Foundations of Inventory Management **Fall 2022, NUS**

Lecture 1: Convex Analysis and Lattice Theory

Lecturer: Zhenyu Hu

1.1 Basic Convex Analysis

A set $\mathcal{X} \subseteq \mathbb{R}^n$ is said to be convex if for any $0 \leq \lambda \leq 1$, and $x, y \in \mathcal{X}$, one has $\lambda x + (1-\lambda)y \in \mathcal{X}$. Convexity is preserved under intersection: for an arbitrary set $\mathcal{A}$, if $\mathcal{X}_\alpha$ is convex for every $\alpha \in \mathcal{A}$, then $\cap_{\alpha \in \mathcal{A}} \mathcal{X}_\alpha$ is also convex.

Example 1.1 (Polyhedron) *A polyhedron described in the form*

$$\{x \in \mathbb{R}^n | Ax \geq b\},$$

where A is an $m \times n$ matrix and $b \in \mathbb{R}^m$ is convex.

Example 1.2 (Positive semidefinite cone) *Let $\mathbb{S}^n$ denote the set of all symmetric $n \times n$ matrices:*

$$\mathbb{S}^n = \{X \in \mathbb{R}^{n \times n} \mid X = X'\}.$$

The positive semidefinite cone $\mathbb{S}^n_+$ is defined as

$$\mathbb{S}^n_+ = \{X \in \mathbb{S}^n \mid X \succeq 0\}.$$

Note that for any $\lambda \in [0,1]$ and $A, B \in \mathbb{S}^n_+$, we have $\lambda A + (1-\lambda)B \in \mathbb{S}^n_+$, and hence $\mathbb{S}^n_+$ is a convex set. As an example, consider

$$\left\{ X \in \mathbb{S}^2 \middle| X = \begin{bmatrix} x & y \\ y & z \end{bmatrix} \succeq 0 \right\},$$

which is the same as

$$\{(x, y, z) \in \mathbb{R}^3 \mid x, z \geq 0, xz - y^2 \geq 0\}.$$

Alternatively, we can express $\mathbb{S}^n_+$ as

$$\mathbb{S}^n_+ = \bigcap_{z \in \mathbb{R}^n} \{X \in \mathbb{S}^n \mid z'Xz \geq 0\},$$

which is the intersection of infinite number of halfspaces and hence convex.

A function $f : \mathbb{R}^n \to \mathbb{R}$ is convex if its domain $\mathrm{dom} f$ is convex and for all $x, y \in \mathrm{dom} f$ and $\lambda \in [0,1]$, we have

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y).$$

We say f is strictly convex if strict inequality holds for all $x \neq y$ and $\lambda \in (0,1)$.

Below are several equivalent characterizations:

First-Order: A differentiable function f is convex if and only if for any $x, y \in \mathrm{dom} f$

$$f(x) + \nabla f(x)'(y - x) \leq f(y).$$

Second-Order: A twice- differentiable function f is convex if and only if for any $x \in \text{dom} f$

$$\nabla^2 f(x) \succeq 0.$$

The following proposition summarizes some commonly used operations that preserve convexity of a function.

Proposition 1.3 **(a)** *If $f_i : \mathbb{R}^n \to \mathbb{R}$ are convex for all $i = 1, ..., m$, then for any $\alpha_i \geq 0$, the function $\sum_{i=1}^m \alpha_i f_i$ is also convex;*

(b) *If $f : \mathbb{R}^m \to \mathbb{R}$ is convex, then for any $m \times n$ matrix A and $b \in \mathbb{R}^m$, the function $g(x) = f(Ax + b)$ is also convex;*

(c) *If $f : \mathbb{R} \to \mathbb{R}$ is convex and nondecreasing and $g : \mathbb{R}^n \to \mathbb{R}$ is convex, then $f(g(x))$ is also convex;*

(d) *Suppose $f(\cdot, \cdot)$ is defined on $\mathbb{R}^n \times \mathbb{R}^m$. If $f(\cdot, y)$ is convex for any given $y \in \mathbb{R}^m$, then for a random vector ξ in $\mathbb{R}^m$, the function defined by $g(x) = \mathbb{E}[f(x, \xi)]$ is convex on $\mathbb{R}^n$.*

