1.3 凸集、凸函数、凸规划

1.3.1 凸集

定义1.3.1 给定非空集合$F \subseteq R^n$,如果$\forall x, y \in F, \alpha \in [0,1]$都有

$\alpha x + (1-\alpha) y \in F$　　　　**称为x与y的凸组合**

那么称F为R^n中的一个凸集。如果凸集为开集，则称为开凸集；若凸集为闭集，则称为闭凸集。

规定： 空集为凸集；

注：单点集与R^n为凸集

例1.3.1 假设$\omega \in R^n \setminus \{0\}, \beta \in R$,证明下面的集合为凸集

1).超平面$H = \{x \in R^n \mid \omega^T x = \beta\}$;

2).闭半空间$\{x \in R^n | \omega^T x \geq \beta\}$和 $\{x \in R^n | \omega^T x \leq \beta\}$;

3). 开半空间$\{x \in R^n \mid \omega^T x > \beta\}$ 和$\{x \in R^n \mid \omega^T x < \beta\}$;

4). 超球$B = \{x \in R^n \mid \|x\| \leq \beta\}$,其中$\beta \geq 0$.

命题1.3.1 假设$F_i \subseteq R^n$为凸集且$\beta_i \in R$(其中$i \in \{1,2,\cdots,p\}$).则下列集合均为凸集

1)交集$F := F_1 \cap F_2 \cap \cdots \cap F_p$

2) 集合$\beta_1 F_1 := \{\beta_1 x \mid x \in F_1\}$;

3)和集$F_1 + F_2 := \{x + y \mid x \in F_1, y \in F_2\}$;

4)集合$\sum_{\text{i}=1}^{p} \beta_i F_i$;

证明： 1)$\forall x, y \in F, \forall \alpha \in [0,1]$

$\forall x, y \in F_i, \alpha x + (1-\alpha) y \in F_i (i = 1, 2, \cdots, p)$

$\alpha x + (1-\alpha) y \in F$

因此F为凸集。

定理1.3.1 非空集合$F \subseteq R^n$为凸集$\Leftrightarrow \forall x_i \in F$及任意满足$\sum_{\mathrm{i}}^{\mathrm{p}} \alpha_{\mathrm{i}} = 1$

的非负实数$\alpha_i (i \in \{1, 2, \cdots p\}$且$p \geq 2$的正整数$)$都有$\sum_{i=1}^{p} \alpha_i x_i \in F$.

证明：由凸集定义得定理的充分性。下用归纳法证明必要性。

$p = 2$时由定义知结论成立；

假设$p = k$时结论成立，下证$p = k + 1$时结论成立。

对$\forall x^i \in F, \sum_{i=1}^{k+1} \alpha_i = 1$的非负实数$\alpha_i (i \in \{1, 2, \cdots, k, k+1\})$, 则

$$\sum_{i=1}^{k+1} \alpha_i x^i = (\sum_{i=1}^{k} \alpha_i x^i) + \alpha_{k+1} x^{k+1} = (1 - \alpha_{k+1})(\sum_{i=1}^{k} \frac{\alpha_i}{1 - \alpha_{k+1}} x^i) + \alpha_{k+1} x^{k+1}$$

由$\sum_{i=1}^{k} \frac{\alpha_i}{1 - \alpha_{k+1}} = 1, \frac{\alpha_i}{1 - \alpha_{k+1}} \geq 0$及归纳假设知$\sum_{i=1}^{k} \frac{\alpha_i}{1 - \alpha_{k+1}} x^i \in F$,

因此$(1 - \alpha_{k+1})(\sum_{i=1}^{k} \frac{\alpha_i}{1 - \alpha_{k+1}} x^i) + \alpha_{k+1} x^{k+1} \in F$,

即$\sum_{i=1}^{k+1} \alpha_i x^i \in F$

有归纳原理知结论成立。

定义1.3.2

假设$F_1,F_2 \subseteq R^n$为两个非空凸集。如果存在非零向量$\omega \in R^n$和实数t, 使得

1)对$\forall x \in F_1, y \in F_2$都有$\omega^{\mathrm{T}} x \geq t$且$\omega^{\mathrm{T}} y \leq t$,则称超平面$\pi$:=$\{\mathrm{x} \in R^n | \omega^T \mathrm{x}=\mathrm{t}\}$分离集合$F_1$和$F_2$;

