集合

x^0的ε-邻域：$N_\varepsilon\left(x^0\right)=\{x \mid \|x-x^0\|<\varepsilon, \varepsilon>0\}$

内点： 设$x^0\in S\subset\Re^n$，若存在$\varepsilon>0$，使得$N_\varepsilon(x^0)\subset S$，

则称x^0为S的一个内点。

补集： 集合S的补集定义为$S^C=\{x \mid x\notin S, x\in\Re^n\}$

开集： 若对$\forall x\in S, x$为内点，则称S为开集。

闭集： 若集合S的补集S^C为开集，则称S为闭集。

有界集 ：若存在正数$M>0$，使得$\forall x\in S, \|x\|\le M$成立，

则称S为有界集。

紧集： 有界闭集称为紧集.

性质：

（1）集合$S \subset \Re^n$是闭集，当且仅当对任意的无穷序列$\{x^k\} \subset S$，若$x^k \to x^*$，则$x^* \in S$。

（2）集合$S \subset \Re^n$是紧集当且仅当对任意的无穷序列$\{x^k\} \subset S$，必存在收敛于S中点的子序列$\{x^{k_i}\}$.

定义：设$S \subseteq E^n$，若对$\forall x^{(1)}, x^{(2)} \in S$及$\forall \lambda \in [0,1]$，都有

$$\lambda x^{(1)} + (1-\lambda) x^{(2)} \in S$$

则称S为凸集。

序列的极限

- 定义：设$\{x^{(k)}\}$是R^n中的一个向量序列，$\bar{x}\in R^n$,如果对任给的$\varepsilon>0$存在正整数K_ε，使得当$k>K_\varepsilon$时就有$\|x^{(k)}-\bar{x}\|<\varepsilon$,则称序列收敛到 $\bar{x}$ 或称序列以$\bar{x}$为极限，记作$\lim\limits_{k\to\infty} x^{(k)}=\bar{x}.$
- 结论：序列若存在极限，则任何子序列有相同的极限，即序列极限是唯一的。
- 定义：设$\{x^{(k)}\}$是R^n中的一个向量序列，如果存在一个子序列$\left\{x^{(k_j)}\right\}$，使得$\lim\limits_{k\to\infty} x^{(k_j)}=\hat{x}$，则称 $\hat{x}$ 是序列$\{x^{(k)}\}$的一个聚点。

- 定义：设$\{x^{(k)}\}$是R^n中的一个向量序列，如果对任给定的ε>0，总存在正整数K_ε，使得当m, $k>K_\varepsilon$是就有$\|x^{(m)}-x^{(k)}\|<\varepsilon$，则称序列$\{x^{(k)}\}$为**Cauchy**序列。

- 定理：设$\{x^{(j)}\}\subset R^n$为**Cauchy**序列，则$\{x^{(j)}\}$的聚点必为极限点。

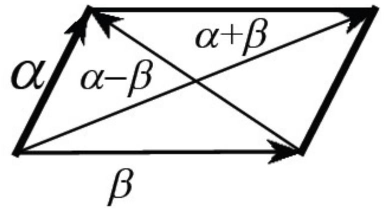
凸集分离定理

定义：

设S_1和S_2是E^n中两个非空集合，$H=\{x \mid p^T x=\alpha\}$为超平面，如果对$\forall x \in S_1$，都有$p^T x \geq \alpha$，对$\forall x \in S_2$，都有$p^T x \leq \alpha$（或情形恰好相反），则称超平面$H$分离集合$S_1$和$S_2$。

定理1：设S为E^n的闭凸集，$y \notin S$，则存在唯一的$\bar{x} \in S$，使得

$$\| y-\bar{x} \| = \inf_{x\in S} \| y-x \| > 0。$$



$\bar{x}$是这一最小距离点$\Leftrightarrow$对$\forall x \in S$，有

$$(y-\bar{x})^T(\bar{x}-x) \ge 0。$$

$$\boxed{\|\alpha+\beta\|^2 + \|\alpha-\beta\|^2 = 2(\|\alpha\|^2 + \|\beta\|^2)}$$

证明：令 $\inf_{x\in S} \| y-x \| = r > 0$

$\Rightarrow \exists$序列$\{x^{(k)}\}, x^{(k)} \in S$,使得$\| y - x^{(k)} \| \to r$。

先证$\{x^{(k)}\}$为Cauchy序列。

$$\| x^{(k)} - x^{(m)} \|^2$$

$$= 2\| x^{(k)} - y \|^2 + 2\| x^{(m)} - y \|^2 - 4\left\| \frac{x^{(k)} + x^{(m)}}{2} - y \right\|^2$$

$$\le 2\| x^{(k)} - y \|^2 + 2\| x^{(m)} - y \|^2 - 4r^2$$

$$\to 0 \quad (当 m \to \infty, k \to \infty)$$

$\therefore \{x^{(k)}\}$为Cauchy序列，$\Rightarrow \{x^{(k)}\}$极限存在，设为$\bar{x}$,

$\because S$为闭集，$\therefore \bar{x} \in S$.

