# 最优化导论　第五次作业

1. 考虑函数 $f(x)=2x_1^2+x_2^2-2x_1x_2+2x_1^3+x_1^4$, 求出其所有一阶稳定点, 并判断它们是否为局部最优点（极小或极大）、鞍点或全局最优点?

**解**　首先计算 $f(x)$ 关于 $x_1, x_2$ 的梯度.

$$
\begin{aligned}
\frac{\mathrm{d}f(x)}{\mathrm{d}x_1} &= 4x_1-2x_2+6x_1^2+4x_1^3, \\
\frac{\mathrm{d}f(x)}{\mathrm{d}x_2} &= 2x_2-2x_1.
\end{aligned}
$$

对于一阶稳定点, 令上述两式为 0 , 我们得到

$$
\begin{aligned}
& x_1 = x_2 \\
& x_1\left(1+3x_1+2x_1^2\right)=0
\end{aligned}
$$

由此得到三个一阶稳定点 $(x_1,x_2)=(0,0),(-1,-1)$ 或 $(-0.5,-0.5)$. 再考虑海□矩阵

$$
\nabla^2 f(x)=\left[\begin{array}{cc} 4+12x_1+12x_1^2 & -2 \\ -2 & 2 \end{array}\right]
$$

有

$$
\begin{aligned}
\nabla^2 f(0,0)=\nabla^2 f(-1,-1) &= \left[\begin{array}{cc} 4 & -2 \\ -2 & 2 \end{array}\right] \\
\nabla^2 f(-0.5,-0.5) &= \left[\begin{array}{cc} 1 & -2 \\ -2 & \frac{2}{2} \end{array}\right].
\end{aligned}
$$

注意到 $(0,0)$ 和 $(-1,-1)$ 处海森矩阵都是正定矩阵. 因此都是局部最优点。

又根据 $f(0,0)=f(-1,-1)=0$ 知这两个点也为全局最优点. 另一方面, $(-0.5,-0.5)$ 处的海恙矩阵为不定矩阵, 因此是一个鞍点.

2.(a) 谈论三次函数 $\mathrm{z}=\boldsymbol{a}\boldsymbol{x}^3+\boldsymbol{b}\boldsymbol{x}^2+\boldsymbol{c}\boldsymbol{x}+\boldsymbol{d}$ 什么情况下是拟凹函数或拟凸函数，什么情况下既非拟凹函数也非拟凸函数。

(b) 对于 $\mathrm{x} \geq 0$ ，有可能对参数施加某些限制，使得函数变成既是拟凹又是拟凸吗?

**解**　(a) 根据定义, 对于任意常数 $k$ ，当且仅当 $\left\{\begin{array}{l} S^{\geq}:=\{x \mid f(x) \geq k\} \\ S^{\leq}:=\{x \mid f(x) \leq k\} \end{array}\right\}$ 是凸集时，函数 $f(x)$ 是 $\left\{\begin{array}{l} \text{拟凹的} \\ \text{拟凸的} \end{array}\right.$

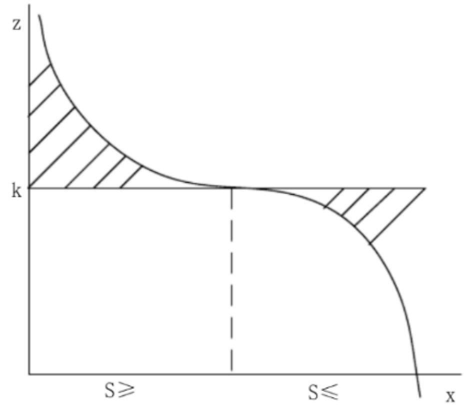


Figure 1: 第二题

由于函数 $z=ax^3+bx^2+cx+d$ 系数不确定, 其图形可能向上倾斜也可能向下倾斜，故不存在任意的 $k$ 值能使上述条件满足，即函数 $z=ax^3+bx^2+cx+d$ 一般情况下既非拟凹函数，也非拟凸函数。

如上图所示, 当函数 $z=ax^3+bx^2+cx+d$ 出现此种情况时, $\mathrm{S}^{\geq}$（虚线左侧的阴影区域）和 $\mathrm{S}\leq$（虚线右侧的阴影区域）均为凹集，即函数 $z=ax^3+bx^2+cx+d$ 既非拟凸也非拟凹。

(b) 要使函数 $z=ax^3+bx^2+cx+d$ 在 $x\geq 0$ 上既是拟凹又是拟凸必须保证函数在 $x\geq 0$ 范围内向右上方倾斜, 且拐点在第一象限。因此需要满足以下条件: (1) 要函数图形向右上方倾斜, 必须保证 $a>0$, 且 $\frac{dz}{dx}=3ax^2+2bx+$ $c$ 的绝对极小值为正, 即 $\min\left(\frac{dz}{dx}\right)=3a\left(\frac{-b}{3a}\right)^2+2b\frac{-b}{3a}+c=\frac{3ac-b^2}{3a}>0$, 亦即 $3ac>b^2$ ; (2) 要使拐点在第一象限, 必须保证 $\frac{dz}{dx}=3ax^2+2bx+c$ 的绝对极小值点出现在 $x>0$ 的区间且截距 $d>0$, 即 $\frac{-b}{3a}>0$, 亦即 $b<0$; (3) 由 $3ac>b^2$ 和 $a>0$ 可知, $c>0$ 必成立;

