最优化

作者: 朱占义

1.1 一些定义

闭函数和下半连续函数的定义

设 $f:\mathbb{R}^n \to \mathbb{R}$ 为广义实值函数,则以下命题等价:

1.f(x) 的 $\forall \alpha, \{x|f(x) \leq \alpha\}$ 为闭集

2.f(x) 是 \mathbb{R}^n 上的下半连续函数 \iff $\liminf f(x) \geq f(x)$

什么是下极限? 什么是下半连续函数

假如一个趋于a的子列有极限,可能有多个极限点。这样的极限点组成一个集合,这个集合的下确界就是f(x)在a处的下极陷 $\lim_{x\to \infty}\inf f(x)$

3. epif为闭集 (f(x)是闭函数的定义)

证明: 1.⇒2.:

通过反证法,假设 $\exists \{x_k\} \to x,$ 并且 $\liminf f(x) < f(x)$. 那么一定 $\exists \alpha$ 使得 $\liminf f(x) < \alpha < f(x)$,所以一定存在 $\{x_k\}$

的子列 $\{x_{ik}\} o x$ 使得 $f(x_{ik}) \le \alpha$,由于命题1., $x_{ik} \in \{x|f(x) \le \alpha\}$,所以 $x \in \{x|f(x) \le \alpha\}$.所以可以知道 $f(x) \le \alpha$

与假设矛盾所以假设不成立

证明2.⇒>3.:

要证明上镜图是一个闭集,即证明 $(x_k, \omega_k) \in epif$ 并且 $(x_k, \omega_k) \to (x, \omega)$. 那么 $(x, \omega) \in epif$,

假设 $(x_k, \omega_k) \in epif$,那么 $f(x_k) \leq \omega_k$,由于它是下半连续函数

$$f(x) \le \lim_{k \to \infty} \inf f(x_k) \le \liminf \omega_k = \omega$$
 (2)

所以 $(x_k, \omega_k) \rightarrow (x, \omega)$

证明3. ⇒ 1.:

想证明 $\{x_k\}\subseteq \{x|f(x)\leq \alpha\}, \exists x_k\to x, 那么x\in \{x|f(x)\leq \alpha\}$

 $(x_k, \alpha) \in epif$, 那么 $(x, \alpha) \in epif$. 因此 $f(x) \leq \alpha$

下面提出上面这个定理的条件放宽一些的情况:

1.假如不是 \mathbb{R}^n 上的闭函数,而是一个集合X上的闭函数。

 $f:X o[-\infty,+\infty], epif\subseteq\mathbb{R}^{n+1}$ 并且 $epif=\{(x,\omega)|x\in X,\omega\in\mathbb{R}, f(x)\leq\omega\}$ 为闭集 $\Leftrightarrow f$ 为闭函数 这时我们对f(x)进行延拓

$$ilde{f}(x) = egin{cases} f(x) & x \in X \\ +\infty & x
eq X \end{cases}$$

(3)

 $epif = epi\tilde{f}(x)$,显然我们知道 $epi\tilde{f}(x)$ 是闭集并且 $\tilde{f}(x)$ 的定义域为 \mathbb{R}^n

1

 $epi\tilde{f}(x)$ 是 \mathbb{R}^n 上的下半连续函数

1

 $\forall \alpha$ $\tilde{f}(x)$ 的水平集 = f(x)的水平集为闭集

f为X上的闭函数⇔f的水平集为闭集

2.假如不是 \mathbb{R}^n 上的下半连续函数,而是在dom f上下半连续

由于前边两个条件是等价的,这个时候我们只要能推导出来满足其中一个,或者找出不满足其中一个的反例

$$f(x)$$
是 $(0, +\infty)$ 上的下半连续函数
$$f(x) = \begin{cases} 1 & x \in (0, +\infty) \\ +\infty & \text{否则} \end{cases}$$
 (4)

显然这个时候epif(x)不为闭集,而且 $\alpha = 2$ 的水平集为 $(0, +\infty)$ 为开集

所以这个时候推不出f(x)是闭函数

倘若继续加条件,即dom f有界+在dom f上下半连续 \Longrightarrow f(x)为闭函数

f,epif,dom f,水平集的凸闭关系

1.f和epif

这个里面前边的要求定义域为凸集

$$f$$
为凸函数 \iff $epif$ 为凸集 f 为闭函数 \iff $epif$ 为闭集

2.f和f的水平集

这个里面定义域是任意的

f为凸函数 $\longrightarrow f$ 的水平集为凸集

$$(从这里我们明白 $\forall \alpha$ 水平集为凸集只是 f 为凸函数的必要条件 $)$ (6)$$

这里我们举个例子 $f = \sqrt{x}$ 对于 $\alpha > 0$, 它的水平集为 $[0, \alpha^2]$, 为凸集; $\alpha < 0$, 为空集也为凸集

$$f$$
为闭集 \iff f 的水平集为闭集 (7)