定理1： 设S为E^n的闭凸集，$y \notin S$，则存在唯一的$\bar{x} \in S$，使得

$$\| y - \bar{x} \| = \inf_{x \in S} \| y - x \| > 0。$$

证明： 假设存在$\hat{x} \in S$，使得$\| y - \bar{x} \| = \| y - \hat{x} \| = r$

$\because S$为凸集，$\bar{x}, \hat{x} \in S$，$\quad \therefore \quad \dfrac{\bar{x} + \hat{x}}{2} \in S.$

$$\therefore \quad r \le \left\| y - \frac{\bar{x} + \hat{x}}{2} \right\| \le \frac{1}{2} \| y - \bar{x} \| + \frac{1}{2} \| y - \hat{x} \| = r$$

$$\Rightarrow \quad \left\| y - \frac{\bar{x} + \hat{x}}{2} \right\| = \frac{1}{2} \| y - \bar{x} \| + \frac{1}{2} \| y - \hat{x} \|$$

$$\Rightarrow \quad y - \bar{x} = \lambda (y - \hat{x}) \quad \lambda \ne 0$$

$$\boxed{\| \alpha + \beta \| \le \| \alpha \| + \| \beta \|}$$

$$\Rightarrow \quad \| y - \bar{x} \| = |\lambda| \| y - \hat{x} \| \Rightarrow \quad \lambda = \pm 1$$

若$\lambda = -1$，则$y = \dfrac{\bar{x} + \hat{x}}{2} \in S$，矛盾

所以$\lambda = 1 \quad \Rightarrow \quad \bar{x} = \hat{x}.$

定理1： $\bar{x}$是这一最小距离点 $\Leftrightarrow$ 对$\forall x \in S$，有

$$(y-\bar{x})^T(\bar{x}-x) \geq 0。$$

证明: “$\Rightarrow$” 假设$\bar{x}$是最小距离点,则对$\forall x \in S$，有

$$\|y-x\|^2 \geq \|y-\bar{x}\|^2.$$

$\because S$是凸集,$\therefore \forall \lambda \in (0,1)$,有$\bar{x}+\lambda(x-\bar{x}) \in S.$

$$\therefore \quad \|y-(\bar{x}+\lambda(x-\bar{x}))\|^2 \geq \|y-\bar{x}\|^2,$$

$$\because \|y-(\bar{x}+\lambda(x-\bar{x}))\|^2 = \|y-\bar{x}\|^2 + \lambda^2\|x-\bar{x}\|^2 - 2\lambda(y-\bar{x})^T(x-\bar{x})$$

$$\therefore \quad \lambda^2\|x-\bar{x}\|^2 - 2\lambda(y-\bar{x})^T(x-\bar{x}) \geq 0$$

$$\Rightarrow \quad \lambda\|x-\bar{x}\|^2 - 2(y-\bar{x})^T(x-\bar{x}) \geq 0$$

令$\lambda \to 0$,得 $\quad -2(y-\bar{x})^T(x-\bar{x}) \geq 0$

$$\therefore \quad (y-\bar{x})^T(\bar{x}-x) \geq 0.$$

定理1： $\bar{x}$是这一最小距离点 $\Leftrightarrow$ 对$\forall x \in S$，有

$$(y-\bar{x})^T(\bar{x}-x) \ge 0。$$

证明：“$\Leftarrow$” 假设对任意的$x \in S, (y-\bar{x})^T(\bar{x}-x) \ge 0$,

则对任意的$x \in S$，有

$$\|y-x\|^2 = \|y-\bar{x}+\bar{x}-x\|^2$$

$$= \|y-\bar{x}\|^2 + \|\bar{x}-x\|^2 + 2(y-\bar{x})^T(\bar{x}-x)$$

$$\ge \|y-\bar{x}\|^2$$

$\therefore$ $\bar{x}$是最小距离点。

定理2： 设S是E^n的非空闭凸集，$y \notin S$，则存在非零向量p及数$\varepsilon > 0$，使得对$\forall x \in S$，有$p^T y \geq \varepsilon + p^T x$.

证明： $\because$ S为闭凸集，$y \notin S$，由定理1，$\exists \bar{x} \in S$，使

$$\|y - \bar{x}\| = \inf_{x \in S} \|y - x\| > 0$$

令$p = y - \bar{x} \neq 0$，$\varepsilon = p^T (y - \bar{x}) = \|y - \bar{x}\|^2 > 0$

$$\because \quad p^T (y - x) = p^T (y - \bar{x} + \bar{x} - x)$$

$$= p^T (y - \bar{x}) + p^T (\bar{x} - x)$$

$$= \varepsilon + (y - \bar{x})^T (\bar{x} - x) \geq \varepsilon$$

$$\therefore \quad p^T y \geq \varepsilon + p^T x.$$

定理3： 设S是E^n的非空凸集，$y \in \partial S$，则存在非零向量 p，使得对$\forall x \in clS$(S的闭包，由S的内点和边界点组成)，有$p^T y \geq p^T x$.

