\delta\leq 2^{-k}$ (in previous parts) $\rightarrow \begin{cases} 2^{-k-1}\leq\delta\leq 2^{-k} \rightarrow \log(\frac{2}{\delta}) \\ (k+1)\geq\log\frac{1}{\delta}\geq kc\end{cases}$

because if $\delta = 2^{-k} \rightarrow$ then

$\log_2(\frac{2}{\delta})\in\mathbb{N} \rightarrow \frac{\lceil 360\log\frac{2}{\delta}\rceil}{360\log\frac{2}{\delta}} = 1$ But this doesn't help and proves $\frac{2}{10}$ Bound.



$\frac{\lceil x\rceil}{x} = f(x)$

$\frac{\lceil x\rceil}{x}\leq 2 : 1<x$

and gives us the $1/10$ Bound that we were looking for.

$$\hookrightarrow \boxed{\mathbb{P}\left[\left|\frac{1}{m}\sum_{v\in B_j}\langle X_i-\mu, v\rangle\right|\geq\frac{r^2}{4}\right]\leq\frac{1}{10}}$$ Q.E.D. ■

---

$\boxed{V \mid \text{IV} \ldots}$

(1) We can use the Bound of the final: $G\sim N(0,I)$

$$\begin{cases}\log N(B^d(r), \|\cdot\|_\Sigma, \varepsilon) \leq C\left(\frac{r\,\mathbb{E}\left[\sqrt{G^T\Sigma G}\right]}{\varepsilon}\right)^2, & \varepsilon = 5r\sqrt{\frac{Tr\{\Sigma\}}{k}} \\ d(u,v) = u^T\Sigma v\end{cases}$$

$$\Rightarrow \log N(S^d(r), \|\cdot\|_\Sigma, \varepsilon) \leq C\left(\frac{r\,\mathbb{E}(\sqrt{G^T\Sigma G})}{5r\sqrt{\frac{Tr\{\Sigma\}}{k}}}\right)^2 = \frac{ck}{25\,Tr\{\Sigma\}}\cdot\left(\mathbb{E}\left\{\sqrt{G^T\Sigma G}\right\}\right)^2$$

$\mathbb{E}(\cdot) = cte = tr\{\Sigma\}$

$$\underset{\text{Jensen}}{\leq} \frac{Ck}{25\,Tr\{\Sigma\}}\,\mathbb{E}\left[G^T\Sigma G\right] = \frac{ck}{25\,tr\{\Sigma\}}\,\mathbb{E}\left[tr\{\underbrace{GG^T}_{I}\Sigma\}\right] = \frac{ck}{25}$$

(ii) Up until now Prove that $\log \mathcal{N}(S^{d-1}(r), \|\cdot\|_2, \varepsilon) \le \frac{ck}{25}$. we wish to Prove that w.p. at least 0.8 of blocks, $\left|\frac{1}{m}\sum_{i\in B_j}\langle X_i-\mu, v\rangle\right| \le \frac{r^2}{4}$.

from Previous Parts we know that

$$\forall v\in\mathbb{R}^d,\ \|v\|_2 = r \to \mathbb{P}\left[\left|\frac{1}{m}\sum_{i\in B_j}\langle X_i-\mu, v\rangle\right| \le \frac{r^2}{4}\right] \le \frac{1}{10} \qquad (4)$$

$$\mathbb{P}\left[\left|\frac{1}{m}\sum_{i\in B_j}\langle X_i-\mu, v\rangle\right| \le \frac{r^2}{4} \text{ for at least } l=\alpha k \text{ blocks}\right]$$

$$= 1-\mathbb{P}\left[\left|\frac{1}{m}\sum_{i\in B_j}\langle X_i-\mu, v\rangle\right| \ge \frac{r^2}{4} \text{ for at most } \overbrace{k-l}^{(1-\alpha)k} \text{ blocks}\right]$$

$$= 1-\mathbb{P}\left[\sum_{j=1}^{k} \underbrace{\overbrace{\mathbb{1}\left\{\left|\frac{1}{m}\sum_{i\in B_j}\langle X_i-\mu, v\rangle\right| \ge \frac{r^2}{4}\right\}}^{A_j}} \ge k-l\right]$$

