

## 二维边值问题有限元离散系统病态问题的病因抑制方法

张 衡<sup>1,2,3</sup>, 张 宏<sup>1,4</sup>, 游文杰<sup>2</sup>, 郑汉垣<sup>5</sup>

<sup>1</sup>无损检测技术福建省高等学校重点实验室(福建技术师范学院), 福建 福清

<sup>2</sup>福建技术师范学院大数据与人工智能学院, 福建 福清

<sup>3</sup>湖南三一工业职业技术学院工程机械学院, 湖南 长沙

<sup>4</sup>福建技术师范学院电子与机械工程学院, 福建 福清

<sup>5</sup>龙岩学院传播与设计学院, 福建 龙岩

收稿日期: 2021年11月13日; 录用日期: 2021年12月8日; 发布日期: 2021年12月15日

### 摘 要

讨论二维Poisson方程边值问题离散系统的病态问题, 针对三角剖分和四角剖分, 基于Lagrange形函数形成有限元离散系统的病态问题, 提出病因抑制方法; 给出该系统的病态结构、病态因子、去病因子; 利用病态因子估计离散系统的条件数; 利用去病因子为最优预条件子, 精准抑制病态发作, 该预条件子的使用, 几乎不增加求解的计算量, 预处理后离散系统保持正定对称, 条件数关于离散系统规模一致有界。

### 关键词

病因抑制方法, 病态结构, 病态因子, 去病因子, 最优预条件子

## A Method of Inhibiting Pathogeny for Ill-Condition Problems of the Finite Element Discrete Systems of 2D Boundary Value Problems

Heng Zhang<sup>1,2,3</sup>, Hong Zhang<sup>1,4</sup>, Wenjie You<sup>2</sup>, Hanyuan Zheng<sup>5</sup>

<sup>1</sup>Key Laboratory of Nondestructive Testing (Fujian Polytechnic Normal University), Fujian Polytechnic Normal University, Fuqing Fujian

<sup>2</sup>School of Big Data and Artificial Intelligence, Fujian Polytechnic Normal University, Fuqing Fujian

<sup>3</sup>Faculty of Engineering Machinery, Hunan SANY Polytechnic College, Changsha Hunan

<sup>4</sup>School of Electronic and Mechanical Engineering, Fujian Polytechnic Normal University, Fuqing Fujian

<sup>5</sup>School of Communication and Design, Longyan University, Longyan Fujian

Received: Nov. 13<sup>th</sup>, 2021; accepted: Dec. 8<sup>th</sup>, 2021; published: Dec. 15<sup>th</sup>, 2021

## Abstract

The ill conditioned problem of the finite element discrete system based on triangular subdivision, tetragonal subdivision and Lagrange shape function for 2D Poisson equation boundary value problem is discussed. The method of inhibiting pathogeny was proposed. The ill condition structure, ill condition factor and ill elimination factor of the system were given; and the condition number of the equation was estimated by the ill conditioned factor. The ill elimination factor was used as the optimal preconditioner to precisely suppress the ill condition. The use of the preconditioner hardly increases the calculation of iteration. After the pretreatment, the equations keep positive definite symmetry and the condition number is uniformly bounded with respect to the scale of equations.

## Keywords

Method of Inhibiting Pathogeny, Ill-Condition Structure, Ill-Condition Factor, Ill Elimination Factor, Optimal Preconditioner

Copyright © 2021 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

偏微分方程的数值求解问题, 通常转化为离散系统——如  $\mathbf{Ax} = \mathbf{b}$  的线性方程组求解[1], 离散系统的病态(高条件数,  $\text{Cond}(\mathbf{A})$  表示  $\mathbf{A}$  的条件数)问题随着求解规模增加而更严重[2], 成为制约大规模求解效率和精度的瓶颈因素, 因此, 在求解过程中, 使用适当的预条件子进行预处理, 将离散系统化成条件数低、易求解的同解方程, 是成功求解的关键[3] [4]。

目前关于病态问题以及预处理技术的理论和方法仍然不完善, 缺乏对病因的研究, 预条件子的设计、预处理效果的评价多用数值模拟的手段进行探索和验证[5]-[11], 产生的预处理方法使用范围小, 可移植性差, 难以找到通用的、最优的预条件子[12], 很多学者甚至认为不存在通用的预处理方法[13] [14] [15]。目前关于病态问题的病因研究鲜见有成果发表。

基于病因的结构分析方法[16] [17] [18] [19] [20], 本文提出离散系统病态问题的病因抑制方法: 确定病因, 精准抑制病因发作。

针对矩形区域上变分形式的二维泊松方程边值问题, 使用有限元方法求解, 分别基于三角剖分和四角剖分, 使用 Lagrange 形函数产生的离散系统, 是稀疏病态方程组。本文使用病因抑制方法解决该方程的病态问题, 通过结构分析[16] [17] [18] [19] [20], 确定病态结构, 分离出病态因子、构造去病因子。对于原发病态和继发病态, 分别提出评估条件数的方法和预处理策略, 定量分析以去病因子作为预条件子的预处理效果(有效性、低成本), 说明了在几乎不显著增加计算量的情况下, 预处理后的条件数关于离散系统规模一致有界, 预处理后离散系统保持正定对称。

## 2. 病态问题的病因抑制方法

### 2.1. 病态因子与去病因子

**定义 1:** 如果矩阵  $Z \in R^{\alpha \times \beta}$ ,  $\beta > \alpha$ ,  $\alpha, \beta$  是整数,  $ZZ^T$  可逆,  $\text{Cond}(ZZ^T)$  是阶数  $\alpha$  的增函数。则称  $Z$  为  $\alpha \times \beta$  病态因子。

**定义 2:** 对于定义 1 中的  $\alpha \times \beta$  病态因子  $Z$ , 如果有可逆矩阵  $H$  满足:  $ZZ^T = HH^T$ , 则称矩阵  $H$  为属于  $Z$  的去病因子[16]。

**命题 1 [16]:** 设  $\beta \geq \alpha$ ,  $Z$  是  $\alpha \times \beta$  病态因子,  $P$  是  $\beta$  阶正定对称矩阵, 则有

- 1).  $\text{Cond}(ZPZ^T) \leq \text{Cond}(P)\text{Cond}(ZZ^T)$ ,
- 2). 如果矩阵  $H$  为属于  $Z$  的去病因子, 则有  $\text{Cond}(H^{-1}ZPZ^TH^{-T}) \leq \text{Cond}(P)$ 。

### 2.2. 矩阵的病态结构

**定义 3:** 可逆矩阵  $A$  的如下结构称为病态结构:

$$A = ZPZ^T + Q \quad (1)$$

其中  $A, Q \in R^{\alpha \times \alpha}$ ,  $P \in R^{\beta \times \beta}$ ,  $\beta > \alpha$ ,  $\alpha, \beta$  是整数,  $Z$  是  $\alpha \times \beta$  病态因子,  $P$  是  $\beta$  阶正定对称矩阵,  $\|ZPZ^T\| \gg \|Q\|$ ,  $\text{Cond}(P)$  关于  $A$  的阶数一致有界,  $\text{Cond}(ZPZ^T)$  是  $A$  的阶数  $\alpha$  的增函数。称  $ZPZ^T$  为  $A$  的病态主体; 也称  $Z$  是  $ZPZ^T$  或者  $A$  的病态因子。