2)对$\forall x \in F_1, y \in F_2$都有$\omega^{\mathrm{T}} x > t$且$\omega^{\mathrm{T}} y < t$,则称超平面$\pi$:=$\{\mathrm{x} \in R^n | \omega^T \mathrm{x}=\mathrm{t}\}$严格分离集合$F_1$和$F_2$;

定理1.3.2 （点与凸集分离定理）

设F为R^n中的非空闭凸集，$x^0 \in R^n$且$x^0 \notin F$,则存在R^n中的超平面严格分离集合F和点$\{x^0\}$.

证明：记$\Omega:=\{y|y=x-x^0, \mathrm{x}\in F\}$。显然,$0\notin\Omega$.

对任意$y^1,y^2\in\Omega$, 存在$x^1,x^2\in F$使得$y^1=x^1-x^0; y^2=x^2-x^0$;

从而对$\forall\alpha\in[0,1], \alpha y^1+(1-\alpha)y^2=\alpha x^1+(1-\alpha)x^2-x^0$

由F为凸集知$\alpha x^1+(1-\alpha)x^2\in F$, 进而$\alpha y^1+(1-\alpha)y^2\in\Omega$.因此$\Omega$为凸集。

对$\forall\delta>0$, 记$N_\delta(0):=\{\mathrm{y}\in R^n|\|y\|_2\le\delta\}$。

选取适当的$\delta>0$, 使得$N_\delta(0)\cap\Omega$为非空有界闭集，

所以连续函数$\|y\|_2$在$N_\delta(0)\cap\Omega$可达到最小值。记其为$\hat{y}$,

则对任意$y\in\Omega$都有$\|\hat{y}\|_2\le\|y\|_2$.

因此对$\forall\varepsilon\in[0,1], \forall y\in\Omega, \|\hat{y}\|_2\le\|\varepsilon y+(1-\varepsilon)\hat{y}\|_2$,

即$\hat{y}^T\hat{y}\le(\varepsilon y+(1-\varepsilon)\hat{y})^T(\varepsilon y+(1-\varepsilon)\hat{y})=\varepsilon^2y^Ty+2\varepsilon(1-\varepsilon)\hat{y}^Ty+(1-\varepsilon)^2\hat{y}^T\hat{y}$

$=\varepsilon^2(y^Ty-2\hat{y}^Ty+\hat{y}^T\hat{y})+2\varepsilon\hat{y}^Ty+\hat{y}^T\hat{y}-2\varepsilon\hat{y}^T\hat{y}$

$=\varepsilon^2\|y-\hat{y}\|^2+2\varepsilon(y-\hat{y})^T\hat{y}+\hat{y}^T\hat{y}.$

从而$\varepsilon^2\|y-\hat{y}\|_2^2+2\varepsilon(y-\hat{y})^T\hat{y}\geq 0.$

由ε的任意性可得$(y-\hat{y})^T\hat{y}\geq 0$，即$y^T\hat{y}\geq\|\hat{y}\|_2^2.$

对$\forall x\in F$,记$y:=x-x^0$,则对$\forall \mathrm{x}\in F$都有$(x-x^0)^T\hat{y}\geq\|\hat{y}\|_2^2$,

即$x^T\hat{y}\geq(x^0)^T\hat{y}+\|\hat{y}\|_2^2.$

令$\omega=\hat{y},t=(x^0)^T\hat{y}+\dfrac{1}{2}\|\hat{y}\|_2^2$,则$\omega^{\mathrm{T}}x>t,\omega^Tx^0<t.$

因此超平面$\pi:=\{\mathrm{x}\in R^n:\omega^T\mathrm{x}=\mathrm{t}\}$可严格分离集合$F$与点$x^0$.

