Chapter 1 Exercises

Gallant Tsao

Exercise 1

Let $x_1, x_2 \in \text{conv}(T)$, and $\lambda \in [0,1]$. Then there exists $j, k \in \mathbb{N}$ such that

$$x_1 = a_1 y_1 + \cdots + a_j y_j, a_i \geq 0, \sum_{i=1}^{j} a_i = 1,$$

$$x_2 = b_1 z_1 + \cdots + b_k z_k, b_I \geq 0, \sum_{i=1}^{k} b_i = 1.$$

Then we get

$$\lambda x_1 + (1-\lambda) x_2 = \lambda \sum_{i=1}^{j} a_i y_i + (1-\lambda) \sum_{i=1}^{k} b_i z_i.$$

From the formulation above,

$$a_i \geq 0 \implies \lambda a_i \geq 0,\ b_i \geq 0 \implies (1-\lambda) b_i \geq 0.$$

Moreover, when summing up the coefficients,

$$\lambda \sum_{i=1}^{j} a_i + (1-\lambda) \sum_{i=1}^{k} b_i = \lambda + (1-\lambda) = 1.$$

Therefore $_1 + (1-\lambda)x_2 \in \text{conv}(T)$. Here we assumed that without loss of generality, there are no shared points between x_1 and x_2. If there were to be shared points, it would not have affected our analysis because each coefficient in the convex combination will still be greater than 0, and also their sum will be 1.

Exercise 2

Let $f_1, \dots, f_m$ be convex functions, and $g : K \to \mathbb{R}$ be defined as

$$g(x) = \max_x (f_1(x), \cdots, f_m(x)).$$

Let $x, y \in K$, and let $\lambda \in [0, 1]$. Then

$$\begin{aligned}
g(\lambda x + (1-\lambda)y) &= \max(f_1(\lambda x + (1-\lambda)y), \cdots, f_m(\lambda x + (1-\lambda)y)) \\
&\leq \max(\lambda f_1(x) + (1-\lambda)f_1(y), \cdots, \lambda f_m(x) + (1-\lambda)f_m(y)) \\
&\leq \max(\lambda f_1(x), \cdots, \lambda f_m(x)) + \max((1-\lambda)f_1(y), \cdots, (1-\lambda)f_m(y)) \\
&= \lambda \max(f_1(x), \cdots, f_m(x)) + (1-\lambda)\max(f_1(y), \cdots, f_m(y)) \\
&= \lambda g(x) + 1 - \lambda)g(y).
\end{aligned}$$

Therefore g is a convex function.

Exercise 3

(a)

($\implies$) Suppose that f is convex. For the base case, when $m = 2$, by the definition of convexity, the statement is true. For the inductive hypothesis, assume that for some $m \in \mathbb{N}$,

$$f\left(\sum_{i=1}^{m} \lambda_i x_i\right) \leq \sum_{i=1}^{m} \lambda_i f(x_i), \lambda_1 \geq 0, \sum_{i=1}^{m} \lambda_i = 1.$$

With $\lambda_j \geq 0, \sum_{j=0}^{m+1} \lambda_j = 1$, without loss of generality assume that $\lambda_{m+1} < 1$ (if not we can switch to another λ that satisfies this condition).

$$\begin{aligned}
f\left(\sum_{i=1}^{m+1} \lambda_i x_i\right) &= f\left((1-\lambda_{m+1}) \sum_{j=1}^{m} \frac{\lambda_j}{1-\lambda_{j+1}} x_j + \lambda_{m+1} x_{m+1}\right) \\
&\leq (1-\lambda_{m+1}) f\left((1-\lambda_{m+1}) \sum_{j=1}^{m} \frac{\lambda_j}{1-\lambda_{j+1}} x_j\right) + \lambda_{m+1} f(x_{m+1}) \quad \text{(Base case)} \\
&\leq (1-\lambda_{m+1}) \sum_{j=1}^{m} \frac{\lambda_j}{1-\lambda_{m+1}} f(x_j) + \lambda_{m+1} f(x_{m+1}) \quad \text{(Inductive step)} \\
&= \sum_{j=1}^{m} \lambda_j f(x_j) + \lambda_{m+1} f(x_{m+1}) \\
&= \sum_{j=1}^{m+1} \lambda_j f(x_j).
\end{aligned}$$

($\impliedby$) Take $m = 2$ and we are done.

