Exercise Sheet – Mathematical Analysis III

Taiyang Xu*

16/10/2025, Week 6

练习 1 (本练习涉及凸区域、凸函数、中值定理和多元 Taylor 公式等内容). 完成以下题目

(1) 设 $D \subseteq \mathbb{R}^n$ 为凸的有界闭区域, f 在 D 上有连续的一阶偏导数。试证明 f 在 D 上满足如下的 Lipschitz 条件: 存在常数 $M > 0$, 使得对任意的 $x, y \in D$, 有

$$|f(x) - f(y)| \leq M|x - y|.$$

(2) 设 $F(x,y,z)$ 在 $\mathbb{R}^3$ 中有连续的一阶偏导数 $\frac{\partial F}{\partial x}$, $\frac{\partial F}{\partial y}$, $\frac{\partial F}{\partial z}$, 并满足不等式:

$$x\frac{\partial F}{\partial x} + y\frac{\partial F}{\partial y} + z\frac{\partial F}{\partial z} \geq \alpha > 0, \quad \forall (x,y,z) \in \mathbb{R}^3.$$

证明: 当 (x,y,z) 沿着曲线 $\Gamma = (x(t), y(t), z(t))$, $x(t) = -\cos t$, $y(t) = \sin t$, $z(t) = t$, $t \geqslant 0$ 趋向 ∞ 时, $F(x,y,z)$ 趋向 $+\infty$.

(3) 设 f 在凸区域 $D \in \mathbb{R}^n$ 上具有连续的一阶偏导数, 则 f 在 D 上为凸函数的充分必要条件是:

$$\forall x, y \in D, \quad f(y) - f(x) \geqslant \nabla f(x) \cdot (y - x).$$

(4) 设 $f(\vec{x})$, $\vec{x} = (x_1, x_2, \ldots, x_n)$ 在凸区域 D 上具有连续的二阶偏导数, 则 f 在 D 上为凸函数的充分必要条件是 Hesse 矩阵 $\left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)_{i,j=1}^n$ 在 D 上的任意点均为半正定矩阵.

解答 2. (1) 由于 D 是凸的有界闭区域，且 f 在 D 上有连续的一阶偏导数，根据微分中值定理，对任意 $x, y \in D$，存在 $\theta \in [0,1]$，使得

$$f(y) - f(x) = \langle \nabla f(x + \theta(y - x)), (y - x) \rangle.$$

由于 D 有界闭，∇f 在 D 上连续, 故根据 Bolzano-Weierstrass 定理, ∇f 在 D 上一致有界。设

$$M := \max_{z \in D} |\nabla f(z)|,$$

*School of Mathematical Sciences, Fudan University, Shanghai 200433, China. Email: tyxu19@fudan.edu.cn

则对任意 $x, y \in D$，由微分中值定理，有

$$f(y) - f(x) = \langle \nabla f(x + \theta(y - x)), (y - x) \rangle,$$

其中 $\theta \in [0, 1]$。因此

$$|f(y) - f(x)| = |\langle \nabla f(x + \theta(y - x)), y - x \rangle| \le |\nabla f(x + \theta(y - x))| \cdot |y - x| \le M|y - x|$$

故 f 满足 Lipschitz 条件。

(2) 定义

$$\Phi(t) := F(-\cos t, \sin t, t).$$

利用微分中值定理, 我们有

$$\Phi(t) = \Phi(0) + \Phi'(\tau)t, \quad \tau \in (0, t).$$

直接计算

$$\begin{aligned}
&\Phi(0) = F(-1, 0, 0), \\
&\Phi'(\tau) = \frac{\partial F}{\partial x}\frac{dx}{dt}\Big|_{t=\tau} + \frac{\partial F}{\partial y}\frac{dy}{dt}\Big|_{t=\tau} + \frac{\partial F}{\partial z}\frac{dz}{dt}\Big|_{t=\tau} = \sin t\frac{\partial F}{\partial x}\Big|_{t=\tau} + \cos t\frac{\partial F}{\partial y}\Big|_{t=\tau} + 1 \cdot \frac{\partial F}{\partial z}\Big|_{t=\tau}, \\
&= x\frac{\partial F}{\partial x}\Big|_{t=\tau} - y\frac{\partial F}{\partial y}\Big|_{t=\tau} + \frac{\partial F}{\partial z}\Big|_{t=\tau} \ge \alpha > 0.
\end{aligned}$$

故而 $\Phi(t) = F(x, y, z) \geqslant F(1, 0, 0) + \alpha t \to +\infty$, $t \to +\infty$.

