1/ HW3-HDP 400101204 محمد [illegible] تمرین سری 3 آمار و احتمالات در ابعاد بالا

سوال 1:

$$\mathbb{E}\left[\sup_{t\in T} X_t\right] = \mathbb{E}[X_0] + \mathbb{E}\left[\sup_{t\in T}\{X_t - X_0\}\right] \underset{|x|\geq x}{\leq} \mathbb{E}\left(\sup_{t\in T}|X_t - X_0|\right)$$

$\mathbb{E}[X_0]=0$

$$\Rightarrow \mathbb{E}\left[\sup_{t\in T} X_t\right] \leq \mathbb{E}\left[\sup_{t\in T}\left|X_t - X_{\pi_N(t)} + \sum_{n=1}^{N} X_{\pi_n(t)} - X_{\pi_{n-1}(t)}\right|\right]$$

Triangle inequality

$$\leq \mathbb{E}\left[\sup_{t\in T}|X_t - X_{\pi_N(t)}|\right] + \sum_{n=1}^{N}\mathbb{E}\left[\sup_{t\in T}|X_{\pi_n(t)} - X_{\pi_{n-1}(t)}|\right]$$

triangle inequality $\rightarrow$ Since $d(t_n, t_{n-1}) \leq d(t,t_n) + d(t, t_{n-1}) \leq 2^{-n} + 2^{-n+1} = 3\times 2^{-n}$

$$\mathbb{E}\left[\sup_{t\in T}|X_{\pi_n(t)} - X_{\pi_{n-1}(t)}|\right] = \mathbb{E}\left[\frac{d(\pi_n(t),\pi_{n-1}(t))}{\lambda}\cdot\frac{\lambda}{d(\pi_n(t),\pi_{n-1}(t))}|X_{\pi_n(t)} - X_{\pi_{n-1}(t)}|\right]$$

$$\leq \frac{3\cdot 2^{-n}}{\lambda}\mathbb{E}\left[\sup_{t\in T}\frac{\lambda(X_{\pi_n(t)} - X_{\pi_{n-1}(t)})}{d(\pi_n(t),\pi_{n-1}(t))}\right] = \frac{3\cdot 2^{-n}}{\lambda}\mathbb{E}\left[\sup_{t\in T}\log\exp(\cdots)\right]$$

Jensen

$$\leq \frac{3\cdot 2^{-n}}{\lambda}\left(\log\left(\mathbb{E}\left(\sup_{t\in T}\exp\left(\frac{\lambda|X_{\pi_n(t)} - X_{\pi_{n-1}(t)}|}{d(\pi_n(t),\pi_{n-1}(t))}\right)\right)\right)\right)$$

$\exp(\sup) = \sup(\exp)$

bigger space

$$\leq \frac{3\cdot 2^{-n}}{\lambda}\log\left[\mathbb{E}\left(\sup_{\substack{t_1\in T_n\\ t_2\in T_{n-1}}}\exp\left(\frac{\lambda|X_{\pi_{n-1}(t)} - X_{\pi_n(t)}|}{d(\pi_n(t),\pi_{n-1}(t))}\right)\right)\right]$$

$$\leq \frac{3\cdot 2^{-n}}{\lambda}\log\mathbb{E}\left[|T_n|\cdot|T_{n-1}|\exp\underbrace{\left(\frac{\lambda|X_{\pi_n(t)} - X_{\pi_{n-1}(t)}|}{d(\pi_n(t),\pi_{n-1}(t))}\right)}_{\psi(\lambda)}\right]$$

$|T_{n-1}| \leq |T_n|$

$$\leq \frac{3\cdot 2^{-n}}{\lambda}\log\mathbb{E}\left[|T_n|^2 e^{\psi(\lambda)}\right] = 3\cdot 2^{-n}\left[\frac{\psi(\lambda) + 2\log|T_n|}{\lambda}\right]$$

$x^2 = e^{2\log x}$

However in 2nd Part of HW2, we proved that:

$$\inf_{\lambda\in\mathbb{R}^+}\left\{\frac{\psi(\lambda)+t}{\lambda}\right\}=\psi^{*-1}(t) \xrightarrow[\forall n]{\text{Thus}} \sup_{\substack{\pi_n,\\ \pi_{n+1}\in T}}\left\{\left|X_{\pi_n(t)}-X_{\pi_{n+1}(t)}\right|\right\}\le 3\cdot 2^{-n}\cdot \psi^{*-1}(2\log|T_n|)$$

$$\Rightarrow \mathbb{E}\left[\sup X_t\right]\le \lim_{N\to\infty}\mathbb{E}\sup_{\substack{t\in T\\ \pi_N(t)\in T}}|X_t-X_{\pi_N(t)}|+\sum_{n=1}^{N}\sup_{\substack{\pi_n(t)\in T\\ \pi_{n-1}(t)}}|X_{\pi_n(t)}-X_{\pi_{n-1}(t)}|$$

