

集合

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

内点: 设 $x^0 \in S \subset \mathbb{R}^n$, 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 $S^C = \{x \mid x \notin S, x \in \mathbb{R}^n\}$

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 $\forall x \in S, \|x\| \leq M$ 成立, 则称 S 为有界集。

紧集: 有界闭集称为紧集。

性质:

- (1) 集合 $S \subset \mathbb{R}^n$ 是闭集, 当且仅当对任意的无穷序列 $\{x^k\} \subset S$, 若 $x^k \rightarrow x^*$, 则 $x^* \in S$ 。
- (2) 集合 $S \subset \mathbb{R}^n$ 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 $\lim_{k \rightarrow \infty} x^{(k)} = \bar{x}$ 。
- 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 $\{x^{(k_j)}\}$ ，使得 $\lim_{k \rightarrow \infty} x^{(k_j)} = \hat{x}$ ，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 $\varepsilon > 0$ ，总存在正整数 K_ε ，使得当 $m, l > K_\varepsilon$ 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

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

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

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

先证 $\{x^{(k)}\}$ 为Cauchy序列。

$$\begin{aligned} & \|x^{(k)} - x^{(m)}\|^2 \\ &= 2\|x^{(k)} - y\|^2 + 2\|x^{(m)} - y\|^2 - 4\left\|\frac{x^{(k)} + x^{(m)}}{2} - y\right\|^2 \\ &\leq 2\|x^{(k)} - y\|^2 + 2\|x^{(m)} - y\|^2 - 4r^2 \\ &\rightarrow 0 \quad (\text{当 } m \rightarrow \infty, k \rightarrow \infty) \end{aligned}$$

$\therefore \{x^{(k)}\}$ 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 $\|y - \bar{x}\| = \|y - \hat{x}\| = r$

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

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

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

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

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

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

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

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

则对任意的 $x \in S$, 有

$$\begin{aligned}\|y - x\|^2 &= \|y - \bar{x} + \bar{x} - x\|^2 \\ &= \|y - \bar{x}\|^2 + \|\bar{x} - x\|^2 + 2(y - \bar{x})^T (\bar{x} - x) \\ &\geq \|y - \bar{x}\|^2\end{aligned}$$

$\therefore \bar{x}$ 是最小距离点。

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 $\bar{x} + \lambda(x - \bar{x}) \in S$.

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

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

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

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

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

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

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

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

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

$$\begin{aligned} \because \quad p^T (y - x) &= p^T (y - \bar{x} + \bar{x} - x) \\ &= p^T (y - \bar{x}) + p^T (\bar{x} - x) \\ &= \varepsilon + (y - \bar{x})^T (\bar{x} - x) \geq \varepsilon \end{aligned}$$

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

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 $y^{(k)} \rightarrow y$.

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 $(p^{(k)})^T y^{(k)} > (p^{(k)})^T x$.

\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 .

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

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

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 $p^T(x - y) \leq 0$.

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

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

(或 $p^T y \geq p^T x$ 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 $S \neq \emptyset$.

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

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

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

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

$S_1 \cap S_2 = \emptyset$, 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 $\{x^{(k)} \in S\} \rightarrow x$, 则存在序列

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

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

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

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

$\Rightarrow x = (x + z) - z \in S_2 - S_1 = S \Rightarrow S$ 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

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

证明: “ \Rightarrow ” 设存在 $y \geq 0$, 使得 $A^T y = c$

$$\text{得 } y^T A = c^T$$

设 \bar{x} 为 $Ax \leq 0, c^T x > 0$ 的一个解,

则有 $A\bar{x} \leq 0, c^T \bar{x} > 0$

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

$\because y \geq 0$, 但 $A\bar{x} \leq 0$

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

Farkas定理: 设 A 为 $m \times n$ 矩阵, c 为 n 维列向量, 则

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

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

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

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

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

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

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

即对任意的 $y \geq 0$, 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 $Ax \leq 0$.