在定义 3 中, 使用  $ZPZ^T$  代替  $A$ , 相对误差为  $\|A - ZPZ^T\| / \|ZPZ^T\| = \|Q\| / \|ZPZ^T\| \ll 1$ , 所以不妨认为  $A \approx ZPZ^T$ ,  $\text{Cond}(A) \approx \text{Cond}(ZPZ^T)$ ,  $Q$  对病态的影响很小, 根据命题 1 结论 1),  $\text{Cond}(A) \approx \text{Cond}(ZPZ^T) \leq \text{Cond}(P)\text{Cond}(ZZ^T)$ , 即可以利用病态因子的条件数, 估计  $A$  的病态。

**定义 4:** 在定义 3 中, 由病态因子表达的病态称为原发性病态, 其他因素表达的病态称为继发性病态。

### 2.3. 病态问题的病因抑制方法

对于方程组

$$Ax = b \quad (2)$$

如果  $A$  有(1)式的病态结构, 使用属于  $Z$  的去病因子  $H$  作为预条件子, 则方程(2)化成

$$H^{-1}AH^{-T}x' = (H^{-1}ZPZ^TH^{-T} + H^{-1}QH^{-T})x' = b' \quad (3)$$

其中  $x' = H^T x$ ,  $b' = H^{-1}b$ 。

根据命题 1 结论 2), 在式(3)中,  $\text{Cond}(H^{-1}AH^{-T}) \approx \text{Cond}(H^{-1}ZPZ^TH^{-T}) \leq \text{Cond}(P)$  这说明属于  $Z$  的去病因子  $H$  抑制或者消除了病态因子  $Z$  的作用, 即消除了原发性病态。因为  $\text{Cond}(P)$  关于  $A$  的阶数一致有界, 因此,  $HH^T$  是最优预条件子[4]。

综上所述, 为解决方程(2)的病态问题, 提出如下病因抑制方法: 通过对病态矩阵  $A$  进行结构分析, 分离出病态因子  $Z$ , 确定如(1)式的病态结构; 一方面可以用病态因子  $Z$  转置积的条件数评价  $A$  的病态程度; 另一方面可以针对病态因子  $Z$  (无须针对  $A$ ) 设计最优预条件子——去病因子  $H$ , 用于消除  $A$  的原发性病态。

病因抑制方法的关键是确定病态因子。

### 2.4. 一个特别的病态因子及其去病因子

设  $\text{diag}(S_1, \dots, S_n)$  表示对角元素为矩阵  $S_i$  ( $1 \leq i \leq n$ ) 的块对角矩阵, 记  $0_\alpha$  是  $\alpha$  维零向量,

$\omega_{i,j}^{(n)} = \sqrt{2/(n+1)} \sin[ij\pi/(n+1)] = \omega_{j,i}^{(n)}$ ,  $\Omega_n = (\omega_{i,j}^{(n)})$  是  $n$  阶正弦变换矩阵,

$\lambda_{i,j}^{(n,m)} = 4\sqrt{1 - \cos^2[i\pi/(2n+2)] \cos^2[j\pi/(2n+2)]}$ ,  $\Lambda_i = \text{diag}(\lambda_{i,1}^{(n,m)}, \dots, \lambda_{i,m}^{(n,m)})$ ,

$$H = (\Omega_n \otimes \Omega_m) \Lambda, \quad \Lambda = \text{diag}(\Lambda_1, \dots, \Lambda_n)$$

$$D^T = (\sigma_0^T, \sigma_1^T, \sigma_2^T, \sigma_3^T), \quad \sigma_0 = (1, 1, -1), \quad \sigma_1 = (-1, 1, 1), \quad \sigma_2 = (1, -1, 1), \quad \sigma_3 = (-1, -1, -1)$$

$Z = (I_n, 0_n) \otimes Z_1 + (0_n, I_n) \otimes Z_0$ ,  $Z_i = (I_m, 0_m) \otimes \sigma_{2i+1} + (0_m, I_m) \otimes \sigma_{2i}$ ,  $i = 0, 1$ , 则有

**命题 2:** 1).  $ZZ^T = HH^T$ ; 2).  $\text{Cond}(ZZ^T) = (\lambda_{n,m}^{(n,m)} / \lambda_{1,1}^{(n,m)})^2$ ; 3).  $\text{Cond}(ZZ^T) = O(\min(n^2, m^2))$ 。

**证明:** 1). 记  $I_n$  是  $n$  阶单位矩阵,  $T_n = (1, 2, 1)_n$  是  $n$  阶三对角矩阵,  $\bar{\Lambda}_n = \text{diag}(\bar{\lambda}_1^{(n)}, \dots, \bar{\lambda}_n^{(n)})$ ,  $\bar{\lambda}_i^{(n)} = 2 \cos i\pi/(2n+2)$ , 则容易验证:  $ZZ^T = 16I_n \otimes I_m - T_n \otimes T_m$ ,  $T_n = \Omega_n \bar{\Lambda}_n \Omega_n$  [21], 因此

$$16I_n \otimes I_m - \bar{\Lambda}_n^2 \otimes \bar{\Lambda}_m^2 = \Lambda^2, \quad T_n \otimes T_m = (\Omega_n \otimes \Omega_m)(\bar{\Lambda}_n^2 \otimes \bar{\Lambda}_m^2)(\Omega_n \otimes \Omega_m)$$

$HH^T = (\Omega_n \otimes \Omega_m)(16I_n \otimes I_m - \bar{\Lambda}_n^2 \otimes \bar{\Lambda}_m^2)(\Omega_n \otimes \Omega_m) = ZZ^T$ . 证毕。

2). 根据命题 2 结论 1),  $ZZ^T$  的特征值为  $(\lambda_{i,j}^{(n,m)})^2$ ,  $1 \leq i \leq n$ ,  $1 \leq j \leq m$ . 所以,

$\text{Cond}(ZZ^T) = \lambda_{\max}(ZZ^T) / \lambda_{\min}(ZZ^T) = (\lambda_{n,m}^{(n,m)} / \lambda_{1,1}^{(n,m)})^2$ . 证毕。

3). 记  $g(x, y) = (x + y + xy) / (1 + x + y)$ ,  $x > 1$ ,  $y > 1$ ,  $\alpha_n = \cot^2 \pi / (2n + 2)$ , 则容易验证:  $\alpha_n = O(n^2)$ ,  $2 \min(x, y) + 2 \geq 2g(x, y) \geq \min(x, y)$ ,  $\text{Cond}(ZZ^T) = g(\alpha_n, \alpha_m)$ , 所以,  $g(x, y) = O(\min(x, y))$ . 结论成立. 证毕。

**命题 3:** 对于任意  $3(m+1)(n+1)$  阶正定对称矩阵  $P$ , 有  $\text{Cond}(ZPZ^T) \leq \text{Cond}(P)\text{Cond}(ZZ^T)$ 。

**证明:** 根据命题 1 结论 1) 可得. 证毕。

根据定义 1、定义 2 和命题 2,  $Z$  是  $mn \times 3(m+1)(n+1)$  病态因子; 矩阵  $H$  是属于  $Z$  的去病因子。