$|T_n|\to$ the covering number

$$\Longrightarrow \mathbb{E}\left[\sup_{t\in T}X_t\right]\le\lim_{N\to\infty}\sum_{n=1}^{N}3\cdot 2^{-n}\psi^{*-1}(2\log|T_n|)\le 2\cdot 3\int_0^\infty\psi^{*-1}(2\log N(T,d,\varepsilon))\,d\varepsilon$$

which finally reduces to $\boxed{\mathbb{E}\left[\sup_{t\in T}X_t\right]\le 6\int_0^\infty\psi^{*-1}(2\log N(T,d,\varepsilon))\,d\varepsilon}$ Q.E.D.

(I) $S=u^TA^TAv=\sum_{i=1}^{n}u^Ta_ia_i^Tv=\sum_{i=1}^{n}\langle a_i,v\rangle\cdot\langle a_i,u\rangle=\sum_{i=1}^{n}z_i$ $\quad =2(\cdot)$

$$\mathbb{P}\left[u^TA^TAv\ge a\right]\overset{\lambda\in\mathbb{R}^+}{\le} e^{-\lambda a}\,\mathbb{E}\left[e^{-\lambda u^TA^TAv}\right]\underset{A\in\mathbb{R}^{m\times n}}{=}e^{-\lambda a}\,\mathbb{E}\left[e^{-\lambda z}\right]^m \qquad A_{ij}\overset{iid}{\sim}N(0,1)$$

$$\Rightarrow \mathbb{E}\left[e^{-\lambda z}\right]=\mathbb{E}\left[e^{-\lambda a^Tuv^Ta}\right]=\frac{1}{(2\pi)^{n/2}}\int_{\mathbb{R}^n}e^{\pm\lambda a^Tuv^Ta}\,e^{-\frac{a^TIa}{2}}\,da$$

$$=\frac{1}{(2\pi)^{n/2}}\int\exp\left(\pm\tfrac{1}{2}a^T\underbrace{(I+\lambda uv^T)}_{\Sigma^{-1}}a\right)da=|I\pm2\lambda uv^T|^{-1/2}$$

$\longrightarrow$ eigenvalues of $I\pm2\lambda uv^T\longrightarrow\begin{cases}\forall\parallel u: 1\pm2\lambda u^Tv\\ \forall\not\parallel u: 1\end{cases}\longrightarrow\det(I\pm2\lambda u^Tv)=1\pm2\lambda u^Tv$

thus $\Longrightarrow \mathbb{E}\left[e^{\pm\lambda z}\right]=(1\mp2\lambda v^Tu)^{-1/2}$ which suggests

$$\mathbb{P}\left[u^TA^TAv\ge a\right]\le\begin{cases}e^{-\lambda a}\left[(1-2\lambda u^Tv)^{-1/2}\right]^m\\ e^{+\lambda a}\left[(1+2\lambda u^Tv)\right]^{-m/2}\end{cases}$$

we can pick the better bound with respect to the given a.

(II) $$\mathbb{P}\left(\left|\tfrac{1}{m}u^TA^TAv-u^Tv\right|\ge\varepsilon u^Tv\right]=\mathbb{P}\left(\tfrac{1}{m}u^TA^TAv-u^Tv\ge\varepsilon u^Tv\right)+\mathbb{P}\left[\tfrac{1}{n}u^TA^TAv-u^Tv\le\varepsilon u^Tv\right]$$

$$=\mathbb{P}\left[u^TA^TAv\ge m(1+\varepsilon)u^Tv\right]+\mathbb{P}\left[u^TA^TAv\le m(1-\varepsilon)u^Tv\right]$$

ادامه سوال ۲:

$$\leq \underbrace{e^{-\lambda_1 m(1+\varepsilon)u^T v}(1-2\lambda_1 u^T v)^{-m/2}}_{f_1} + e^{\lambda_2 m(1-\varepsilon)u^T v}(1+2\lambda_2 u^T v)^{-m/2}$$

$$\inf_{\lambda_1} f_1(\lambda) = \inf_{\lambda \geq 0} \left\{ e^{-\lambda a}(1-2\lambda u^T v)^{-m/2} \right\} \rightarrow \frac{\partial f_1}{\partial \lambda} = -a e^{-\lambda a}(1-2\lambda u^T v)^{-m/2} + e^{-\lambda a}(1-2\lambda u^T v)^{-\frac{m}{2}-1} v^T u \overset{\circ}{=}$$

$$\Rightarrow m v^T u = a(1-2\lambda u^T v) \rightarrow \lambda^* = \frac{a - m u^T v}{2a v^T u} \rightarrow a = (1+\varepsilon)u^T v \rightarrow \lambda^* = \frac{\varepsilon}{2(1+\varepsilon)u^T v}$$

