# Stat450-1 2024Fall
## HW2

Tuorui Peng[1]

### Exercise 1

1.(a)

$\Rightarrow$ We prove by contradiction. Suppose $T$ is unbounded, we can find a sequence $t_n \in T$ such that $\|t_n\| \to \infty$. Then we consider the set $T \cup (-T) = \{t\}_{t\in T} \cup \{-t\}_{t\in T}$, using the symmetry of $g \sim \mathcal{N}(0, I)$, we have

$$w(T) = w(-T) \geq \frac{1}{2} w(T \cup (-T))$$

and we have $T_n := \{t_n, -t_n\} \subseteq T \cup (-T)$. For such $T_n$ we notice that

$$\begin{aligned} w(T_n) =& \mathbb{E}\left[\sup_{t\in\{t_n,-t_n\}} \langle t, g\rangle\right] \\ =& \mathbb{E}\left[\left|\mathcal{N}(0, \|t_n\|^2)\right|\right] \\ \geq& \|t_n\| \, \mathbb{E}\left[|\mathcal{N}(0,1)|\right] \\ =& \|t_n\| \sqrt{2/\pi} \end{aligned}$$

so using $T_n \subset T \cup (-T)$

$$w(T) \geq \frac{1}{2} w(T \cup (-T)) \geq \frac{1}{2} w(T_n) \geq \frac{1}{2} \|t_n\| \sqrt{2/\pi} \to \infty$$

thus if $T$ is unbounded, $w(T)$ is unbounded. Here we have the contradiction so if $w(T)$ is finite, $T$ is bounded.

$\Leftarrow$ If $T$ is bounded, say $\|t\| \leq \tau$, $\forall t \in T$, we have

$$\begin{aligned} w(T) =& \mathbb{E}\left[\sup_{t\in T} \langle t, g\rangle\right] \\ \leq& \mathbb{E}\left[\sup_{t\in T} \|t\| \cdot \|g\|\right] \\ \leq& \mathbb{E}\left[\tau \|g\|\right] \\ \leq& \tau \mathbb{E}\left[\sqrt{\chi_n^2}\right] \\ \leq& \tau \sqrt{\mathbb{E}\left[\chi_n^2\right]} \\ =& \tau\sqrt{n} < \infty \end{aligned}$$

[1] TuoruiPeng2028@u.northwestern.edu

1.(b)

Using the unitarity of $g \sim Ug \sim \mathcal{N}(0, I)$, we have

$$\begin{aligned}
w(UT + y) =& \mathbb{E}\left[\sup_{t\in T} \langle Ut + y, g\rangle\right] \\
=& \mathbb{E}\left[\sup_{t\in T} \left\langle U'(Ut + y), g\right\rangle\right] \\
=& \mathbb{E}\left[\sup_{t\in T} \langle t + Uy, g\rangle\right] \\
=& \mathbb{E}\left[\sup_{t\in T} \langle t, g\rangle + \langle Uy, g\rangle\right] \\
=& \mathbb{E}\left[\sup_{t\in T} \langle t, g\rangle\right] + \mathbb{E}\left[\langle Uy, g\rangle\right] \\
=& w(T) + \langle Uy, 0\rangle \\
=& w(T)
\end{aligned}$$

1.(c)

For any $t_1, t_2 \in T$ and $\alpha \in \mathbb{R}$, we have $t_\alpha := \alpha t_1 + (1-\alpha)t_2 \in \text{conv}(T)$. And we have for any $g \in \mathbb{R}$

$$\langle t_\alpha, g\rangle = \alpha \langle t_1, g\rangle + (1-\alpha)\langle t_2, g\rangle \leq \sup_{t\in\{t_1,t_2\}} \langle t, g\rangle$$

the above statement holds for any $t_1, t_2 \in T$, thus prove the statement (because for any $\tilde{t} \in \text{conv}(T)$ we can always find such $t_1, t_2 \in T$ and $\alpha$ that $\tilde{t} = \alpha t_1 + (1-\alpha)t_2$).