证明： $\because S$是凸集，$\therefore clS$是闭凸集。

$\because y \in \partial S$，则存在序列$\{y^{(k)}\} \notin clS$，使得$y^{(k)} \to y$.

对每个点$y^{(k)}$，由定理2，存在单位向量$p^{(k)}$,

使得对每个$x \in clS$，有$\left(p^{(k)}\right)^T y^{(k)} > \left(p^{(k)}\right)^T x$.

$\because$ 序列$\{p^{(k)}\}$有界(单位向量)，$\therefore$ 存在收敛的子序列$\{p^{(k_j)}\}$，其极限为单位向量p.

$\because \left(p^{(k_j)}\right)^T y^{(k_j)} > \left(p^{(k_j)}\right)^T x$对每个$x \in clS$成立，

$\therefore$ 令$k_j \to \infty$，得到$p^T y \geq p^T x$，$\forall x \in clS$.

推论4: 设S是E^n的非空凸集，$y \notin S$，则存在非零向量 p，使得对$\forall x \in clS$(S的闭包，由S的内点和边界点组成)，有$p^T(x-y) \le 0$.

定理5: 设S_1和S_2是E^n的两个非空凸集，$S_1 \cap S_2 = \varnothing$，则存在非零向量$p$，使得

$$\inf\{\, p^T x \mid x \in S_1\} \ge \sup\{\, p^T x \mid x \in S_2\}.$$

(或 $p^T y \ge p^T x$ 对$\forall y \in S_1, \forall x \in S_2$成立)

证明: 设$S = S_2 - S_1 = \{x^{(2)} - x^{(1)} \mid x^{(1)} \in S_1, x^{(2)} \in S_2\}$

$\because$ S_1, S_2是非空凸集,

$\therefore$ S是凸集且$S \ne \varnothing$.

$\because$ $S_1 \cap S_2 = \varnothing$, $\therefore$ $0 \notin S$

$\Rightarrow$ 存在$p \ne 0$, 对$\forall x \in S$, 有$p^T(x-0) \le 0$

$\Rightarrow$ $p^T x^{(2)} \le p^T x^{(1)}$

定理6: 设S_1和S_2是E^n的两个非空闭凸集，S_1有界，且

$S_1 \cap S_2 = \varnothing$，则存在非零向量$p$和$\varepsilon > 0$，使得

$$\inf\{p^T x \mid x \in S_1\} \geq \varepsilon + \sup\{p^T x \mid x \in S_2\}.$$

(或 $p^T y \geq \varepsilon + p^T x$ 对$\forall y \in S_1, \forall x \in S_2$成立)

证明　设$S = S_2 - S_1 = \{x^{(2)} - x^{(1)} \mid x^{(1)} \in S_1, x^{(2)} \in S_2\}$

则S是非空凸集且$0 \notin S$.

假设存在序列$\left\{x^{(k)} \in S\right\} \to x$,则存在序列

$\left\{y^{(k)} \in S_2\right\}$和$\left\{z^{(k)} \in S_1\right\}$,使得$x^{(k)}=y^{(k)} - z^{(k)}, k = 1, 2, \cdots$

$\because S_1$是有界闭集,$\therefore \left\{z^{(k)}\right\}$存在收敛的子列，不妨设

$z^{(k)} \to z \in S_1$, 则有$y^{(k)}=x^{(k)} + z^{(k)} \to x + z$

$\because S_2$是闭集,$\therefore x + z \in S_2$

$\Rightarrow \quad x = (x + z) - z \in S_2 - S_1 = S \Rightarrow S$是闭集.

由定理2，结论成立.

Farkas定理：设A为$m\times n$矩阵，c为n维列向量，则$Ax\le 0$，$c^T x>0$有解的充要条件是$A^T y=c$，$y\ge 0$无解。