So: $$\boxed{\mathbb{P}\left[u^T A^T A v \geq m(1+\varepsilon)u^T v\right] \leq e^{-\frac{m\varepsilon}{2}}\left(1-\frac{\varepsilon}{1+\varepsilon}\right)^{-m/2} = \left((1+\varepsilon)e^{-\varepsilon}\right)^{m/2}} \quad (*)$$

likewise we can do the same thing for $\inf_{\lambda_2} f_2(\lambda_2)$.

$$\arg\min_{\lambda_2} f_2(\lambda_2) = \lambda^* = \frac{a - m u^T v}{2a v^T u} \Rightarrow a = u(1+\varepsilon)u^T v \rightarrow \lambda^* = \frac{\varepsilon}{2(1-\varepsilon)u^T v}$$

$$\mathbb{P}\left(u^T A^T A v \geq m(1+\varepsilon)u^T v\right) \leq e^{-\frac{m\varepsilon}{2}} \cdot \left(1-\frac{\varepsilon}{1+\varepsilon}\right)^{-m/2} = \left((1+\varepsilon)e^{-\varepsilon}\right)^{m/2}$$

$$\boxed{\mathbb{P}\left[u^T A^T A v \geq m(1-\varepsilon)u^T v\right] \leq e^{-\frac{m\varepsilon}{2}}\left(1-\frac{\varepsilon}{1+\varepsilon}\right)^{-\frac{m}{2}} = \left((1+\varepsilon)e^{-\varepsilon}\right)^{m/2}} \quad (**)$$

(*), (**) implies that $\mathbb{P}\left[\left|\frac{1}{m}u^T A^T A v - u^T v\right| \geq \varepsilon u^T v\right] \leq \left((1+\varepsilon)e^{-\varepsilon}\right)^{m/2} + \left((1-\varepsilon)e^{\varepsilon}\right)^{m/2}$

since $\varepsilon \geq 0 \rightarrow (1+\varepsilon)e^{-\varepsilon} \geq (1-\varepsilon)e^{+\varepsilon} \rightarrow \mathbb{P}(\ldots \geq \ldots) \leq 2\left[(1+\varepsilon)e^{-\varepsilon}\right]^{\frac{m}{2}}$

Now since $(1+\varepsilon)e^{-\varepsilon} \leq e^{-\frac{\varepsilon^2}{2}+\frac{\varepsilon^3}{3}} \rightarrow \left((1+\varepsilon)e^{-\varepsilon}\right)^{\frac{m}{2}} \leq e^{-m\left(\frac{\varepsilon^2}{4}-\frac{\varepsilon^3}{6}\right)}$

which finally suggests that

Q.E.D. $$\boxed{\mathbb{P}\left[\left|\frac{u^T A^T A v}{m} - u^T v\right| \geq \varepsilon u^T v\right] \leq 2e^{-m\left(\frac{\varepsilon^2}{4}-\frac{\varepsilon^3}{6}\right)}}$$

سؤال 3:

$T = \left\{ \frac{e_k}{\sqrt{1+\log k}}, k \in [n] \right\}$ $\quad W(T) = \mathbb{E}\left[\sup_{t\in T} \{t^T X\}\right]$; $X \sim N(0, I_n)$

$Y_t = t^T X$

$\rightarrow \mathbb{E}\left[e^{\lambda(Y_t - Y_s)}\right] = \mathbb{E}\left[e^{\lambda(t-s)^T X}\right] \leq \mathbb{E}\left[e^{\frac{\lambda^2}{2}\|t-s\|_2^2}\right] \rightarrow$ it's $\|t-s\|_2^2$ – subG $\hookrightarrow Y_t - Y_s$

So we do an ε-covering where $\varepsilon \geq \frac{1}{\sqrt{\log k}} \rightarrow k \geq e^{\frac{1}{\varepsilon^2}}$

$k \geq N(\varepsilon, T, d) \geq P(2\varepsilon, T, d) \rightarrow$ if we do ε-packing $\rightarrow k \geq e^{\frac{1}{4\varepsilon^2}}$

therefore $\rightarrow \int_0^\infty \sqrt{\log N(T, d, \varepsilon)}\, d\varepsilon \geq \int_1^\infty \sqrt{\frac{1}{4\varepsilon^2}}\, d\varepsilon = \frac{1}{2}\int_0^\infty \frac{d\varepsilon}{\varepsilon} = \infty$ which doesn't converge.

therefore $\int_0^\infty \sqrt{\log N(T, d, \varepsilon)}\, d\varepsilon$ is not a good bound!