from (4) we know that this event ($\mathbb{1}\{-\}=1$) happens with at most 1/10 of probability, thus: $\mathbb{E}\left[\sum_{j=1}^{k}\underbrace{\mathbb{1}\{-\}}_{A_j}\right] \le \frac{k}{10}$ ( $\sum_{j=1}^{k} A_j$ is $\frac{k}{4}$ subgaussian )

$$\mathbb{P}\left(\sum_{j=1}^{k} A_j \ge t + \frac{k}{10}\right] \le \mathbb{P}\left(\sum_{j=1}^{k} A_j - \underbrace{\mathbb{E}A_j}_{\le \frac{k}{10}} \ge t\right] \le e^{-\frac{2t^2}{k^2}}$$

$\mathbb{P}[A \ge t] \le \mathbb{P}(A \ge t+\varepsilon)$ $\varepsilon>0$

$V$ is our set of coverage $|V| = \mathcal{N}(-)$

Now let $t = 0.1k \to$

$$\mathbb{P}\left[\left|\frac{1}{m}\sum_{i\in B_j}\langle X_i-\mu, v\rangle\right| \le \frac{r^2}{4} \text{ for at least } 0.8k \text{ Blocks } \forall v\in V\right] \ge 1-\sum_{v\in V}\left(1-\mathbb{P}\left[\left|\frac{1}{m}\sum_{i\in B_j}\langle X_i-\mu, v\rangle\right| \le \frac{r^2}{4} \text{ for at least } \underbrace{l}_{\alpha k} \text{ Blocks}\right]\right)$$

$$\ge 1-\sum_{v\in V} e^{-\frac{k}{50}} = 1-|\mathcal{N}(-)|\,e^{-\frac{k}{50}} \le 1-e^{\frac{k}{360}}e^{-\frac{k}{50}}$$

$\left(\frac{1}{50}+\frac{1}{360}\right)^{-1} \approx 43 \longrightarrow \ge 1-e^{-\frac{k}{40}} \ge 1-e^{-\frac{k}{360}} \ge 1-e^{-\frac{1}{720}\left\lceil 360\log\frac{2}{\delta}\right\rceil}$

$\frac{1}{40} > \frac{1}{360} > \frac{1}{720}$ Rewriting $k = \left\lceil 360\log\frac{2}{\delta}\right\rceil$

assuming that $\left(360\log\frac{2}{\delta}\right) > 1$, we can say that $\frac{k}{720} \ge \frac{2\log\frac{2}{\delta}}{2} = \log\frac{2}{\delta}$

$$\mathbb{P}[-] \le 1-e^{-\frac{1}{360}\times\frac{1}{2}\times 2\log\left(\frac{2}{\delta}\right)} = 1-e^{-\log\frac{2}{\delta}} = 1-\frac{\delta}{2}.$$

Q.E.D. ■

So we proved

$$\mathbb{P}\left[ \left| \frac{1}{m} \sum_{i \in B_j} \langle X_i - \mu, v \rangle \right| \le \frac{r^2}{4} \ \text{for at least } 0.8 \text{ of blocks} \right] \ge 1 - \frac{\delta}{2}$$

---

(iii) $$\underbrace{\sup_{x \in S^d(r)} \left\{ \frac{1}{k} \sum_{j=1}^{k} \mathbb{1}\left\{ \left| \frac{1}{m} \sum_{i \in B_j} \langle X_i - \mu, x - v_x \rangle \right| \ge \frac{r^2}{4} \right\} \right.}_{Y} \le \frac{1}{10} \quad \text{w.p. } 1 - \frac{\delta}{2}$$

we can see that $Y$ is subgaussian. let $\hat{B}_i = \{X_j \in B_i\}$. so $Y$ is a function of $\hat{B}_i$'s. we'll show that $Y$ is finite difference:

only one of these $\mathbb{1}(-)$ indicator functions can alter so

$$Y(\hat{B}_1', \hat{B}_{2:n}) - Y(\hat{B}_1, \hat{B}_{2:n}) = \sup_{x \in S^d(r)}(-) - \sup_{x \in \mathbb{R}^d}(-)$$

