Math Problem Set 5
Open Source Macroeconomics Laboratory Boot Camp

Ruby Zhang

July 18, 2017

7.1 Take any two points $\mathbf{v}, \mathbf{u} \in \mathrm{conv}(S)$. Then $\mathbf{v}$ and $\mathbf{u}$ are convex combinations of elements of S, so let's write them as $\mathbf{v} = \alpha_1\mathbf{x}_1 + \cdots + \alpha_k\mathbf{x}_k, \mathbf{v} = \beta_1\mathbf{y}_1 + \cdots + \beta_n\mathbf{y}_n$ where $\mathbf{x}_i, \mathbf{y}_j \in S, k, n \in \mathbb{N}, \alpha_i, \beta_j \geq 0, \alpha_1 + \cdots + \alpha_k = \beta_1 + \cdots + \beta_n = 1$. Then for any $\lambda \in [0,1]$, we have that

$$\begin{aligned}\lambda\mathbf{v} + (1-\lambda)\mathbf{u} &= \lambda(\alpha_1\mathbf{x}_1 + \cdots + \alpha_k\mathbf{x}_k) + (1-\lambda)(\beta_1\mathbf{y}_1 + \cdots + \beta_n\mathbf{y}_n) \\ &= \lambda\alpha_1\mathbf{x}_1 + \cdots + \lambda\alpha_k\mathbf{x}_k + (1-\lambda)\beta_1\mathbf{y}_1 + \cdots + (1-\lambda)\beta_n\mathbf{y}_n)\end{aligned}$$

Note that

$$\begin{aligned}\lambda\alpha_1 + \cdots + \lambda\alpha_k + (1-\lambda)\beta_1 + \cdots + (1-\lambda)\beta_n &= \lambda(\alpha_1 + \cdots + \alpha_k) + (1-\lambda)(\beta_1 + \cdots + \beta_n) \\ &= \lambda + (1-\lambda) = 1\end{aligned}$$

By definition, the result is also a convex combination of elements of S, which means that any convec combination of $\mathbf{v}$ and $\mathbf{u}$ are in $\mathrm{conv}(S)$, thus making $\mathrm{conv}(S)$ a convex set.

7.2 (i) Take any two elements $\mathbf{x_1}, \mathbf{x_2}$ in the hyperplane $P = \{\mathbf{x} \in V | \langle \mathbf{a}, \mathbf{x} \rangle = b\}$. For all $\lambda \in [0,1]$, we have that:

$$\begin{aligned}\langle \mathbf{a}, \lambda\mathbf{x_1} + (1-\lambda)\mathbf{x_2} \rangle &= \lambda\langle \mathbf{a}, \mathbf{x-1} \rangle + (1-\lambda)\langle \mathbf{a}, \mathbf{x_2} \rangle \\ &= \lambda b + (1-\lambda)b = b\end{aligned}$$

Therefore, any convex combination of two points in P is still in hyperplane P so hyperplanes are convex.

(ii) Take any two elements $\mathbf{x_1}, \mathbf{x_2}$ in the half-space $H = \{\mathbf{x} \in V | \langle \mathbf{a}, \mathbf{x} \rangle \leq b\}$. For all $\lambda \in [0,1]$, we have that:

$$\begin{aligned}\langle \mathbf{a}, \lambda\mathbf{x_1} + (1-\lambda)\mathbf{x_2} \rangle &= \lambda\langle \mathbf{a}, \mathbf{x-1} \rangle + (1-\lambda)\langle \mathbf{a}, \mathbf{x_2} \rangle \\ &\leq \lambda b + (1-\lambda)b = b\end{aligned}$$

Therefore, any convex combination of two points in H is still in half-space H so half-spaces are convex.

7.4 (i)

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

(ii) Suppose we have that $\langle\mathbf{x}-\mathbf{p},\mathbf{p}-\mathbf{y}\rangle \le 0 \quad \forall\mathbf{y}\in C$. Since the inner product is always positive, we have that:

$$\begin{aligned}\|\mathbf{x}-\mathbf{y}\|^2 &= \|\mathbf{x}-\mathbf{p}\|^2+\|\mathbf{p}-\mathbf{y}\|^2+2\langle\mathbf{x}-\mathbf{p},\mathbf{p}-\mathbf{y}\rangle\\ &\ge \|\mathbf{x}-\mathbf{p}\|^2+\|\mathbf{p}-\mathbf{y}\|^2\\ &\ge \|\mathbf{x}-\mathbf{p}\|^2\\ \therefore \|\mathbf{x}-\mathbf{y}\| &> \|\mathbf{x}-\mathbf{p}\| \quad \forall\mathbf{y}\in C,\mathbf{y}\neq\mathbf{p}\end{aligned}$$