However, it's obvious that $W(T) < \infty$, when $n \to \infty$: $W(T) < \infty$

$$W(T) = \mathbb{E}\left[\sup_{t\in T} t^T X\right] = \mathbb{E}\left[\max_{1\leq k\leq n}\left\{\frac{X_k}{\sqrt{1+\log k}}\right\}\right] \leq \int_0^\infty \mathbb{P}\left[\max_k \frac{|X_k|}{\sqrt{1+\log k}} \geq t\right] dt$$

$$\leq \int_0^T 1\, dt + \int_T^\infty \mathbb{P}\left[\max_k \frac{X_k}{\sqrt{1+\log k}} \geq t\right] dt \underset{\text{Union Bound}}{\leq} T + \int_T^\infty \sum_{k=1}^N \mathbb{P}\left[|X_k| \geq t\sqrt{1+\log k}\right] dt$$

$$\leq T + \sum_{k=1}^N \int_T^\infty 2e^{-\frac{t^2}{2}(1+\log k)}\, dt \leq T + c\sum_{k=1}^N e^{-\frac{T^2}{2}\log k} < \infty$$

Q.E.D. ■

$c = e^{-\frac{T^2}{2}}$ $\quad \left(\forall t \geq T \cdot e^{-T^2/2} \geq e^{-\frac{t^2}{2}}\right)$

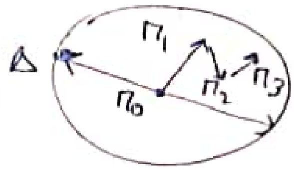
now, $T + c\sum_{k=1}^N e^{-\frac{T^2}{2}\log k}$ is Bounded and it's far better than $\int_0^\infty \sqrt{\log N(\varepsilon, T, d)}\, d\varepsilon$ which is not bounded and not tight.

$X_t - X_s \sim d(s,t)$ subg $\qquad \Delta = \operatorname{diam}(T) = \sup_{s,t} d(s,t)$ $\qquad$: 4 ...

since in the optimal case we can cover a the set with only a point and Δ-covering, like the thing that Dr. Yasser did, we define:

$n \in \{0,1,2,\dots\} \to \varepsilon_n = 2^{-n} \cdot \Delta$: T_n : $\Delta \cdot 2^{-n}$ covering and $\pi_n(t)$ is our quantizer

$$\Rightarrow |X_t - X_s| = \left| X_s - X_{\pi_N(s)} + \sum_{n=1}^{N} \left(X_{\pi_n(s)} - X_{\pi_{n-1}(s)} \right) - \sum_{n=1}^{N} \left(X_{\pi_n(t)} - X_{\pi_{n-1}(t)} \right) - \left(X_t - X_N(t) \right) \right|$$



triangle inequality $\leq \sum_{\sigma \in \{s,t\}} \left(|X_\sigma - X_{\pi_N(\sigma)}| + \sum_{n=1}^{N} |X_{\pi_n(\sigma)} - X_{\pi_{n-1}}(\sigma)| \right)$

$$\Rightarrow \sup_{t,s \in T} \{ |X_t - X_s| \} \leq \sup_{t,s \in T} \left\{ \sum_{\sigma \in \{s,t\}} |X_\sigma - X_{\pi_N(\sigma)}| + \sum_{n=1}^{N} |X_{\pi_n(\sigma)} - X_{\pi_{n-1}}(\sigma)| \right\}$$

Probabilistic chaining

$$\leq 2 \left(\sup_{t \in T} \underbrace{|X_t - X_{\pi_N(t)}|}_{\Delta \cdot 2^{-n} \sim \text{subg}} + \sum_{n=1}^{N} \sup_{t \in T} \underbrace{|X_{\pi_n(t)} - X_{\pi_{n-1}}(t)|}_{3\Delta 2^{-n} \sim \text{subg}} \right)$$

$$\mathbb{P}\left[\sup_{t \in T} |X_{\pi_n(t)} - X_{\pi_{n-1}}(t)| < 6\Delta \cdot 2^{-n} \sqrt{\log |N_n|} + 3\Delta 2^{-n} u_i \right] \geq 1 - e^{-\frac{u_i^2}{2}} \quad (\text{def})$$

$$\Rightarrow \mathbb{P}\left(\nexists n \cdot \sup_{t \in T} |X_{\pi_n(t)} - X_{\pi_{n-1}}(t)| \leq 6 \cdot \Delta 2^{-n} \sqrt{\log |N_n|} \right] \geq 1 - \mathbb{P}[\exists n \; \dots]$$

Union Bound

$$\geq 1 - \sum_{i=1}^{N} \mathbb{P}\left[\sup_{t \in T} \{\dots\} \geq 6\Delta 2^{-n} \sqrt{\log |N_n|} \right] \geq 1 - \sum_{i=1}^{n} e^{-\frac{u_i^2}{2}}$$