### 3. 二维泊松方程边值问题的 Lagrange 有限元离散系统

#### 3.1. Lagrange 有限元离散格式

使用有限元方法, 求解如下变分形式的二维泊松方程边值问题[22]

$$\iint_D [p(\nabla u \cdot \nabla v) + quv] dx dy + \int_{\partial D} \sigma puv ds - \iint_D f v dx dy = 0, \quad \forall v \quad (4)$$

其中  $\nabla u = (\partial u / \partial x, \partial u / \partial y)$ ,  $p = p(x, y) \geq p_{\min} > 0$ ,  $q = q(x, y) \geq 0$ ,  $\sigma = \sigma(x, y) \geq 0$ ,  $f = f(x, y)$ ,  $(x, y) \in G = [a_1, b_1] \times [a_2, b_2] \subset R^2$ 。

对  $G$  做剖分:  $a_1 = x_0 < x_1 < \dots < x_n < x_{n+1} = b_1$ ,  $a_2 = y_0 < y_1 < \dots < y_m < y_{m+1} = b_2$ , 记  $h_{i,x} = x_i - x_{i-1}$ ,  $h_{j,y} = y_j - y_{j-1}$ ,  $e_{i,j} = [x_{i-1}, x_i] \times [y_{j-1}, y_j]$ ,  $u_{i,j} = u(x_i, y_j)$ ,  $v_{i,j} = v(x_i, y_j)$ 。  $0 \leq i \leq n+1$ ,  $0 \leq j \leq m+1$ ,  $h_{\min} = \min_{i,j} \{h_{i,x}, h_{j,y}\}$ ,  $h = \max_{i,j} \{h_{i,x}, h_{j,y}\}$ , 则(4)式变成

$$\sum_{j=1}^{m+1} \sum_{i=1}^{n+1} \left\{ \iint_{e_{i,j}} [p(\nabla u \cdot \nabla v) + quv] dx dy + \int_{\partial e_{i,j}} \sigma puv ds - \iint_{e_{i,j}} f v dx dy \right\} = 0, \quad \forall v \quad (5)$$

取  $(\xi, \eta)$  为自然坐标 [23],  $(\xi, \eta) \in e = [0, 1] \times [0, 1]$ , 在  $e_{ij}$  上, 做变换  $x = x_i(\xi) = x_{i-1} + \xi h_{i,x}$ ,  $y = y_j(\eta) = y_{j-1} + \eta h_{j,y}$  及其逆变换  $\xi = (x - x_{i-1}) / h_{i,x}$ ,  $\eta = (y - y_{j-1}) / h_{j,y}$ , 记  $X_{ij} = (x_i(\xi), y_j(\eta))$ ,  $p_{ij} = p(X_{ij})$ ,  $q_{ij} = q(X_{ij})$ ,  $\sigma_{ij} = \sigma(X_{ij})$ ,  $f_{ij} = f(X_{ij})$ ,  $J_{ij} = h_{i,x} h_{j,y}$ , 则直接验证可以得到

**命题 4:** 1)  $\partial(x_i(\xi), y_j(\eta)) / \partial(\xi, \eta) = \text{diag}(h_{i,x}, h_{j,y})$ ;

$$2) \partial(\xi_i(x), \eta_j(y)) / \partial(x, y) = \text{diag}(h_{i,x}^{-1}, h_{j,y}^{-1});$$

$$3) J_{ij} = |\partial(x_i(\xi), y_j(\eta)) / \partial(\xi, \eta)| = O(h^2);$$

证明：直接验证。证毕

### 3.1.1. 基于四角剖分的 Lagrange 有限元离散格式

在  $e_{ij}$  上, 取 2 次 Lagrange 形函数为:  $N_{\alpha,\beta} = N_{\alpha,\beta}(\xi, \eta) \in P_2[\xi, \eta]$ ,  $(\xi, \eta) \in e$ ,  $\{\alpha, \beta\} \subset \{0, 1\}$ , 满足:  $N_{\alpha,\beta}(\alpha, \beta) = 1$ ,  $N_{\alpha,\beta}(\xi, \eta) = 0$ ,  $(\xi, \eta) \neq (\alpha, \beta)$ ,  $\{\xi, \eta\} \subset \{0, 1\}$ 。

$$\text{记 } \bar{u}_{ij} = (u_{i-1,j-1}, u_{i-1,j}, u_{i,j-1}, u_{i,j})^T, \quad \bar{v}_{ij} = (v_{i-1,j-1}, v_{i-1,j}, v_{i,j-1}, v_{i,j})^T, \quad U_{ij} = \bar{u}_{ij}^T N, \quad V_{ij} = \bar{v}_{ij}^T N$$

$$e = (0, 0, 0, 1)^T, \quad N = (N_{0,0}, N_{0,1}, N_{1,0}, N_{1,1})^T, \quad \bar{N} = \text{diag}(N_{0,0}, N_{0,1}, N_{1,0}, N_{1,1}), \quad \mathbf{M}_{ij} = \nabla \tilde{N}|_{e_{ij}}$$

$$\tilde{N} = \tilde{N} \tilde{N}^T - \frac{1}{4} \begin{pmatrix} N_{0,0} + N_{0,1} & N_{0,0} & N_{0,1} \\ N_{0,0} & N_{0,0} + N_{1,0} & N_{1,0} \\ N_{0,1} & N_{1,0} & N_{1,0} + N_{0,1} \end{pmatrix}, \quad \tilde{N} = \frac{1}{2} \begin{pmatrix} N_{0,0} + N_{0,1} \\ N_{0,0} + N_{1,0} \\ N_{1,0} + N_{0,1} \end{pmatrix} = \begin{pmatrix} \tilde{N}_1 \\ \tilde{N}_2 \\ \tilde{N}_3 \end{pmatrix}$$

则有:

**命题 5:** 1).  $D\tilde{N} = e - N$ ; 2).  $\mathbf{M}_{ij} = (\partial\tilde{N}/\partial\xi, \partial\tilde{N}/\partial\eta) \text{diag}(h_{i,x}^{-1}, h_{j,y}^{-1})$ ; 3).  $\|\mathbf{M}_{ij}\| = O(h^{-1})$ ; 4).  $U_{ij} V_{ij} = \bar{v}_{ij}^T N N^T \bar{u}_{ij} = \bar{v}_{ij}^T (D\tilde{N} D^T + \bar{N}) \bar{u}_{ij}$ ; 5).  $\nabla V_{ij} \cdot \nabla U_{ij} = \bar{v}_{ij}^T \mathbf{D} \mathbf{M}_{ij} \mathbf{M}_{ij}^T \mathbf{D}^T \bar{u}_{ij}$ 。

证明: 1), 2), 3). 利用形函数的归一性, 即  $N_{0,0} + N_{0,1} + N_{1,0} + N_{1,1} = 1$  [24], 直接验证即得。