1.(d)

- We have

$$\begin{aligned}
w(T + S) =& \mathbb{E}\left[\sup_{r\in T+S} \langle r, g\rangle\right] \\
=& \mathbb{E}\left[\sup_{t\in T, s\in S} \langle t + s, g\rangle\right] \\
=& \mathbb{E}\left[\sup_{t\in T} \langle t, g\rangle + \sup_{s\in S} \langle s, g\rangle\right] \\
=& w(T) + w(S)
\end{aligned}$$

- We have

$$
\begin{aligned}
w(aT) =& \mathbb{E}\left[\sup_{r\in aT}\langle r, g\rangle\right]\\
=& \mathbb{E}\left[\sup_{t\in T}\langle at, g\rangle\right]\\
=& \begin{cases} a\mathbb{E}\left[\sup_{t\in T}\langle t, g\rangle\right] & a\geq 0\\ -a\mathbb{E}\left[\sup_{t\in T}\langle t, g\rangle\right] & a<0\end{cases}\\
=& \left|a\right| w(T)
\end{aligned}
$$

1.(e)

$w(T) = \dfrac{1}{2}w(T) + \dfrac{1}{2}w(-T) = \dfrac{1}{2}w(T-T)$ is trivial using the previous results. Thus we have

$$
w(T) = \frac{1}{2}w(T-T) = \frac{1}{2}\mathbb{E}\left[\sup_{x,y\in T}\langle x-y, g\rangle\right]
$$

1.(f)

It suffices the consider a centered (in the sense that $\|t\|\leq b = \mathrm{diam}(T)/2$) set $T$ due to previous arguments that $w(T)$ is invariant under unitary and translation. We have

*(i)* Already proved in (a). Assume $\tilde{t}$ being the 'boundary' point of $T$, i.e. $\left\|\tilde{t}\right\| = b$, we have

$$
w(T) \geq \frac{1}{2}w(T\cup(-T)) \geq \frac{1}{2}w(\{\tilde{t}, -\tilde{t}\}) \geq \frac{1}{2}\left\|\tilde{t}\right\|\sqrt{2/\pi} = b\sqrt{2/\pi} = \frac{\mathrm{diam}(T)}{\sqrt{2\pi}}
$$

*(ii)* Already proved in (b).

$$
\begin{aligned}
w(T) =& \mathbb{E}\left[\sup_{t\in T}\langle t, g\rangle\right]\\
\leq& \mathbb{E}\left[\sup_{t\in T}\|t\|\cdot\|g\|\right]\\
\leq& \mathbb{E}\left[b\left\|g\right\|\right]\\
\leq& b\mathbb{E}\left[\sqrt{\chi^2_n}\right]\\
\leq& b\sqrt{\mathbb{E}\left[\chi^2_n\right]}\\
=& \frac{\sqrt{n}}{2}\mathrm{diam}(T)
\end{aligned}
$$

Together we have

$$
\frac{\mathrm{diam}(T)}{\sqrt{2\pi}} \leq w(T) \leq \frac{\sqrt{n}}{2}\mathrm{diam}(T)
$$

## Exercise 2

Proof for

$$\sqrt{\frac{2}{\pi}\sum_{i=1}^n a_i^2} \overset{(i)}{\leq} \sqrt{\frac{2}{\pi}}\mathcal{R}(T) \overset{(ii)}{\leq} \mathcal{G}(T) \overset{(iii)}{\leq} \sqrt{\sum_{i=1}^n a_i^2}$$

*(i)* Note that $T$ is symmetric w.r.t. the origin, so the sup is always achieved at $\operatorname{sgn}(t_i) = \varepsilon_i$:

$$\begin{aligned}
\mathcal{R}(T) =& \mathbb{E}\left[\sup_{t:\sum t_i^2/a_i^2\leq 1} \sum_{i=1}^n \varepsilon_i t_i\right] \\
=& \sup_{t:\sum t_i^2/a_i^2\leq 1} \|t\|_1 \\
=& \sup_{\tilde{t}\in\mathbb{S}^{n-1}} \sum_{i=1}^n |a_i\tilde{t}_i|
\end{aligned}$$

The supermum is achieved at $\tilde{t} \propto (a_1, \cdots, a_n)$, so we have $\mathcal{R}(T) = \sqrt{\sum_{i=1}^n a_i^2}$.

