

偏微分方程数值解: 第二次作业

林立康 25210180078

November 13, 2025

相关代码可从 https://github.com/LiHuaGz/Spectral_Method.git 获取.

1.(a)

$$D_N(-x) = \frac{\sin(-(N + \frac{1}{2})x)}{\sin(-\frac{x}{2})} = \frac{\sin((N + \frac{1}{2})x)}{\sin(\frac{x}{2})} = D_N(x).$$

$$\begin{aligned} D_N\left(\frac{1}{2} + x\right) &= \frac{\sin((N + \frac{1}{2})(\frac{1}{2} + x))}{\sin(\frac{1}{2}(\frac{1}{2} + x))} \\ &= \frac{\sin((N + \frac{1}{2}) \cdot \frac{1}{2}) \cos((N + \frac{1}{2})x) + \cos((N + \frac{1}{2}) \cdot \frac{1}{2}) \cdot \sin((N + \frac{1}{2})x)}{\sin(\frac{1}{2} \cdot \frac{1}{2}) \cdot \cos(\frac{1}{2}x) + \cos(\frac{1}{2} \cdot \frac{1}{2}) \sin(\frac{1}{2}x)} \\ &= \frac{\sin((N + \frac{1}{2}) \cdot \frac{1}{2}) \cos(-(N + \frac{1}{2})x) + \cos((N + \frac{1}{2}) \cdot \frac{1}{2}) \cdot \sin(-(N + \frac{1}{2})x)}{\sin(\frac{1}{2} \cdot \frac{1}{2}) \cdot \cos(-\frac{1}{2}x) + \cos(\frac{1}{2} \cdot \frac{1}{2}) \sin(-\frac{1}{2}x)} \\ &= D_N\left(\frac{1}{2} - x\right). \end{aligned}$$

1.(b)

$$\begin{aligned} \int_0^{2\pi} D_N(x) dx &= \int_0^{2\pi} (1 + 2 \sum_{k=1}^N \cos kx) dx \\ &= \int_0^{2\pi} dx + 2 \sum_{k=1}^N \int_0^{2\pi} \cos kx dx \\ &= 2\pi. \end{aligned}$$

$$\begin{aligned}
\int_0^{2\pi} D_N^2(x) dx &= \int_0^{2\pi} (1 + 2 \sum_{k=1}^N \cos kx)^2 dx \\
&= \int_0^{2\pi} dx + 4 \int_0^{2\pi} \sum_{k=1}^N \cos kx dx + 4 \int_0^{2\pi} (\sum_{k=1}^N \cos kx)^2 dx \\
&= 2\pi + 4 \sum_{j,k=1}^N \int_0^{2\pi} \cos jx \cdot \cos kx dx \\
&= 2\pi + 4 \sum_{k=1}^N \int_0^{2\pi} \cos^2 kx dx \\
&= 2\pi + 4 \sum_{k=1}^N \int_0^{2\pi} \frac{\cos 2kx + 1}{2} dx \\
&= 2\pi + 4N\pi.
\end{aligned}$$

因此,

$$\frac{1}{2\pi} \int_0^{2\pi} D_N(x) dx = 1, \quad \frac{1}{2\pi} \int_0^{2\pi} D_N^2(x) dx = 2N + 1.$$

1.(c)

当 $x \in [0, \pi]$ 时, $\sin(\frac{x}{2}) \geq \frac{2}{\pi} \cdot \frac{x}{2} = \frac{x}{\pi}$, 故

$$|D_N(x)| = \frac{|\sin((N + \frac{1}{2})x)|}{\sin(x/2)} \leq \frac{1}{\sin(x/2)} \leq \frac{\pi}{x}, \quad x \in (0, \pi].$$

可知 $\int_0^{2\pi} |D_N(x)| dx = 2 \int_0^\pi |D_N(x)| dx$, 则

$$\begin{aligned}
\int_0^{2\pi} |D_N(x)| dx &= 2 \int_0^\pi |D_N(x)| dx \\
&\leq 2 \left(\int_0^{\frac{1}{2N+1}} (2N+1) dx + \int_{\frac{1}{2N+1}}^\pi \frac{\pi}{x} dx \right) \\
&= 2 \left((2N+1) \cdot \frac{1}{2N+1} + \pi \ln \frac{\pi}{\frac{1}{2N+1}} \right) \\
&= 2(1 + \pi \ln(2N+1)) \\
&\leq C \ln N, \quad N \geq 2,
\end{aligned}$$

其中 C 为常数.