4). 只需验证  $N N^T = D\tilde{N} D^T + \bar{N}$ 。

5).  $\nabla N|_{e_{ij}} = \mathbf{D} \mathbf{M}_{ij}$ ,  $\nabla V_{ij} \cdot \nabla U_{ij} = \bar{v}_{ij}^T \nabla N (\nabla N)^T|_{e_{ij}} \bar{u}_{ij} = \bar{v}_{ij}^T \mathbf{D} \mathbf{M}_{ij} \mathbf{M}_{ij}^T \mathbf{D}^T \bar{u}_{ij}$ 。证毕

记

$$\begin{aligned} \mathbf{Q}_{i,j} &= \iint_e q_{ij} \bar{N} J_{ij} d\xi d\eta + \int_{\partial e} \sigma_{ij} p_{ij} \bar{N} L_{ij} ds = \text{diag}(q_{i,j}^{(1)}, q_{i,j}^{(2)}, q_{i,j}^{(3)}, q_{i,j}^{(4)}) \\ \mathbf{P}_{i,j} &= \iint_e (p_{ij} \mathbf{M}_{ij} \mathbf{M}_{ij}^T + q_{ij} \bar{N}) J_{i,j} d\xi d\eta + \int_{\partial e} \sigma_{ij} p_{ij} \bar{N} L_{ij} ds \\ b_{i,j} &= \iint_e f_{ij} J_{ij} N d\xi d\eta = (b_{i,j}^{(1)}, b_{i,j}^{(2)}, b_{i,j}^{(3)}, b_{i,j}^{(4)})^T \end{aligned}$$

则有:

**命题 6:**  $\|\mathbf{P}_{i,j}\| = O(\|J_{ij} \mathbf{M}_{ij} \mathbf{M}_{ij}^T\|) = O(1)$ ,  $\|\mathbf{Q}_{i,j}\| = O(\|J_{ij}\|) = O(h^2)$ 。

证明: 根据命题 4 和命题 5 结论 2), 3) 可得。证毕

在  $e_{ij}$  上, 用  $U_{ij}, V_{ij}$  取代  $u, v$ , 则得到问题(4)的基于四角剖分的有限元离散格式[22]:

$$\iint_{e_{ij}} [p(\nabla u \cdot \nabla v) + quv] dx dy + \int_{\partial e_{ij}} \sigma p u v ds - \iint_{e_{ij}} f v dx dy = \bar{v}_{i,j}^T [(D\mathbf{P}_{i,j} D^T + \mathbf{Q}_{i,j}) \bar{u}_{i,j} - b_{i,j}] = 0$$

从而(5)式变成

$$\sum_{i=1}^{n+1} \sum_{j=1}^{m+1} \bar{v}_{i,j}^T (D\mathbf{P}_{i,j} D^T + \mathbf{Q}_{i,j}) \bar{u}_{i,j} = \sum_{i=1}^{n+1} \sum_{j=1}^{m+1} \bar{v}_{i,j}^T b_{i,j} \tag{6}$$

### 3.1.2. 基于三角剖分的 Lagrange 有限元离散格式

将  $e_{ij}$  剖分为两个三角形单元: 以  $X_{i,j}, X_{i,j+1}, X_{i+1,j+1}$  为顶点的三角形单元记为  $e_{ij}^{(1)}$ , 以  $X_{i,j}, X_{i+1,j}, X_{i+1,j+1}$  为顶点的三角形单元记为  $e_{ij}^{(2)}$ , 上述变换下  $e^{(1)} = \{(\xi, \eta) \in e, \xi \leq \eta\}$  为  $e_{ij}^{(1)}$  的像,  $e^{(2)} = \{(\xi, \eta) \in e, \xi \geq \eta\}$  为  $e_{ij}^{(2)}$  的像, 在  $e_{ij}^{(\alpha)}$  上取 Lagrange 形函数  $N_{\beta}^{(\alpha)} = N_{\beta}^{(\alpha)}(\xi_1^{(\alpha)}, \xi_2^{(\alpha)}, \xi_3^{(\alpha)})$ ,  $\beta = 1, 2, 3$ ,  $\xi_1^{(\alpha)}, \xi_2^{(\alpha)}, \xi_3^{(\alpha)}$  是形函数的

自然坐标, 则有  $\xi_1^{(\alpha)} + \xi_2^{(\alpha)} + \xi_3^{(\alpha)} = 1$ ,  $\alpha = 1, 2$  [23].

$$\begin{aligned} \text{记 } \mathbf{N}^{(1)} &= (N_1^{(1)}, 0, N_2^{(1)}, N_3^{(1)})^T, \mathbf{N}^{(2)} = (N_1^{(2)}, N_2^{(2)}, 0, N_3^{(2)})^T, \bar{\mathbf{N}} = \text{diag}(\mathbf{N}^{(1)} + \mathbf{N}^{(2)}), \tilde{\mathbf{N}}^{(1)} = (N_1^{(1)}, 0, N_2^{(1)})^T, \\ \tilde{\mathbf{N}}^{(2)} &= (N_1^{(2)}, N_2^{(2)}, 0)^T, \hat{\mathbf{N}} = \tilde{\mathbf{N}}^{(1)}\tilde{\mathbf{N}}^{(1)T} + \tilde{\mathbf{N}}^{(2)}\tilde{\mathbf{N}}^{(2)T} - \text{diag}(\tilde{\mathbf{N}}^{(1)} + \tilde{\mathbf{N}}^{(2)}), \bar{\mathbf{u}}_{ij} = (u_{i-1,j-1}, u_{i-1,j}, u_{i,j-1}, u_{i,j})^T, \\ \bar{\mathbf{v}}_{ij} &= (v_{i-1,j-1}, v_{i-1,j}, v_{i,j-1}, v_{i,j})^T, \\ U_{ij} &= \bar{\mathbf{u}}_{ij}^T(\mathbf{N}^{(1)}, \mathbf{N}^{(2)}), V_{ij} = \bar{\mathbf{v}}_{ij}^T(\mathbf{N}^{(1)}, \mathbf{N}^{(2)}) \end{aligned}$$

$$\bar{\mathbf{D}} = \frac{1}{2} \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}, \mathbf{M}_{ij}^{(\alpha)} = \left( \frac{\partial \tilde{\mathbf{N}}^{(\alpha)}}{\partial \xi}, \frac{\partial \tilde{\mathbf{N}}^{(\alpha)}}{\partial \eta} \right) \begin{pmatrix} h_{i,x}^{-1} & 0 \\ 0 & h_{j,y}^{-1} \end{pmatrix}, \bar{\mathbf{M}}_{ij}^{(\alpha)} = \left( \frac{\partial \mathbf{N}^{(\alpha)}}{\partial x}, \frac{\partial \mathbf{N}^{(\alpha)}}{\partial y} \right) \Bigg|_{e_{ij}^{(\alpha)}}, \alpha = 1, 2$$

则有:

- 命题 7:** 1).  $\mathbf{D}\bar{\mathbf{D}}\tilde{\mathbf{N}}^{(\alpha)} = \mathbf{e} - \mathbf{N}^{(\alpha)}$ ; 2).  $\|\mathbf{M}_{ij}^{(\alpha)}\| = O(h^{-1})$ ;  
3).  $\bar{\mathbf{M}}_{ij}^{(\alpha)} = -\mathbf{D}\bar{\mathbf{D}}\mathbf{M}_{ij}^{(\alpha)}$ ; 4).  $U_{ij}V_{ij} = \bar{\mathbf{v}}_{ij}^T[\mathbf{D}\bar{\mathbf{D}}\tilde{\mathbf{N}}\bar{\mathbf{D}}^T\mathbf{D}^T + \bar{\mathbf{N}}]\bar{\mathbf{u}}_{ij}$ ;  
5).  $\nabla V_{ij} \cdot \nabla U_{ij} = \bar{\mathbf{v}}_{ij}^T\mathbf{D}\bar{\mathbf{D}}(\mathbf{M}_{ij}^{(1)}\mathbf{M}_{ij}^{(1)T} + \mathbf{M}_{ij}^{(2)}\mathbf{M}_{ij}^{(2)T})\bar{\mathbf{D}}^T\mathbf{D}^T\bar{\mathbf{u}}_{ij}$ ;

证明: 1), 2), 利用形函数的归一性, 即  $N_1^{(\alpha)} + N_2^{(\alpha)} + N_3^{(\alpha)} = 1$  [24], 直接验证即得.

3) 根据 1), 可验证

$$\bar{\mathbf{M}}_{ij}^{(\alpha)} = \left( \frac{\partial \mathbf{N}^{(\alpha)}}{\partial x}, \frac{\partial \mathbf{N}^{(\alpha)}}{\partial y} \right) \Bigg|_{e_{ij}^{(\alpha)}} = -\mathbf{D}\bar{\mathbf{D}} \left( \frac{\partial \tilde{\mathbf{N}}^{(\alpha)}}{\partial x}, \frac{\partial \tilde{\mathbf{N}}^{(\alpha)}}{\partial y} \right) \Bigg|_{e_{ij}^{(\alpha)}} = -\mathbf{D}\bar{\mathbf{D}}\mathbf{M}_{ij}^{(\alpha)}$$

4).  $U_{ij}V_{ij} = \bar{\mathbf{v}}_{ij}^T(\mathbf{N}^{(1)}\mathbf{N}^{(1)T} + \mathbf{N}^{(2)}\mathbf{N}^{(2)T})\bar{\mathbf{u}}_{ij}$ , 容易验证  $\mathbf{N}^{(1)}\mathbf{N}^{(1)T} + \mathbf{N}^{(2)}\mathbf{N}^{(2)T} = \mathbf{D}\bar{\mathbf{D}}\tilde{\mathbf{N}}\bar{\mathbf{D}}^T\mathbf{D}^T + \bar{\mathbf{N}}^{(1)} + \bar{\mathbf{N}}^{(2)}$ , 所以结论成立.

5).  $\nabla V_{ij} \cdot \nabla U_{ij} = \bar{\mathbf{v}}_{ij}^T[\bar{\mathbf{M}}_{ij}^{(1)}\bar{\mathbf{M}}_{ij}^{(1)T} + \bar{\mathbf{M}}_{ij}^{(2)}\bar{\mathbf{M}}_{ij}^{(2)T}]\bar{\mathbf{u}}_{ij}$ , 利用结论 3)即可验证. 证毕

记

$$\begin{aligned} \bar{b}_{i,j} &= \iint_{e^{(\alpha)}} f_{ij} J_{ij}(\mathbf{N}^{(1)} + \mathbf{N}^{(2)}) d\xi d\eta = (\bar{b}_{i,j}^{(1)}, \bar{b}_{i,j}^{(2)}, \bar{b}_{i,j}^{(3)}, \bar{b}_{i,j}^{(4)})^T \\ \mathbf{Q}_{i,j}^{(\alpha)} &= \iint_{e^{(\alpha)}} q_{ij} \bar{\mathbf{N}}^{(\alpha)} J_{ij} d\xi d\eta + \int_{\partial e^{(\alpha)}} \sigma_{ij} p_{ij} \bar{\mathbf{N}}^{(\alpha)} L_{ij} ds \\ \mathbf{P}_{i,j}^{(\alpha)} &= \iint_{e^{(\alpha)}} (p_{ij} \mathbf{M}_{ij}^{(\alpha)} \mathbf{M}_{ij}^{(\alpha)T} + q_{ij} \bar{\mathbf{N}}^{(\alpha)}) J_{i,j} d\xi d\eta + \int_{\partial e^{(\alpha)}} \sigma_{ij} p_{ij} \bar{\mathbf{N}}^{(\alpha)} L_{ij} ds, \alpha = 1, 2 \\ \bar{\mathbf{Q}}_{i,j} &= \mathbf{Q}_{i,j}^{(1)} + \mathbf{Q}_{i,j}^{(2)} = \text{diag}(\bar{q}_{i,j}^{(1)}, \bar{q}_{i,j}^{(2)}, \bar{q}_{i,j}^{(3)}, \bar{q}_{i,j}^{(4)}) \\ \bar{\mathbf{P}}_{i,j} &= \bar{\mathbf{D}}[\mathbf{P}_{i,j}^{(1)} + \mathbf{P}_{i,j}^{(2)}]\bar{\mathbf{D}}^T \end{aligned}$$

则有:

**命题 8:**  $\|\bar{\mathbf{P}}_{i,j}\| = O(\|J_{ij}(\mathbf{M}_{ij}^{(1)}\mathbf{M}_{ij}^{(1)T} + \mathbf{M}_{ij}^{(2)}\mathbf{M}_{ij}^{(2)T})\|) = O(1)$ ,  $\|\bar{\mathbf{Q}}_{i,j}\| = O(\|J_{ij}\|) = O(h^2)$ .

**证明:** 根据命题 4 和命题 7 结论 2), 3)可得. 证毕

在  $e_{ij}$  上, 用  $U_{ij}, V_{ij}$  作为  $u, v$  的近似, 则也可以得到问题(4)的基于三角剖分的有限元离散格式:

$$\iint_{e_{ij}} [p(\nabla u \cdot \nabla v) + quv] dx dy + \int_{\partial e_{ij}} \sigma p u v ds - \iint_{e_{ij}} f v dx dy = \bar{\mathbf{v}}_{i,j}^T [(\mathbf{D}\bar{\mathbf{P}}_{i,j}\mathbf{D}^T + \bar{\mathbf{Q}}_{i,j})\bar{\mathbf{u}}_{i,j} - \bar{b}_{i,j}] = 0$$

从而(5)式变成

$$\sum_{i=1}^{n+1} \sum_{j=1}^{m+1} \tilde{v}_{i,j}^T (D\bar{P}_{i,j}D^T + \bar{Q}_{i,j}) \bar{u}_{i,j} = \sum_{i=1}^{n+1} \sum_{j=1}^{m+1} \tilde{v}_{i,j}^T \bar{b}_{i,j} \tag{7}$$