$(*)$

$$\geq 1 - \sum_{i=1}^{\infty} e^{-\frac{u_i^2}{2}} = 1 - 2 e^{\frac{-u^2}{2}} \frac{e^{-2}}{1 - e^{-2}} \geq 1 - \frac{1}{3} e^{-u^2/2}$$

$e^x \geq 0$

from AM-GM $\to u_i^2 = u^2 + 4i \leq (u + 2\sqrt{i})^2$

thus we've proven that with probability of at least "$1-\delta$" where $\delta = \frac{1}{3} e^{-u^2/2}$

$$\mathbb{P}\left[\sup_{t \in T} |X_s - X_t| \leq 2 \sum_{n=1}^{\infty} 6\Delta 2^{-n} \sqrt{\log |N_n|} + 3\Delta 2^{-n} u_i \right) \geq 1 - \frac{1}{3} e^{-\frac{u^2}{2}}$$

اثبات ۴۰۱:

Now since $2\sum_{n=1}^{\infty} 6\Delta 2^{-n} + 3\Delta 2^{-n} u_i = 12 \underbrace{\sum_{n=1}^{\infty} \Delta 2^{-n}\sqrt{\log N(2^{-n}\Delta)}}_{\le \left(2\int_0^{\infty} \cdots\right)\times 12} + 3\Delta 2^{-n} u_i$

therefore we get

$\Rightarrow$

$$\sup_{s,t\in T}\left\{|X_t - X_s|\right\} \le 24\int_0^{\infty}\sqrt{\log N(T,d,\varepsilon)}\,d\varepsilon + 3\Delta\sum_{n=1}^{\infty} 2^{-n}(u + 2\sqrt{n})$$

$$\le 24\int_0^{\infty}\sqrt{\log N(T,d,\varepsilon)}\,d\varepsilon + 3u\ \mathrm{diam}(T) + 6\sum_{n=1}^{\infty} 3\Delta 2^{-n}\sqrt{n}$$

$$\le C_1\int_0^{\infty}\sqrt{\log N(T,d,\varepsilon)} + C_2 u\ \mathrm{diam}(T) + C_3\ \mathrm{diam}(T)$$

In order to cover the set T with $\mathrm{diam}\ T$, we only need one element from the set T. Thus

$$N\left(T,d,\frac{\mathrm{diam}(T)}{2}\right) \ge 2 \rightarrow 2^{-n}\Delta \le c\sqrt{\log N(T,d,\Delta 2^{-n})}\ 2^{-n}\Delta$$

$$\Rightarrow \sum_{n=1}^{\infty}\Delta 2^{-n} \le c\sum_{n=1}^{\infty}\sqrt{\log N(T,d,\Delta 2^{-n})} \le c'\int_0^{\infty}\sqrt{\log N(T,d,\varepsilon)}\,d\varepsilon$$

Riemann Integral by $\lim N \to \infty$

It suffices if $\Delta 2^{-n} = o(1)$

Thus finally we've proven that

$$\exists C:\ \sup_{t,s}|X_t - X_s| \le C\left[\int_0^{\infty}\sqrt{\log N(T,d,\varepsilon)}\,d\varepsilon + u\ \mathrm{diam}(T)\right]$$

with probability of at least $1 - \frac{1}{3}e^{-\frac{u^2}{2}}$ Q.E.D.

which is better than $1 - 7e^{-u^2/2}$

Scanned by CamScanner

$$\mathbb{E}\left[e^{\lambda(X_t - X_s)}\right] \le \exp\left(\frac{\lambda^2 d(s,t)^2}{2}\right) \qquad |X_t - X_s| \le C \cdot d(s,t)$$

· Again, we consider the same approach, we take the $\delta \cdot 2^{-n}$ covering, which suggests:

$$\sup_{t\in T}\{X_t\} = \sup_{t\in T}\left\{X_t - X_{\pi_L(t)} + \sum_{i=k+1}^{L} X_{\pi_i(t)} - X_{\pi_{i-1}(t)} + X_0\right\}$$

Triangle

$$\le \sup_{t\in T}\{X_t - X_{\pi_L(t)}\} + \sum_{i=k+1}^{L} \sup_{t\in T}\{X_{\pi_i(t)} - X_{\pi_{i-1}(t)}\} + X_0 \le C \cdot \sup_{t\in T} d(t, \pi_L(t))$$

