Lecture 19

§1 epigraph

1、epigraph（上镜图）

1° $C=\{(x,y): y \geq f(x)\}$ 被称为函数 f 的 epigraph

2° 通常用 epi f 来表示.

2、性质一

If f is a convex function, then epi$f$ is a convex set

（若 f 为凸函数，则 epi$f$ 为凸集）

3、性质二

If epi$f$ is a convex set, then f is a convex function

（若 epi$f$ 为凸集，则 f 为凸函数）

证明：

- 任取两点 $(\vec{x}_1, f(\vec{x}_1))$ 与 $(\vec{x}_2, f(\vec{x}_2))$
- 在两点连线上取一点，

$$\lambda(\vec{x}_1, f(\vec{x}_1)) + (1-\lambda)(\vec{x}_2, f(\vec{x}_2))$$
$$= (\lambda\vec{x}_1 + (1-\lambda)\vec{x}_2,\ \lambda f(\vec{x}_1) + (1-\lambda) f(\vec{x}_2))$$

- 由凸集定义，有

$$\lambda f(\vec{x}_1) + (1-\lambda) f(\vec{x}_2) \geq f(\lambda\vec{x}_1 + (1-\lambda)\vec{x}_2)$$ （凸函数定义式）

§2 Operations that preserve convexity

1、Composition

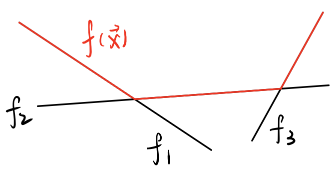
若 f 为凸函数，则 $g(\vec{x}) = f(a\vec{x}+b) - c$ 也为凸函数

证明：

$$
\begin{aligned}
& g(\lambda\vec{x}_1 + (1-\lambda)\vec{x}_2) \\
= & f[a(\lambda\vec{x}_1 + (1-\lambda)\vec{x}_2) + b] - c \\
= & f[\lambda(a\vec{x}_1+b) + (1-\lambda)(a\vec{x}_2+b)] - c \\
\leq & \lambda f(a\vec{x}_1+b) + (1-\lambda) f(a\vec{x}_2+b) - c \\
= & \lambda\cdot(f(a\vec{x}_1+b)-c) + (1-\lambda)\cdot(f(a\vec{x}_2+b)-c) \\
= & \lambda\cdot g(\vec{x}_1) + (1-\lambda) g(\vec{x}_2)
\end{aligned}
$$

2、Elementwise maximum

若 $f_1, \cdots\cdots, f_m$ 为凸函数，则 $f(\vec{x}) = \max\{f_1(\vec{x}), \cdots\cdots, f_m(\vec{x})\}$ 也为凸函数



证明1:

取任意点 $\vec{x}_1, \vec{x}_2 \in \mathrm{dom}(f)$, $\lambda \in [0,1]$. 则

$$
\begin{aligned}
& f(\lambda \vec{x}_1 + (1-\lambda)\vec{x}_2) \\
= & f_j(\lambda \vec{x}_1 + (1-\lambda)\vec{x}_2) \quad \text{for some } j \in \{1, \cdots, m\} \\
\le & \lambda f_j(\vec{x}_1) + (1-\lambda) f_j(\vec{x}_2) \\
\le & \lambda \max\{f_1(\vec{x}_1), \cdots, f_m(\vec{x}_1)\} + (1-\lambda) \cdot \max\{f_1(\vec{x}_2), \cdots, f_m(\vec{x}_2)\} \\
= & \lambda f(\vec{x}_1) + (1-\lambda) f(\vec{x}_2)
\end{aligned}
$$

证明2:

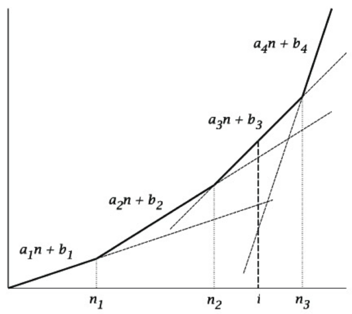
- $\mathrm{epi}\, f = \bigcap_{i=1}^{m} (\mathrm{epi} f_i)$
- 由 $\mathrm{epi} f_i$ 为凸集, 得 $\mathrm{epi} f$ 为凸集
- 因此 f 为凸函数

例1: Piecewise-linear functions

The maximum of L linear functions:

$$f(\vec{x}) = \max\{a_1^T \vec{x} + b_1, \cdots, a_L^T \vec{x} + b_L\}$$

where $a_i^T \vec{x} = \Sigma_m (a_i)_m \cdot x_m$



注: 并非所有 piecewise function 均 convex, 对多个凸函数取 maximum 后得到的函数 convex

例2: Distance to the farthest point of a set C

$$f(x) = \sup_{y \in C} \|x - y\|$$

For L2-distance, $\|y - x\| = \sqrt{\Sigma_i (y_i - x_i)^2}$

例3:

Example

- A seller wants to improve the product quality.
- If he improves the quality to q, the cost he incurs is $k\,C(q)$.
- Given q, his profit is $R(q)$.
- So, the seller's maximum profit is

$$\pi(k) = Max_q\, R(q) - k\, C(q)$$

- $f_q(k) = R(q) - k\,C(q)$, which is linear (convex) in k.
- As $\pi(k) = Max_q\, f_q(k)$, $\pi(k)$ is convex in k.
- Given any q, $f_q(k)$ is decreasing in k, thus $\pi(k)$ is decreasing in k.

(可以把 $-C(q)$ 视作例1中的 a_i, $R(q)$ 视作 b_i, 不过此处下标为 q 而非 i)

3. Nonnegative weighted sums

若 $f_1, f_2, \cdots, f_n$ 为凸函数，且 $w_1, \cdots, w_n > 0$，则 $f = w_1 f_1 + \cdots + w_n f_n$ 为凸函数

证明：

- 对任意 $\vec{x}, \vec{y}$，$0 \le \lambda \le 1$ 和 m

$$f_m(\lambda \vec{x} + (1-\lambda)\vec{y}) \le \lambda f_m(x) + (1-\lambda) f_m(y)$$

- 在不等式两侧同乘 w_m，再对 m 求和，得

$$w_1 \cdot f_1(\lambda \vec{x} + (1-\lambda)\vec{y}) + \cdots + w_n \cdot f_n(\lambda \vec{x} + (1-\lambda)\vec{y})$$
$$\le \lambda (w_1 f(\vec{x}) + \cdots + w_n f(\vec{x})) + (1-\lambda) \cdot (w_1 f(\vec{y}) + \cdots + w_n f(\vec{y}))$$

- 得：

$$f(\lambda \vec{x} + (1-\lambda)\vec{y}) \le \lambda f(\vec{x}) + (1-\lambda) f(\vec{y})$$

推广：

若 $f(x,y)$ 对任意 $y \in A$ 都 convex in x，且对每个 $y \in A$，都有 $w(y) \ge 0$，则

$$g(x) = \int_A w(y) f(x,y)\, dy$$

也 convex in x（给出积分存在）

注：可将 $g(x)$ 理解为 $\sum_{y \in A} w(y) \cdot dy \cdot f_y(x)$，其中将 $w(y) \cdot dy$ 视作整体，其大于0

例：证明：若 f 为凸函数（可能不可微），$x > 0$，则 $F(x) = \frac{1}{x} \int_0^x f(t)\, dt$ 也为凸函数

- set $t = xz$ and let z be the integration variable, then the integration domain is changed to $[0, 1]$

$$\begin{aligned} F(x) &= \frac{1}{x} \int_0^x f(t)\, dt \\ &= \frac{1}{x} \int_0^1 f(xz)\, d(xz) \\ &= \frac{1}{x} \int_0^1 f(xz) \cdot x\, dz \\ &= \int_0^1 f(xz)\, dz \end{aligned}$$

4. Minimization

若 f is convex in 向量 $(\vec{x}, \vec{y})$，且 C 为一个非空凸集，则函数

$$g(\vec{x}) = \inf_{\vec{y} \in C} f(\vec{x}, \vec{y})$$

也为凸函数

注：注意区分 maximization 与 minimization 的不同之处：

maximization：f is convex（对至少一个变量 convex 即可）

minimization：f is convex in 向量 $(\vec{x}, \vec{y})$（对所有变量均 convex）

不是对所有的 minimization forms 都成立。

证明：

对任意 $\vec{x}_1, \vec{x}_2 \in \operatorname{dom} g$，令 $\varepsilon > 0$，则存在 $\vec{y}_1, \vec{y}_2 \in C$，使得

$f(\vec{x}_i, \vec{y}_i) \le g(\vec{x}_i) + \varepsilon$ for $i = 1, 2$，令 $\theta \in [0, 1]$，有

$$
\begin{aligned}
g(\theta x_1+(1-\theta)x_2) &= \inf_{\vec{y}\in C} f(\theta\vec{x}_1+(1-\theta)\vec{x}_2,\vec{y}) \\
&\le f(\theta\vec{x}_1+(1-\theta)\vec{x}_2,\ \theta\vec{y}_1+(1-\theta)\vec{y}_2) \\
&\le \theta f(x_1,y_1)+(1-\theta)f(x_2,y_2) \\
&\le \theta g(x_1)+(1-\theta)g(x_2)+\varepsilon
\end{aligned}
$$