$$\le \frac{1}{k} \sum_{j=1}^{k} \mathbb{1}\left\{ \left| \underbrace{\frac{1}{m} \sum_{i \in B_j} \langle X_i - \mu, \tilde{x} - \tilde{v}_x \rangle}_{H_i} \right| \ge \frac{r^2}{4} \right\} - \frac{1}{k} \sum_{j=1}^{k} \mathbb{1}\{--\} \le \frac{1}{k}$$

so $Y$ is finite difference with $c_1 = c_2 = \dots = c_n = 1/k \to \frac{\sum c_i^2}{4}$ subgaussian $= \frac{1}{4k}$

so $Y \sim \frac{1}{4k}$ - subgaussian.

$\hookrightarrow \mathbb{P}[Y - \mathbb{E}Y \ge t] \le e^{-2kt^2}$

let $j \sim \mathrm{Unif}\{1, \dots, k\}$

$$\Rightarrow Y = \frac{1}{k} \sum_{j=1}^{k} (--) = \mathbb{E}_{j \sim \mathrm{Unif}\{1..k\}}\left[ \mathbb{1}\left[ H_i \ge \frac{r^2}{4} \right] \right] = \mathbb{P}_{j \sim \mathrm{Unif}}\left[ H_i \ge \frac{r^2}{4} \right]$$

$$Y \underset{\text{markov}}{\le} \frac{4\,\mathbb{E}_{j \sim \mathrm{Unif}}[H_j]}{r^2} = \frac{4}{r^2 k} \sum_{j=1}^{K} \underbrace{|\langle Z_j - \mu, x - v_x \rangle|}_{f(Z_j)}$$

the idea of chaining over ($\sup \mathbb{E}$)

$$\mathbb{E}Y = \frac{4}{r^2} \mathbb{E}\left[ \sup_{f \in \mathcal{F}} \left\{ \mathbb{E}_{emp}[f(Z_j)] \right\} \right] = \frac{4}{r^2} \mathbb{E}\left[ \sup_{f \in \mathcal{F}} \left\{ \mathbb{E}_{emp}(f(Z_j)) - \mathbb{E}[f(Z_j')] + \mathbb{E}[f(Z_j')] \right\} \right]$$

$$\le \frac{4}{r^2} \left( \underbrace{\mathbb{E}\left[ \sup_{f \in F} \mathbb{E}_{emp} f(Z) - \mathbb{E} f(Z') \right]}_{\mathrm{I}} + \underbrace{\sup_{f \in F} \mathbb{E} f(Z')}_{\mathrm{II}} \right)$$

(II) the first term:

$$\frac{4}{r^2} \sup_{f \in F} [\mathbb{E} f(Z')] = \frac{4}{r^2} \sup_{x \in S^d(r)} \left\{ \mathbb{E} |\langle Z - \mu, x - v_x \rangle| \right\} \underset{\text{Jensen}}{\le} \frac{4}{r^2} \sup_{x \in S^d(r)} \left\{ \sqrt{\mathbb{E} |\langle Z - \mu, x - v_x \rangle|^2} \right\}$$

$$\le \frac{4}{r^2} \sup_{x \in S^d(r)} \left\{ \frac{1}{\sqrt{m}} \|x - v_x\|_\Sigma \right\} = \frac{4}{r^2 \sqrt{m}} \sup_x \underbrace{\|x - v_x\|_\Sigma}_{\le \varepsilon} \le \frac{4\varepsilon}{r^2 \sqrt{m}}$$

they will be on the same direction

due to the $\varepsilon$-covering by $N_\varepsilon(-)$

However, we know that $\varepsilon = 5r\sqrt{\frac{Tr\{\Sigma\}}{k}}$

Thus: $\text{(I)} \to \frac{4}{r^2} \sup \mathbb{E}[f(z')] \leq \frac{4\varepsilon}{r^2\sqrt{m}} = \frac{20}{r\sqrt{m}\,k}\sqrt{Tr\{\Sigma\}}$

$r = \max\left\{400\sqrt{\frac{Tr(\Sigma)}{N}},\ 240\sqrt{\frac{\|\Sigma\|\log(2/\delta)}{N}}\right\} \geq 400\frac{\sqrt{Tr(\Sigma)}}{\sqrt{N}}$