证明："$\Rightarrow$" 设存在$y\ge 0$，使得$A^T y=c$

得 $y^T A=c^T$

设$\bar{x}$为$Ax\le 0, c^T x>0$的一个解，

则有$A\bar{x}\le 0, c^T\bar{x}>0$

$$\Rightarrow \quad y^T A\bar{x}=c^T\bar{x}>0 \qquad (1)$$

$\because \quad y\ge 0$, 但$A\bar{x}\le 0$

$\therefore \quad y^T A\bar{x}\le 0$ 与（1）矛盾。

Farkas定理：设A为$m \times n$矩阵，c为n维列向量，则

$Ax \le 0$，$c^T x > 0$有解的充要条件是$A^T y = c$，$y \ge 0$无解。

“$\Leftarrow$” 设$A^T y = c, y \ge 0$无解，令

$$S = \{z \mid z = A^T y, y \ge 0\}, \text{则} c \notin S$$

可以证明S为闭凸集，由定理2，$\exists x \ne 0, \varepsilon > 0,$

使得对$\forall z \in S,$有$x^T c \ge \varepsilon + x^T z$

$\because \varepsilon > 0, \quad \therefore \quad x^T c > x^T z$

$\Rightarrow c^T x > z^T x = y^T Ax$

即对任意的$y \ge 0$，有$c^T x > y^T Ax$ （1）

令$y = 0$，得$c^T x > 0$

$\because c^T x$为一定数，y的分量可取任意大，

$\therefore$由（1），必有$Ax \le 0.$

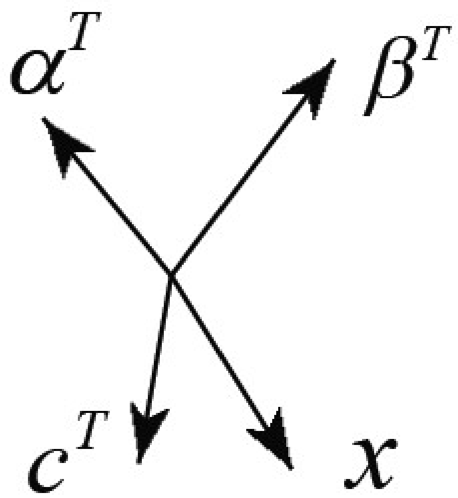
既非零向量x是$Ax \le 0$，$c^T x > 0$的解。

Farkas定理：设A为$m\times n$矩阵，c为n维列向量，则 $Ax\le 0$，$c^T x>0$有解的充要条件是$A^T y=c$，$y\ge 0$无解。

$$\max c^T x$$
$$s.t.\quad Ax\le 0$$

$$\min 0^T y$$
$$s.t.\quad A^T y=c$$
$$y\ge 0$$

$$A=\begin{pmatrix}\alpha^T\\ \beta^T\end{pmatrix}$$



Farkas 定理中说明 S={z|z=Ay, y≥0}是闭集

设 $A_{m\times n}$， $e_1,e_2,\ldots,e_n$ 是 R^n 的自然基，令 $f_j=Ae_j, 1\le j\le n$，

则 $S=\{z \mid z=Ay=y_1f_1+y_2f_2+\cdots+y_nf_n, y_i\ge 0\}$。

以下对 ***n*** 归纳。

***n*=1** 时，显然成立。

假设对 ***n*≥1**，结论成立，考虑

$$S=\{z \mid z=Ay=y_1f_1+y_2f_2+\cdots+y_{n+1}f_{n+1}, y_i\ge 0\}$$

若 $f_1,f_2,\ldots,f_{n+1}$ 线性无关，则序列 $\{z_k=Ay^{(k)}\}$ 收敛当且仅当对所有的 **1≤*j*≤*n*+1** 序列 $\{y_j^{(k)}\}$ 收敛，所以 ***S*** 的闭集。

假设$f_1,f_2,\ldots,f_{n+1}$线性相关，则存在不全为零的数（且其中至少有一个为正数）$\alpha_1,\alpha_2,\ldots,\alpha_{n+1}$使得

$$\sum_{i=1}^{n+1}\alpha_i f_i = 0$$
。

对任意的$z \in S = \{z \mid z = Ay = \sum_{i=1}^{n+1} y_i f_i, y_i \geq 0\}$，令

$$t(y) = \min_{\substack{1 \leq j \leq n+1 \\ \alpha_j > 0}} \frac{y_j}{\alpha_j} = \frac{y_{i(y)}}{\alpha_{i(y)}} \geq 0$$

则$y_{i(y)} - t(y)\alpha_{i(y)} = 0$，且

$$z = \sum_{i=1}^{n+1} y_i f_i = \sum_{i=1}^{n+1} (y_i - t(y)\alpha_i) f_i$$

$$= \sum_{\substack{1 \leq i \leq n+1 \\ i \neq i(y)}} (y_i - t(y)\alpha_i) f_i$$

其中$y_j - t(y)\alpha_j \geq 0$（**1≤*j*≤*n*+1**）。

对任意的 $\mathbf{1}\leqslant \boldsymbol{j}\leqslant \boldsymbol{n}+\mathbf{1}$，令

$$S_i=\left\{z \mid z=Ay=\sum_{\substack{1\le j\le n+1\\ j\ne i}} y_j f_j, y_i \ge 0\right\}$$

则 $S=\bigcup_{i=1}^{n+1} S_i$ ，由归纳假设，S_i 是闭集，所以 $\boldsymbol{S}$ 是闭集。

Gordan定理：设A为$m\times n$矩阵，那么$Ax<0$有解的充要条件是不存在非零向量$y\geq 0$，使得$A^T y=0$。

证明：

"$\Rightarrow$"　设存在$\bar{x}$，使得$A\bar{x}<0$

若存在非零向量$y\geq 0,$使得$A^T y=0$

则有$y^T A=0,\quad \Rightarrow y^T A\bar{x}=0$

$\because A\bar{x}<0$

$\therefore y$的各分量不可能为非负数,与$y\geq 0$矛盾.