### 3.2. 二次 Lagrange 有限元离散系统

记  $\mathbf{v}_i = (v_{i,1}, \dots, v_{i,m})^T$ ,  $\tilde{\mathbf{v}}_i = (v_{i,0}, \mathbf{v}_i^T, v_{i,m+1})^T$ ,  $\mathbf{v} = (\mathbf{v}_1^T, \dots, \mathbf{v}_n^T)^T$ ,  $\tilde{\mathbf{v}} = (\tilde{\mathbf{v}}_0^T, \dots, \tilde{\mathbf{v}}_{n+1}^T)^T$   
 $\mathbf{u}_i = (u_{i,1}, \dots, u_{i,m})^T$ ,  $\tilde{\mathbf{u}}_i = (u_{i,0}, \mathbf{u}_i^T, u_{i,m+1})^T$ ,  $\mathbf{u} = (\mathbf{u}_1^T, \dots, \mathbf{u}_n^T)^T$ ,  $\tilde{\mathbf{u}} = (\tilde{\mathbf{u}}_0^T, \dots, \tilde{\mathbf{u}}_{n+1}^T)^T$   
 $\mathbf{Q}_i^{(s)} = \text{diag}(q_{i,2}^{(2s-1)} + q_{i,1}^{(2s)}, \dots, q_{i,m+1}^{(2s-1)} + q_{i,m}^{(2s)})$ ,  $\tilde{\mathbf{Q}}_i^{(s)} = \text{diag}(q_{i,1}^{(2s-1)}, \mathbf{Q}_i^{(s)}, q_{i,m+1}^{(2s)})$ ,  $s = 1, 2$   
 $\tilde{\mathbf{Q}} = \text{diag}(\tilde{\mathbf{Q}}_1^{(1)}, \tilde{\mathbf{Q}}_2^{(1)} + \tilde{\mathbf{Q}}_1^{(2)}, \dots, \tilde{\mathbf{Q}}_{n+1}^{(1)} + \tilde{\mathbf{Q}}_n^{(2)}, \tilde{\mathbf{Q}}_{n+1}^{(2)})$ ,  $\mathbf{Q} = \text{diag}(\mathbf{Q}_2^{(1)} + \mathbf{Q}_1^{(2)}, \dots, \mathbf{Q}_{n+1}^{(1)} + \mathbf{Q}_n^{(2)})$   
 $\mathbf{b}_i^{(s)} = (b_{i,2}^{(2s-1)} + b_{i,1}^{(2s)}, \dots, b_{i,m+1}^{(2s-1)} + b_{i,m}^{(2s)})^T$ ,  $\tilde{\mathbf{b}}_i^{(s)} = (b_{i,1}^{(2s-1)}, \mathbf{b}_i^{(s)T}, b_{i,m+1}^{(2s)})^T$ ,  $s = 1, 2$   
 $\tilde{\mathbf{b}} = (\tilde{\mathbf{b}}_1^{(1)T}, \tilde{\mathbf{b}}_2^{(1)T} + \tilde{\mathbf{b}}_1^{(2)T}, \dots, \tilde{\mathbf{b}}_{n+1}^{(1)T} + \tilde{\mathbf{b}}_n^{(2)T}, \tilde{\mathbf{b}}_{n+1}^{(2)T})^T$ ,  $\mathbf{b} = (\mathbf{b}_2^{(1)T} + \mathbf{b}_1^{(2)T}, \dots, \mathbf{b}_{n+1}^{(1)T} + \mathbf{b}_n^{(2)T})^T$   
 $\mathbf{P}_i = \text{diag}(\mathbf{P}_{i,1}, \dots, \mathbf{P}_{i,m+1})$ ,  $\mathbf{P} = \text{diag}(\mathbf{P}_1, \dots, \mathbf{P}_{n+1})$ , 则有:

**命题 9:** 离散格式(6)的矩阵形式为:

$$\tilde{\mathbf{v}}^T [(\bar{\mathbf{Z}}\bar{\mathbf{P}}\bar{\mathbf{Z}}^T + \tilde{\mathbf{Q}})\tilde{\mathbf{u}} - \tilde{\mathbf{b}}] = 0 \tag{8}$$

(8)式中  $\bar{\mathbf{Z}} = (\mathbf{I}_{n+1}, \mathbf{0}_{n+1})^T \otimes \mathbf{Z}_0 + (\mathbf{0}_{n+1}, \mathbf{I}_{n+1})^T \otimes \mathbf{Z}_1$ ,  $\mathbf{Z}_i = (\mathbf{I}_{m+1}, \mathbf{0}_{m+1})^T \otimes \sigma_{2i} + (\mathbf{0}_{m+1}, \mathbf{I}_{m+1})^T \otimes \sigma_{2i+1}$ ,  $i = 0, 1$ 。

证明: 容易验证:

$$\begin{aligned} \sum_{i=1}^{n+1} \sum_{j=1}^{m+1} \tilde{v}_{ij}^T D\bar{P}_{ij}D^T \bar{u}_{ij} &= \sum_{i=1}^{n+1} (\tilde{v}_{i-1}^T, \tilde{v}_i^T) \begin{pmatrix} \bar{\mathbf{Z}}_0 \\ \bar{\mathbf{Z}}_1 \end{pmatrix} \mathbf{P}_i (\bar{\mathbf{Z}}_0^T, \bar{\mathbf{Z}}_1^T) \begin{pmatrix} \tilde{\mathbf{u}}_{i-1} \\ \tilde{\mathbf{u}}_i \end{pmatrix} = \tilde{\mathbf{v}}^T \bar{\mathbf{Z}}\bar{\mathbf{P}}\bar{\mathbf{Z}}^T \tilde{\mathbf{u}} \\ \sum_{i=1}^{n+1} \sum_{j=1}^{m+1} \tilde{v}_{ij}^T \mathbf{Q}_{i,j} \bar{u}_{ij} &= \sum_{i=1}^{n+1} (\tilde{v}_{i-1}^T \tilde{\mathbf{Q}}_i^{(1)} \tilde{\mathbf{u}}_{i-1} + \tilde{v}_i^T \tilde{\mathbf{Q}}_i^{(2)} \tilde{\mathbf{u}}_i) = \tilde{\mathbf{v}}^T \tilde{\mathbf{Q}} \tilde{\mathbf{u}} \\ \sum_{i=1}^{n+1} \sum_{j=1}^{m+1} \tilde{v}_{ij}^T \mathbf{b}_{ij} &= \sum_{i=1}^{n+1} (\tilde{v}_{i-1}^T \tilde{\mathbf{b}}_i^{(1)} + \tilde{v}_i^T \tilde{\mathbf{b}}_i^{(2)}) = \tilde{\mathbf{v}}^T \tilde{\mathbf{b}} \end{aligned}$$

从而将问题(4)的离散格式(6)写成公式(8)。证毕

根据式(8), 有

$$(\bar{\mathbf{Z}}\bar{\mathbf{P}}\bar{\mathbf{Z}}^T + \tilde{\mathbf{Q}})\tilde{\mathbf{u}} = \tilde{\mathbf{b}} \tag{9}$$