引例1.3.1（Farkas引理） 设$A \in R^{m \times n}, b \in R^{n}$,则不等式组

$Ax \le 0, b^T x > 0(1)$

$A^T y = b, y \ge 0(2)$

有且仅有一组有解。

证明：假设(2)有解，即存在$y \in R^m$使得$A^T y = b, y \ge 0.$

若(1)有解即存在$x \in R^n$使得$Ax \le 0,$则

$b^T x = (A^T y)^T x = y^T Ax \le 0.$

从而(2)有解则(1)必无解。

若(2)无解下证(1)有解。

记$\Omega := \{ z : A^T y \mid y \ge 0 \}$则$\Omega \subseteq R^n$为非空闭凸集且$b \notin \Omega.$

由定理1.3.2知，存在$\omega \in R^n, t \in R$使得

$\omega^T b > t, \omega^T z < t, \forall z \in \Omega.$由$0 \in \Omega$知$t > 0.$

$t > \omega^T z = \omega^T (A^T y) = y^T A\omega, y \ge 0.$

由y的任意性知，$A\omega \le 0, b^T w > 0.$从而$\omega$为(1)的解。

定理1.3.3 设p和q是两个非负整数，$u^0,u^1,\cdots,u^p,v^1,\cdots v^q\in R^n$,则等式与不等式组

$$d^Tu^0<0,\ d^Tu^i=0(i\in\{1,2,\cdots,p\}),\ d^Tv^i\geq 0(i\in\{1,2,\cdots,q\})(1)$$

无解

$\Leftrightarrow$ 存在实数$\alpha_i(i\in\{1,2,\cdots,p\})$和非负实数$\beta_i(i\in\{1,2,\cdots,\})$使得;

$$u^0=\sum_{i=1}^{p}\alpha_iu^i+\sum_{i=1}^{q}\beta_iv^i(2)$$

证明：$d^Tu^i=0\Leftrightarrow d^Tu^i\geq 0,(-d)^Tu^i\geq 0,(i=1,2,\cdots,p);$

令$A=(-u^1,\cdots,-u^p,u^1,\cdots,u^p,-v^1,\cdots,-v^q)^T$,则（1）可写成

$Ad\leq 0,(-u^0)^Td>0.$

由$Farkas$引理,(1)无解 $\Leftrightarrow$ 存在$(x,y,z)\in R_+^p\times R_+^p\times R_+^q$使得

$$-u^0=A^T\begin{pmatrix}x\\y\\z\end{pmatrix}=\sum_{i=1}^{p}x_i(-u^i)+\sum_{i=1}^{p}y_iu^i+\sum_{i=1}^{q}z_i(-v^i).$$

令$\alpha_i:=x_i-y_i(i=1,2,\cdots,p),\beta_i:=z_i\geq 0(i=1,\cdots,q)$则

$$u^0=\sum_{i=1}^{p}\alpha_iu^i+\sum_{i=1}^{q}\beta_iv^i.$$

1.3.2 凸函数

1 定义：设$F \subseteq R^n$为非空凸集，给定函数$f: F \to R$,

(1) 若对任意的$x, y \in F$及任意的$\alpha \in [0,1]$有:

$$f(\alpha x + (1-\alpha)y) \le \alpha f(x) + (1-\alpha) f(y)$$

则称函数f为凸集F上的凸函数

（2）若对任意的$x, y \in F$且$x \ne y$,及任意的$\alpha \in (0,1)$有:

$$f(\alpha x + (1-\alpha)y) < \alpha f(x) + (1-\alpha) f(y)$$

则称函数f为凸集F上的严格凸函数

(3) 若存在常数$c > 0$,使得对任意的$x, y \in F$及任意的$\alpha \in (0,1)$有：

$$f(\alpha x + (1-\alpha)y) \le \alpha f(x) + (1-\alpha) f(y) - c\alpha(1-\alpha)\|x-y\|^2$$

则称函数f为凸集F上的强凸函数（一致凸函数）

例. 给定向量$c \in R^n$.证明线性函数$f(x) = c^T x$即是凸函数又是凹函数.