“$\Leftarrow$”（证等价命题）即若$Ax<0$无解，则存在非零向量$y\geq 0$,使得$A^T y=0$.设$Ax<0$无解，令

$S_1=\{z \mid z=Ax, x\in E^n\} \quad S_2=\{z \mid z<0\}$

$\because Ax<0$无解, $\therefore S_1\cap S_2=\varnothing$,由凸集分离定理知，

存在非零向量y，使得对$\forall x\in E^n, \forall z\in S_2$，有

$$y^T Ax \geq y^T z \qquad (1)$$

特别地，当$x=0$时，有$y^T z\leq 0$。

$\because z<0$，它的分量可取任意负数，$\therefore y\geq 0$

在（1）中令$z\to 0$，则对$\forall x\in E^n$,有

$$y^T Ax\geq 0 \qquad (2)$$

令$x=-A^T y$，代入（2），得$-y^T AA^T y\geq 0$

即$-\left\|A^T y\right\|\geq 0, \quad \therefore \quad A^T y=0.$

因此，存在非零向量$y\geq 0$,使得$A^T y=0.$

Gordan定理： 设A为$m\times n$矩阵，那么$Ax<0$有解的充要条件是不存在非零向量$y\geq 0$，使得$A^T y=0$。

证明：

"$\Leftarrow$"设$A^T y=0, y\geq 0 (y\neq 0)$无解，令

$$S=\left\{z \mid z=A^T y, y\geq 0\right\}，则0\notin S.$$

可以证明S是闭凸集，则由点与凸集的分离定理，

$\exists x\neq 0, \varepsilon>0$，使得对$\forall z\in S$，都有$x^T 0\geq \varepsilon+x^T z$

$\because \varepsilon>0, \therefore$ 对$\forall z\in S$，都有$x^T z<0$

$\because z=A^T y, \therefore$ 对$\forall y\geq 0$，有$y^T Ax<0$成立，

$\Rightarrow \quad Ax<0,$

即存在x，使$Ax<0$有解.

证法正确吗？

Gordan定理: 设A为$m\times n$矩阵，那么$Ax<0$有解的充要条件是不存在非零向量$y\geq 0$，使得$A^T y=0$。

证明：（$\Leftarrow$）：设 $A^T y=0, y\geq 0 (y\neq 0)$ 无解，

令 $S=\{z|z=A^T y, y\geq \delta\}$(其中$\delta>0$), 则 $0\notin S$.

令 $S_1=\{z|z=A^T y, y\geq 0\}$，$S_1$ 是一个闭集。

设序列 $z^{(k)}\in S$ 且 $\lim z^{(k)}=\bar{z}$，则$\exists y^{(k)}\geq\delta$,使得

$$z^{(k)}-A^T\delta=A^T(y^{(k)}-\delta)\in S_1$$

令$z_1^{(k)}=z^{(k)}-A^T\delta$，$\because z^{(k)}$ 有极限，$\therefore z_1^{(k)}$ 也有极限，

且$\lim z_1^{(k)}=\bar{z}-A^T\delta$

$\because z_1^{(k)}\in S_1$ 且 S_1为闭集 $\therefore \bar{z}-A^T\delta\in S_1$ 故 $\exists\bar{y}\geq 0$，使得

$\bar{z}-A^T\delta=A^T\bar{y}$ 即 $\bar{z}=A^T(\delta+\bar{y})\in S$, 故$S$为闭集。

由点与凸集的分离定理，

$\exists x \neq 0, \varepsilon > 0$，使得$\forall z \in S$，都有 $x^T 0 \geq \varepsilon + x^T z$

$\because \varepsilon > 0$， $\therefore \forall z \in S$，都有$x^T z < 0$

$\because z = A^T y$， $\boxed{\therefore \forall y \geq \delta\text{，都有}y^T Ax < 0\text{成立，}}$

$\boxed{\text{从而可以推出 } Ax < 0\text{，}}$ 即存在x，使得 $Ax < 0$有解.

这个结论不成立！

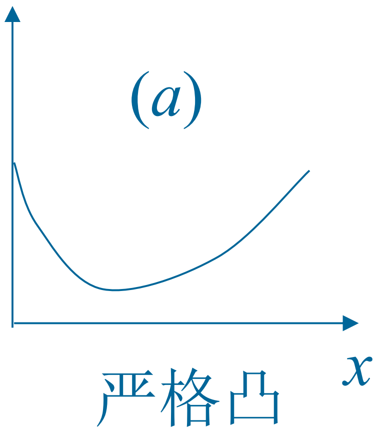
凸函数

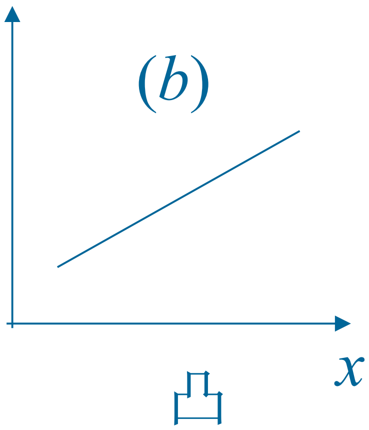
凸函数： 设S是E^n中的非空凸集，$f(x)$是定义在S上的实函数，如果对于每一对x_1，$x_2 \in S$及每一个a，$0 \le a \le 1$,都有

