OSM Lab 2017 **Name:** HUNG HO
Math Problem Set #5 **Email:** hqdhftw@uchicago.edu

7.1. Consider $\mathbf{x},\mathbf{y} \in \text{conv}(S)$. We can write

$$\mathbf{x} = \sum_{i=1}^{k} \lambda_k \mathbf{x}_i$$
$$\mathbf{y} = \sum_{i=1}^{l} \gamma_i \mathbf{y}_i.$$

Where $\sum_{i=1}^{k} \lambda_i = \sum_{i=1}^{l} \gamma_i = 1$, and $\mathbf{x}_i, \mathbf{y}_i \in S$ for all i. For any $\lambda \in (0,1$, we have

$$\lambda \mathbf{x} + (1-\lambda)\mathbf{y} = \sum_{i=1}^{k+l} \theta_i \mathbf{z}_i.$$

where $\theta_i = \lambda\lambda_i$ for $i \leq k$, $\theta_i = (1-\lambda)\gamma_{i-k}$ for $k+1 \leq i \leq k+l$, $\mathbf{z}_i = \mathbf{x}_i$ for $i \leq k$ and $\mathbf{z}_i = \mathbf{y}_{\mathbf{i-k}}$ for $k+1 \leq i \leq k+l$. Since $\mathbf{z}_i \in S$ for all i, we deduce that $\lambda\mathbf{x} + (1-\lambda)\mathbf{y} \in \text{conv}(S)$, thus $\text{conv}(S)$ is convex.

7.2.

(i) Let $H = \{\mathbf{x} \mid \langle \mathbf{a}, \mathbf{x} \rangle = \mathbf{b}\ \}$ be a hyperplane. For $\mathbf{x},\mathbf{y} \in H$ and $\lambda \in (0,1)$, we have

$$\langle a, \lambda\mathbf{x} + (1-\lambda)\mathbf{y} \rangle = \lambda \langle a, \mathbf{x} \rangle + (1-\lambda) \langle a, \mathbf{y} \rangle = \lambda \mathbf{b} + (1-\lambda)\mathbf{b} = \mathbf{b}.$$

This shows that $\lambda\mathbf{x} + (1-\lambda)\mathbf{y} \in H$ for all $\mathbf{x},\mathbf{y} \in H, \lambda \in (0,1)$, and therefore H is convex.

(ii) The proof is completely similar to that of part (i), except that we change every "$=$" into "$\leq$".

7.4.

(i) Note that since $C \subseteq \mathbb{R}^n$, we have $\langle \mathbf{a}, \mathbf{b} \rangle = \langle \mathbf{b}, \mathbf{a} \rangle$ for all $\mathbf{a},\mathbf{b}$, and thus

$$\|\mathbf{a}+\mathbf{b}\|^2 = \langle \mathbf{a}+\mathbf{b}, \mathbf{a}+\mathbf{b} \rangle = \|\mathbf{a}\|^2 + \|\mathbf{b}\|^2 + 2\langle \mathbf{a}, \mathbf{b} \rangle.$$

Choose $\mathbf{a} = \mathbf{x} - \mathbf{p}, \mathbf{b} = \mathbf{p} - \mathbf{y}$, then we have the desired result.

(ii) Using part (i), we know that

$$\|\mathbf{x}-\mathbf{y}\|^2 \geq \|\mathbf{x}-\mathbf{p}\|^2 + \|\mathbf{p}-\mathbf{y}\|^2$$

This show that if (7.14) holds, then unless $\|\mathbf{p}-\mathbf{y}\| = 0$, or $\mathbf{p} = \mathbf{y}$, we have $\|\mathbf{x}-\mathbf{y}\| > \|\mathbf{x}-\mathbf{p}\|$.

(iii) Using part (i) and the fact that $\mathbf{p} - \mathbf{z} = \lambda(\mathbf{y} - \mathbf{z})$, we have

$$\begin{aligned} \|\mathbf{x}-\mathbf{z}\|^2 &= \|\mathbf{x}-\mathbf{p}\|^2 + \|\mathbf{p}-\mathbf{z}\|^2 + 2\langle \mathbf{x}-\mathbf{p}, \mathbf{p}-\mathbf{z} \rangle \\ &= \|\mathbf{x}-\mathbf{p}\|^2 + \lambda^2 \|\mathbf{y}-\mathbf{p}\|^2 + 2\lambda \langle \mathbf{x}-\mathbf{p}, \mathbf{p}-\mathbf{y} \rangle. \end{aligned}$$

(iv) Since $\mathbf{p}$ is a projection of $\mathbf{x}$ onto C, we know that $\|\mathbf{x}-\mathbf{p}\|^2 \leq \|\mathbf{x}-\mathbf{z}\|^2$. This means that for all $\lambda \in (0,1)$, we have

$$\lambda \|\mathbf{y}-\mathbf{p}\|^2 + 2\langle \mathbf{x}-\mathbf{p}, \mathbf{p}-\mathbf{y} \rangle \geq 0.$$

If there exists $\mathbf{y}$ such that $\langle \mathbf{x}-\mathbf{p}, \mathbf{p}-\mathbf{y} \rangle < 0$, then we can choose λ small enough such that $\lambda \|\mathbf{y}-\mathbf{p}\|^2 + 2\langle \mathbf{x}-\mathbf{p}, \mathbf{p}-\mathbf{y} \rangle < 0$, a contradiction, hence $\langle \mathbf{x}-\mathbf{p}, \mathbf{p}-\mathbf{y} \rangle \geq 0$ for all $\mathbf{y}$.