$\xrightarrow{N=mk}$ $\boxed{\text{(I)} \leq \frac{20}{400}\sqrt{\frac{Tr(\Sigma)}{mk}}\sqrt{\frac{mk}{Tr(\Sigma)}} = \frac{1}{20}}$ (V)

$f(z) = |\langle z-\mu, x-v_x\rangle|$

(II) the second term:

$$\mathbb{E}_z\left[\sup_{f\in F} \mathbb{E}_{\sup}[f(z)] - \mathbb{E}_{z'}[f(z')]\right] \leq \mathbb{E}_{z,z'}\left[\sup_{f\in F} \frac{1}{k}\sum_{j=1}^{k} f(z_j) - f(z'_j)\right]$$

$f(z) - f(z') = \langle z - z', x - v_x\rangle$ due to linear functions w.r.t $z$.

Now let $\vec{\varepsilon}_j \sim Rad(\frac{1}{2}) = \begin{cases} +1 & w.p.\ 1/2 \\ -1 & w.p.\ 1/2 \end{cases} \longrightarrow \leq \mathbb{E}_{z,\varepsilon}\left[\sup_{f\in F} \frac{1}{k}\sum_{j=1}^{k} \varepsilon_j f(z_j)\right]$

expanding $f(z)$

$$= \mathbb{E}_z\mathbb{E}_\varepsilon\left[\sup_{x\in S^d(r)} \left|\frac{1}{k}\sum_{j=1}^{k} \left(\langle (z_j-\mu)\varepsilon_j, x - v_x\rangle\right)\right|\right]$$

$$= \mathbb{E}_z\mathbb{E}_\varepsilon\left[\sup_{x\in S^d(r)} \left|\langle \underbrace{\sum_{j=1}^{k} \frac{(z_j-\mu)}{k}}_{T}, x - v_x\rangle\right|\right]$$

writing all rows $j$ from 1 to $k$ to form the matrix $T$

$T(s,:) = \frac{\vec{z}_s - \vec{\mu}}{k}$

$$= \mathbb{E}_z\mathbb{E}_\varepsilon\left[\sup_{x\in S^d(r)} |T(\vec{x} - \vec{v}_x)|\right]$$

$\leq \|T\|_2 \cdot \underbrace{\|x - v_x\|_2}_{\leq 2r} \leq 2r \cdot \|T\|$

$$\leq 2r\,\mathbb{E}_{z,\varepsilon}\left[\|T\|_2\right]$$

$\|T\|_2 = tr\{TT^T\}$, Jensen $\leq 2r\sqrt{\mathbb{E}_{z,\varepsilon}\left[tr\{TT^T\}\right]}$

$\varepsilon$ has no effect so we ignore $\mathbb{E}_{\vec{\varepsilon}}$

$$= 2r\sqrt{\mathbb{E}_z\left[\frac{1}{k}\sum_{j=1}^{k}(z_j-\mu)^2\right]} = 2r\sqrt{\mathbb{E}_x\left[\frac{1}{N}\left(\sum_{i=1}^{N}(X^i-\mu^i)\right)^2\right]} \leq 2r\sqrt{\frac{Tr\{\Sigma\}}{N}}$$

$$\text{(II)} \leq \frac{4\times2}{r^2}\left(2r\sqrt{\frac{Tr\{\Sigma\}}{N}}\right) = \frac{2\times8}{r}\sqrt{\frac{Tr\{\Sigma\}}{N}} \leq \frac{8\times2}{400} = \frac{1}{50}$$ $1/25$ (^)

$\frac{400\,Tr\{\Sigma\}}{N} \leq r = \max\left\{400\sqrt{\frac{Tr\{\Sigma\}}{N}},\ 240\sqrt{\frac{\|\Sigma\|\log(2/\delta)}{N}}\right\}$

$\Rightarrow \mathbb{E}Y \le \frac{1}{20} + \frac{2}{50} = \frac{9}{100} \Rightarrow \mathbb{P}[Y - \frac{9}{100} \ge t] \le e^{-2kt^2}$