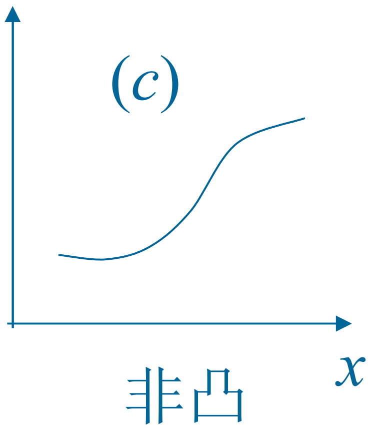
$$f(ax_1+(1-a)x_2) \le a f(x_1)+(1-a)f(x_2)$$

则称函数$f(x)$为S上的凸函数．上式中，若$\le$变为$<$，则称为严格凸函数。

若$-f(x)$为S的凸函数，则称$f(x)$为S上的凹函数.







凸函数性质

(1) 设$f_1(x)$，$f_2(x)$是凸集S上的凸函数，则函数$f_1(x)+f_2(x)$在S上也是凸函数。

(2) 设$f(x)$是凸集S上的凸函数，则对任意的$a\geq 0$，函数$af(x)$是凸的。

推广：设$f_1(x)$，$f_2(x)$, …, $f_k(x)$是凸集S上的凸函数，$a_i\geq 0$,则$a_1f_1(x)+a_2f_2(x)+ \ldots + a_kf_k(x)$也是凸集$S$上的凸函数.

(3) 设$f(x)$是凸集S上的凸函数，对每一个实数c，则集合

$$S_c=\{x \mid x\in S，f(x)\leq c\}$$是凸集。

(4)设S是E^n中的非空凸集，f是定义在S上的凸函数，则f在S上的局部极小点是整体极小点，且极小点的集合是凸集.

证明：设$\bar{x}$是f在S中的局部极小点，即存在$\bar{x}$的$\varepsilon > 0$邻域$N_\varepsilon(\bar{x})$

使得对$\forall x \in S \cap N_\varepsilon(\bar{x})$，有$f(x) \geq f(\bar{x})$。

若$\bar{x}$不是整体极小点，则$\exists x^{(0)} \in S$使$f(\bar{x}) > f(x^{(0)})$,

$\because$ S是凸集，$\therefore$ 对$\forall \lambda \in (0,1)$有$\lambda x^{(0)} + (1-\lambda)\bar{x} \in S$,

$\because$ f是S上的凸函数，

$$\therefore\ f(\lambda x^{(0)} + (1-\lambda)\bar{x}) \leq \lambda f(x^{(0)}) + (1-\lambda) f(\bar{x})$$

$$< \lambda f(\bar{x}) + (1-\lambda) f(\bar{x}) = f(\bar{x})$$

当λ取得充分小时，可使$\lambda x^{(0)} + (1-\lambda)\bar{x} \in S \cap N_\varepsilon(\bar{x})$,

这与$\bar{x}$为局部极小点矛盾，

$\therefore$ $\bar{x}$是f在S上的整体极小点。

f在S上的极小值也是它在S上的最小值。

极小点集合为$\Gamma_\alpha = \{x \mid x \in S, f(x) \leq \alpha\}$,

则由性质（3），Γ_α为凸集。

凸函数的判别

梯度： $\nabla f(x)=\left(\frac{\partial f(x)}{\partial x_1},\cdots,\frac{\partial f(x)}{\partial x_n}\right)^T$

Hesse矩阵：

$$\nabla^2 f(x)=\begin{bmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \cdots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(x)}{\partial x_2 \partial x_1} & \cdots & \frac{\partial^2 f(x)}{\partial x_2 \partial x_n} \\ \cdots\cdots\cdots\cdots & & \\ \frac{\partial^2 f(x)}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f(x)}{\partial x_n^2} \end{bmatrix}$$

$$f(x)=\frac{1}{2}x^T Ax+b^T x+c$$

$$\nabla f(x)=Ax+b$$

$$\nabla^2 f(x)=A$$

方向导数

设$x^0 \in E^n$, $p \in E^n$, $p \neq 0$，则函数$f(x)$在点x^0关于方向p的方向导数定义为：

$$\frac{\partial f(x^0)}{\partial p} = \lim_{t \to 0^+} \frac{f(x^0 + tp) - f(x^0)}{t}.$$