1.(d)

$$\begin{aligned}
|\phi(x)| &\leq \frac{1}{2\pi} \int_0^{2\pi} |\phi(t)| |D_N(x-t)| dt \\
&\leq \frac{1}{2\pi} \left(\int_0^{2\pi} |\phi(t)|^2 dt \right)^{\frac{1}{2}} \cdot \left(\int_0^{2\pi} |D_N(x-t)|^2 dt \right)^{\frac{1}{2}} \\
&\leq \frac{1}{\sqrt{2\pi}} \left(\frac{1}{2\pi} \int_0^{2\pi} |\phi(t)|^2 dt \right)^{\frac{1}{2}} \cdot \left(\int_0^{2\pi} |D_N(x-t)|^2 dt \right)^{\frac{1}{2}} \\
&= \frac{1}{\sqrt{2\pi}} \|\phi\| \sqrt{2\pi(2N+1)} \\
&= \sqrt{2N+1} \|\phi\|, \quad x \in [0, 2\pi].
\end{aligned}$$

其中, $\int_0^{2\pi} |D_N(x-t)|^2 dt = \int_{x-2\pi}^x |D_N(y)|^2 dy = 2\pi(2N+1)$.

因此,

$$\|\phi\|_\infty = \max_{x \in [0, 2\pi]} |\phi(x)| \leq \sqrt{2N+1} \|\phi\|.$$

1.(e)

可知,

$$\begin{aligned}
\sum_{N=1}^{M-1} D_N(x) &= \frac{1}{\sin \frac{x}{2}} \sum_{N=1}^{M-1} \sin((N + \frac{1}{2})x) \\
&= \frac{1}{2i \sin \frac{x}{2}} \sum_{N=1}^{M-1} (e^{i(Nx + \frac{x}{2})} - e^{-i(Nx + \frac{x}{2})}),
\end{aligned}$$

当 $x \notin \pi\mathbb{Z}$ 时,

$$\begin{aligned}
\sum_{N=1}^{M-1} (e^{i(Nx + \frac{x}{2})} - e^{-i(Nx + \frac{x}{2})}) &= e^{i\frac{x}{2}} \sum_{n=1}^{M-1} e^{inx} - e^{-i\frac{x}{2}} \sum_{n=1}^{M-1} e^{-inx} \\
&= e^{i\frac{x}{2}} \cdot \frac{e^{ix} - e^{iMx}}{1 - e^{ix}} - e^{-i\frac{x}{2}} \cdot \frac{e^{-ix} - e^{-iMx}}{(1 - e^{-ix})} \\
&= \frac{e^{i\frac{x}{2}}}{1 - e^{ix}} (e^{ix} - e^{iMx}) + \frac{e^{i\frac{x}{2}}}{1 - e^{ix}} (e^{-ix} - e^{-iMx}) \\
&= \frac{e^{i\frac{x}{2}}}{1 - e^{ix}} [e^{ix} + e^{-ix} - (e^{iMx} + e^{-iMx})] \\
&= \frac{1}{e^{-i\frac{x}{2}} - e^{i\frac{x}{2}}} [e^{ix} + e^{-ix} - (e^{iMx} + e^{-iMx})] \\
&= \frac{1}{2i \sin(-\frac{x}{2})} (2 \cos x - 2 \cos Mx),
\end{aligned}$$

所以

$$\begin{aligned}\sum_{N=1}^{M-1} D_N(x) &= \frac{1}{\sin^2 \frac{x}{2}} \cdot \frac{(1 - 2 \sin^2 \frac{x}{2}) - (1 - 2 \sin^2 \frac{Mx}{2})}{2} \\ &= \frac{1}{\sin^2 \frac{x}{2}} (-\sin^2 \frac{x}{2} + \sin^2 \frac{Mx}{2}) \\ &= -1 + \frac{\sin^2 \frac{Mx}{2}}{\sin^2 \frac{x}{2}}.\end{aligned}$$

从而

$$\frac{1}{M} \sum_{N=0}^{M-1} D_N(x) = \frac{\sin^2 \frac{Mx}{2}}{M \sin^2 \frac{x}{2}}, \quad x \notin \pi \mathbb{Z}.$$

当 $x \in \pi \mathbb{Z}$ 时, 结论显然也成立.

2.(a)

(1) 理论分析

易知 $f(x) = e^{\sin(8x)}(-8 \cos(x) \cos(8x) + 64(2 + \sin(x))(\sin(8x) - \cos^2(8x)) + \sin^2(x))$.

对 $[a, b]$ 做划分: $a = x_0 < x_1 < \dots < x_N = b$, 其中 $x_k = \frac{(b-a)k}{N} + a$, $k = 0, 1, \dots, N$.

将 $u(x)$ 做周期延拓, 并记 $v(x) = p(x)u'(x)$. 则原方程可写为:

$$-v'(x) + q(x)u(x) = f(x), \quad x \in (a, b),$$

且 $v(a) = v(b)$, $u(a) = u(b)$.