*(ii)* We have for any $T$:

$$\begin{aligned}
\mathcal{G}(T) =& \mathbb{E}_g\left[\sup_{t\in T}\sum_{i=1}^n g_i t_i\right] \\
=& \mathbb{E}_{g,\varepsilon}\left[\sup_{t\in T}\sum_{i=1}^n \varepsilon_i |g_i| t_i\right] \\
\geq& \mathbb{E}_{\varepsilon}\left[\sup_{t\in T}\sum_{i=1}^n \varepsilon_i \mathbb{E}_g[|g_i|] t_i\right] \\
=& \sqrt{\frac{2}{\pi}}\mathcal{R}(T)
\end{aligned}$$

*(iii)* We have

$$\begin{aligned}
\mathcal{G}(T) =& \mathbb{E}_g\left[\sup_{t:\sum t_i^2/a_i^2\leq 1}\sum_{i=1}^n g_i t_i\right] \\
=& \mathbb{E}_g\left[\sup_{\tilde{t}\in\mathbb{S}^{n-1}}\sum_{i=1}^n a_i g_i \tilde{t}_i\right]
\end{aligned}$$

similarly, the supermum is achieved at $\tilde{t} \propto (a_1 g_i, \cdots, a_n g_n)$, then

$$\begin{aligned}
\mathcal{G}(T) =& \mathbb{E}_g\left[\sqrt{\sum_{i=1}^n a_i^2 g_i^2}\right] \\
\leq& \sqrt{\mathbb{E}_g\left[\sum_{i=1}^n a_i^2 g_i^2\right]} \\
=& \sqrt{\sum_{i=1}^n a_i^2}
\end{aligned}$$

## Exercise 3

1. Construct the covering set sequence $\{\mathcal{N}(T,\varepsilon_i)\}_{i=k}^{K}$, with $\varepsilon_i$ and $k, K$ chosen as follows:

$$\begin{aligned}\varepsilon_i =& 2^{-i}\\ k :& 2^{-k} \leq \operatorname{diam}(T) \leq 2^{-k+1}\\ K :& 2^{-K-1} \leq \frac{\kappa w(T)}{\sqrt{n}} \leq 2^{-K}\end{aligned}$$

for some pre-determined small $\kappa$. i.e. we have $\operatorname{diam}(T) \sim \varepsilon_k \overset{\varepsilon_i=2^{-i}}{\searrow} \varepsilon_K \sim \dfrac{\kappa w(T)}{\sqrt{n}}$.

Using the covering sequence, for each given $t \in T$, we can define maps $\pi_i(t)$ as :

$$\pi_i(t) = t_i \in \mathcal{N}(T,\varepsilon_i),\ s.t.\ \pi_i(t) \in \mathcal{N}(t,\varepsilon_{i-1})$$

Then we have

$$\begin{aligned}w(T) =& \mathbb{E}\left[\sup_{t\in T} \langle t, g\rangle\right]\\ \leq& \sum_{i=k}^{K} \mathbb{E}\left[\sup_{t\in T}\langle \pi_i(t)-\pi_{i-1}(t), g\rangle\right] + \mathbb{E}\left[\sup_{t\in T}\langle t-\pi_K(t), g\rangle\right]\end{aligned}$$

2. For $t-\pi_K(t)$ term, since $K$ satisfy a $\leq \dfrac{w(T)}{\sqrt{n}}$ covering, we have

$$\mathbb{E}\left[\sup_{t\in T}\langle t-\pi_K(t), g\rangle\right] \leq \|t-\pi_K(t)\|\,\mathbb{E}\left[\|g\|\right] \leq \varepsilon_K\sqrt{n} \leq \kappa w(T)$$