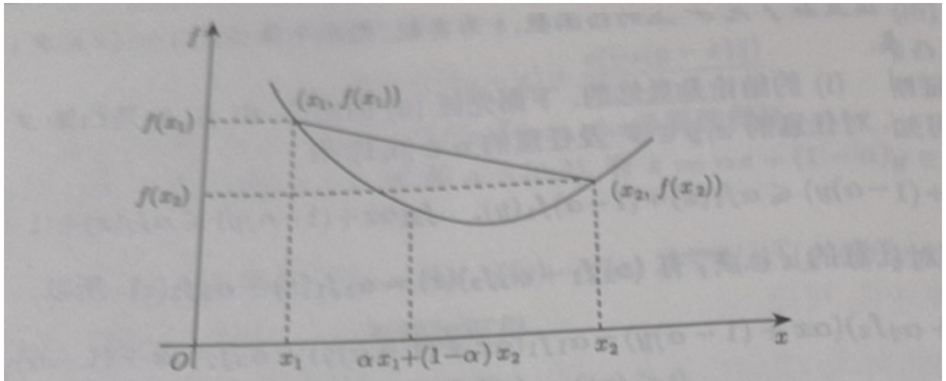
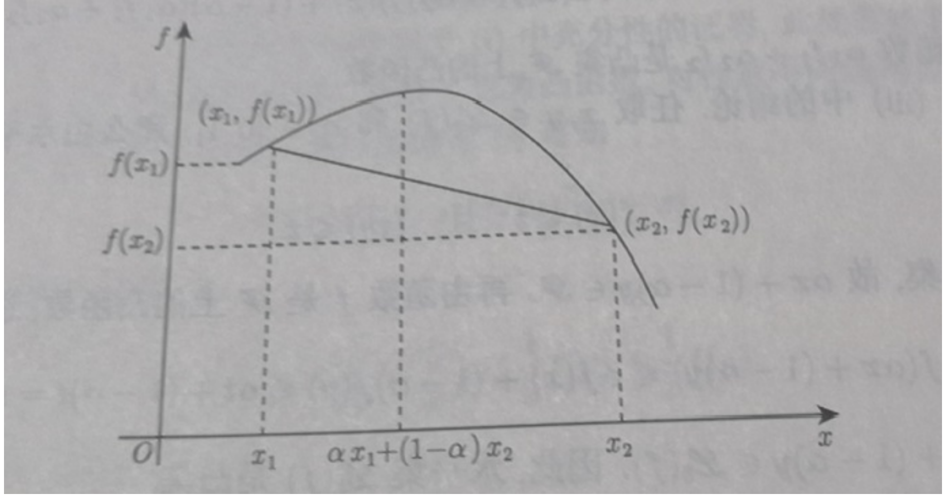

f
$(x_1, f(x_1))$
$f(x_1)$
$f(x_2)$
$(x_2, f(x_2))$
O
x_1
$\alpha x_1+(1-\alpha)x_2$
x_2
x

图 1-1


f
$(x_1, f(x_1))$
$f(x_1)$
$f(x_2)$
$(x_2, f(x_2))$
O
x_1
$\alpha x_1+(1-\alpha)x_2$
x_2
x

凸函数的性质

命题1.3.2

设$F \subseteq R^n$是凸集，则：

(1) 函数f是F上的（严格）凸函数$\Leftrightarrow -f$是F上的（严格）凹函数;

(2) 设函数f_1, f_2是F上的凸函数, 实数$\alpha_1, \alpha_2 \geq 0$，则函数$\alpha_1 f_1 + \alpha_2 f_2$是$F$上的凸函数;

(3) 设函数f是F上的凸函数, t为实数，则水平集

$$L_t(f) = \{x \in F : f(x) \leq t\}$$

是凸集.