$v'(x_k)$ 的二阶有限差分格式为:

$$v'(x_k) = \frac{v(x_k + \frac{h}{2}) - v(x_k - \frac{h}{2})}{h} + O(h^2), \quad k = 1, \dots, N-1.$$

由于

$$\begin{aligned}v(x_k + \frac{h}{2}) &= p(x_k + \frac{h}{2})u'(x_k + \frac{h}{2}) = p(x_k + \frac{h}{2}) \cdot \frac{u(x_{k+1}) - u(x_k)}{h} + O(h^2), \\ v(x_k - \frac{h}{2}) &= p(x_k - \frac{h}{2})u'(x_k - \frac{h}{2}) = p(x_k - \frac{h}{2}) \cdot \frac{u(x_k) - u(x_{k-1})}{h} + O(h^2),\end{aligned}$$

故

$$v'(x_k) \approx \frac{p(x_k + \frac{h}{2}) \cdot \frac{u(x_{k+1}) - u(x_k)}{h} - p(x_k - \frac{h}{2}) \cdot \frac{u(x_k) - u(x_{k-1})}{h}}{h}, \quad k = 0, 1, \dots, N-1.$$

将上述格式代入原方程, 可得线性方程组:

$$-\frac{p(x_k + \frac{h}{2}) \cdot \frac{u(x_{k+1}) - u(x_k)}{h} - p(x_k - \frac{h}{2}) \cdot \frac{u(x_k) - u(x_{k-1})}{h}}{h} + q(x_k)u(x_k) = f(x_k), \quad k = 0, 1, \dots, N-1.$$

整理得:

$$\begin{cases} -\frac{p_{k-\frac{1}{2}}}{h^2}u_{k-1} + \left(\frac{p_{k+\frac{1}{2}} + p_{k-\frac{1}{2}}}{h^2} + q_k\right)u_k - \frac{p_{k+\frac{1}{2}}}{h^2}u_{k+1} = f_k, & k = 0, 1, \dots, N-1, \\ u_N = u_0, \quad u_{-1} = u_{N-1}, \end{cases}$$

其中 $p_{k \pm \frac{1}{2}} = p(x_k \pm \frac{h}{2})$, $q_k = q(x_k)$, $f_k = f(x_k)$, $u_k = u(x_k)$.

(2) 数值结果

h 与 L^2 、 H^1 误差的关系如图 1 所示. 由图可见, L^2 误差 (斜率: 2.02) 和 H^1 误差 (斜率: 1.94) 的拟合直线斜率均非常接近 2, 这表明算法的收敛阶为 2, 与理论分析结果一致.

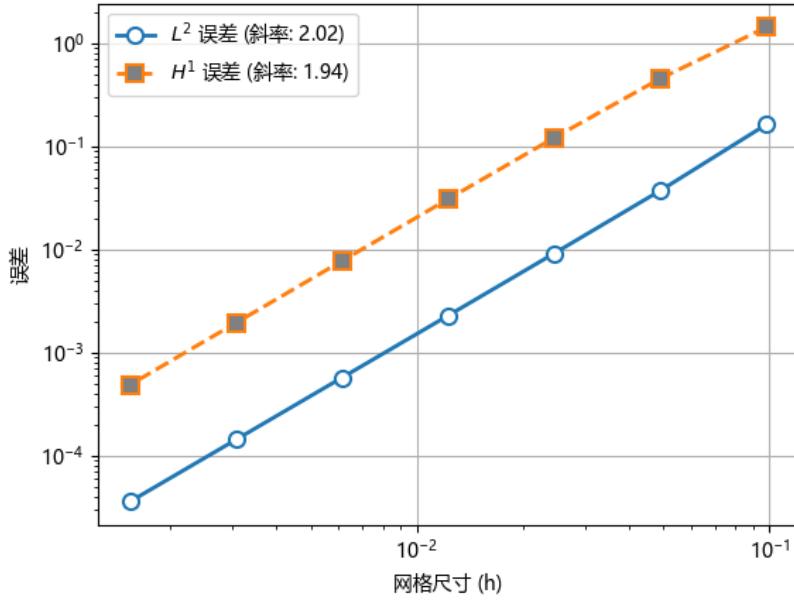


Figure 1: 有限差分法 h 与 L^2 、 H^1 误差的关系

2.(b)

(1) Fourier-Galerkin 谱方法

记 $X_N = \text{span} \{e^{ikx} : k = -N/2, \dots, N/2 - 1\}$, $u_N = \sum_{k=-N/2}^{N/2-1} \tilde{u}_k e^{ikx}$, 则 Fourier-Galerkin 谱方法即找出 $u_N \in X_N$, 使得

$$\int_0^{2\pi} [-(p(x)u'_N(x))' + q(x)u_N(x) - f(x)]e^{-imx} dx = 0, \quad m = -N/2, \dots, N/2 - 1. \quad (1)$$