3.f和domf

$$f$$
为凸函数 $\longrightarrow dom f$ 为凸集 (8)

$$f(x) = \begin{cases} \sqrt{x} & , x \in [0, +\infty) \\ +\infty & , otherwise \end{cases}$$
(9)

有效域为 $[0,+\infty)$ 为凸集,但是f(x)并不是凸函数

所以有效域为凸推不处出凸函数, 但是根据凸函数的定义就要求这个函数的有效域是一个凸集

$$f$$
为闭函数 $\Leftrightarrow dom f$ 为闭集 (10)

$$f = \frac{1}{x}$$
, 是一个闭函数,但是有效域是 $(0, +\infty)$ 并不是闭集 (11)

$$f(x) = \begin{cases} 0 & , x \in (1,2) \\ 1 & x \in \{1,2\} \end{cases}$$
 (12)

显然函数并不是闭函数,但是有效域是闭集

4.epif和domf

$$dom f$$
为凸/闭集 $\iff epif$ 为凸/闭集 (13)

$$dom f = \{x | f(x) < +\infty\}, epif = \{(x, \omega), f(x) \le \omega\}$$

$$\tag{14}$$

上镜图可以投影到 \mathbb{R}^n (可以通过线性变换得到有效域)线性变换后还是凸集/闭集

1.2一些保凸、闭性的运算

非负线性组合的保凸保闭研究

$$f_i:\mathbb{R}^n o(\infty,+\infty], \lambda_i\geq 0, i=1,2,\cdots,m$$

$$g(x)=\sum_{i=1}^m\lambda_if_i(x) \tag{15}$$

那么就有 $f_i(x)$ 为凸函数 $\Rightarrow g(x)$ 为凸函数 $f_i(x)$ 为闭函数 $\Rightarrow g(x)$ 为闭函数

证明:

非负的线性组合仍然是凸函数,这个显然

下面证明闭函数 $f_i(x)$ 为闭函数所以它在 \mathbb{R}^n 上下半连续

$$\liminf f_i(x_k) \ge f_i(x)$$

$$\liminf g(x) = \liminf \sum_{i=1}^m \lambda_i f_i(x) \ge \sum_{i=1}^m \lambda_i \liminf f_i(x) \ge \sum_{i=1}^m \lambda_i f_i(x) = g(x)$$
(16)

因此我们也证明出来了g(x)在 \mathbb{R}^n 上下半连续,那么g(x)为闭函数

显然 $epif_i(x)$ 为闭集/凸集 $\Rightarrow g(x)$ 为闭集/凸集

上下确界的保凸性,保闭性运算

$$f_i: \mathbb{R}^n \to (\infty, +\infty], \lambda_i \ge 0, i = 1, 2, \cdots, m$$
$$g(x) = \sup f_i(x)$$
(17)

不难发现 $\operatorname{epi}(q) = \bigcap_i \operatorname{epi}(f_i)$,然而交集是保凸并且保闭的运算

可微凸函数

 $C\subseteq\mathbb{R}^n$ 上的凸集, $f:\mathbb{R}^n o\mathbb{R}$ 且f为可微函数

1.

$$f$$
在 C 上为凸函数 $\Leftrightarrow f(x) \ge f(y) + \langle \nabla f(y), x - y \rangle$ 对于 $\forall x, y \in \mathbb{C}$
 f 在 C 上为严格凸函数 $\Leftrightarrow f(x) > f(y) + \langle \nabla f(y), x - y \rangle$ 对于 $\forall x, y \in \mathbb{C}$ (18)

"⇐"即要证明

$$\alpha f(x) + (1 - \alpha)f(y) \ge f(\alpha x + (1 - \alpha)y)$$
对于 $f(x) \ge f(z) + \langle \nabla f(z), x - z \rangle$

$$f(y) \ge f(z) + \langle \nabla f(z), y - z \rangle$$

$$\Leftrightarrow z = \alpha x + (1 - \alpha)y,$$
下面我们得到 $\alpha f(x) + (1 - \alpha)f(y) \ge \alpha f(z) + (1 - \alpha)f(z) + 0 = f(z)$

"⇒"不妨令z>x

于是我们构造出下面割线斜率的值 $g(x)=rac{f(\alpha z+(1-lpha)x)-f(x)}{lpha(z-x)}$

显然α越大斜率就越大

$$g(\alpha) = \frac{f(\alpha z + (1 - \alpha)x) - f(x)}{\alpha(z - x)}$$
$$g(1) \ge g(0)$$
$$\frac{f(z) - f(x)}{z - x} \ge \nabla f(x)$$
 (20)

1.3凸包与仿射包

(下面介绍的无论是凸包,还是仿射包,还是生成锥我们都要研究它保闭性吗?保凸性吗?)