$d(t, \pi_L(t)) \le 2^{-L} = 2\delta \rightarrow \delta = 2^{-L-1}$

$$\Rightarrow \sup_{t\in T} X_t \le 2\delta C + \sum_{i=k+1}^{L} \sup_{t\in T}\{|X_{\pi_i(t)} - X_{\pi_{i-1}(t)}|\} \qquad \text{as we did before:}$$

$$\xrightarrow{\forall i} \mathbb{E}\left[\sup_{t\in T} X_{\pi_i(t)} - X_{\pi_{i-1}(t)}\right] \le \mathbb{E}\left[\sup_{t_1,t_2} \underbrace{X_{t_1} - X_{t_2}}_{\delta(\pi_i, \pi_{i-1}) \sim 3\cdot 2^{-i}\text{ subg}}\right] \le \sqrt{2(3\cdot 2^{-i})^2 \log|N_i||N_{i-1}|} \le (6\,2^{-i})\sqrt{\log|N_i|}$$

$$\Rightarrow \sum_{i=k+1}^{L} \sup_{t\in T}\{X_{\pi_i(t)} - X_{\pi_{i-1}(t)}\} \le 6\sum_{i=k+1}^{L} 2^{-i}\sqrt{\log|N_i|} \le 12\int_\delta^\infty \sqrt{\log N(\varepsilon, T, d)}\, d\varepsilon$$

$$\xrightarrow{\forall \delta} \mathbb{E}\left\{\sup_{t\in T} X_t\right\} \le 2\delta\,\mathbb{E}[C] + 12\int_\delta^\infty \sqrt{\log N(\varepsilon, T, d)}\, d\varepsilon$$

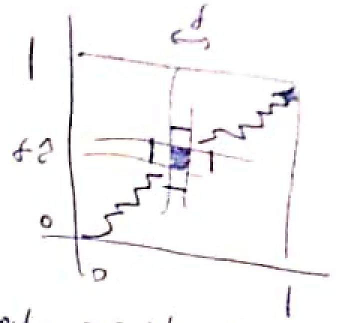
we can find the tightest bound by applying infimum over the R.H.S. of the inequality, which suggests

$$\boxed{\mathbb{E}\left[\sup_{t\in T} X_t\right] \le \inf_{\delta} \left\{2\delta\,\mathbb{E}[C] + 12\int_\delta^{+\infty} \sqrt{\log N(T, d, \varepsilon)}\, d\varepsilon\right\}}$$

Q.E.D.

I $|f(x)-f(y)| \leq \|x-y\|_\infty$ $\{X_1 - X_n\} \overset{iid}{\sim} P_X$ $: 60\frac{1}{2}$

$$W_1(P_n,P) = \sup_{f \in 1\text{-lip}} \left\{ \mathbb{E}_{P_n}[f(X)] - \mathbb{E}_P[f(X)] \right\}$$



we did as we did in the class for functions, we make a grid of width & length δ. So now we want to find the covering functions with δ-widthed grid $\longrightarrow N(\varepsilon, \cdot) \leq \left(\frac{2}{\varepsilon}\right) \cdot \prod_{i=1}^{\lceil 1/\delta \rceil - 1} 3^{2(i+1)} \cdot 5 = \left(\frac{2}{\varepsilon}\right) \cdot 5^{\frac{1}{\delta}-1} \, 3^{\sum_{h=1}^{\frac{1}{\delta}-1} 2h}$

$\leq \left(\frac{2}{\varepsilon}\right) 5^{1/\varepsilon} \, 3^{1/\varepsilon^2}$

$\log N(\varepsilon, -) \leq (\log 3)\left(\frac{1}{\varepsilon}\right)^2 + (\log 5)\left(\frac{1}{\varepsilon}\right) + \log \frac{2}{\varepsilon}$

$\leq c \cdot \left(\frac{1}{\varepsilon}\right)^2$

each square is neighboring 5 squares.

$\longrightarrow N(\varepsilon) \leq \exp\left(c\left(\frac{1}{\varepsilon}\right)^2\right)$

II

$$N(\varepsilon) \geq N_{\text{one dimensional}} \geq \prod_{i=1}^{\lceil \frac{1}{\delta} \rceil - 1} 3^{i-1}$$

$\hookrightarrow \int_0^1 \sqrt{\log N(\varepsilon)}\, d\varepsilon \geq \int_0^1 c\left(\frac{1}{\varepsilon}+1\right) d\varepsilon$ which is clearly diverging. $= \infty$

Q.E.D.

III let $Z_f = \frac{1}{n}\sum_{i=1}^n f(X_i) - \mathbb{E}f(X_i)$

here we do the exact same thing as we did in lectures!