7.6. Consider any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $\lambda \in [0,1]$ such that $f(\mathbf{x}), f(\mathbf{y}) \leq c$. We have

$$f(\lambda \mathbf{x} + (1-\lambda)\mathbf{y}) \leq \lambda f(\mathbf{x}) + (1-\lambda) f(\mathbf{y}) \leq \lambda c + (1-\lambda)c = c.$$

This shows that $f(\lambda \mathbf{x} + (1-\lambda)\mathbf{y}) \leq c$, and thus $\lambda \mathbf{x} + (1-\lambda)\mathbf{y} \in S = \{\mathbf{x} \in \mathbb{R}^n \mid f(\mathbf{x}) \leq c\}$, or S is convex.

7.7. Consider $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $\lambda \in [0,1]$. We have

$$\begin{aligned} f(\lambda \mathbf{x} + (1-\lambda)\mathbf{y}) &= \sum_{i=1}^{k} \lambda_i f_i(\lambda \mathbf{x} + (1-\lambda)\mathbf{y}) \\ &\leq \sum_{i=1}^{k} [\lambda \lambda_i f_i(\mathbf{x}) + (1-\lambda)\lambda_i f_i(\mathbf{y})] \\ &= \lambda \sum_{i=1}^{k} \lambda_i f_i(\mathbf{x}) + (1-\lambda) \sum_{i=1}^{k} \lambda_i f_i(\mathbf{y}) \\ &= \lambda f(\mathbf{x}) + (1-\lambda) f(\mathbf{y}). \end{aligned}$$

This shows that f is convex.

7.13. Suppose there exist $\mathbf{x}, \mathbf{y}$ suh that $f(\mathbf{x}) > f(\mathbf{y})$. Suppose further that f is bounded above by a. We have:

$$f(\mathbf{y}) < f(\mathbf{x}) \leq \lambda f\left(\frac{\mathbf{x} - (1-\lambda)\mathbf{y}}{\lambda}\right) + (1-\lambda) f(\mathbf{y}) \leq \lambda a + (1-\lambda) f(\mathbf{y}).$$

This means $f(\mathbf{y}) < \lambda a + (1-\lambda) f(\mathbf{y})$ for all $\lambda \in [0,1]$. Let $\lambda \to 0^+$, we have a contradiction, thus f is not abounded above.

7.20. Since f and $-f$ are convex, we have

$$\begin{aligned} f(\lambda \mathbf{x} + (1-\lambda)\mathbf{y}) &\leq \lambda f(\mathbf{x}) + (1-\lambda) f(\mathbf{y}) \\ -f(\lambda \mathbf{x} + (1-\lambda)\mathbf{y}) &\leq -\lambda f(\mathbf{x}) - (1-\lambda) f(\mathbf{y}). \end{aligned}$$

Therefore,

$$f(\lambda \mathbf{x} + (1-\lambda)\mathbf{y}) = \lambda f(\mathbf{x}) + (1-\lambda) f(\mathbf{y}). \quad (1)$$

for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n, \lambda \in [0,1]$. Denote $f(\mathbf{0}) = \mathbf{c}$ and $L(\mathbf{x}) = f(\mathbf{x}) - \mathbf{c}$. Choose $\mathbf{y} = 0$. We know that $L(\lambda \mathbf{x}) = \lambda L(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^n, \lambda \in [0,1]$. Again, change $\mathbf{x}$ to $\dfrac{\mathbf{x}}{\lambda}$, we deduce that $L(\mathbf{x}) = \lambda L\left(\frac{\mathbf{x}}{\lambda}\right)$. Since $\dfrac{1}{\lambda}$ can take any values on $[1, +\infty]$ for $\lambda \in [0,1]$, we deduce that $L(a\mathbf{x}) = aL(\mathbf{x})$ for all $a \in [0, +\infty]$.
Next, from substitute $L = f + \mathbf{c}$ into (1), we also have $L(\lambda \mathbf{x} + (1-\lambda)\mathbf{y}) = \lambda L(\mathbf{x}) + (1-\lambda) L(\mathbf{y})$. Now change $\mathbf{y}$ to $\frac{\lambda \mathbf{y}}{1-\lambda}$, we deduce that

$$L(\lambda \mathbf{x} + \lambda \mathbf{y}) = \lambda L(\mathbf{x}) + (1-\lambda) L(\lambda \frac{\mathbf{y}}{\mathbf{1} - \lambda}).$$

Using the fact that $L(a\mathbf{y}) = aL(\mathbf{y})$ for all $a \geq 0$, and divide both sides by λ, we deduce that

$$L(\mathbf{x} + \mathbf{y}) = L(\mathbf{x}) + L(\mathbf{y}). \quad (2)$$

Now it remains to show that $L(a\mathbf{x}) = aL(\mathbf{x})$ for all $a < 0$. From (2), let $\mathbf{y} = -\mathbf{x}$, we have $L(-\mathbf{x}) = L(\mathbf{0}) - L(\mathbf{x}) = -L(\mathbf{x})$. Hence, for $a < 0$, so that $-a > 0$, we have

$$L(a\mathbf{x}) = -L(-a\mathbf{x}) = -aL(\mathbf{x}).$$

Therefore, L is a linear transformation, thus f is affine.

7.21. $\mathbf{x}^*$ is a local minimizer of $\phi \circ f$ if and only if $\phi(f(\mathbf{x}^*)) \leq \phi(f(\mathbf{x}))$ for all $\mathbf{x}$ in a neighborhood of $\mathbf{x}^*$. Since ϕ is strictly increasing, this is also equivalent to $f(\mathbf{x}^*) \leq f(\mathbf{x})$ for all $\mathbf{x}$ in a neighborhood of $\mathbf{x}^*$, or $\mathbf{x}^*$ is a local minimizer of f (subject to the same constraints).