因为上式对 $\forall \varepsilon>0$ 均成立，有

$$g(\theta x_1+(1-\theta)x_2) \le \theta g(x_1)+(1-\theta)g(x_2)$$

We prove this by verifying Jensen's inequality for $x_1,\ x_2 \in \mathbf{dom}\, g$. Let $\epsilon > 0$. Then there are $y_1,\ y_2 \in C$ such that $f(x_i, y_i) \le g(x_i) + \epsilon$ for $i = 1,\ 2$. Now let $\theta \in [0,1]$. We have

$$
\begin{aligned}
g(\theta x_1 + (1-\theta)x_2) &= \inf_{y\in C} f(\theta x_1 + (1-\theta)x_2, y) \\
&\le f(\theta x_1 + (1-\theta)x_2, \theta y_1 + (1-\theta)y_2) \\
&\le \theta f(x_1,y_1) + (1-\theta) f(x_2,y_2) \\
&\le \theta g(x_1) + (1-\theta) g(x_2) + \epsilon.
\end{aligned}
$$

Since this holds for any $\epsilon > 0$, we have

$$g(\theta x_1 + (1-\theta)x_2) \le \theta g(x_1) + (1-\theta) g(x_2).$$

例:

Example 3.16 *Distance to a set.* The distance of a point x to a set $S \subseteq \mathbf{R}^n$, in the norm $\|\cdot\|$, is defined as

$$\mathbf{dist}(x,S) = \inf_{y\in S} \|x-y\|.$$

The function $\|x-y\|$ is convex in (x,y), so if the set S is convex, the distance function $\mathbf{dist}(x,S)$ is a convex function of x.

案例分析 1:

Pricing and Revenue Management

- Consider a seller selling a product. He observes from past data that demand can be modeled as a linear function of price p: D(p)=a-bp, where a>0 and b>0 are parameters. The price he can charge is restricted in $\left(0,\frac{a}{b}\right)$. What's the optimal price to charge in order to maximize his revenue?

- Maximize $\mathrm{R(p)} = \mathrm{pD(p)} = ap - bp^2$
- Subject to $0 < p < \frac{a}{b}$.

Sol: Feasible set is a convex set

- Any interval (l,b) or [l,b] is convex.

Objective function is concave

- Note this is a maximization problem! Concavity rather than convexity of the objective function is required.
- $\mathrm{R(p)} = \mathrm{pD(p)} = ap - bp^2$
- FOC: $\mathrm{R(p+s)} \le R(p) + R'(p)s$
- SOC: $R''(p) \le 0$

- $\mathrm{R'(p)} = 0 \longrightarrow \mathrm{p} = \frac{\mathrm{a}}{2b} \in \left(0,\frac{a}{b}\right)$

案例分析2:

The Newsvendor Problem

- A newsboy needs to decide how many newspapers to buy each morning in order to generate the most expected sales profit. The demand D for newspapers in a day is random, with $f(\cdot)$ being its pdf and $F(\cdot)$ being its CDF. He needs to buy each newspaper at price c, and can sell it at price p.

Sol: · 令 purchase decision 为 q. 则 objective 为

$$\text{Maximize } E[\text{profit}] = E[p\cdot \min(q, D)] - cq$$

· 先求销售量的期望

$$E[\min(q, D)] = \int_0^q x\cdot f(x)dx + q(1-F(q))$$

· 由此求得利润的期望

$$E[\text{profit}] = p\int_0^q xf(x)dx + pq(1-F(q)) - cq$$

· 判断 E[profit] 是否 concave (用 SOC)

$$\frac{\partial}{\partial q}E[\text{profit}] = pqf(q) + p(1-F(q)) - pqF'(q) - c$$
$$= p(1-F(q)) - c$$
$$\frac{\partial^2}{\partial q^2}E[\text{profit}] = pF'(q) \le 0$$

· 令 $\frac{\partial}{\partial q}E[\text{profit}] = 0$ 得:

$$q^* = F^{-1}(1-\frac{c}{p})$$

- Let the newsboy's purchase decision be q. His objective is

$$\text{Maximize } \mathrm{E}[\text{Profit}] = E[p\min(q, D)] - cq.$$

We can write

$$\mathrm{E}[\min(\mathrm{q}, \mathrm{D})] = \int_0^q xf(x)dx + q(1 - F(q)).$$

So

$$\mathrm{E}[\text{Profit}] = \mathrm{p}\int_0^q xf(x)dx + pq(1 - F(q)) - cq.$$

Is E[Profit] concave?