(iii) Suppose $\mathbf{z}=\lambda\mathbf{y}+(1-\lambda)\mathbf{p}$ for $\lambda\in[0,1]$. Then we have:

$$\begin{aligned}\|\mathbf{x}-\mathbf{z}\|^2 &= \langle\mathbf{x}-\mathbf{z},\ \mathbf{x}-\mathbf{z}\rangle\\ &=\langle\mathbf{x}-\mathbf{z},\mathbf{x}-\lambda\mathbf{y}-(1-\lambda)\mathbf{p}\rangle\\ &=\langle\mathbf{x}-\mathbf{z},\mathbf{x}-\mathbf{p}\rangle+\lambda\langle\mathbf{x}-\mathbf{z},\mathbf{p}-\mathbf{y}\rangle\\ &=\langle\mathbf{x}-\mathbf{p},\mathbf{x}-\mathbf{p}\rangle+\lambda\langle\mathbf{p}-\mathbf{y},\mathbf{x}-\mathbf{p}\rangle+\lambda\langle\mathbf{x}-\mathbf{p},\mathbf{p}-\mathbf{y}\rangle+\lambda^2\langle\mathbf{p}-\mathbf{y},\mathbf{p}-\mathbf{y}\rangle\\ &=\|\mathbf{x}-\mathbf{p}\|^2+2\lambda\langle\mathbf{x}-\mathbf{p},\mathbf{p}-\mathbf{y}\rangle+\lambda^2\|\mathbf{y}-\mathbf{p}\|^2\end{aligned}$$

(iv) If $\mathbf{p}$ is a projection of $\mathbf{x}$ onto convex set C, then by definition $\|\mathbf{x}-\mathbf{p}\|\le\|\mathbf{x}-\mathbf{y}\| \quad \forall\mathbf{y}\in C$. Since C is convex, $\mathbf{z}=\lambda\mathbf{y}+(1-\lambda)\mathbf{p}\in C \quad \forall\mathbf{y}\in C,\lambda\in[0,1]$ and we have that $\|\mathbf{x}-\mathbf{p}\|\le\|\mathbf{x}-\mathbf{z}\|$. Thus, $0\le\|\mathbf{x}-\mathbf{z}\|-\|\mathbf{x}-\mathbf{p}\|=2\lambda\langle\mathbf{x}-\mathbf{p},\mathbf{p}-\mathbf{y}\rangle+\lambda^2\|\mathbf{y}-\mathbf{p}\|^2$ from part (iii). Since $0\le\lambda$, we have that $0\le2\langle\mathbf{x}-\mathbf{p},\mathbf{p}-\mathbf{y}\rangle+\lambda\|\mathbf{y}-\mathbf{p}\|^2$.

$\Longrightarrow$

Suppose that a point $\mathbf{p}$ is a projection of $\mathbf{x}$ onto convex set C. From part (iv), we know that $0\le2\langle\mathbf{x}-\mathbf{p},\mathbf{p}-\mathbf{y}\rangle+\lambda\|\mathbf{y}-\mathbf{p}\|^2 \quad \forall\lambda\in[0,1]$. Then the statement holds true for $\lambda=0$, in which case $0\le2\langle\mathbf{x}-\mathbf{p},\mathbf{p}-\mathbf{y}\rangle\Longrightarrow0\le\langle\mathbf{x}-\mathbf{p},\mathbf{p}-\mathbf{y}\rangle$.

$\Longleftarrow$

Suppose we have that $\langle\mathbf{x}-\mathbf{p},\mathbf{p}-\mathbf{y}\rangle\le0 \quad \forall\mathbf{y}\in C$. According to part(ii), $\|\mathbf{x}-\mathbf{y}\|>\|\mathbf{x}-\mathbf{p}\| \quad \forall\mathbf{y}\in C,\mathbf{y}\neq\mathbf{p}$. By definition, $\mathbf{p}$ is the projection of $\mathbf{x}$ onto convex set C.

7.6 Take any two points $\mathbf{x_1},\mathbf{x_2}\in A=\{\mathbf{x}\in\mathbb{R}^n|f(\mathbf{x}\le c)\}$. Since f is a convec function, for any $\lambda\in[0,1]$, we have that

$$f(\lambda\mathbf{x_1}+(1-\lambda)\mathbf{x_2})\le\lambda f(\mathbf{x_1})+(1-\lambda)f(\mathbf{x_2})\le\lambda c+(1-\lambda)c=c$$

Therefore, any convex combination of two points in A is still in A, so set A is convex.

7.7 Let $f(\mathbf{x}) = \sum_{i=1}^{k} \alpha_i f_i(\mathbf{x})$ where $\alpha_i \in \mathbf{R}_+, f_i : C \to \mathbb{R}, f_i$ convex $\quad \forall 1 \leq i \leq k$ and C convex. Then $f : C \to \mathbb{R}$. For any two points $\mathbf{x_1}, \mathbf{x_2} \in C$ and $\lambda \in [0,1]$, we have that

$$\begin{aligned} f(\lambda\mathbf{x_1} + (1-\lambda)\mathbf{x_2}) &= \sum_{i=1}^{k} \alpha_i f_i(\lambda\mathbf{x_1} + (1-\lambda)\mathbf{x_2}) \\ &\leq \lambda \sum_{i=1}^{k} \alpha_i f_i(\mathbf{x_1}) + (1-\lambda) \sum_{i=1}^{k} \alpha_i f_i(\mathbf{x_2}) \\ &= \lambda f(\mathbf{x_1}) + (1-\lambda) f(\mathbf{x_2}) \\ \therefore \quad & f(\mathbf{x}) \text{ is a convex function} \end{aligned}$$