(3) $\Rightarrow$: 假设 f 是凸函数。对任意 $x, y \in D$，对任意 $\lambda \in [0, 1]$，由凸性有

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

即

$$f(x + \lambda(y - x)) - f(x) \leqslant \lambda(f(y) - f(x)).$$

由于 f 在 D 上具有连续的一阶偏导数, 根据多元函数的中值定理,

$$f(x + \lambda(y - x)) - f(x) = \nabla f(x + \theta\lambda(y - x)) \cdot (\lambda(y - x)), \quad \theta \in (0, 1).$$

因此, 对 $\lambda \ne 0$, 有

$$\nabla f(x + \theta\lambda(y - x)) \cdot (y - x) \leqslant f(y) - f(x).$$

令 $\lambda \to 0^+$, 则 $x + \theta\lambda(y - x) \to x$, 由 ∇f 的连续性, 有最终结果。
也可以利用 Taylor 公式 (而非中值定理) 来证明. 对 $f(x + \lambda(y - x))$ 在 x 做 Taylor 展开:

$$f(x + \lambda(y - x)) = f(x) + \nabla f(x) \cdot (\lambda(y - x)) + \mathcal{O}(||\lambda(y - x)||^2), \quad \theta \in (0, 1).$$

$\Leftarrow$: 假设对任意 $x,y\in D$，都有 $f(y)-f(x)\geqslant \nabla f(x)\cdot(y-x)$。我们需要证明 f 是凸函数。取任意 $x,y\in D$，任意 $\lambda\in[0,1]$，令 $z=\lambda y+(1-\lambda)x\in D$。由假设，对 x,z 有

$$f(x)-f(z)\geqslant \nabla f(z)\cdot(x-z).$$

同理，对 y,z 有

$$f(y)-f(z)\geqslant \nabla f(z)\cdot(y-z).$$

将两式分别按乘上 λ 和 $1-\lambda$ 再相加得到

$$\begin{aligned}\lambda f(x)+(1-\lambda)f(y)&\geqslant \lambda f(z)+\lambda\nabla f(z)\cdot(x-z)+(1-\lambda)f(z)+(1-\lambda)\nabla f(z)\cdot(y-z)\\&=f(z)+\nabla f(z)\cdot(\lambda(x-z)+(1-\lambda)(y-z))\\&=f(z)+\nabla f(z)\cdot(\lambda x+(1-\lambda)y-z)\\&=f(z).\end{aligned}$$

因此 f 是凸函数。

(4) $\Leftarrow$: $\forall x,y\in D$, $\exists z=x+\theta(y-x)$, 使得 (Lagriange 余项型的 Taylor 展开)

$$f(y)=f(x)+\nabla f(x)\cdot(y-x)+\frac{1}{2}\left(\sum_{k=1}^{n}(y_k-x_k)\frac{\partial}{\partial x_k}\right)^2 f(z).$$

事实上, 你会发现

$$\begin{aligned}&\left(\sum_{k=1}^{n}(y_k-x_k)\frac{\partial}{\partial x_k}\right)^2 f(z)\\&=\sum_{i,j=1}^{n}(y_i-x_i)(y_j-x_j)\frac{\partial^2 f}{\partial x_i\partial x_j}(z)\\&=(y-x)^{\mathrm{T}}H_f(z)(y-x).\end{aligned}$$

由于 $H_f(z)$ 为半正定矩阵, 则

$$(y-x)^{\mathrm{T}}H_f(z)(y-x)\geqslant 0.$$

因此 $f(y)-f(x)\geqslant \nabla f(x)\cdot(y-x)$. 由 (3) 可知 f 是凸函数。

$\Rightarrow$: 如若不然. 假设 H_f 在某点 $x\in D$ 不是半正定矩阵, 则 $\exists v\in\mathbb{R}^n$, 使得

$$v^{\mathrm{T}}H_f(x)v<0,$$

这里 $v=(v_1,\dots,v_n)^{\mathrm{T}}\neq 0$. 另一方面, 由 Taylor 公式，对 $f(x+tv)$ 在 x 做 Taylor 展开 $(t\to 0)$, 则

$$\begin{aligned}f(x+tv)&=f(x)+\nabla f(x)\cdot(tv)+\frac{1}{2}(tv)^{\mathrm{T}}H_f(x)(tv)+o(||tv||^2)\\&=f(x)+t\nabla f(x)\cdot v+\frac{t^2}{2}\left[v^{\mathrm{T}}H_f(x)v+o(1)\right].\end{aligned}$$

当 t 充分小时, $v^{\mathrm{T}}H_f(x)v+o(1)<0$, 故

$$f(x+tv)\leqslant f(x)+\nabla f(x)\cdot tv.$$

这与 (3) 中的充分条件矛盾.

练习 3 (本题训练 Lagrange 乘子法求极值问题). (1) 求解以下约束问题

(a) 求函数 $f(x,y,z)=xy+yz+zx$ 在约束条件 $xyz=1$, $x,y,z>0$ 下的条件极值.