证明：

(2)由f_1, f_2是凸集F上的凸函数知，对$\forall x, y \in F, \forall \alpha \in [0,1]$

$$f_1(\alpha x + (1-\alpha)y) \le \alpha f_1(x) + (1-\alpha) f_1(y);$$

$$f_2(\alpha x + (1-\alpha)y) \le \alpha f_2(x) + (1-\alpha) f_2(y);$$

由$(\alpha_1 f_1 + \alpha_2 f_2)(z) = \alpha_1 f_1(z) + \alpha_2 f_2(z)$得

$$\begin{aligned}
&(\alpha_1 f_1 + \alpha_2 f_2)(\alpha x + (1-\alpha)y) \\
&= \alpha_1 f_1(\alpha x + (1-\alpha)y) + \alpha_2 f_2(\alpha x + (1-\alpha)y) \\
&\le \alpha_1 \alpha f_1(x) + \alpha_1(1-\alpha) f_1(y) + \alpha_2 \alpha f_2(x) + \alpha_2(1-\alpha) f(y) \\
&= \alpha(\alpha_1 f_1 + \alpha_2 f_2)(x) + (1-\alpha)(\alpha_1 f_1 + \alpha_2 f_2)(y)
\end{aligned}$$

(3)$\forall x, y \in L_t(f), \forall \alpha \in [0,1], f(x) \le t, f(y) \le t.$

由F为凸集，$\alpha x + (1-\alpha)y \in F$

$$f(\alpha x + (1-\alpha)y) \le \alpha f(x) + (1-\alpha) f(y) \le t$$

所以$L_t(f)$为凸集。

两个判别准则

定理1.3.4 （一阶判别定理）

设函数f在凸集$F \subseteq R^n$上可微，则：

(1) f在F上为凸函数 $\Leftrightarrow$ 对任意的$x, y \in F$有:

$$f(y)-f(x) \geq \nabla f(x)^T (y-x).$$

(2) f在F上为严格凸函数 $\Leftrightarrow$ 对任意不同的$x, y \in F$有:

$$f(y)-f(x) > \nabla f(x)^T (y-x).$$

证明： （定理 1.3.4）

（1）$\Rightarrow \forall x, y \in F, \forall \alpha \in [0,1]$

$f(\alpha y + (1-\alpha)x) \le \alpha f(y) + (1-\alpha) f(x)$,即

$f(x + \alpha(y-x)) \le f(x) + \alpha[f(y) - f(x)]$.

由一阶泰勒展开式得：

$f(x+\alpha(y-x)) = f(x) + \alpha\nabla f(x)^T (y-x) + o(\|\alpha(y-x)\|)$.

所以$f(y) - f(x) \ge \nabla f(x)^T (y-x) + \dfrac{o(\|\alpha(y-x)\|)}{\alpha}$.

令$\alpha \to 0$得$f(y) - f(x) \ge \nabla f(x)^T (y-x)$.

$\Leftarrow \forall x, y \in F, \forall \alpha \in [0,1]$有$z := \alpha x + (1-\alpha) y \in F$。

$f(x) - f(z) \ge \nabla f(z)^T (x-z)$;

$f(y) - f(z) \ge \nabla f(z)^T (y-z)$;

两式分别乘α，（$1-\alpha$）并相加得：$\alpha f(x) + (1-\alpha) f(y) - f(z) \ge 0$

即$f(\alpha x + (1-\alpha) y) \le \alpha f(x) + (1-\alpha) f(y)$.

因此f在F上是凸函数。

(2) 充分性类似与 (1) 的证明。

"$\Rightarrow$" $f(x)$为严格凸函数则$f(x)$必为凸函数，从而

$\forall x, y \in F$且$x \neq y$, 令$z = \frac{1}{2}(x+\mathrm{y})$则$z \in F$且

$$f(z) - f(x) \geq \nabla f(x)^T (z - x).$$

$$f(z) = f(\frac{1}{2}x + \frac{1}{2}\mathrm{y}) < \frac{1}{2}f(x) + \frac{1}{2}f(\mathrm{y}).$$

所以$\frac{1}{2}f(x) + \frac{1}{2}f(\mathrm{y}) > f(x) + \nabla f(x)^T (z - x).$

整理即得

$$f(y) - f(x) > \nabla f(x)^T (y - x).$$

定理1. 3. 5 （二阶判别定理）

设在开凸集$F \subseteq R^n$内函数f二阶可微，则：

(1) f在F内为凸函数$\Leftrightarrow$任意的$x \in F$，$\nabla^2 f(x)$半正定.

(2) 若$\forall x \in F$，$\nabla^2 f(x)$正定，则f在F内为严格凸函数.