凸组合:

$$(\alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3), \sum \alpha_i = 1, \alpha_i \in [0, 1]$$
 (21)

仿射组合:

$$(\alpha_1x_1+\alpha_2x_2+\alpha_3x_3), \sum \alpha_i=1, \alpha_i\in\mathbb{R}$$
 (22)

1.凸包: (convex hull) $X \subseteq \mathbb{R}^n$

$$conv(X) = \{ 包含X的最小的凸集 \}$$
 $= \{ 包含X的所有凸集的交 \}$
 $= \{ X中所有元素的凸组合 \}$
(23)

1) 对于保凸性:显然的

2) 对于保闭性,可以举出下面的这个反例

例子:
$$X = \left\{ (x_1, x_2) \mid x_1 \geq \frac{1}{x_1}, \; x_1, x_2 > 0 \right\} \cup \left\{ (0, 0) \right\}$$

显然在这个例子中,X是一个闭集,但是它的凸包是一个第一象限包含原点但不含坐标轴的集合,显然是一个开集

那我们知道了凸包不能保持集合的闭性,但是如果这个闭集是有界的呢?(直接记住这个结论)

$$X$$
为紧集 $\Rightarrow conv(X)$ 为紧集 (24)

2.生成锥: (cone)

 $cone(X): \{X$ 中所有元素的非负组合 $\}$

$$cone(X) = \{\sum_{i=1}^{m} \alpha_i x_i | x_i \in X\}$$
 (25)

1)对于保凸性: (容易证明)

2) 对于保闭性: (可以举下面的反例)

$$\overline{B}((0,1),1) = \{(x,y) \in \mathbb{R}^2 \mid x^2 + (y-1)^2 \le 1\}$$
 (26)

显然它的生成锥是X轴上方的全体,不包含x轴但包含原点,我们发现它的生成锥不是一个闭集

3.仿射包: (affine hull)

仿射包的性质:

1.仿射包的单调性

2仿射包的幂等性

仿射包、闭包、凸包的关系: remark:

1) $conv(X) \subseteq aff(X)$

$$2)aff(X) = aff(cl(X)) = aff(conv(X))$$
(可以自己证明一下)

aff(X) = aff(conv(X)):

首先证明 $aff(X) \subseteq aff(conv(X))$

$$X \subseteq conv(X)$$
, 由于仿射包的单调性
$$aff(X) \subseteq aff(conv(X))$$
 (28)

其次证 $aff(X) \supseteq aff(conv(X))$

由仿射包和凸包的定义不难发现,
$$conv(X) \subseteq aff(X)$$
 所以 $aff(conv(X)) \subseteq aff(aff(x))$ (29)

aff(X) = aff(cl(X))

首先证明 $aff(X) \subseteq aff(cl(X))$

$$X \subseteq cl(X)$$
, 由于仿射包具有单调性
$$aff(X) \subseteq aff(cl(X))$$
 (30)

其次证 $aff(X) \supseteq aff(cl(X))$

由于仿射包为闭集,
$$cl(X)$$
是包含 X 的最小闭集,即包含 X 的所有闭集的交集
$$\mathbb{R} \times cl(X) \subseteq aff(X), \text{ 所以} aff(cl(X)) \subseteq aff(aff(X)) = aff(X)$$

 $3)\dim(X) = \dim(aff(X))$

1.4相对内点、相对内部

相对内点:

$$C\subseteq \mathbb{R}^n$$
且 $C
eq \Phi, x\in C$ $B(x,arepsilon)\cap aff(C)\subseteq C$ 那么就称 x 是相对于 $aff(C)$ 的内点 x 为 C 的相对内点 (32)

(一定要记住无论是内点还是相对内点,他都属于这个集合C)

相对内部:由C中所有相对内点构成的集合 relative interior riC

相对内部和内部的关系:

$$intC = \{x \in C \mid \exists \delta > 0 \text{ s.t. } B(x, \delta) \subseteq C\}$$

$$riC = \{x \in C \mid \exists \delta > 0 \text{ s.t. } B(x, \delta) \cap \text{aff} C \subseteq C\}$$

$$\# \angle intC \subseteq riC$$
(33)

证明: 任取intC中的元素x, 如果 $B(x,\delta_0)\subseteq C$, 可得 $B(x,\delta_0)\cap affC\subseteq C$