Concavity of profit function

- $\mathrm{E}[\text{Profit}] = \mathrm{p}\int_0^q xf(x)dx + pq(1 - F(q)) - cq.$

- SOC: $\frac{\partial}{\partial q}\mathrm{E}[\text{profit}] = pqf(q) + pq(-F'(q)) + p[1 - F(q)] - c = p[1 - F(q)] - c$

  $\frac{\partial^2}{\partial q^2}\mathrm{E}[\text{profit}] = p[-F'(q)]$ ≤ 0 (note F() is an increasing function)

$\frac{\partial}{\partial q}E[Profit] = 0$: $q^* = F^{-1}\left(\frac{p-c}{p}\right)$.

案例分析 3

Flair Furniture Company's problem

- Flair Furniture Company produces inexpensive tables and chairs. The production process for each is similar in that both require a certain number of labor hours in the carpentry department and a certain number of labor hours in the painting department.
- Each table takes 3 hours of carpentry work and 2 hours of painting work.
- Each chair requires 4 hours of carpentry and 1 hour of painting.
- During the current month, 2,400 hours of carpentry time and 1,000 hours of painting time are available.

- The marketing department wants Flair to make no more than 450 new chairs this month because there is a sizable existing inventory of chairs.
- However, because the existing inventory of tables is low, the marketing department wants Flair to make at least 100 tables this month.
- Each table sold results in a profit contribution of $7, and each chair sold yields a profit contribution of $5. How should Flair Furniture Company make its production plan?

Sol:

Flair Furniture Company's decision variables

- Two decision variables:
  - the number of tables (to be) produced (T)
  - the number of chairs (to be) produced (C)

When writing the formulation on paper, it is convenient to express the decision variables using notations that are easy to understand. For example, the number of tables to be produced can be denoted by names such as T, and the number of chairs to be produced can be denoted by names such as C.

Flair Furniture Company's objective function

- Flair Furniture's problem is to determine the best possible combination of tables and chairs to manufacture this month in order to attain the maximum profit.

$$\begin{aligned}\text{Profit} = {} & (\$7 \text{ profit per table}) \times (\text{number of tables produced}) \\ & + (\$5 \text{ profit per chair}) \times (\text{number of chairs produced})\end{aligned}$$

- The objective function is

$$\text{Maximize } \$7T + \$5C$$

Flair Furniture Company's constraints

- "Each table takes 3 hours of carpentry work and 2 hours of painting work. Each chair requires 4 hours of carpentry and 1 hour of painting. During the current month, 2,400 hours of carpentry time and 1,000 hours of painting time are available."
- Carpentry time limit:

$$3T + 4C \leq 2400$$

- Painting time limit:

$$2T + C \leq 1000$$

- "The marketing department wants Flair to make no more than 450 new chairs this month because there is a sizable existing inventory of chairs. However, because the existing inventory of tables is low, the marketing department wants Flair to make at least 100 tables this month."
- Marketing-specifed constraint for chairs:

$$C \leq 450$$

- Marketing-specifed constraint for tables:

$$T \geq 100$$

- Nonnegativity constraints!
- number of tables produced ≥ 0:

$$T \geq 0$$

- number of chairs produced ≥ 0:

$$C \geq 0$$

The LP Model for Flair Furniture Company

The complete LP model for Flair's problem can be restated as follows:

$$\text{Maximize profit} = \$7T + \$5C$$

subject to the constraints

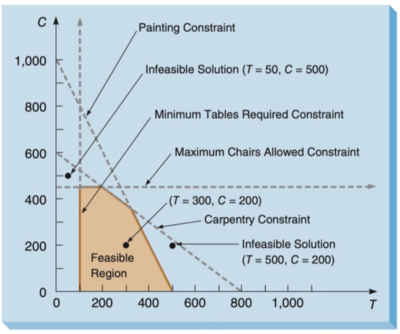
$$\begin{aligned}
3T + 4C &\leq 2{,}400 && \text{(carpentry time)}\\
2T + 1C &\leq 1{,}000 && \text{(painting time)}\\
C &\leq 450 && \text{(maximum chairs allowed)}\\
T &\geq 100 && \text{(minimum tables required)}\\
T, C &\geq 0 && \text{(nonnegativity)}
\end{aligned}$$

Integers actually

Graphical method

- The graphical method works only when there are two decision variables, but it provides valuable insight into how larger problems are solved.

Graphical Representation of Constraints



Drawing level lines