由于

$$\begin{aligned}
& \int_0^{2\pi} (p(x)u'_N(x))' e^{-imx} dx \\
&= p(x)u'_N(x)e^{-imx} \Big|_0^{2\pi} - \int_0^{2\pi} p(x)u'_N(x)(-ime^{-imx}) dx \\
&= im \int_0^{2\pi} p(x)u'_N(x)e^{-imx} dx \\
&= im \int_0^{2\pi} p(x) \left(\sum_{k=-N/2}^{N/2-1} ik\tilde{u}_k e^{ikx} \right) e^{-imx} dx \\
&= \sum_{k=-N/2}^{N/2-1} -mk\tilde{u}_k \int_0^{2\pi} p(x)e^{i(k-m)x} dx, \\
&= \sum_{k=-N/2}^{N/2-1} -2\pi mk\tilde{u}_k \hat{p}_{m-k},
\end{aligned}$$

且

$$\begin{cases} \int_0^{2\pi} q(x)u_N(x)e^{-imx} dx = 2\pi \sum_{k=-N/2}^{N/2-1} \hat{q}_{m-k} \tilde{u}_k, \\ \int_0^{2\pi} f(x)e^{-imx} dx = 2\pi \hat{f}_m, \end{cases} \quad (2)$$

故原问题可化为求解线性方程组:

$$\sum_{k=-N/2}^{N/2-1} mk\hat{p}_{m-k} \tilde{u}_k + \sum_{k=-N/2}^{N/2-1} \hat{q}_{m-k} \tilde{u}_k = \hat{f}_m, \quad m = -N/2, \dots, N/2 - 1. \quad (3)$$

(2) Fourier-Galerkin 伪谱方法

事实上, 式 (1) 也即

$$-(\widehat{p(x)u'_N(x)})'_m + \widehat{q(x)u_N(x)}_m = \hat{f}_m, \quad m = -N/2, \dots, N/2 - 1,$$

由式 (3) 可知, 这可以视为作用在 $\tilde{u} = [\tilde{u}_{-N/2}, \dots, \tilde{u}_{N/2-1}]^T$ 上的线性算子. 将等式左端第一项和第二项对应的线性算子分别记为 $\mathcal{L}_1, \mathcal{L}_2$, 则 Fourier-Galerkin 伪谱方法即找出 \tilde{u} 使得

$$\mathcal{L}_1(\tilde{u}) + \mathcal{L}_2(\tilde{u}) = \hat{f}. \quad (4)$$

若直接使用式 (3) 左端的卷积形式计算 $\mathcal{L}_1(\tilde{u})$, 则每次计算都需要 $O(N^2)$ 的时间复杂度. 为了提高计算效率, 使用伪谱方法计算 $\mathcal{L}_1(\tilde{u})$.

使用高斯消元法时的计算流程如下: (1) 设 FFT 和 IFFT 对应矩阵分别为 $\mathbf{F}, \mathbf{F}^{-1}$, 设 $P = \text{diag}(p(x_0), \dots, p(x_{N-1}))$, $Q = \text{diag}(q(x_0), \dots, q(x_{N-1}))$, (2) 计算 \mathcal{L}_1 算子对应的矩阵

$$\mathbf{L}_1 = -i\mathbf{m} \cdot \mathbf{F} \cdot P \cdot \mathbf{F}^{-1} \cdot i\mathbf{m},$$

其中 $\mathbf{m} = (0, 1, \dots, N/2 - 1, 0, -N/2 + 1, \dots, 0)$; (3) 计算 \mathcal{L}_2 算子对应的矩阵

$$\mathbf{L}_2 = \mathbf{F} \cdot Q \cdot \mathbf{F}^{-1};$$

(4) 求解线性方程组 $(-i\mathbf{m} \cdot \mathbf{F} \cdot P \cdot \mathbf{F}^{-1} \cdot i\mathbf{m} + \mathbf{F} \cdot Q \cdot \mathbf{F}^{-1})\tilde{u} = \hat{f}$.

使用共轭梯度法求解时的计算流程如下: (1) 设当前迭代点的 \tilde{u}_N 已知, 通过 IFFT 计算 $u'_N(x_j)$, 其中 $x_j = \frac{2\pi j}{N}, j = 0, 1, \dots, N - 1$ 为配点; (2) 逐点相乘得到 $p(x_j)u'_N(x_j)$; (3) 通过 FFT 计算 $p(x)u'_N(x)$ 的傅里叶系数, 并对其乘以 $i\mathbf{m}$ 得到 $-(p(x)u'_N(x))'$ 的傅里叶系数. $\mathcal{L}_2(\tilde{u})$ 的计算类似.

(3) 数值结果

使用高斯消元法求解时, N 与 L^2 误差的关系如图 2 所示. 可见 L^2 误差随 N 增加而呈指数级快速下降, 在 $N \approx 500$ 之后, 误差达到了约 10^{-12} 量级, 这代表了由双精度浮点数运算所决定的机器精度极限. 该结果验证了 Fourier 伪谱方法的谱精度. 随后误差略有上升, 这可能是由累积误差所致.