既非零向量 x 是 $Ax \leq 0, c^T x > 0$ 的解。

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

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

$\therefore y$ 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, $\therefore S_1 \cap S_2 = \emptyset$, 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

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

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 $z \rightarrow 0$, 则对 $\forall x \in E^n$, 有

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

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

即 $-\|A^T y\|^2 \geq 0$, $\therefore A^T y = 0$.

因此, 存在非零向量 $y \geq 0$, 使得 $A^T y = 0$.

逆否命题

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

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

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,

$\Rightarrow Ax < 0$,

即存在 x , 使 $Ax < 0$ 有解.

这个证法不正确, 但是

为啥不正确我忘记了==

证法正确吗?

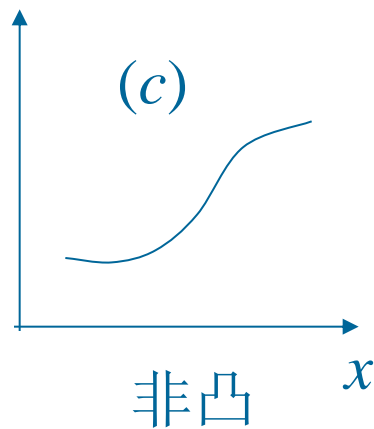
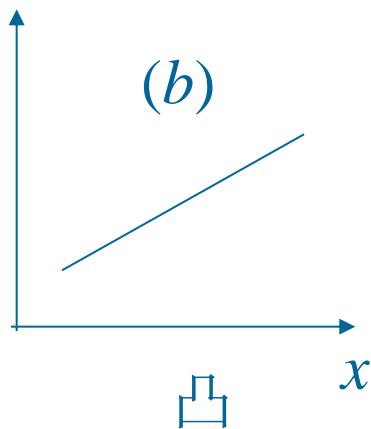
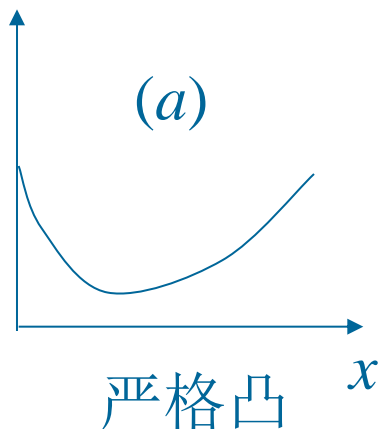
凸函数

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 $a, 0 \leq a \leq 1$, 都有

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

则称函数 $f(x)$ 为 S 上的凸函数. 上式中, 若 \leq 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 $a_1f_1(x)+a_2f_2(x)+\dots+a_kf_k(x)$ 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

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

(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,

$$\begin{aligned}\therefore f(\lambda x^{(0)} + (1 - \lambda)\bar{x}) &\leq \lambda f(x^{(0)}) + (1 - \lambda)f(\bar{x}) \\ &< \lambda f(\bar{x}) + (1 - \lambda)f(\bar{x}) = f(\bar{x})\end{aligned}$$

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

Hesse矩阵:

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

$$f(x) = \frac{1}{2} x^T A x$$

$$+ b^T x + c$$

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

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

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:

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

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

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

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

令 $\lambda \rightarrow 0^+$, 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 $\lambda = \frac{1}{2}$, 则 $y = \frac{1}{2}x^{(1)} + \frac{1}{2}x^{(2)} \in S$ 且