相对边界: clC/riC

相对边界
$$\cup$$
 相对内部 $=$ 闭包 (34)

例1:

$$\{(x,y)|x\in[1,2],y=1\}\tag{35}$$

它的相对内部就是不包含两个端点的线段

例2: 假如平面上一个实心的正方形,那么它的相对内部是什么?如果只有四条边呢? 实心的话相对内部就是不含边界的这个正方形,如果只有四条边的话相对内部是空集

需要注意以下所有的性质针对的都是非空凸集

非空凸集的相对内部一定非空, 但是内部可能是空集

1.4.1线段原则、仿射包和相对内部的关系、相对内点的等价定义

c是非空凸集:

1) $\forall x \in riC, \forall y \in clC.$ 则x,y两点连线线段之间的所有点(不含y)均在riC中 $(\alpha x + (1-\alpha)y \in riC$ 且 $\alpha \in (0,1]$) (线段原则)这个证明不要求掌握

仿射包和相对内部的关系

2) aff(C) = aff(riC) (自己的仿射包等于相对内部的仿射包)

 $aff(riC) \subseteq aff(C)$:

显然, $riC \subseteq C$ 由仿射包的性质 $aff(riC) \subseteq aff(C)$

 $aff(riC) \supseteq aff(C)$:

 $\forall x \in C, y \in riC$ 根据线段原则令 $z = \alpha y + (1-\alpha)x \in riC, \alpha \in (0,1]$ 那么z和y都属于riC

分情况:
$$1)x = y \in riC \subseteq aff(riC)$$

$$2)x = \frac{1}{1-\alpha}z - \frac{\alpha}{1-\alpha}y; \alpha \in (0,1)$$
 并且 $y, z \in riC$
 $x \in aff(riC)$ (36)

所以 $C \subseteq aff(riC)$, 由仿射包的性质 $aff(C) \subseteq aff(aff(riC)) \subseteq aff(riC)$

 $\exists x \in riC \Leftrightarrow \forall y \in C \ \exists \gamma > 1$ 使得 $x + (\gamma - 1)(x - y) \in riC$

证明: ←

$$z = x + (\gamma - 1)(x - y) \in riC$$
并且 $y \in C \subseteq clC$
$$x = \frac{1}{r}z + \frac{r - 1}{r}y$$
在 y, z 连线的线段上 (37)

 \Rightarrow

$$x \in riC$$
, 于是日 ε 使得 $B(x,\varepsilon) \cap aff(C) \subseteq C$
 $\forall y \in C, \gamma - 1 = \frac{\varepsilon}{2}|x - y|, \varepsilon^* = \frac{\varepsilon}{4}$
 $B(x + (\gamma - 1)(x - y), \varepsilon^*) \subseteq B(x,\varepsilon)$ $\gamma x + (1 - \gamma)y \in aff(C)$
 $B(x + (\gamma - 1)(x - y), \varepsilon^*) \cap aff(C) \subseteq C$
 $x + (\gamma - 1)(x - y) \in riC$ (38)

1.4.2 常值性定理

 $X\subseteq\mathbb{R}^n$ 是非空凸集, $f:X\to\mathbb{R}$ 为凹函数, $X^*=\{x^*|f(x^*)=\inf_{x\in X}f(x)\};X^*\cap riX
eq \phi$ 那么f(x)在X上为常值函数 (自己理解:凹函数在相对内部取不到最小值)

证明:

取
$$x^* \in X^* \cap riX$$
, 那么 $\forall y \in X$ 得 $z = x^* + (\gamma - 1)(x^* - y) \in riX$

$$x^* = \frac{\gamma - 1}{\gamma}y + \frac{1}{\gamma}z; \text{且}y$$
和 z 均为 X 得相对内点
由于 $f(x)$ 为凹函数, $f(x^*) \ge \frac{\gamma - 1}{\gamma}f(y) + \frac{1}{\gamma}f(z)$
由于 $x^* \in X^*, \frac{\gamma - 1}{\gamma}f(y) + \frac{1}{\gamma}f(z) \ge f(x^*)$

$$f(y) = f(x^*)$$
为定值

例子1:

$$\min c^T x + b$$
 $x \in X(X$ 为凸集) (40)

那么我们就能知道最小值点一定不会取在X的相对内部

例子2:

$$\|A\| = \sup_{\|x\| \le 1} \|Ax\|_2 = -\inf_{\|x\| \le 1} - \|Ax\|_2$$

注意到在这里面 $\|x\| \le 1$ 是凸集, $-\|Ax\|_2$ 是凹函数

所以最小值肯定不能在内部取,那么 $-\inf_{\|x\| \le 1} - \|Ax\|_2 = -\inf_{\|x\| = 1} - \|Ax\|_2 = \sup_{\|x\| = 1} \|Ax\|_2$ (也是一种A的2范数的定义