证明：定理1.3.5

(1) $\Rightarrow \forall x \in F, 0 \neq y \in R^n$,由$F$是开集知存在$\varepsilon > 0$使得

$\forall \alpha \in (-\varepsilon, \varepsilon), x + \alpha y \in F.$

由定理1.3.4,$f(x+\alpha y) \geq f(x) + \alpha \nabla f(x)^T y.$

由$Taylor$展式得，

$$f(x+\alpha y) = f(x) + \alpha \nabla f(x)^T y + \frac{1}{2}\alpha^2 y^T \nabla^2 f(x) y + o(\|\alpha y\|^2).$$

所以$y^T \nabla^2 f(x) y + \dfrac{o(\|\alpha y\|^2)}{2\alpha^2} \geq 0.$

令$\alpha \to 0$得$y^T \nabla^2 f(x) y \geq 0$.由$y$的任意性得$\nabla^2 f(x)$正半定。

$\Leftarrow \forall x, y \in F$及$\nabla^2 f(x)$对称正半定得

$$f(y) = f(x) + \nabla f(x)^T (y-x) + \frac{1}{2}(y-x)^T \nabla^2 f(\xi)(y-x)$$

$\geq f(x) + \nabla f(x)^T (y-x)$,其中$\xi = x + t(y-x), t \in (0,1).$

因此f在F上是凸函数。

定理1.3.6 （强凸函数的判定定理）

设 $f: R^n \to R$ 二次连续可微，则：

f是强凸函数 $\Leftrightarrow \nabla^2 f(x)$一致正定

注：$\nabla^2 f(x)$一致正定：

即存在常数$c > 0$使得：$d^T \nabla^2 f(x) d \geq c\|d\|^2, (\forall x, d \in R^n$）

1.3.3 凸规划

1．定义：设$F \subseteq R^n$为凸集，$f: F \to R$为凸函数，则称

$$\min_{x \in F} f(x) \text{为凸规划问题。}$$

2　凸规划的性质：

定理　1.3.7

(1)　凸规划问题的任一局部最优解x^*为其全局最优解；

(2)　凸规划问题的最优解集S为凸集；

(3)　若函数f为非空凸集F上的严格凸函数，且凸规划问题存在全局最优解，则其全局最优解唯一。

证明:

(1)(反证法)假设x^*是凸优化问题的局部但非全局最优解，则至少存在一个$y^* \in F$使得$f(y^*) < f(x^*)$.

由函数f为凸函数，F为凸集得，对任意$\alpha \in [0,1]$,

$$\begin{cases} \alpha y^* + (1-\alpha)x^* \in F; \\ f(\alpha y^* + (1-\alpha)x^*) \le \alpha f(y^*) + (1-\alpha) f(x^*) < f(x^*). \end{cases}$$

当$\alpha \to 0^+$时，$\alpha y^* + (1-\alpha)x^* \to x^*$.

因此在充分靠近x^*时$f(\alpha y^* + (1-\alpha)x^*) < f(x^*)$这与$x^*$为局部最优点矛盾。

(2)由空集与单元素集合为凸集，不妨设凸规划问题的最优解集S至少含两个元素. 假设$\mathrm{x}^*, y^* \in S$,则$\forall \alpha \in [0,1]$有$\mathrm{x}^*, y^* \in F$且

$$f(\alpha x^* + (1-\alpha)y^*) \le \alpha f(x^*) + (1-\alpha) f(y^*) = f(x^*) = \min_{x\in F} f(x).$$

因此$\alpha x^* + (1-\alpha)y^* \in S$.

(3)(反证法)假设全局最优点不唯一，则$\exists x^*, y^* \in S$且$x^* \neq y^*$.

显然$x^*, y^* \in F$且$f(x^*) = f(y^*)$.

$\forall \alpha \in [0,1], \alpha x^* + (1-\alpha) y^* \in S$且

$f(\alpha x^* + (1-\alpha) y^*) < \alpha f(x^*) + (1-\alpha) f(y^*) = f(x^*)$

这与x^*为全局最优点矛盾，所以凸规划问题有唯一最优解。