Example 1.4 *Are the following functions convex/concave?*

1. $\sum_{i=1}^n x_i^2 - \sqrt{\sum_{i=1}^n x_i}$.
2. *Let $\mathcal{N} = \{1, ..., n\}$. Consider the function $\sum_{\mathcal{S} \subseteq \mathcal{N}} \sqrt{\sum_{i \in \mathcal{S}} x_i}$.*
3. $e^{\sum_{i=1}^n x_i^2}$.
4. $\mathbb{E}[(D - \sum_{i=1}^n \xi_i x_i)^+]$, *where D and $\xi_i, i = 1, ..., n$ are random variables.*
5. $x \cdot y$.

The following well-known result points to the central role of convexity in optimization.

Proposition 1.5 *Suppose $\mathcal{X}$ is a convex set, and f is continuously differentiable. If f is convex on $\mathcal{X}$ and $\nabla f(x^*) = 0$ for some $x^* \in \mathcal{X}$, then x^* minimizes f over $\mathcal{X}$.*

Proof: Consider any $y \in \mathcal{X}$. By the first-order characterization of convexity, we have

$$f(y) \geq f(x^*) + \nabla f(x^*)'(y - x^*) = f(x^*).$$

■

Optimization sometimes also preserves convexity as the next two results show.

Proposition 1.6 *Let $\mathcal{Y}$ be an arbitrary non-empty set, and $g(\cdot, y)$ is a convex function on a convex set $\mathcal{X}$ for each $y \in \mathcal{Y}$. Then,*

$$f(x) = \sup_{y \in \mathcal{Y}} g(x, y)$$

is a convex function on $\mathcal{X}$.

Proof: Let $x, x' \in \mathcal{X}$ and $\lambda \in [0,1]$. Then

$$\begin{aligned} f(\lambda x + (1-\lambda)x') &= \sup_{y \in \mathcal{Y}} g(\lambda x + (1-\lambda)x', y) \\ &\leq \sup_{y \in \mathcal{Y}} \{\lambda g(x,y) + (1-\lambda) g(x', y)\} \\ &\leq \lambda \sup_{y \in \mathcal{Y}} g(x,y) + (1-\lambda) \sup_{y \in \mathcal{Y}} g(x', y) \\ &= \lambda f(x) + (1-\lambda) f(x'). \end{aligned}$$

■

Proposition 1.6 simply states that the maximum of a number of convex functions is still convex. The minimum of convex functions, however, is in general not convex. Yet, by imposing a stronger condition of joint convexity, one can still preserve convexity under minimization.

Proposition 1.7 *Let $\mathcal{X}$ be a convex set, $\mathcal{Y}(x)$ be a nonempty set for every $x \in \mathcal{X}$. Suppose the set $\mathcal{C} = \{(x,y) | x \in \mathcal{X}, y \in \mathcal{Y}(x)\}$ is convex, and $g(x,y)$ is a convex function on $\mathcal{C}$. Then, if the function defined by*

$$f(x) = \inf_{y \in \mathcal{Y}(x)} g(x,y)$$

is proper, i.e., $f(x) > -\infty, \forall x \in \mathcal{X}$, then $f(x)$ is a convex function on $\mathcal{X}$.

Proof: Let $x, x' \in \mathcal{X}$ and $\lambda \in [0,1]$. By definition of infimum, for any $\delta > 0$, there exist $y \in \mathcal{Y}(x)$ and $y' \in \mathcal{Y}(x')$ such that

$$\begin{aligned} g(x,y) &\leq f(x) + \delta \\ g(x', y') &\leq f(x') + \delta. \end{aligned}$$

Then,

$$\begin{aligned} \lambda f(x) + (1-\lambda) f(x') &\geq \lambda g(x,y) + (1-\lambda) g(x', y') - \delta \\ &\geq g(\lambda x + (1-\lambda)x', \lambda y + (1-\lambda) y') - \delta \\ &\geq f(\lambda x + (1-\lambda)x') - \delta, \end{aligned}$$

where the second inequality uses the join-concavity of g and the third inequality uses the fact that $\mathcal{C}$ is convex (which then implies $\lambda y + (1-\lambda)y' \in \mathcal{Y}(\lambda x + (1-\lambda)x')$). By letting $\delta \to 0$, we then arrive at the conclusion. ■