我们用$Df(x^0; p)$表示f在点x^0关于方向p的方向导数。

方向导数通常用下面的公式计算：

$$Df(x^0; p) = \nabla f(x^0)^T p.$$

定理(一阶充要条件): 设S是E^n中非空开凸集，$f(x)$是定义在S上的可微函数，则$f(x)$为凸函数的充要条件是对任意两点$x^{(1)},x^{(2)}\in S$，有

$$f(x^{(2)})\geq f(x^{(1)})+\nabla f(x^{(1)})^T(x^{(2)}-x^{(1)});$$

$f(x)$为严格凸函数的充要条件是对任意互不相同两点$x^{(1)},x^{(2)}\in S$，有

$$f(x^{(2)})> f(x^{(1)})+\nabla f(x^{(1)})^T(x^{(2)}-x^{(1)}).$$

证明: “$\Rightarrow$” 设f是S上的凸函数，则对$\forall x^{(1)},x^{(2)}\in S$及$\lambda\in(0,1)$，有

$$f(\lambda x^{(2)}+(1-\lambda)x^{(1)})\leq\lambda f(x^{(2)})+(1-\lambda)f(x^{(1)})$$

即$$\frac{f(x^{(1)}+\lambda(x^{(2)}-x^{(1)}))-f(x^{(1)})}{\lambda}\leq f(x^{(2)})-f(x^{(1)})$$

令$\lambda\to 0^+$，由f的可微性，得f在点$x^{(1)}$关于方向$x^{(2)}-x^{(1)}$的方向导数

$$Df(x^{(1)};x^{(2)}-x^{(1)})=\nabla f(x^{(1)})^T(x^{(2)}-x^{(1)})\leq f(x^{(2)})-f(x^{(1)}),$$

$$\Rightarrow f(x^{(2)})\geq f(x^{(1)})+\nabla f(x^{(1)})^T(x^{(2)}-x^{(1)}).$$

"$\Rightarrow$" 当f是S上的严格凸函数时，对$\forall x^{(1)}, x^{(2)} \in S$及$x^{(1)} \neq x^{(2)}$，

取$\lambda = \dfrac{1}{2}$，则$y = \dfrac{1}{2}x^{(1)} + \dfrac{1}{2}x^{(2)} \in S$且

$$f(y) = f\left(\frac{1}{2}x^{(1)} + \frac{1}{2}x^{(2)}\right) < \frac{1}{2}f(x^{(1)}) + \frac{1}{2}f(x^{(2)}).$$

$\because$ f是凸函数，

$$\therefore \quad f(y) \geq f(x^{(1)}) + \nabla f(x^{(1)})^T (y - x^{(1)}).$$

$$\therefore \quad \frac{1}{2}f(x^{(1)}) + \frac{1}{2}f(x^{(2)}) > f(x^{(1)}) + \nabla f(x^{(1)})^T (y - x^{(1)})$$

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

$$\therefore \quad f(x^{(2)}) > f(x^{(1)}) + \nabla f(x^{(1)})^T (x^{(2)} - x^{(1)}).$$

“$\Leftarrow$” 设对$\forall x^{(1)}, x^{(2)} \in S$，有

$$f(x^{(2)}) \geq f(x^{(1)}) + \nabla f(x^{(1)})^T (x^{(2)} - x^{(1)})。$$

$\forall \lambda \in (0,1)$，令$y = \lambda x^{(1)} + (1-\lambda) x^{(2)}$，则$y \in S$。

由假设，对$x^{(1)}, y \in S$及$x^{(2)}, y \in S$有

$$f(x^{(1)}) \geq f(y) + \nabla f(y)^T (x^{(1)} - y)$$

$$f(x^{(2)}) \geq f(y) + \nabla f(y)^T (x^{(2)} - y)$$

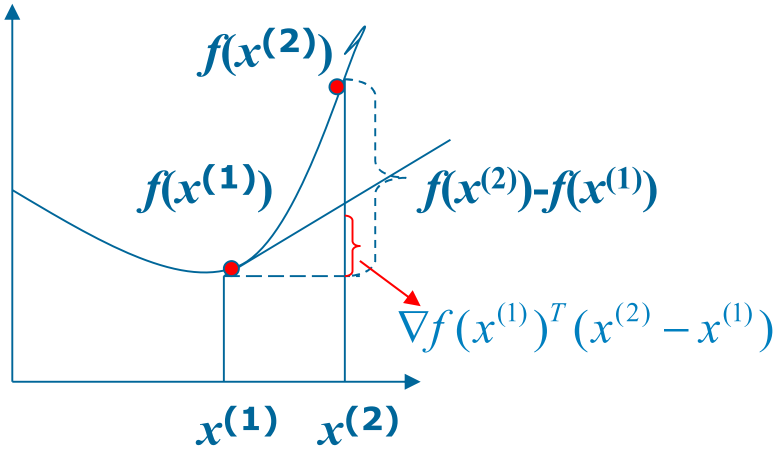
$$\therefore \quad \lambda f(x^{(1)}) + (1-\lambda) f(x^{(2)})$$