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

$\therefore f$ 是凸函数,

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

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

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

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

“ \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

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

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

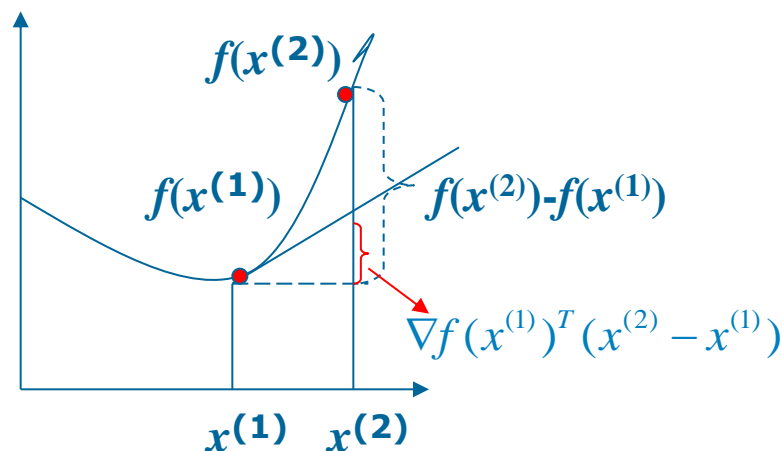
$$\therefore \lambda f(x^{(1)}) + (1 - \lambda)f(x^{(2)})$$

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

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

$\therefore f$ 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 $\bar{x} + \lambda x \in S$ 。

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

$\because f$ 在点 $\bar{x} \in S$ 二次可微,

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

其中 $\lim_{\lambda \rightarrow 0} a = 0$ 。

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

两边除以 λ^2 后, 令 $\lambda \rightarrow 0$, 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 $\forall x, \bar{x} \in S$,
由带Lagrange余项的二阶Taylor展开式, 得

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

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

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

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

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

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

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

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

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

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

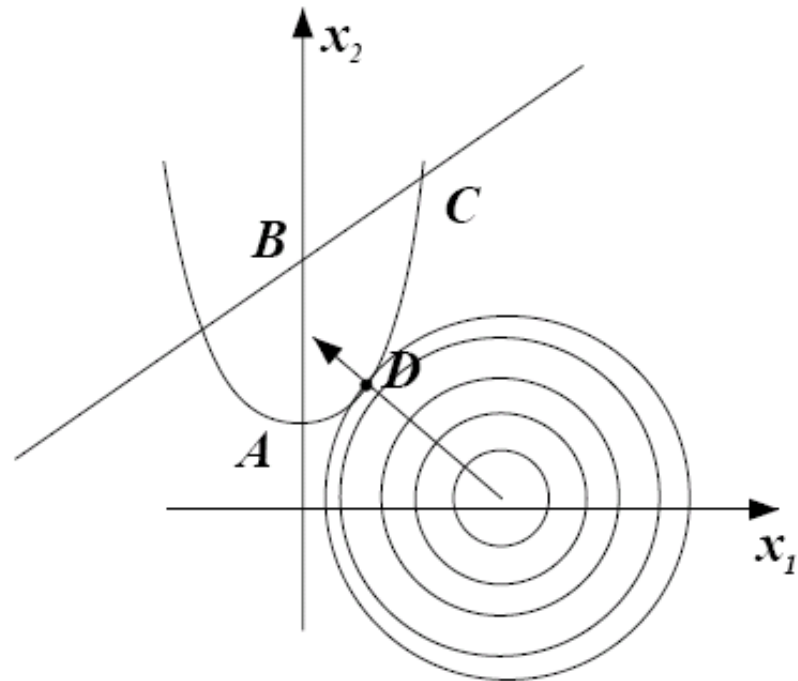
$\Rightarrow x^*$ 为全局极小点。

$$\min x_1^2 + x_2^2 - 4x_1 + 4$$

$$s.t. \quad x_1 - x_2 + 2 \geq 0$$

$$-x_1^2 + x_2 - 1 \geq 0$$

$$x_1, x_2 \geq 0$$



凸规划

- 凸规划：求凸函数在凸集上的极小点。

$$\min f(x)$$

$$s.t. \quad g_i(x) \geq 0, i = 1, \dots, m$$

$$h_j(x) = 0, j = 1, \dots, l$$

若 $f(x)$ 是凸函数， $g_i(x)(i = 1, \dots, m)$ 是凹函数，
 $h_j(x)(j = 1, \dots, l)$ 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。