(b)

By the definition given for X,, let

$$P(X = x_i) = p_i, i = 1, \ldots, n,\ p_i \geq 0, \sum_{i=1}^{n} p_i = 1.$$

We can directly see from our construction that

$$\mathbb{E}[X] = \sum_{i=1}^{n} p_i x_i.$$

Then from part (a),

$$f(\mathbb{E}[X]) = f\left(\sum_{i=1}^{n} p_i x_i\right) \leq \sum_{i=1}^{n} p_i f(x_i) = \mathbb{E}[f(X)].$$

Exercise 4

Let $x \in \mathrm{conv}(T)$. Then for some $m \in \mathbb{N}$,

$$x = \lambda_1 z_1 + \cdots + \lambda_m z_m,\ \ \lambda_i \geq 0, \sum_{i=1}^{m} \lambda_i = 1.$$

Then by Jensen's Inequality from Exercise 3,

$$f(x) = f\left(\sum_{i=1}^{m} \lambda_i z_i\right) \leq \sum_{i=1}^{m} \lambda_i f(z_i) \leq \sup_i f(z_i).$$

Therefore we get

$$\sup_{x \in \mathrm{conv}(T)} f(x) \leq \sup_{x \in T} f(x).$$

The other side ("$\geq$") is obvious because $T \subseteq \mathrm{conv}(T)$. Therefore we get the equality.

Exercise 5

We'll proceed via proof by induction. For the base case when $n = 1$, let $x \in [-1.1]$. Then x can be written as a combination via

$$x = \frac{1+x}{2} \cdot 1 + \frac{1-x}{2} \cdot (-1).$$

For the inductive step, assume if $x \in [-1, 1]^n$, $x \in \operatorname{conv}(\{-1, 1\}^n)$. Now let's consider $x \in [-1, 1]^{n+1} = (x_1, \cdots, x_{n+1})$. For a fixed value of $x_{n+1} \in [-1, 1]$, from the induction hypothesis, $x \in \operatorname{conv}(\{-1, 1\}^n)$. Then

$$\begin{aligned} & x x_1, \cdots, x_{n+1}) \\ & = \frac{1 + x_{n+1}}{2}(x_1, \cdots, x_n, 1) + \frac{1 - x_{n+1}}{2}(x_1, \cdots, x_n, -1). \end{aligned}$$

Therefore x is a convex combination of points from a convex combination (we can acheive that via normalizing), hence $x \in \operatorname{conv}(\{-1, 1\}^{n+1})$ so $[-1, 1]^n \subseteq \operatorname{conv}(\{-1, 1\}^n)$.
For the other side of the proof, let $x \in \operatorname{conv}(\{-1, 1\}^n)$. Then $\exists m \leq 2^n$ such that

$$x = \lambda_1 z_1 + \cdot + \lambda_m z_m, z_i \in \operatorname{conv}(\{-1, 1\}^n), \lambda_i \geq 0, \sum_{i=1}^{m} \lambda_1 = 1.$$

Each entry x_1 satisfies $-1 \leq x_i \leq 1$ and equality occurs when all corresponding entries in z_i are either 1 or -1, hence $\operatorname{conv}(\{-1, 1\}^n) \subseteq [-1, 1]^n$.
Finally we conclude that $\operatorname{conv}(\{-1, 1\}^n) = [-1, 1]^n$.

Exercise 6

Let $x \in B_1^n$ so $\sum_{i=1}^n |x_i| \leq 1$. Define the following sets:

$$I_+ = \{i \in \{1, \cdots, n\} : x_i > 0\}, I_- = \{i \in \{1, \cdots, n\} : x_i < 0.$$

Without loss of generality, assume either $|I_+| > 0$ or $I_- > 0$. If both are zero, x has to be the origin, which finding a convex combination from the standard bases vectors would be very easy. Define

$$\lambda_{i+} = \begin{cases} |x_i| & \text{if } \ i \in I_+, \\ 0 & \text{if } \ i \in I_-, \\ \frac{1}{2(|I_-|+|I_+|)}\left(1 - \sum_{i \in I_- \cup I_+} |x_i|\right) & \text{otherwise} \end{cases},$$

$$\lambda_{i+} = \begin{cases} |x_i| & \text{if } \ i \in I_-, \\ 0 & \text{if } \ i \in I_+, \\ \frac{1}{2(|I_-|+|I_+|)}\left(1 - \sum_{i \in I_- \cup I_+} |x_i|\right) & \text{otherwise} \end{cases}.$$