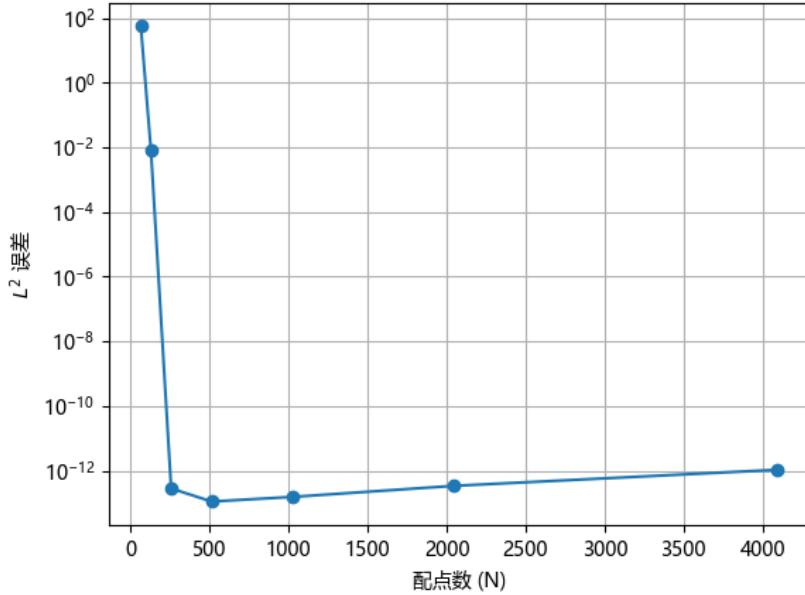


Figure 2: 伪谱方法 N 与 L^2 误差的关系 (高斯消元法)

2.(c)

使用共轭梯度法求解 (容忍精度 10^{-10}), N 与 L^2 误差的关系如图 4 所示. 从图中可以观测到: L^2 误差从 $N = 100$ 时的 10^{-1} 量级迅速降至 $N \approx 250$ 时的 10^{-7} 量级以下, 这表明该伪谱方法对于求解光滑周期解问题具有谱精度; 当 N 增加到约 500 之后, L^2 误差不再下降, 这可能是由于 $\mathcal{L}_1 + \mathcal{L}_2$ 不是自伴算子, 从而共轭梯度法不能保证收敛; 后面的误差上升可能是由累积误差所致 (也可能是我的实现方式不对?).

达到 10 位精度时共轭梯度法所需循环次数与 N 的关系如图 3 所示。由图可知，CG 迭代次数与 N 近似成正比（例如 $N = 1000$ 时迭代约 1000 次， $N = 4000$ 时迭代约 4000 次），这说明共轭梯度法并没有因为提前达到 10 位精度而提前终止，应证了上述“可能由于 $\mathcal{L}_1 + \mathcal{L}_2$ 不是自伴算子，导致使用共轭梯度法不能保证收敛”。