综上所述, 使函数 $z=ax^3+bx^2+cx+d$ 在 $x\geq 0$ 上既是拟凹又是拟凸的条件为:

$$a,c,d>0,\quad b<0,\quad 3ac>b^2$$

3. 设函数 $f:\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}$ 满足如下性质：当固定 $z\in\mathbb{R}^m$ 时，$f(x,z)$ 关于 $x$ 为凸函数；当固定 $x\in\mathbb{R}^n$ 时，$f(x,z)$ 关于 $z$ 是凹函数，则称 $f$ 为凸 - 凹函数.

(a) 设 $f$ 二阶可导，试利用海瑟矩阵 $\nabla^2 f$ 给出 $f$ 为凸 - 凹函数的一个二阶条件；

(b) 设 $f$ 为凸 - 凹函数且可微, 且在点 $(\bar{x},\bar{z})$ 处满足 $\nabla f(\bar{x},\bar{z})=0$, 求证：对任意 $x$ 和 $z$ ，如下鞍点性质成立：

$$f(\bar{x},z)\leqslant f(\bar{x},\bar{z})\leqslant f(x,\bar{z})$$

进一步证明 $f$ 满足极小 - 极大性质：

$$\sup_z\inf_x f(x,z)=\inf_x\sup_z f(x,z).$$

(c) 设 $f$ 可微但不一定是凸 - 凹函数, 且在点 $(\bar{x},\bar{z})$ 处满足鞍点性质

$$f(\bar{x},z)\leqslant f(\bar{x},\bar{z})\leqslant f(x,\bar{z}),\quad\forall x,z$$

求证: $\nabla f(\bar{x}, \bar{z}) = 0$.

**解** (a) 设 $f$ 的海瑟矩阵是

$$\nabla^2 f = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix},$$

其中

$$A_{11} \in \mathbb{R}^{n\times n}, A_{12} \in \mathbb{R}^{n\times m}, A_{21} \in \mathbb{R}^{m\times n}, A_{22} \in \mathbb{R}^{m\times m}.$$

则 $\nabla_x^2 f = A_{11}, \nabla_z^2 f = A_{22}$. 若固定 $z$ 的 $f$ 关于 $x$ 凸, 则 $A_{11}$ 半正定；若固定 $x$ 时 $f$ 关于 $z$ 凹，则 $A_{22}$ 半负定. 这就是由海瑟矩阵给出的二阶条件。

(b) 先证明第一个不等式. 由于 $f(x,z)$ 关于 $x$ 是凸函数, 利用凸函数的性质有

$$f(x, \bar{z}) \geqslant f(\bar{x}, \bar{z}) + \nabla_x f(\bar{x}, \bar{z})(x - \bar{x}) = f(\bar{x}, \bar{z}),$$

同理可得左半边的不等式. 再证明 $\sup_z \inf_x f(x,z) = f(\bar{x}, \bar{z})$ 。首先

$$\inf_x f(x,z) \leqslant f(\bar{x}, z) \leqslant f(\bar{x}, \bar{z}),$$

故

$$\sup_z \inf_x f(x,z) \leqslant f(\bar{x}, \bar{z}),$$

又

$$\sup_z \inf_x f(x,z) \geqslant \inf_x f(x, \bar{z}) \geqslant f(\bar{x}, \bar{z}),$$

因此 $\sup_z \inf_x f(x,z) = f(\bar{x}, \bar{z})$ 。同理得 $\inf_x \sup_z f(x,z) = f(\bar{x}, \bar{z})$, 因此等式 $^{zx}$ 得证。

(c) 只需分别证明 $\nabla_x f(\bar{x}, \bar{z}) = 0$ 和 $\nabla_z f(\bar{x}, \bar{z}) = 0$. 用反证法证明. 先假设 $\nabla_x f(\bar{x}, \bar{z}) \neq 0$, 取向量 $v = \left(\nabla_x f(\bar{x}, \bar{z})^{\mathrm{T}}, 0\right)^{\mathrm{T}}$ ，对 $f(\bar{x} + tv, \bar{z})$ （其中 $t \neq 0$ ）在 $(\bar{x}, \bar{z})$ 处展开一阶, 得

$$f(\bar{x} + tv, \bar{z}) = f(\bar{x}, \bar{z}) + t \left\|\nabla_x f(\bar{x}, \bar{z})\right\|^2 + \mathcal{O}\left(t^2\right).$$

若取 $t < 0$ 且绝对值足够小, 就有 $f(\bar{x} + tv, \bar{z}) < f(\bar{x}, \bar{z})$, 与题设矛盾. 因此假设不成立, $\nabla_x f(\bar{x}, \bar{z}) = 0$. 同理可得 $\nabla_z f(\bar{x}, \bar{z}) = 0$. 综上，命题成立。

**课本习题：** 《Convex Optimization》4.2(a), 4.4(a)(b)(c), 4.7(a)(b)(c), 4.11(a)(b), 4.22