1.2 Lattice Theory

The motivation of this section stems from the comparative statics analysis of the following generic parametric optimization problem: for any $x \in \mathcal{X}$, let

$$\begin{aligned} f(x) &= \min_{y \in \mathcal{Y}(x)} g(x,y) \\ y^*(x) &\in \arg\min_{y \in \mathcal{Y}(x)} g(x,y), \end{aligned} \tag{1.1}$$

where $\mathcal{Y}(x)$ is a constraint set that possibly depends on the parameter x (we use $\mathcal{Y}$ instead if the constraint set is independent of x). How does $y^*(x)$ in (1.1) change when x changes? A common method in conducting such analysis relies on the convexity of the problem (1.1) and the implicit function theorem. That is, suppose $\mathcal{Y} = \mathbb{R}$, and $g(x, y)$ is strictly convex in y and twice continuously differentiable on $\mathbb{R}^2$. Then, $y^*(x)$ is the unique solution to the equation

$$\frac{\partial g(x, y)}{\partial y} = 0.$$

Implicit function theorem says that for an implicit function $y(x)$ defined by the equation

$$h(x, y) = 0,$$

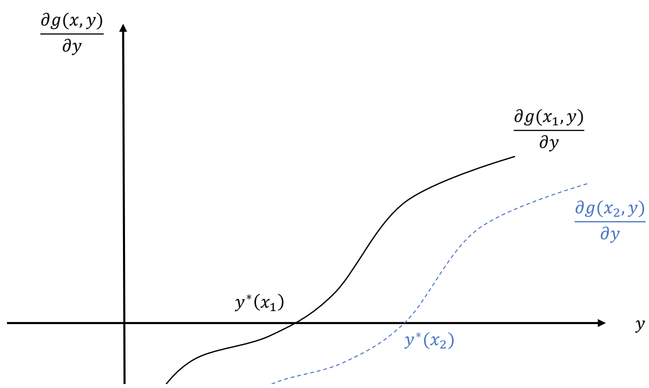
where $h(x, y)$ is assumed to be continuously differentiable on $\mathbb{R}^2$ and $\frac{\partial h(x,y)}{\partial y} \neq 0$, the derivative of the implicit function can be computed as

$$\frac{dy(x)}{dx} = -\frac{\partial h/\partial x}{\partial h/\partial y}.$$

Here, with $h(x, y) = \frac{\partial g(x,y)}{\partial y}$, we then have

$$\frac{dy^*(x)}{dx} = -\frac{\partial^2 g/\partial x \partial y}{\partial^2 g/\partial y^2}.$$

Strict convexity implies that $\partial^2 g/\partial y^2 > 0$. It follows that $y^*(x)$ is increasing in x if and only if $\partial^2 g/\partial x \partial y \leq 0$. The figure also illustrates that $y^*(x)$ is increasing in x if $\partial g/\partial y$ is decreasing in x.



It seems from the above flow of argument that convexity and differentiability properties are rather important for establishing comparative statics. However, are these conditions necessary? Consider the following transformation of problem (1.1). Let $z = \phi(y)$, where ϕ is an arbitrary (possibly discontinuous), strictly increasing function, and $\mathcal{Z} = \{z | z = \phi(y), y \in \mathcal{Y}\}$. We can then define $h(x, z) = g(x, \phi^{-1}(z))$ and consider the problem

$$z^*(x) \in \arg\min_{z \in \mathcal{Y}} h(x, z).$$

Clearly, we must have $z^*(x) = \phi(y^*(x))$. It follows that $y^*(x)$ is increasing in x if and only if $z^*(x)$ is increasing in x. However, with an arbitrary transformation ϕ, $h(x, z)$ may not be even continuous in z, let alone convexity and differentiability.[1] Hence, one should develop conditions that are independent of order preserving transformations, which is the purpose of lattice theory—a theory about order.

[1] This insightful argument is provided in Milgrom and Shannon (1994) who also conduct a thorough analysis of comparative statics.