将(9)式中与边界网格点的值  $u_{0,j}, u_{n+1,j}, u_{i,0}, u_{i,m+1}$  相关的部分移到右端, 就得到问题(4)基于四角剖分的有限元离散系统:

$$\mathbf{A}\mathbf{u} = \mathbf{b} \tag{10}$$

式(10)中, 刚度矩阵为

$$\mathbf{A} = \mathbf{Z}\mathbf{P}\mathbf{Z}^T + \mathbf{Q} \tag{11}$$

类似地, 可以根据离散格式(7)得到问题(4)基于三角剖分的有限元离散系统:

$$\bar{\mathbf{A}}\mathbf{u} = (\bar{\mathbf{Z}}\bar{\mathbf{P}}\bar{\mathbf{Z}}^T + \bar{\mathbf{Q}})\mathbf{u} = \bar{\mathbf{b}} \tag{12}$$

式(12)中  $\bar{\mathbf{P}}_i = \text{diag}(\bar{\mathbf{P}}_{i,1}, \dots, \bar{\mathbf{P}}_{i,m+1})$ ,  $\bar{\mathbf{P}} = \text{diag}(\bar{\mathbf{P}}_1, \dots, \bar{\mathbf{P}}_{n+1})$

$$\bar{Q}_i^{(s)} = \text{diag}(\bar{q}_{i,2}^{(2s-1)} + \bar{q}_{i,1}^{(2s)}, \dots, \bar{q}_{i,m+1}^{(2s-1)} + \bar{q}_{i,m}^{(2s)}), \quad \bar{b}_i^{(s)} = (\bar{b}_{i,2}^{(2s-1)} + \bar{b}_{i,1}^{(2s)}, \dots, \bar{b}_{i,m+1}^{(2s-1)} + \bar{b}_{i,m}^{(2s)})^T, \quad s=1,2$$

$$\bar{Q} = \text{diag}(\bar{Q}_2^{(1)} + \bar{Q}_1^{(2)}, \dots, \bar{Q}_{n+1}^{(1)} + \bar{Q}_n^{(2)}), \quad \bar{b} = (\bar{b}_2^{(1)T} + \bar{b}_1^{(2)T}, \dots, \bar{b}_{n+1}^{(1)T} + \bar{b}_n^{(2)T})^T$$

不论是三角剖分, 还是四角剖分, 离散系统的总刚度矩阵都具有式(1)的形式, 这个形式有利于评估条件数, 便于设计预条件子和算法。

#### 4. 对有限元离散系统的病因抑制方法及其预处理效果

根据命题 6、8, (10)和(12)式中, 分别有  $\|ZPZ^T\| \gg \|Q\|$  和  $\|Z\bar{P}Z^T\| \gg \|\bar{Q}\|$ , 所以在总刚度矩阵  $A$  和  $\bar{A}$  中,  $ZPZ^T$  和  $Z\bar{P}Z^T$  分别是  $A$  与  $\bar{A}$  的主体,  $\text{Cond}(A) \approx \text{Cond}(ZPZ^T)$ ,  $\text{Cond}(\bar{A}) \approx \text{Cond}(Z\bar{P}Z^T)$ ; 矩阵  $P$  和  $\bar{P}$  的元素由  $p = p(x, y)$  的振幅, 以及网格大小决定, 几乎不受  $m, n$  影响, 所以  $\text{Cond}(P)$  和  $\text{Cond}(\bar{P})$  关于  $A$  与  $\bar{A}$  的阶数都是一致有界的, 根据定义 1、3 和命题 2, (10)和(12)式分别说明了刚度矩阵  $A$  与  $\bar{A}$  都具有(1)式的病态结构。

(10)和(12)式中,  $Q$  和  $\bar{Q}$  对刚度矩阵的病态影响很小。根据命题 3, 由病态因子  $Z$  表达的原发病态是离散系统的主要病态,  $P, Q$  和  $\bar{P}, \bar{Q}$  等表达的继发病态是散系统的次要病态, 原发病态是微分算子造成的, 是本质的, 离散精度越高, 刚度矩阵的阶数越大, 刚度矩阵条件数就越大, 原发病态越严重; 继发病态来自数值方法和问题(4)本身, 是非本质的, 可以随着应用问题的不同而不同, 可以在应用中调整。

综上所述, (10)和(12)说明, 基于不同剖分的离散系统, 有类似的病态结构和相同的病态因子  $Z$ , 即有相同的原发病态。

使用  $Z$  的去病因子  $H$  作为预条件子, 对方程(10)、(12)进行预处理, 其中方程(10)化成

$$H^{-1}AH^{-T}\tilde{u} = (H^{-1}ZPZ^TH^{-T} + H^{-1}QH^{-T})\tilde{u} = \tilde{b} \quad (13)$$

其中  $\tilde{u} = H^T u$ ,  $\tilde{b} = H^{-1}b$ 。

根据命题 1 的结论 2), 有  $\text{Cond}(H^{-1}AH^{-T}) \approx \text{Cond}(H^{-1}ZPZ^TH^{-T}) \leq \text{Cond}(P)$ , 因此,  $HH^T$  是最优预条件子[5]。

使用  $Z$  的去病因子  $H$  作为预条件子, 对方程(12)进行预处理, 也有类似的结果。

$H, H^{-1}$  有简单确定的结构, 都是对角矩阵与离散正弦变换矩阵的直积的积, 因此预条件子的构造以及预处理计算不会明显增加计算成本, 其中预处理需要的计算是矩阵  $H^T, H^{-1}$  与向量的乘积, 相当于对向量实施快速离散正弦变换, 每次需要的计算操作数为  $O(nm \log_2(nm))$ , 所以每个迭代步的计算量没有显著增加。预处理后, 方程(13)的系数矩阵仍然是正定对称矩阵, 所以仍然可以与经典 Krylov 子空间方法等求解方法相结合, 形成预处理求解方法[4]。对方程(12)进行同样的预处理, 也有类似的结果。

#### 5. 例证

对于积分形式的二维泊松方程边值问题(4) [22], 如果  $\sigma = q = 0$ ,  $p$  为常数, 取 2 次 Lagrange 形函数为:  $N_1^{(1)} = 1 - \eta$ ,  $N_2^{(1)} = \xi$ ,  $N_3^{(1)} = \eta - \xi$ ,  $N_1^{(2)} = 1 - \xi$ ,  $N_2^{(2)} = \xi - \eta$ ,  $N_3^{(2)} = \eta$ , 并取  $h_x = h_{i,x}$ ,  $h_y = h_{j,y}$ , 记  $\bar{h} = h_y h_x^{-1}$ ,  $B_n = (-1, 2, -1)_n$ ,  $C_n = (0, 1, 1)_n$ , 则

$$P_{i,j} = \frac{p}{8} \begin{pmatrix} 2\bar{h} + \bar{h}^{-1} & \bar{h}^{-1} & \\ \bar{h}^{-1} & 2\bar{h}^{-1} & \bar{h}^{-1} \\ & \bar{h}^{-1} & 2\bar{h} + \bar{h}^{-1} \end{pmatrix}$$



$$\begin{aligned}
 \mathbf{P} &= \frac{p}{8} \mathbf{I}_{mm} \otimes \begin{pmatrix} 2\bar{h} + \bar{h}^{-1} & \bar{h}^{-1} & & \\ & \bar{h}^{-1} & 2\bar{h}^{-1} & \bar{h}^{-1} \\ & & \bar{h}^{-1} & 2\bar{h} + \bar{h}^{-1} \end{pmatrix} \\
 \mathbf{A} = \mathbf{ZPZ}^T &= \frac{p}{2} \left[ 2\bar{h} \begin{pmatrix} \mathbf{B}_n & & & \\ & \ddots & & \\ & & \mathbf{B}_n & \\ & & & \mathbf{B}_n \end{pmatrix} + \bar{h}^{-1} \begin{pmatrix} 4\mathbf{I}_n & -\mathbf{C}_n & & \\ -\mathbf{C}_n^T & \ddots & \ddots & \\ & \ddots & \ddots & -\mathbf{C}_n \\ & & -\mathbf{C}_n^T & 4\mathbf{I}_n \end{pmatrix} \right] \\
 \text{Cond}(\mathbf{ZPZ}^T) &\leq \text{Cond}(\mathbf{P})\text{Cond}(\mathbf{Z}\mathbf{Z}^T)
 \end{aligned}$$