3. For $\pi_i(t)-\pi_{i-1}(t)$, which is $\varepsilon_{i-1}$ bounded and there are at most $N_2(T,\varepsilon_{i-1})$ possible choices of $\pi_i(t)-\pi_{i-1}(t)$, we can use maximal inequality for sub-Gaussian random variables to get

$$\mathbb{E}\left[\sup_{t\in T}\langle \pi_i(t)-\pi_{i-1}(t), g\rangle\right] \lesssim \varepsilon_{i-1}\sqrt{N_2(T,\varepsilon_{i-1})} \leq s(T)$$

and there are $\sim (K-k)$ terms in the sum, so we have

$$\begin{aligned}w(T) \leq& C(K-k)s(T) + \mathbb{E}\left[\sup_{t\in T}\langle t-\pi_k(t), g\rangle\right]\\ \leq& C(K-k)s(T) + \kappa w(T)\\ \Rightarrow& w(T) \leq \frac{C(K-k)}{1-\kappa}s(T) \lesssim ks(T)\end{aligned}$$

4. Now we analyze the scale of $K-k$: Note that we chose $2^{-k} \sim \operatorname{diam}(T)$ and $2^{-K} \sim \dfrac{\kappa w(T)}{\sqrt{n}}$. Thus

$$K-k = -\log_2\frac{\kappa w(T)}{\sqrt{n}\operatorname{diam}(T)} \leq \log_2\frac{\sqrt{2\pi n}}{\kappa} \lesssim \log_2 n$$

To summarize, we have

$$w(T) \lesssim s(T)\log(n)$$

## Exercise 4 MJW 8.3

We first show Courant–Fischer variational representation of eigenvalue, given in Exercise 8.1. We have

$$\min_{\mathbb{V}\in\mathcal{V}_{j-1}} \max_{x\in\mathbb{S}^{n-1}\cap\mathbb{V}^\perp} \langle Qx, x\rangle = \min_{\mathbb{V}\in\mathcal{V}_{j-1}} \max_{\nu} \left\langle \sum_{i=j}^n \nu_i v_i, Q\sum_{i=j}^n \nu_i v_i \right\rangle$$

in which $\{v_i\}$ is an orthonormal basis of $\mathbb{V}$. The transformation between $\{v_i\}$ and $\{q_i\}$ is denoted $V = QP$, then

$$\begin{aligned}
\min_{\mathbb{V}\in\mathcal{V}_{j-1}} \max_{\nu} \left\langle \sum_{i=j}^n \nu_i v_i, Q\sum_{i=j}^n \nu_i v_i \right\rangle &= \min_{\mathbb{V}\in\mathcal{V}_{j-1}} \max_{\nu} \left\langle \sum_{i=j}^n \nu_i \sum_{k=1}^n P_{ki} q_k, Q\sum_{i=j}^n \nu_i \sum_{k=1}^n P_{ki} q_k \right\rangle \\
&= \min_{\mathbb{V}\in\mathcal{V}_{j-1}} \max_{\nu} \sum_{i,\tilde{i}} \sum_{k,\tilde{k}} \nu_i\nu_{\tilde{i}} P_{ki} P_{\tilde{k}\tilde{i}} \langle q_k, Qq_{\tilde{k}}\rangle \\
&= \min_{\mathbb{V}\in\mathcal{V}_{j-1}} \max_{\nu} \sum_{i,\tilde{i}} \sum_{k,\tilde{k}} \nu_i\nu_{\tilde{i}} P_{ki} P_{\tilde{k}\tilde{i}} \delta_{k,\tilde{k}} \gamma_k \\
&= \min_{\mathbb{V}\in\mathcal{V}_{j-1}} \max_{\nu} \sum_{i,\tilde{i}} \sum_{k} \nu_i\nu_{\tilde{i}} P_{ki} P_{k\tilde{i}} \gamma_k \\
&= \min_{P\in SO(n)} \max_{\|\nu\|=1,\nu\in\mathbb{R}^{n-j+1}} \sum_k \gamma_k (\sum_{i=j}^n P_{ki}\nu_i)^2
\end{aligned}$$

note that we have

$$\sum_{k=1}^n (\sum_{i=j}^n P_{ki}\nu_i)^2 = \sum_{i,j} \nu_i\nu_j \sum_{k=1}^n P_{ki}P_{kj} = \sum_{i,j} \nu_i\nu_j\delta_{ij} = \|\nu\|^2 = 1$$

thus the above equation has max reached when

$$\arg\min_k \sum_{i=j}^n P_{ki}\nu_i = 1$$

and then the min is reached when the first $j-1$ rows of $P$ is of shape

$$\begin{bmatrix} \tilde{P}_{(j-1)\times(j-1)} & 0 \end{bmatrix}$$

and the extreme value is $\gamma_j$.