Let $\mathcal{X} \subseteq \mathbb{R}^n$.[2] For any $x, y \in \mathcal{X}$, the meet of x and y, denoted as $x \wedge y$ is defined by

$$x \wedge y = (\min\{x_1, y_1\}, \min\{x_2, y_2\}, ..., \min\{x_n, y_n\}),$$

and the joint of x and y, denoted as $x \vee y$ is defined by

$$x \vee y = (\max\{x_1, y_1\}, \max\{x_2, y_2\}, ..., \max\{x_n, y_n\}).$$

A set $\mathcal{X}$ is a lattice if for any $x, y \in \mathcal{X}$, we have $x \wedge y, x \vee y \in \mathcal{X}$. Note that a lattice is not necessarily convex and a convex set is not necessarily a lattice as the following example shows.

Example 1.8 *Are the following sets lattices?*

1. *Any subset of a real line* $\mathbb{R}$.
2. $\mathcal{X} = \{(1,1), (1,2), (2,1), (2,2)\} \subseteq \mathbb{R}^2$.
3. $\{(x_1, x_2) \in \mathbb{R}^2 | x_1 + x_2 \leq 1, x_1, x_2 \geq 0\}$.

Given a lattice $\mathcal{X} \subseteq \mathbb{R}^n$, a function $f : \mathcal{X} \to \mathbb{R}$ is called submodular if

$$f(x \wedge y) + f(x \vee y) \leq f(x) + f(y).$$

We say f is supermodular if $-f$ is submodular.

A related and more intuitive concept is decreasing differences. A function $f : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ has decreasing differences in (x, y) if for any $y' \geq y$

$$f(x, y') - f(x, y)$$

is decreasing in x. Note that the definition here implies

$$f(x, y') - f(x, y) \geq f(x', y') - f(x', y),$$

for any $x' \geq x, y' \geq y$, which is the same as

$$f(x', y) - f(x, y) \geq f(x', y') - f(x, y').$$

That is, $f(x', y) - f(x, y)$ is also decreasing in y. We say $f : \mathbb{R}^n \to \mathbb{R}$ has decreasing differences if it has decreasing differences in (x_i, x_j) for any pair of indices i, j.

Theorem 1.9 *A function $f : \mathbb{R}^2 \to \mathbb{R}$ is submodular (supermodular) if and only if f has decreasing (increasing) differences.*

Proof: Suppose first that f is submodular. For any $x'_1 \geq x_1$ and $x'_2 \geq x_2$, we have

$$\begin{aligned} f(x_1, x'_2) - f(x_1, x_2) &= f(x_1, x'_2) - f(x_1 \wedge x'_1, x_2 \wedge x'_2) \\ &\geq f(x_1 \vee x'_1, x_2 \vee x'_2) - f(x'_1, x_2) \\ &= f(x'_1, x'_2) - f(x'_1, x_2). \end{aligned}$$

That is, $f(x_1, x'_2) - f(x_1, x_2)$ is decreasing in x_1.

[2] All the concepts developed here can be generalized to an arbitrary partially ordered set $\mathcal{X}$. One is referred to Topkis (1998) for such general theory.

Suppose now that f has decreasing differences. By definition, submodularity requires for any $x, y \in \mathbb{R}^2$

$$f(x_1, x_2) - f(x_1 \wedge y_1, x_2 \wedge y_2) \geq f(x_1 \vee y_1, x_2 \vee y_2) - f(y_1, y_2).$$

Indeed,

$$\begin{aligned} f(x_1, x_2) - f(x_1 \wedge y_1, x_2 \wedge y_2) &= f(x_1, x_2) - f(x_1, x_2 \wedge y_2) + f(x_1, x_2 \wedge y_2) - f(x_1 \wedge y_1, x_2 \wedge y_2) \\ &\geq f(x_1 \vee y_1, x_2) - f(x_1 \vee y_1, x_2 \wedge y_2) + f(x_1, y_2) - f(x_1 \wedge y_1, y_2). \end{aligned}$$

Note that regardless of whether $x_1 \leq y_1$ or $x_1 \geq y_1$, we must have

$$f(x_1, y_2) - f(x_1 \wedge y_1, y_2) = f(x_1 \vee y_1, y_2) - f(y_1, y_2),$$

and similarly we have

$$f(x_1 \vee y_1, x_2) - f(x_1 \vee y_1, x_2 \wedge y_2) = f(x_1 \vee y_1, x_2 \vee y_2) - f(x_1 \vee y_1, y_2).$$