$\longrightarrow \mathbb{E}[W_1(P_n,P)] = \mathbb{E}\left[\sup_{f \in F} Z_f\right]$

$\Longrightarrow |Z_f - Z_g| = \left|\frac{1}{n}\sum_{i=1}^n (f(X_i) - g(X_i)) - \mathbb{E}_P[f(X)] + \mathbb{E}_P[g(X)]\right|$

$\leq \mathbb{E}_P\left[|f(X) - g(X)|\right] + \frac{1}{n}\sum_{i=1}^n |f(X_i) - g(X_i)| \leq \overbrace{2\|f-g\|_\infty}^{d(f,g)}$

$\Rightarrow \mathbb{E}[W_1(P_n,P)] = \mathbb{E}\left[\sup_{f \in F} Z_f\right] \leq \inf_\delta \left\{ 4\delta\sqrt{n} + 12\int_\delta^\infty \sqrt{\log N\left(F, \frac{\|\ \|_\infty}{\sqrt{n}}, \varepsilon\right)}\, d\varepsilon \right\}$

$\leq \inf_{\delta \geq 0} \left\{ 4\delta\sqrt{n} + \frac{12}{\sqrt{n}} \int_\delta^2 \frac{c}{\varepsilon}\, d\varepsilon \right\} \leq c' \inf_{\delta \geq 0} \left\{ \delta\sqrt{n} + \frac{1}{\sqrt{n}} \int_\delta^\infty \frac{d\varepsilon}{\varepsilon} \right\}$

$= \frac{c'}{\sqrt{n}}\left(1 + \log(2\sqrt{n})\right) \leq \boxed{c'' \frac{\log n}{\sqrt{n}}}$ Q.E.D.

Scanned by CamScanner

7)/i

$A = \{X_1 - , X_m\}$ $X_i \sim N(0, I_n)$ $m \geq \exp(cn)$, by definition:

$CH(A) = \{y \in \mathbb{R}^n, \exists a_1 - a_m \in [0,1] : \sum a_i = 1, y = \sum a_i X_i\}$

$$f(u) = \sup_{x \in CH(A)} u^T X = \max_{x \in A} u^T X = \max_{1 \leq i \leq n} u^T X_i = Y_u$$ which is a process with index u

let $r_+ = \sup_{\|u\|=1} \{Y_u\}$, $r_- = \inf_{\|u\|=1} \{Y_u\}$, now consider an ε-covering on the surface of B_1 $\rightarrow$ $r_+ \leq \max_{1 \leq i \leq n} u^{*T} X_i \leq \max_{1 \leq i \leq n} u^{*T} X_i + \varepsilon |X_i|$

$$r_+ \leq \max_{1 \leq i \leq n} \{u_\varepsilon^{*T} X_i\} + \varepsilon \cdot \max_{1 \leq i \leq n} |X_i| \longrightarrow r_+ \leq \frac{1}{1-\varepsilon} \max u_\varepsilon^T X_i$$

$$\Rightarrow \mathbb{P}[r_+ \geq t] \leq \mathbb{P}\left[\frac{\max_{u \in N_\varepsilon} \max_{1 \leq i \leq n} \{u^T X_i\}}{1-\varepsilon} \geq t\right] = \mathbb{P}\left[\max_{u,i} u^T X_i \geq (1-\varepsilon)t\right]$$

$$\leq m |N_\varepsilon| e^{-\frac{t^2(1-\varepsilon)^2}{2}} \leq m \left(\frac{3}{\varepsilon}\right)^n \cdot e^{-\frac{t^2(1-\varepsilon)^2}{2}}$$

$$\Rightarrow \mathbb{P}[r_+ \geq t] \leq \exp\left(\log m + n \log \frac{3}{\varepsilon} - \frac{1}{2}(1-\varepsilon)^2 t^2\right)$$ $\Big|$ $m \geq e^{cn}$, $n \leq \frac{1}{c} \log m$

$$\mathbb{P}[r_+ \geq t]\Big|_{t = \frac{\log m}{1-\varepsilon}} \leq \exp\left(\log m + n \log \frac{3}{\varepsilon} - \frac{1}{2}\frac{\cancel{(1-\varepsilon)^2}}{\cancel{(1-\varepsilon)^2}}(\log m)^2\right)$$

let $\varepsilon = 3e^{-\alpha c}$

$$\Rightarrow \mathbb{P}[r_+ \geq \log m] \leq \exp\left(\left(1 + \frac{1}{c}\log \frac{3}{\varepsilon}\right)\log m - \frac{1}{2}\log^2(m)\right) = \exp\left(\log m\left(1 + \alpha - \frac{1}{2}\log m\right)\right)$$

$\longrightarrow$ $\log m \gg 2(1+\alpha)$ as for large enough m ~~we get~~ will suffice.