7.13 Let $f : \mathbb{R}^n \to \mathbb{R}$ be convex and bounded above by M. Suppose there exists two points $\mathbf{x_1} \neq \mathbf{x_2} \in \mathbb{R}^n$ such that $f(\mathbf{x_1}) > f(\mathbf{x_2})$. Then for all $\lambda \in [0,1]$, we have that

$$\begin{aligned} f(\mathbf{x_1}) &= f(\lambda \frac{\mathbf{x_1} - (1-\lambda)\mathbf{x_2}}{\lambda} + (1-\lambda)\mathbf{x_2}) \\ &\leq \lambda f(\frac{\mathbf{x_1} - (1-\lambda)\mathbf{x_2}}{\lambda}) + (1-\lambda) f(\mathbf{x_2}) \\ \frac{f(\mathbf{x_1}) - (1-\lambda) f(\mathbf{x_2})}{\lambda} &\leq f(\frac{\mathbf{x_1} - (1-\lambda) f(\mathbf{x_2})}{\lambda}) < M \\ \therefore \frac{f(\mathbf{x_1}) - f(\mathbf{x_2})}{\lambda} &+ f(\mathbf{x_2}) < M \end{aligned}$$

However, the last statement is a contradiction since as $\lambda \to 0$, the expression approaches infinity, contradicting the fact that f is bounded above.

7.20 Note that if $f(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}$ is affine, then we can express it as $f(\mathbf{x}) = L(\mathbf{x}) + c$ so for any $\mathbf{x_1}, \mathbf{x_2} \in \mathbb{R}^n$ and $a, b \in \mathbb{R}$ we have that

$$\begin{aligned} f(a\mathbf{x_1} + b\mathbf{x_2}) &= L(a\mathbf{x_1} + b\mathbf{x_2}) + c && (1) \\ &= aL(\mathbf{x_1}) + bL(\mathbf{x_2}) + c && (2) \\ \therefore f(a\mathbf{x_1} + b\mathbf{x_2}) &= af(\mathbf{x_1}) + bf(\mathbf{x_2}) + (1-a-b)c && (3) \end{aligned}$$

Now let us suppose that $f, -f$ are convex and f is NOT affine. Then there exist $\mathbf{x_1}, \mathbf{x_2} \in \mathbb{R}^n$ such that equation (3) does not hold for all $a, b \in \mathbb{R}$. Since f is convex, for any $\lambda \in [0,1]$ we have that:

$$f(\lambda\mathbf{x_1} + (1-\lambda)\mathbf{x_2}) \leq \lambda f(\mathbf{x_1}) + (1-\lambda) f(\mathbf{x_2})$$
$$f(\lambda\mathbf{x_1} + (1-\lambda)\mathbf{x_2}) \geq -\lambda f(\mathbf{x_1}) - (1-\lambda) f(\mathbf{x_2})$$

Similarly, since $-f$ is convex, we also have that:

$$-f(\lambda\mathbf{x_1} + (1-\lambda)\mathbf{x_2}) \leq -\lambda f(\mathbf{x_1}) - (1-\lambda) f(\mathbf{x_2})$$

Therefore, $f(\lambda\mathbf{x_1} + (1-\lambda)\mathbf{x_2}) = \lambda f(\mathbf{x_1}) + (1-\lambda) f(\mathbf{x_2})$. This contradicts the fact that $\mathbf{x_1}, \mathbf{x_2}$ does not satisfy equation (3) for all scalars in $\mathbb{R}$. Therefore, f is affine.

7.21 $\Longrightarrow$

Suppose $\mathbf{x}^*$ is a local minimizer for the problem with objective function f. Since ϕ is strictly increasing, then for all $\mathbf{x} \neq \mathbf{x}^*$ that satisfy the constraints, $f(\mathbf{x}^*) \leq f(\mathbf{x}) \Longrightarrow \phi \circ f(\mathbf{x}^*) \leq \phi \circ f(\mathbf{x})$ so by definition, $\mathbf{x}^*$ is also a minimizer for the objective function $\phi \circ f$.

$\Longleftarrow$

Suppose $\mathbf{x}^*$ is a local minimizer for the problem with objective function $\phi \circ f$ but not a local minimzer for the problem with objective function f. Then there exists $\mathbf{x_0}$ in the neighborhood of $\mathbf{x}^*$ that satisfies the constraints and $f(\mathbf{x_0}) \leq f(\mathbf{x}^*) \Longrightarrow \phi \circ f(\mathbf{x_0}) \leq \phi \circ f(\mathbf{x}^*)$, which contradicts the fact that $\mathbf{x}^*$ is a local minimizer for $\phi \circ f$. Thus, $\mathbf{x}^*$ is also a local minimizer for f.