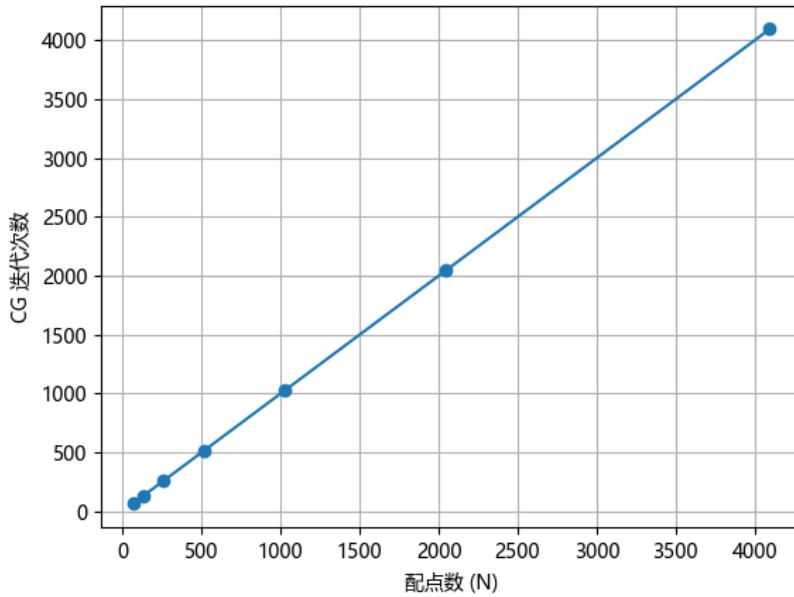


Figure 3: 共轭梯度法所需循环次数与 N 的关系

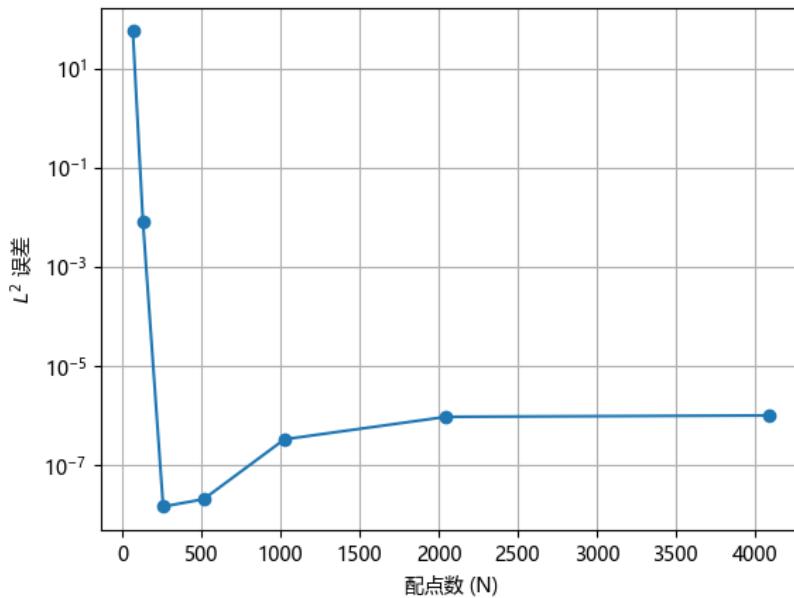


Figure 4: 伪谱方法 N 与 L^2 误差的关系 (共轭梯度法)

2.(d)

(1) 理论分析

Fourier 配点法即要找近似解 $u_N(x) = \sum_{k=-N/2}^{N/2-1} \tilde{u}_k e^{ikx}$, 使得在所有配点 x_j 上原微分方程成立:

$$-p'(x_j)u'_N(x_j) - p(x_j)u''_N(x_j) + q(x_j)u_N(x_j) = f(x_j), \quad j = 0, 1, \dots, N-1.$$

记 $\mathbf{u} = [u(x_0), \dots, u(x_{N-1})]^T$ 为在配点 x_j 上的函数值向量, $\mathbf{f} = [f(x_0), \dots, f(x_{N-1})]^T$, $\mathbf{P} = \text{diag}(p(x_0), \dots, p(x_{N-1}))$, $\mathbf{P}' = \text{diag}(p'(x_0), \dots, p'(x_{N-1}))$, $\mathbf{Q} = \text{diag}(q(x_0), \dots, q(x_{N-1}))$.

记一阶和二阶微分矩阵分别为 $D^{(1)}$ 和 $D^{(2)}$, 则原问题可表示为矩阵形式:

$$(-\mathbf{P}'D^{(1)} - \mathbf{P}D^{(2)} + \mathbf{Q})\mathbf{u} = \mathbf{f}.$$

(2) 数值结果

使用高斯消元法求解时, N 与 L^2 误差的关系如图 5 所示. 可见 L^2 误差随 N 增加而呈指数级快速下降, 在 $N \approx 500$ 之后, 误差达到了 10^{-11} 量级, 这代表了由双精度浮点数运算所决定的机器精度极限. 该结果验证了 Fourier 配点法的谱精度. 同时说明了在该题中配点法的精度可能不如上述的 Galerkin 伪谱方法.

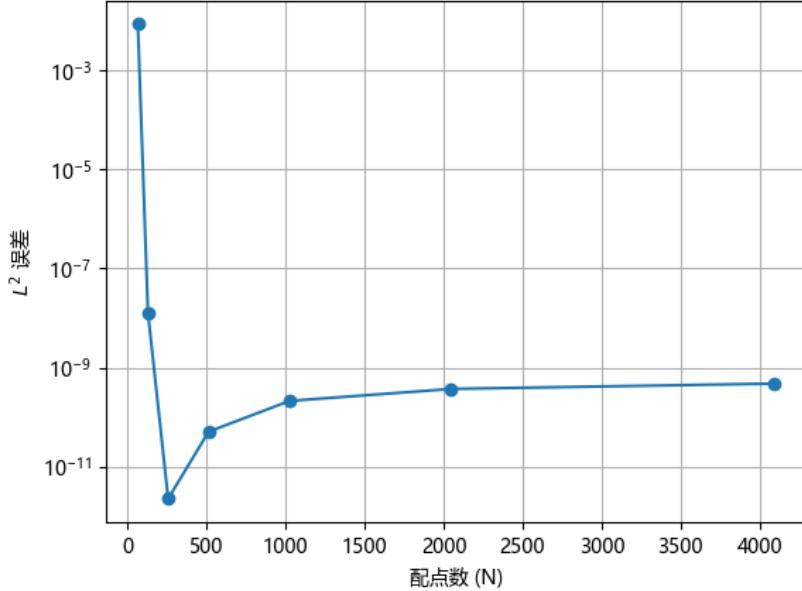


Figure 5: 配点法 N 与 L^2 误差的关系 (高斯消元法)

2.(e)

达到 10 位精度时共轭梯度法所需循环次数与 N 的关系如图 6 所示. 使用共轭梯度法求解时, N 与 L^2 误差的关系如图 7 所示. 图 7 中, 初期误差的指数下降证明了方法的

精度, 而最终误差趋向于 10^{-5} , 高于图 5 中的机器精度极限 (10^{-11}), 可能是因为求解时使用了共轭梯度法, 而 $-\mathbf{P}'D^{(1)} - \mathbf{P}D^{(2)} + \mathbf{Q}$ 不是厄米特阵 ($D^{(1)}$ 不是对称阵), 导致共轭梯度法不能保证收敛到机器精度.