$$\geq f(y) + \nabla f(y)^T (\lambda x^{(1)} + (1-\lambda) x^{(2)} - y)$$

$$= f(y) = f(\lambda x^{(1)} + (1-\lambda) x^{(2)})$$

$\therefore \quad f$是凸函数。

几何意义



$f(x)$是凸函数当且仅当任意点处的切线增量不超过函数的增量。

推论：设S是E^n中的凸集，$\bar{x}\in S, f(x)$是定义在E^n上的凸函数，且在点$\bar{x}$可微，则对任意$x\in S$，有

$$f(x)\geq f(\bar{x})+\nabla f(\bar{x})^T(x-\bar{x})$$

定理(二阶充要条件)：设S是E^n中非空开凸集，$f(x)$是定义在S上的二次可微函数，则$f(x)$为凸函数的充要条件是对任意$x \in S$，$f(x)$在x处的$Hessian$矩阵$\nabla^2 f(x)$是半正定的。

证明:"$\Rightarrow$"设f是S上的凸函数, 对任意$\bar{x} \in S$

$\because S$是开集，则对$\forall x \in E^n$，$\exists \delta > 0$使当$\lambda \in (0, \delta)$，有$\bar{x} + \lambda x \in S$。

$$\therefore f(\bar{x} + \lambda x) \geq f(\bar{x}) + \lambda \nabla f(\bar{x})^T x \qquad (1)$$

$\because f$在点$\bar{x} \in S$二次可微,

$$\therefore f(\bar{x} + \lambda x) = f(\bar{x}) + \lambda \nabla f(\bar{x})^T x + \frac{1}{2}\lambda^2 x^T \nabla^2 f(\bar{x}) x + \lambda^2 \| x \|^2 a \qquad (2)$$

其中$\lim\limits_{\lambda \to 0} a = 0$。

由(1),(2)得，$\frac{1}{2}\lambda^2 x^T \nabla^2 f(\bar{x}) x + \lambda^2 \| x \|^2 a \geq 0$。

两边除以λ^2后，令$\lambda \to 0$，得，$x^T \nabla^2 f(\bar{x}) x \geq 0$。

定理(二阶充要条件)：设S是E^n中非空开凸集，$f(x)$是定义在S上的二次可微函数，则$f(x)$为凸函数的充要条件是对任意$x\in S$，有f在x处的$Hessian$矩阵$\nabla^2 f(x)$半正定。

"$\Leftarrow$"设$\nabla^2 f(x)$在任意点$x\in S$半正定，对$\forall x,\overline{x}\in S$，

由带Lagrange余项的二阶$Taylar$展开式，得

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

其中$\xi=\lambda\overline{x}+(1-\lambda)x$，$\lambda\in(0,1)$

因为S是凸集，所以$\xi\in S$，又$\nabla^2 f(x)$半正定，

$$\therefore \frac{1}{2}(x-\overline{x})^T\nabla^2 f(\xi)(x-\overline{x})\geq 0$$

$$\Rightarrow f(x)\geq f(\overline{x})+\nabla f(\overline{x})^T(x-\overline{x})$$

$\therefore f$是凸函数。

定理：设S是E^n中非空开凸集，$f(x)$是定义在S上的二次可微函数，如果对任意$x \in S$，有f在x处的$Hessian$矩阵$\nabla^2 f(x)$正定，则$f(x)$为严格凸函数。

对二次函数$f(x) = \dfrac{1}{2}x^T Ax + bx + c$，

若A正定，则$f(x)$为严格凸函数；

若A半正定，则$f(x)$为凸函数。

例：判断下列函数是否为凸函数.

(1) $f(x_1, x_2) = x_1^2 - 2x_1x_2 + x_2^2 + x_1 + x_2$

(2) $f(x_1, x_2) = x_1 e^{-(x_1 + x_2)}$.

定理：设$f(x)$是定义在凸集S上的可微凸函数，若$\exists x^* \in S$，使对$\forall x \in S$，都有

$$\nabla f(x^*)^T(x - x^*) \geq 0,$$

则x^*是$f(x)$在凸集S上的全局极小点。

证明：对$\forall x \in S$，因为$f(x)$是凸函数，所以有

$$f(x) \geq f(x^*) + \nabla f(x^*)^T(x - x^*)$$

$$\because \nabla f(x^*)^T(x - x^*) \geq 0$$

$$\therefore f(x) \geq f(x^*)$$

$\Rightarrow$ x^*为全局极小点。

凸规划

- 凸规划：求凸函数在凸集上的极小点。

$$\begin{aligned} \min \quad & f(x) \\ s.t. \quad & g_i(x) \geq 0, i = 1, \cdots, m \\ & h_j(x) = 0, j = 1, \cdots, l \end{aligned}$$

若$f(x)$是凸函数，$g_i(x)(i=1,\cdots,m)$是凹函数，$h_j(x)(j=1,\cdots,l)$是线性函数，则原问题为凸规划。

性质：凸规划的局部极小点就是整体极小点，
且极小点的集合为凸集。