Using the representation we have $\forall i, j \in [n]$:

$$\gamma_i(A) = \min_{\mathbb{V}\in\mathcal{V}_{i-1}} \max_{x\in\mathbb{S}^{n-1}\cap\mathbb{V}^\perp} \langle Ax, x\rangle$$

$$\gamma_j(B) = \min_{\mathbb{V}\in\mathcal{V}_{j-1}} \max_{x\in\mathbb{S}^{n-1}\cap\mathbb{V}^\perp} \langle Bx, x\rangle$$

denote the corresponding $\mathbb{V}$ as $\mathbb{V}_A$ and $\mathbb{V}_B$. The vector being $x_A$ and $x_B$ respectively. Note that we have $\dim(\mathbb{V}_A \cup \mathbb{V}_B) \le i+j-2$, i.e. $\mathcal{V}_{i-1}\cup\mathcal{V}_{j-1} \subset \mathcal{V}_{i+j-2}$, we would have

$$\begin{aligned}
\gamma_{i+j-1}(A+B) &= \min_{\mathbb{V}\in\mathcal{V}_{i+j-2}} \max_{x\in\mathbb{S}^{n-1}\cap\mathbb{V}^\perp} \langle (A+B)x, x\rangle \\
&\le \min_{\mathbb{V}\in\mathcal{V}_{i+j-2}} \max_{x\in\mathbb{S}^{n-1}\cap\mathbb{V}^\perp} \langle Ax, x\rangle + \max_{x\in\mathbb{S}^{n-1}\cap\mathbb{V}^\perp} \langle Bx, x\rangle \\
&\le \min_{\mathbb{V}\in\mathcal{V}_{i+j-2}} \langle Ax_A, x_A\rangle + \langle Bx_B, x_B\rangle \\
&\le \gamma_i(A) + \gamma_j(B)
\end{aligned}$$

Taking $j=1,\ A=Q,\ B=R$ we have

$$\gamma_i(Q)-\gamma_i(R)\leq\gamma_1(Q-R)=|||Q-R|||_2$$

## Exercise 5 MJW 7.9

5.(a)

For any $x,y\in\mathbb{L}_0(k)$ and $\alpha\in[0,1]$ we have

$$\alpha x+(1-\alpha)y\in\mathbb{B}_2(1)$$

which is trivial because$\mathbb{B}_2(1)$ itself is convex. Then it's left to show $\alpha x+(1-\alpha)y\in\mathbb{B}_1(\sqrt{k})$. Notice that

$$\begin{aligned}\left\|\alpha x+(1-\alpha)y\right\|_1\leq&\alpha\left\|x\right\|_1+(1-\alpha)\left\|y\right\|_1\\ \leq&\alpha\sqrt{k}+(1-\alpha)\sqrt{k}=\sqrt{k}\end{aligned}$$

5.(b)

It suffices to prove that

$$\sup\{x\cdot a:a\in\mathbb{L}_1(k)\}\leq2\sup\{x\cdot a:a\in\mathrm{conv}\mathbb{L}_0(k)\}$$

We have

$$\sup\{x\cdot a:a\in\mathbb{L}_1(k)\}=\min\{\left\|x\right\|_2,\sqrt{k}\left\|x\right\|_\infty\}$$

and

$$\begin{aligned}\sup\{x\cdot a:a\in\mathrm{conv}\mathbb{L}_0(k)\}=&\sup\{x\cdot a:a\in\mathbb{L}_0(k)\}\\ =&\sqrt{\sum k\text{ largest }x_i^2}\end{aligned}$$

For simplicity we consider $\left\|x\right\|_2=1$. It suffices to show that

$$2\sqrt{\sum k\text{largest }x_i^2}\geq\min\{1,\sqrt{k}\left\|x\right\|_\infty\}$$

Not completed.