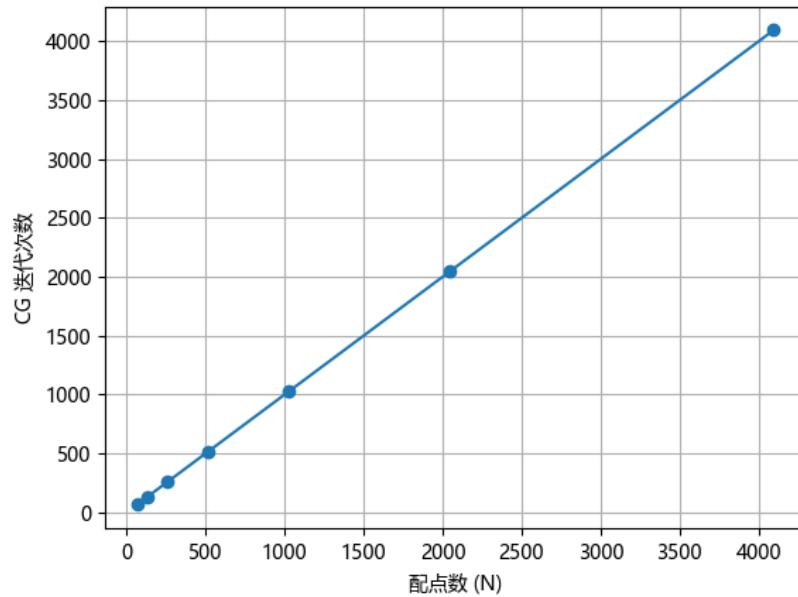


Figure 6: 共轭梯度法所需循环次数与 N 的关系

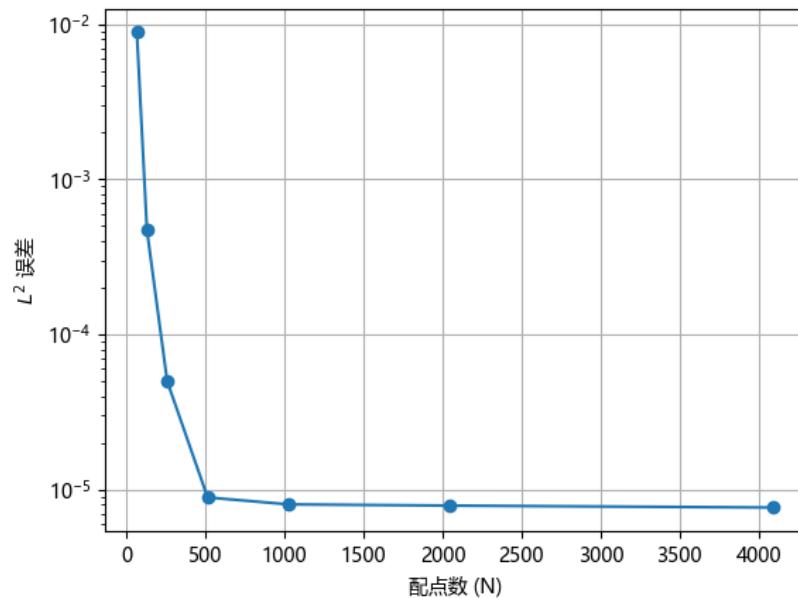


Figure 7: 配点法 N 与 L^2 误差的关系 (共轭梯度法)

3.

(1) 理论分析

将解 $u(x, t)$ 展开为截断的傅里叶级数:

$$u(x, t) \approx u_N(x, t) = \sum_{k=-N/2}^{N/2-1} \hat{u}_k(t) e^{ikx},$$

记

$$R_N(x, t) = \frac{\partial u_N}{\partial t} - \epsilon \frac{\partial^2 u_N}{\partial x^2} - u_N \frac{\partial u_N}{\partial x},$$

原问题化为找 u_N , 使得

$$\int_0^{2\pi} R_N(x, t) \cdot e^{-imx} dx = 0, \quad m = -N/2, \dots, N/2 - 1,$$