It then follows that

$$\begin{aligned} f(x_1, x_2) - f(x_1 \wedge y_1, x_2 \wedge y_2) &\geq f(x_1 \vee y_1, x_2 \vee y_2) - f(x_1 \vee y_1, y_2) + f(x_1 \vee y_1, y_2) - f(y_1, y_2) \\ &= f(x_1 \vee y_1, x_2 \vee y_2) - f(y_1, y_2). \end{aligned}$$

■

One can also use similar argument to prove the equivalence between submodularity and decreasing differences for n-dimensional function (see also Theorem 2.2.2 in Simchi-Levi et al., 2013). For functions defined on $\mathcal{X} \times \mathcal{Y}$, however, since we do not require $\mathcal{X}, \mathcal{Y}$ here to be lattices, the two concepts are different. With additional conditions on smoothness, we then have the following equivalent characterizations for submodular functions defined on $\mathbb{R}^n$:

First-Order: A differentiable function $f : \mathbb{R}^n \to \mathbb{R}$ is submodular if and only if $\frac{\partial f(x)}{\partial x_i}$ is decreasing in x_j for any pair of indices i, j.

Second-Order: A twice-differentiable function f is submodular if and only if

$$\frac{\partial^2 f(x)}{\partial x_i \partial x_j} \leq 0, \ \forall i, j.$$

Returning to problem (1.1), one can see that besides conditions on $g(x, y)$, one clearly also needs some conditions on $\mathcal{Y}(x)$ to guarantee the monotone property of $y^*(x)$. For example, if $\mathcal{Y}(x)$ is a singleton set: $\mathcal{Y}(x) = \{y | y = \psi(x)\}$ for some function ψ, then we must have $y^*(x) = \psi(x)$ and $y^*(x)$ is increasing if and only if ψ is increasing. When $\mathcal{Y}(x)$ is not a singleton set, we also require certain "increasing" property which we define next.

A set-valued function $\mathcal{Y}(x)$ is called increasing in x if for any $x \leq x'$ and any $y \in \mathcal{Y}(x), y' \in \mathcal{Y}(x')$, we have $y \wedge y' \in \mathcal{Y}(x)$ and $y \vee y' \in \mathcal{Y}(x')$. In this sense, we interpret $\mathcal{Y}(x')$ to be "larger" than $\mathcal{Y}(x)$. Note, however, that this is different from the concept of set inclusion as the following examples illustrate.

Example 1.10 *Are the following set functions increasing?*

1. *Given an increasing function $\psi(x)$, let $\mathcal{Y}(x) = \{y \in \mathbb{R} | y = \psi(x)\}$.*
2. *$\mathcal{Y}(x) = [x, +\infty) = \{y \in \mathbb{R} | y \geq x\}$.*

3. $\mathcal{Y}(x) = [-x, +\infty) = \{y \in \mathbb{R} | y + x \geq 0\}$.

4. $\mathcal{Y}(x) = \{y \in \mathbb{R}^2 | y_1 + y_2 \geq x\}$.

From the definition, one can also see that if $\mathcal{Y}(x)$ is increasing in x, then $\mathcal{Y}(x)$ is a lattice for any $x \in \mathcal{X}$.

With all the concepts set in place, we arrive at the following result on monotone comparative statics.

Theorem 1.11 *Consider problem* (1.1)*. Suppose that $\mathcal{Y}(x)$ is increasing in x, and $g(x, y)$ is submodular in y for any fixed $x \in \mathcal{X}$ and has decreasing differences in (x, y). Then,*

$$\mathcal{Y}^*(x) := \arg \min_{y \in \mathcal{Y}(x)} g(x, y),$$

is also an increasing set function in x on $\{x \in \mathcal{X} | \mathcal{Y}^*(x) \neq \emptyset\}$.