Q.E.D

1. ~~[...]~~

$$\mathbb{P}\left[\|A\|_{op} \geq c(\sqrt{p}+\sqrt{q}+t)\right] \leq e^{-t^2}$$

$$\Rightarrow \mathbb{E}\left[\underbrace{\|A\|_{op}}_{\geq 0}\right] = \int_0^\infty \mathbb{P}\left[\|A\|_{op} \geq t\right]dt = \int_0^{c\sigma(\sqrt{p}+\sqrt{q})} \mathbb{P}\left[\|A\|_{op} \geq t\right]dt + \int_{c\sigma(\sqrt{p}+\sqrt{q})}^{\infty} \mathbb{P}\left[\|A\|_{op} \geq t\right]dt$$

$$\leq \int_0^{c\sigma(\sqrt{p}+\sqrt{q})} 1\,dt + \int_0^\infty \mathbb{P}\left[\|A\|_{op} \geq c\sigma(\sqrt{p}+\sqrt{q}+t)\right] \cdot c\sigma\,dt$$ shifting

$$\leq c\sigma(\sqrt{p}+\sqrt{q}) + c\sigma\frac{\sqrt{\pi}}{2} \leq c'\sigma(\sqrt{p}+\sqrt{q}) \rightarrow \boxed{\mathbb{E}(\|A\|_{op}) \lesssim \sigma\sqrt{p}+\sigma\sqrt{q}}$$

Now we formulate $\mathbb{E}\|A\|_{op}$ by $\mathbb{E}\sup$ — by:

$$\Rightarrow \|A\|_{op} = \sup_{(u,v)\in B_m\times B_n} \{u^TAv\} = \sup_{(u,v)\in B_m\times B_n} \{\sum u_iv_jA_{ij}\}$$

$$\Rightarrow \mathbb{E}\left[e^{\lambda u^TAv}\right] = \mathbb{E}\left[e^{\sum u_iv_jA_{ij}}\right] = \prod_{ij} e^{\lambda u_iv_jA_{ij}} \leq e^{\frac{\lambda^2\sigma^2}{2}u_i^2v_j^2} = \exp\left(\frac{\lambda^2\sigma^2}{2}\sum_{ij}u_i^2v_j^2\right)$$

$$\leq \exp\left(\frac{\lambda^2\sigma^2}{2}\|u\|_2^2\|v\|_2^2\right) = e^{\frac{\lambda^2\sigma^2}{2}}$$

$\forall (u,v)\in B_m\times B_n \longrightarrow u^TAv \sim \sigma^2$ subg

In the class we get:

$$\mathbb{P}\left[\sup_{u,v\in B_m\times B_n}\{u^TAv\} \geq t\right] \leq \mathbb{P}\left[\max_{u,v\in B_m\times B_n}\{u^TAv\} \geq (1-2\varepsilon)t\right] \leq |N_1||N_2| \cdot e^{\frac{-t^2}{2\sigma^2}(1-2\varepsilon)^2}$$

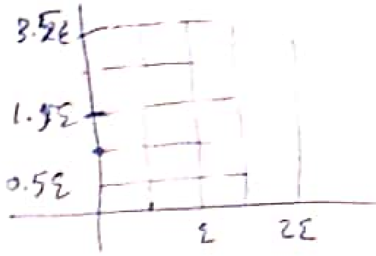
We know that Balls in $\mathbb{R}^d$ can be covered by $(\frac{3}{\varepsilon})^d$. thus

$|N_1| \leq (\frac{3}{\varepsilon})^n$, $|N_2| \leq (\frac{3}{\varepsilon})^m$ which leads to:

$$\mathbb{P}\left[\|A\|_{op} \geq t\right] \leq \exp\left[\underbrace{-\frac{t^2(1-2\varepsilon)^2}{2\sigma^2} + \log(\tfrac{3}{\varepsilon})(m+n)}_{\leq -u^2}\right]$$

$$\Rightarrow \boxed{\mathbb{P}\left[\|A\|_{op} \geq c\sigma(\sqrt{m}+\sqrt{n}+u)\right] \leq e^{-u^2}}$$ Q.E.D. ∎

we need to introduce a covering. if we take a 2-dim grid :20/2
with $\frac{\varepsilon}{2}$ distances. Now if we're given a function, we need to give a function that covers the given function



the number of functions:

$$\rightarrow \left(\frac{1}{(\varepsilon/2)} \right)^{(1/(\varepsilon/2))} = \left(\frac{2}{\varepsilon} \right)^{\frac{2}{\varepsilon}}$$

this is a good ε-coverage. thus $N(\varepsilon, \mathcal{F}, \|\cdot\|_\infty) \le \left(\frac{2}{\varepsilon}\right)^{\frac{2}{\varepsilon}}$.

Since in each step we have three choices, we get

$$N(\mathcal{F}, \|\cdot\|_\infty, \varepsilon) = \frac{1}{\varepsilon} \times 3^{\lceil \frac{2}{\varepsilon} \rceil}$$

since the function we introduced aren't 1-lipschitz, these are exterior coverage. so due to $N(\mathcal{F}, \|\cdot\|_\infty, \varepsilon) \le N^{ext}(\mathcal{F}, \|\cdot\|_\infty, \varepsilon/2)$. Q.E.D. ∎