也即找 $u \in X_N = \text{span}\{e^{ikx} : k = -N/2, \dots, N/2 - 1\}$, 使得

$$\frac{d\hat{u}_k}{dt} = \epsilon \left(\widehat{\frac{\partial^2 u}{\partial x^2}} \right)_k + \left(\widehat{u \frac{\partial u}{\partial x}} \right)_k, \quad k = -N/2, \dots, N/2 - 1, \quad (5)$$

这里的 $\widehat{(\cdot)}_k$ 表示“括号内项的第 k 个傅里叶系数”. 记 $\hat{N}_k = \widehat{(u \frac{\partial u}{\partial x})}_k$, 则上式可写为

$$\frac{d\hat{u}_k}{dt} = -\epsilon k^2 \hat{u}_k + \hat{N}_k, \quad k = -N/2, \dots, N/2 - 1. \quad (6)$$

计算非线性项 $\hat{N}_k = \widehat{(u \frac{\partial u}{\partial x})}_k$ 的流程: (1) 通过 FFT 计算 $\frac{\partial u}{\partial x}$, (2) 逐点相乘 $u * \frac{\partial u}{\partial x}$, (3) 再次使用 FFT 计算乘积的傅里叶系数.

注: 实际计算使用的是式 (5) 的转化形式:

$$\frac{d\tilde{u}_k^{fft}}{dt} = \left(\epsilon \widehat{\frac{\partial^2 u}{\partial x^2}} + \widehat{u \frac{\partial u}{\partial x}} \right)_k^{fft}, \quad k = -N/2, \dots, N/2 - 1. \quad (7)$$

流程: (1) 已知 $t = t_n$ 时刻的 $\tilde{u}_k^{fft}(t_n)$, 通过 IFFT 计算 $u(x_j, t_n)$, $\frac{\partial u}{\partial x}(x_j, t_n)$ 和 $\frac{\partial^2 u}{\partial x^2}(x_j, t_n)$; (2) 逐点相乘得到 $u \frac{\partial u}{\partial x}(x_j, t_n)$; (3) 通过 FFT 计算式 (7) 右端的傅里叶系数; (4) 使用四阶 Runge-Kutta 法求解 ODE 系统, 得到 $t = t_{n+1}$ 时刻的 $\tilde{u}_k^{fft}(t_{n+1})$.

(2) 数值结果

$t = 1$ 时刻的数值结果如图 8 所示, 其中 $\epsilon = 0.03$, $N_x = N_t = 128$. (x, t) 与 $u(x, t)$ 的关系如图 9 所示.

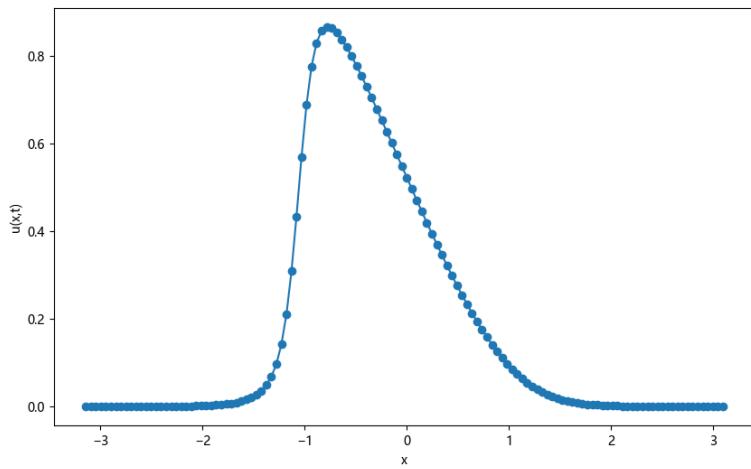


Figure 8: 时刻 $t = 1$ 的数值解

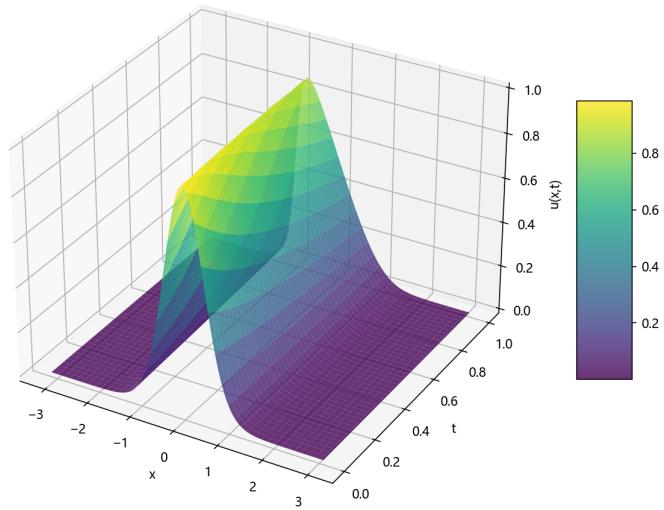


Figure 9: (x, t) 与 $u(x, t)$ 的关系