Proof: For any $x', x \in \mathcal{X}$ with $x' \geq x$ such that $\mathcal{Y}^*(x')$ and $\mathcal{Y}^*(x)$ are non-empty, let $y' \in \mathcal{Y}^*(x'), y \in \mathcal{Y}^*(x)$. Since $\mathcal{Y}(x)$ is increasing, we have $y \wedge y' \in \mathcal{Y}(x)$ and $y \vee y' \in \mathcal{Y}(x')$. From the optimality of y, y', we have

$$\begin{aligned} & g(x, y \wedge y') - g(x, y) \geq 0, \\ \text{and} \quad & g(x', y') - g(x', y \vee y') \leq 0. \end{aligned}$$

By submodularity in y, we have

$$0 \leq g(x, y \wedge y') - g(x, y) \leq g(x, y') - g(x, y \vee y').$$

Using decreasing differences in (x, y), we further have

$$g(x, y') - g(x, y \vee y') \leq g(x', y') - g(x', y \vee y') \leq 0.$$

Hence, we must have $y \wedge y' \in \mathcal{Y}^*(x)$ and $y \vee y' \in \mathcal{Y}^*(x')$. ■

Some remarks follow.

- When the solution to problem (1.1) is unique, i.e., $\mathcal{Y}^*(x) = \{y^*(x)\}$, Theorem 1.11 implies that $y^*(x)$ is increasing in x.
- When the set $\mathcal{C} = \{(x, y) | x \in \mathcal{X}, y \in \mathcal{Y}(x)\}$ is a lattice, an often used stronger sufficient condition is to require $g(x, y)$ to be submodular in (x, y) on $\mathcal{C}$, which then implies decreasing differences in (x, y).
- The conditions of submodularity and decreasing differences can be further relaxed to quasi-submodularity and single-crossing property (see Milgrom and Shannon, 1994).
- With some additional compactness and continuity assumptions, one can ensure the non-emptiness and compactness of $\mathcal{Y}^*(x)$ and guarantee a monotone selection: by selecting $y^*(x)$ to be either the greatest or the least element in $\mathcal{Y}^*(x)$, we have $y^*(x)$ to be increasing as well.

Finally, in dynamic problems, in addition to monotonicity of $y^*(x)$, we are also interested in whether $f(x)$ is submodular as well. The following result provides conditions for this.

Theorem 1.12 *Let $\mathcal{C}$ be a lattice in $\mathbb{R}^n \times \mathbb{R}^m$ and $g(\cdot, \cdot) : \mathcal{C} \to \mathbb{R}$ is submodular on $\mathcal{C}$. For any $x \in \mathbb{R}^n$, let $\mathcal{Y}(x) = \{y \in \mathbb{R}^m | (x, y) \in \mathcal{C}\}$ and $\mathcal{X} = \{x \in \mathbb{R}^n | \mathcal{Y}(x) \neq \emptyset\}$. Then, $\mathcal{X}$ is a lattice and*

$$f(x) = \min_{y \in \mathcal{Y}(x)} g(x, y),$$

is submodular on $\mathcal{X}$.

Proof: Note that $\mathcal{X}$ is the projection of $\mathcal{C}$ onto the x-coordinates. For any $x, x' \in \mathcal{X}$, by definition there exist y, y' such that $(x, y), (x', y') \in \mathcal{C}$. Since $\mathcal{C}$ is a lattice, we have $(x \wedge x', y \wedge y'), (x \vee x', y \vee y') \in \mathcal{C}$. Hence, $y \wedge y' \in \mathcal{Y}(x \wedge x')$ and $y \vee y' \in \mathcal{Y}(x \vee x')$, which implies $x \wedge x', x \vee x' \in \mathcal{X}$.

Now let $y \in \mathcal{Y}^*(x)$ and $y' \in \mathcal{Y}^*(x')$. We then have

$$\begin{aligned} f(x) + f(x') &= g(x, y) + g(x', y') \\ &\geq g(x \wedge x', y \wedge y') + g(x \vee x', y \vee y') \\ &\geq f(x \wedge x') + f(x \vee x'). \end{aligned}$$

■

References

Milgrom, P. and C. Shannon (1994). Monotone comparative statics. *Econometrica: Journal of the Econometric Society*, 157–180.

Simchi-Levi, D., X. Chen, and J. Bramel (2013). *The Logic of Logistics: Theory, Algorithms, and Applications for Logistics Management.* Springer Science & Business Media.

Topkis, D. M. (1998). *Supermodularity and complementarity.* Princeton university press.