(b) 求函数 $f(x,y,z)=\sin x\sin y\sin z$ 在约束条件 $x+y+z=\frac{\pi}{2}$, $x,y,z>0$ 下的条件极值.

(c) 求二次型 $f(x)=\sum_{i,j=1}^n a_{ij}x_ix_j$ (其中 $a_{ij}=a_{ji}$) 在单位球面 $\sum_{i=1}^n x_i^2=1$ 上的极值.

(2) 求椭球 $\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1$ 的最大体积的内接长方体.

(3) 设 f 是 $\mathbb{R}$ 上的非负连续函数且 $\int_{-\infty}^{\infty}f(x)dx=1$. 假设区间 $[a,b]$ 是使得 $\int_a^b f(x)dx=\frac{1}{2}$ 的长度最短的区间. 证明 $f(a)=f(b)$.

(4) 求函数 $f(x_1,\dots,x_n)=\sum_{k=1}^n x_k\log x_k$ 在约束 $x_1+\cdots+x_n=a$ (其中 $a>0$) 下的最大值.

(5) 给定正实数 $\alpha_1,\dots,\alpha_n$ 满足 $\alpha_1+\alpha_2+\cdots+\alpha_n=1$. 利用 Lagrange 乘数法证明: 对任意正实数 $x_1,\dots,x_n$, 有

$$\prod_{k=1}^n x_k^{\alpha_k}\le\sum_{k=1}^n\alpha_kx_k.$$

解答 4. (1) (a) 使用拉格朗日乘数法. 定义拉格朗日函数:

$$L(x,y,z,\lambda)=xy+yz+zx-\lambda(xyz-1).$$

求偏导数并令为零:

$$\frac{\partial L}{\partial x}=y+z-\lambda yz=0,\quad \frac{\partial L}{\partial y}=x+z-\lambda xz=0,\quad \frac{\partial L}{\partial z}=x+y-\lambda xy=0.$$

由这些方程可得:

$$\lambda=\frac{y+z}{yz}=\frac{1}{z}+\frac{1}{y},\quad \lambda=\frac{x+z}{xz}=\frac{1}{z}+\frac{1}{x},\quad \lambda=\frac{x+y}{xy}=\frac{1}{y}+\frac{1}{x}.$$

比较可得 $x=y=z$. 代入约束 $xyz=1$ 得 $x^3=1$, 所以 $x=y=z=1$. 此时 $f(1,1,1)=3$. 由 AM-GM 不等式, $xy+yz+zx\ge 3\sqrt[3]{(xy)(yz)(zx)}=3\sqrt[3]{(xyz)^2}=3$, 等号当且仅当 $x=y=z$. 因此, 条件极小值为 3.

(b) 使用拉格朗日乘数法. 定义拉格朗日函数:

$$L(x,y,z,\lambda)=\sin x\sin y\sin z-\lambda\left(x+y+z-\frac{\pi}{2}\right).$$

求偏导数并令为零:

$$\frac{\partial L}{\partial x}=\cos x\sin y\sin z-\lambda=0,\quad \frac{\partial L}{\partial y}=\sin x\cos y\sin z-\lambda=0,\quad \frac{\partial L}{\partial z}=\sin x\sin y\cos z-\lambda=0.$$

由于 $\sin x,\sin y,\sin z>0$, 有 $\cos x\sin y=\sin x\cos y$ 即 $\tan x=\tan y$, 所以 $x=y$. 同理 $y=z$, 故 $x=y=z$. 代入约束 $x+y+z=\frac{\pi}{2}$ 得 $3x=\frac{\pi}{2}$, 所以 $x=y=z=\frac{\pi}{6}$. 此时 $f\left(\frac{\pi}{6},\frac{\pi}{6},\frac{\pi}{6}\right)=\left(\sin\frac{\pi}{6}\right)^3=\left(\frac{1}{2}\right)^3=\frac{1}{8}$. 当 $x,y,z\to 0^+$ 时 $f\to 0$, 因此条件极大值为 $\frac{1}{8}$.

(c) 二次型 $f(x)=x^TAx$, 其中 A 为对称矩阵. 在约束 $x^Tx=1$ 下, 由 Rayleigh-Ritz 定理, $f(x)$ 的极值为矩阵 A 的特征值. 具体地, 最小值为 A 的最小特征值, 最大值为 A 的最大特征值. 因此, 条件极值为 A 的所有特征值.