1.4.3 相对内部和闭包的关系

C是一个非空凸集:

 $\mathbf{1}.riC=ri(clC)$ (一个集合的相对内部等于这个集合闭包的相对内部)

证明:
$$ri(C) \subseteq ri(clC)$$

$$C \subseteq clC$$
, 由仿射包的单调性 $ri(C) \subseteq ri(clC)$ (42)

 $ri(C) \supseteq ri(clC)$

 $\forall x \in ri(clC)$,根据相对内部的等价定义, $\forall y \in clC \ \exists \gamma > 1$ 使得 $z = x + (\gamma - 1)(x - y) \in ri(clC)$ $y \in clC, z \in ri(clC)$ $x = \frac{1}{\gamma}z + \frac{\gamma - 1}{\gamma}y($ 根据线段原则), $x \in clC$ (43)

因此我们知道 $ri(clC) \subset clC$

2.clC = cl(riC)

证明: $cl(riC) \subseteq clC$:

由于闭包也具有单调性,
$$riC \subseteq C$$
则 $cl(riC) \subseteq clC$ (44)

 $cl(riC) \supseteq clC$:

下边这三个没必要可以会, 会用上边的来推导就可以了

1) riD = riC

2)clD = clC

3) $riC \subseteq D \subseteq clC$

这个的证明用上面结论以后过于简单就不证了

1.4.4 闭包、相对内部线性变换下的性质

补充知识:对于连续函数开集的原像是开集

连续函数可以把紧集映射成紧集

c是非空凸集 $A \in \mathbb{R}^{m \times n}$

$$AclC \subseteq cl(AC)$$

(当 C 集合有界时取等) (46)

e.g.比方说 $y=\frac{1}{x}$ 这个函数的上镜图组成的集合,A把它映射到x轴上。那么Acl(C)是不包含原点的x轴正半轴,cl(AC)是包含原点的x轴正半轴

$$AriC = ri(AC) \tag{47}$$

1.4.5 闭包、相对内部有关集合的交、和的性质

(没有证明,知识举了例子辅助理解)

$$ri(A \cap B) \supseteq riA \cap riB$$
 (48)

e.g. A:在平面上x轴的非负半轴

B:在平面上x轴的非正半轴

那么:riA 是平面上x轴的正半轴 riB 是平面上x轴的负半轴 $riA \cap riB = \phi$

而 $A \cap B = (0,0)$,单点集是非空凸集,那么它的相对内部也是非空

$$cl(A \cap B) \subseteq clA \cap clB \tag{49}$$

e.g. A:在平面上x轴的正半轴

B:在平面上x轴的负半轴

那么clA在平面上x轴的非负半轴,clB在平面上x轴的非正半轴; $clA \cap clB = \{(0,0)\}$

而 $cl(A \cap B) = \phi$

$$ri(C_1) + ri(C_2) = ri(C_1 + C_2)$$

$$cl(C_1) + cl(C_2) \subseteq cl(C_1 + C_2)$$
(50)

e.g.
$$C_1 = \left\{ (x_1, x_2) \mid x_1 \geq rac{1}{x_1}, \; x_1, x_2 > 0
ight\}$$
 C_2 是y轴

那么 $cl(C_1)+cl(C_2)$ 是第一、四象限不包含y轴的点的集合

 $cl(C_1+C_2)$ 是第一、四象限包含y轴的点的集合

1.5 回收锥 (recession cone)

回收方向 回收锥: (recession direction)

$$C\subseteq\mathbb{R}^n$$
的非空凸集, $\forall x\in C$, $\forall \alpha\geq 0$ $\exists y$ 使得 $x+\alpha y\in C$ 那么 y 被称为集合 C 的回收方向 由所有回收方向组成的集合称为 C 的回收锥 $R_C=\{y\mid \forall x\in C, \forall \alpha\geq 0, x+ay\in C\}$

$$X^* = \{x^* | f(x^*) = \inf_{x \in X} f(x)\}$$

= $\cap \{x \in X | f(x) \le r_k, r_k \downarrow \inf_{x \in X} f(x)\}$ (52)

若f(x)是个闭凸函数,那么水平集仍然为闭凸集。

对于满足 $c_k \supseteq c_{k+1}$:

在这种情况下,这些闭凸集的交是空集;但是如果这些集合中存在某个 c_k 是紧集,那么他们的交就不是空集