Then,

$$x = \sum_{i=1}^n \lambda_{i+} e_i + \lambda_{i-}(-e_i), \ \lambda_{i+}, \lambda_{i-} \geq 0, \ \sum_{i=1}^n \lambda_{i+} + \lambda_{i-} = 1.$$

Hence $x \in \text{conv}(\{\pm e_1, \cdots, \pm e_n\})$ so $B_1^n \subseteq \text{conv}(\{\pm e_1, \cdots, \pm e_n\})$.
Now let $x \in \text{conv}(\{\pm e_1, \cdots, \pm e_n\})$. Then $\exists \lambda_{i+}, \lambda_{i-} \geq 0$ and summing to 1 such that

$$\begin{aligned} x &= \lambda_{1+} e_1 + \cdots + \lambda_{n+} e_n + \lambda_{1-}(-e_1) + \cdots + \lambda_{n-}(-e_n) \\ &\leq |\lambda_{1+} e_1| + \cdot + |\lambda_{n+} e_n| + |\lambda_{1-} e_1| + |\lambda_{n-} e_n| \\ &= \sum_{i=1}^n |\lambda_{i+}| + |\lambda_{i-}| \\ &= 1. \end{aligned}$$

Therefore $x \in B_1^n$ so $\text{conv}(\{\pm e_1, \cdots, \pm e_n\}) \in B_1^n$. We conclude that $B_1^n = \text{conv}(\{\pm e_1, \cdots, \pm e_n\})$.

Exercise 7

Denote E_i =event that freshman i has no friends, X =number of freshman. Then we are bounding

$$\begin{aligned}
\sum_{n=0}^{\infty} P\left(\bigcup_{i=1}^{X} E_i \middle| X = n\right) P(x = n) &= \sum_{n=0}^{\infty} P\left(\bigcup_{i=1}^{n} E_i\right) P(x = n) \\
&\leq \sum_{n=1}^{\infty} \frac{\lambda^n e^{-\lambda}}{n!} \sum_{i=1}^{n} P(E_i) \\
&= \sum_{n=1}^{\infty} \frac{\lambda^n e^{-\lambda}}{n!} \cdot n(1-p)^{n-1} \\
&= \sum_{n=1}^{\infty} \frac{\lambda^n e^{-\lambda}}{(n-1)!} (1-p)n - 1 \\
&= \lambda \sum_{n=0}^{\infty} \frac{\lambda^n e^{-\lambda}}{n!} (1-p)^n \\
&= \lambda e^{-p\lambda}.
\end{aligned}$$

From the question, since $p \geq 2 \ln \lambda / \lambda$,

$$\lambda e^{-p\lambda} \leq \lambda e^{-2 \ln \lambda} = \frac{1}{\lambda}.$$

Exercise 8

Exercise 9

Let E_i = the event that student i has no friends, and $B = \{\text{there exists a friendless student}\}$. We are bounding the probability

$$P(B) = P\left(\bigcup_{i=1}^{n} E_i\right) \leq \sum_{i=1}^{n} P(E_i) = n(1-p_n)^{n-1}$$

Now when we take the limit,

$$\lim_{n\to\infty} n(1-p_n)^{n-1} < \lim_{n\to\infty} n\left(\frac{(1+\varepsilon)\ln n}{n}\right)^{n-1} \to 0 \text{ as } n \to \infty.$$

Exercise 10

(a)

Proving this statement is equivalent of proving

$$\mathbb{E}[|X|^p]^{q/p} \leq \mathbb{E}[|X|^q].$$

The function $f(x) = x^{q/p}$ is convex because $q/p \geq 1$. Then by Jensen's Inequality,

$$\mathbb{E}[|X|^p]^{q/p} \leq \mathbb{E}[(|X|^p)^{q/p}] = \mathbb{E}[|X|^q].$$

(b)

Let $q < \infty$, and $a = \frac{p+q}{2}$. Let X be the random variable with pdf

$$f_X(x) = \frac{a}{x^{a+1}},\ x \geq 1.$$

Then we have that

$$\|X\|_{L^p} = \int_1^\infty x^p \cdot \frac{a}{x^{a+1}}\ dx = \int_1^\infty a x^{(p-q)/2-1}\ dx < \infty \quad (\frac{p-q}{2} - 1 < -1),$$

$$\|X\|_{L^q} = \int_1^\infty x^q \cdot \frac{a}{x^{a+1}}\ dx = \int_1^\infty a x^{(q-p)/2-1}\ dx = \infty \quad (\frac{q-p}{2} - 1 \geq -1),$$