(2) 假设长方体与坐标轴平行, 设其一顶点为 (x,y,z) 且 $x,y,z>0$, 则体积为 $V=8xyz$. 需在约束 $\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1$ 下最大化 V. 使用拉格朗日乘数法. 定义拉格朗日函数:

$$L(x,y,z,\lambda)=8xyz-\lambda\left(\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}-1\right).$$

求偏导数并令为零:

$$\frac{\partial L}{\partial x}=8yz-\lambda\frac{2x}{a^2}=0,\quad \frac{\partial L}{\partial y}=8xz-\lambda\frac{2y}{b^2}=0,\quad \frac{\partial L}{\partial z}=8xy-\lambda\frac{2z}{c^2}=0.$$

解得:

$$\lambda=\frac{4yza^2}{x},\quad \lambda=\frac{4xzb^2}{y},\quad \lambda=\frac{4xyc^2}{z}.$$

比较可得:

$$\frac{x^2}{a^2}=\frac{y^2}{b^2}=\frac{z^2}{c^2}.$$

设 $\frac{x^2}{a^2}=\frac{y^2}{b^2}=\frac{z^2}{c^2}=k$, 代入约束得 $3k=1$, 所以 $k=\frac{1}{3}$. 于是 $x=\frac{a}{\sqrt{3}}$, $y=\frac{b}{\sqrt{3}}$, $z=\frac{c}{\sqrt{3}}$. 最大体积为:

$$V=8xyz=8\cdot\frac{a}{\sqrt{3}}\cdot\frac{b}{\sqrt{3}}\cdot\frac{c}{\sqrt{3}}=\frac{8abc}{3\sqrt{3}}.$$

(3) 考虑最小化区间长度 $b-a$ under constraint $\int_a^b f(x)dx=\frac{1}{2}$. 定义拉格朗日函数:

$$L(a,b,\lambda)=b-a-\lambda\left(\int_a^b f(x)dx-\frac{1}{2}\right).$$

求偏导数并令为零:

$$\frac{\partial L}{\partial a}=-1-\lambda\left(-f(a)\right)=-1+\lambda f(a)=0\implies \lambda f(a)=1,$$

$$\frac{\partial L}{\partial b} = 1 - \lambda f(b) = 0 \implies \lambda f(b) = 1.$$

因此 $f(a) = f(b)$.

(4) 使用拉格朗日乘数法. 定义拉格朗日函数:

$$L(x_1, \ldots, x_n, \lambda) = \sum_{k=1}^{n} x_k \log x_k - \lambda \left(\sum_{k=1}^{n} x_k - a \right).$$

求偏导数并令为零:

$$\frac{\partial L}{\partial x_k} = \log x_k + 1 - \lambda = 0 \implies \log x_k = \lambda - 1 \implies x_k = e^{\lambda - 1}.$$

所以所有 x_k 相等, 设 $x_k = c$, 则 $nc = a$, 所以 $c = \frac{a}{n}$. 此时函数值为:

$$f = \sum_{k=1}^{n} \frac{a}{n} \log \frac{a}{n} = a \log \frac{a}{n}.$$

但函数 $g(x) = x \log x$ 是凸函数 (因为 $g''(x) = \frac{1}{x} > 0$), 因此 $\sum x_k \log x_k$ 是凸函数, 在单纯形约束下, 最小值在内部点 $x_k = \frac{a}{n}$ 处达到, 最大值在边界处达到. 当某个 $x_k = a$ 且其余为 0 时, 函数值为 $a \log a$. 由于 $a \log a > a \log \frac{a}{n}$ for $n > 1$, 且当 $x_k \to 0^+$ 时 $x_k \log x_k \to 0$, 因此最大值为 $a \log a$.

(5) 考虑在约束 $\sum_{k=1}^{n} \alpha_k x_k = S$ (常数) 下最大化 $F(x_1, \ldots, x_n) = \prod_{k=1}^{n} x_k^{\alpha_k}$. 取对数, 定义拉格朗日函数:

$$L(x_1, \ldots, x_n, \lambda) = \sum_{k=1}^{n} \alpha_k \log x_k - \lambda \left(\sum_{k=1}^{n} \alpha_k x_k - S \right).$$

求偏导数并令为零:

$$\frac{\partial L}{\partial x_k} = \frac{\alpha_k}{x_k} - \lambda \alpha_k = 0 \implies \frac{1}{x_k} = \lambda \implies x_k = \frac{1}{\lambda}.$$

所以所有 x_k 相等, 设 $x_k = c$. 代入约束:

$$\sum_{k=1}^{n} \alpha_k c = c \sum_{k=1}^{n} \alpha_k = c = S \implies c = S.$$

此时 $F = \prod_{k=1}^{n} c^{\alpha_k} = c^{\sum \alpha_k} = c = S$. 所以最大值是 S. 因此对于任意 x_k, 有

$$\prod_{k=1}^{n} x_k^{\alpha_k} \leq \sum_{k=1}^{n} \alpha_k x_k.$$