(v) (∧) $\mathbb{P}[Y \ge \frac{1}{10}] = \mathbb{P}[Y - \frac{9}{100} > t]\Big|_{t=\frac{1}{100}} \le e^{-2k \cdot 10^{-4}}$

However we also knew $k = \left\lceil 360 \log(2/\delta) \right\rceil$. $\hookrightarrow e^{-2\lceil 360 \log\frac{2}{\delta}\rceil \cdot 10^{-4}}$

$\mathbb{P}[Y \ge \frac{1}{10}] \le c\, e^{-\frac{2(360)\log\frac{2}{\delta}}{10000}} \Rightarrow \mathbb{P}[Y \le \frac{1}{10}] \ge 1 - \underbrace{c\, e^{-\frac{720}{10^4}\log(\frac{2}{\delta})}}_{\frac{c\delta}{2}}$

$\Rightarrow$ Unfortunately our Bound for this Part was not that tight

we couldn't Prove the $1-\frac{\delta}{2}$, instead we proved $1 - \frac{c\delta}{2}$. $c \ne 1$

~~$\mathbb{E}Y(\frac{\delta}{10}) \dots \mathbb{P}[\dots] \le e^{-\frac{k}{50}} \le e^{-\frac{360\log(2/\delta)}{50}} \le \frac{\delta}{2}$~~

---

(iv) $\mathbb{P}\left[ \left| \frac{1}{m} \sum_{i=1}^{m} \langle X_i - \mu, v \rangle \right| \le \frac{r^2}{4} \right] \ge 1-\delta$

$\hookrightarrow$ then · less than 0.3 (at most) of blocks must have $\left| \frac{1}{m} \sum_{i=1}^{m} \langle X_i - \mu, v \rangle \right| \ge \frac{r^2}{4}$

we know that $\mathbb{P}\left[ \frac{1}{k} \sum_{j=1}^{k} \mathbb{1}\left[ \frac{1}{m} \sum_{i=1}^{m} \langle X_i - \mu, x - v_x \rangle \right| \ge \frac{r^2}{4} \right] \le \frac{1}{10}$

So to get 0.7 of blocks, we can say that since $1 - 0.7 = 0.3 = 0.1 + 0.2$

, we can consider two sets. one with the size $0.1k$ and the other with cardinality $0.2k$. thus the conditions for median will be.

$$\begin{cases} \forall j \in A: |A| < \frac{2}{10}k : \left| \frac{1}{m} \sum_{i=1}^{m} \langle X_i - \mu, v \rangle \right| \ge \frac{r^2}{4} \quad (*) \\ \forall j \in B: |B| < \frac{1}{10}k : \left| \frac{1}{m} \sum_{j=1}^{m} \langle X_i - \mu, v \rangle \right| \ge \frac{r^2}{4} \quad (**) \end{cases}$$

$\Rightarrow \{ \forall j \in \bar{J_1} \cap \bar{J_2} : \overline{(*)}\ \&\ \overline{(**)}$ holds. $\rightarrow \left| \frac{1}{m} \sum - \right| \le \frac{r^2}{4}, \left| \frac{1}{m} \sum - \right| \le \frac{r^2}{4}$

$$\begin{cases} \left| \frac{1}{m} \sum_{i \in B_j} \langle \vec{X_i} - \vec{\mu}, \vec{v} \rangle \right| \le \frac{r^2}{4} + \frac{r^2}{4} = \frac{r^2}{2} \\ |\bar{J_1} \cap \bar{J_2}| = |\overline{(J_1 \cup J_2)}| = k - |J_1 \cup J_2| \ge k - |J_1| - |J_2| = 0.7k \end{cases}$$

$|A \cup B| \le |A| + |B|$

$\Rightarrow$ the idea

therefore: $\Rightarrow$ $\boxed{\mathbb{P}\left[ \begin{matrix} \left| \frac{1}{m} \sum_{i \in B_j} \langle X_i - \mu, x \rangle \right| \le \frac{r^2}{2} \\ \forall x \in \mathbb{R}^d, \text{ for at least } 0.7k \text{ blocks} \end{matrix} \right] \ge 1 - \delta}$

Q.E.D. ∎