$\text{Cond}(\mathbf{H}^{-1}\mathbf{A}\mathbf{H}^{-T}) \leq \text{Cond}(\mathbf{P}) \leq 4 \max(\bar{h}, \bar{h}^{-1}) / \min(\bar{h}, \bar{h}^{-1})$ ，即预处理后，矩阵的条件数一致有界。

特别地，如果取  $h_x = h_y$ ，则  $\bar{h} = 1$ ， $\text{Cond}(\mathbf{H}^{-1}\mathbf{A}\mathbf{H}^{-T}) \leq 4$ ，即预处理后，矩阵的条件数小于 4，因此，去病因子  $\mathbf{H}$  是最优的预条件子。

## 6. 结论

1) 二维泊松方程边值问题基于三角、四角剖分的 Lagrange 有限元离散系统虽然不同，其系数矩阵却有类似的病态结构，有相同的病态因子，它的病态由原发病态和继发病态组成，原发性病态是由微分算子引起的，由病态因子表达，是主要的、本质的，不受剖分方法影响，可以使用相同的去病因子进行预处理；继发病态是由数值方法和边值问题的系数函数引起，是次要的、非本质的，可以在应用中调整。

2) 去病因子是最优预条件子[5]，精准针对病态因子起作用，可消除原发性病态，几乎将矩阵的条件数降为常数，相当于精准地抑制了病态因子的致病作用。去病因子的特别结构，使得预处理过程基本不增加计算操作数，并且预处理后矩阵保持正定对称，有利于与其他求解方法结合。

3) 病态因子及其去病因子是相对独立的，与边值问题中函数  $p$ ， $q$ ， $f$  等没有关系，不论三角剖分，还是四角剖分，都不影响病态因子和去病因子，但是剖分方法、步长可以影响继发病态。

4) 去病因子的设计是针对病态因子(无须针对整个刚度矩阵)，只要从刚度矩阵中分离出病态因子，就可以得到去病因子，因此去病因子的获取难度有限。

综上所述，与目前主流的预处理方法相比，本文提出的病因抑制方法，具有难度低，不增加计算成本，精准抑制原发病态等特点。

## 基金项目

国家自然科学基金资助项目(61473329, 61601125)。

## 参考文献

- [1] Zhong X.J. (1988) The Convergence of Krylov Subspace Methods for Large Nonsymmetric Linear Systems. *Acta Mathematica Sinica*, **14**, 507-518. <https://doi.org/10.1007/BF02580408>
- [2] 吴勃英, 王德明, 丁效华, 李道华. 数值分析原理[M]. 北京: 科学出版社, 2003: 71-101.
- [3] 张科. 两类线性方程组的预处理技术及数值求解方法[D]: [博士学位论文]. 上海: 上海大学, 2014: 7-9.
- [4] 李荣华. 偏微分方程数值解法[M]. 第2版. 北京: 高等教育出版社, 2010: 139-151, 13-15.
- [5] 于春肖, 苑润浩, 穆运峰. 新预处理 ILUCG 法求解稀疏病态线性方程组[J]. 数值计算与计算机应用, 2014, 35(1): 21-27.
- [6] Xue, Q.F., Gao, X.B. and Liu, X.G. (2014) Comparison Theorems for a Class of Preconditioned AOR Iterative Methods. *Journal of Mathematics*, **34**, 448-460.
- [7] 潘春平, 马成荣, 曹文方, 王红玉. 一类预条件 AOR 迭代法的收敛性分析[J]. 数学杂志, 2013, 33(3): 479-484.

- [8] 罗芳. L-矩阵的多参数预条件 AOR 迭代法[J]. 数学的实践与认识, 2013, 43(15): 277-282.
- [9] 吴建平, 赵军, 马怀发, 宋君强, 张卫民, 李晓梅. 一般稀疏线性方程组的因子组合型并行预条件研究[J]. 计算机应用与软件, 2012, 29(5): 6-9, 108.
- [10] 李继成, 蒋耀林. 预条件 IMGS 迭代方法的比较定理[J]. 数学物理学报, 2011, 31(4): 880-886.
- [11] Pan, C.P. (2011) A New Effective Preconditioned Gauss-Seidel Iteration Method. *Journal on Numerical Methods and Computer Applications*, **32**, 267-273.
- [12] 张文生. 科学计算中的偏微分方程有限差分法[M]. 北京: 高等教育出版社, 2006: 205-209.
- [13] Bai, Z.Z. and Li, G.Q. (2003) Restrictively Preconditioned Conjugate Gradient Methods for Systems of Linear Equations. *IMA Journal of Numerical Analysis*, **23**, 561-580. <https://doi.org/10.1093/imanum/23.4.561>
- [14] Wang, Z.Q. (2015) Restrictively Preconditioned Chebyshev Method for Solving Systems of Linear Equations. *Journal of Engineering Mathematics*, **93**, 61-76. <https://doi.org/10.1007/s10665-014-9694-5>
- [15] Shen, H.L., Shao, X.H. and Zhang, T. (2012) Preconditioned Iterative Methods for Solving Weighted Linear Least Squares Problems. *Applied Mathematics and Mechanics*, **33**, 375-384. <https://doi.org/10.1007/s10483-012-1557-x>
- [16] 张衡, 郑汉垣. 二维泊松方程边值问题有限差分方程的病态结构和最优预条件子[J]. 福建师范大学福清分校学报, 2018(5): 1-6.
- [17] 张衡. 一维问题  $m$  次 lagrange 形函数有限元方程的条件数与预处理[J]. 石河子大学学报(自然科学版), 2017, 35(4): 518-521.
- [18] 张衡. 两点边值问题 2 次 lagrange 形函数有限元方程的条件数和预处理[J]. 福州大学学报(自然科学版), 2017, 45(5): 617-622.
- [19] 张衡. 两点边值问题 3 次 lagrange 形函数有限元方程的条件数和预处理[J]. 计算力学学报, 2017, 34(5): 672-676.
- [20] 张衡, 郑汉垣. 两点边值问题网格方程病态机理和预处理[J]. 福州大学学报(自然科学版), 2019, 47(3): 296-299, 306.
- [21] 吕涛, 石济民, 林振宝. 区域分解算法——偏微分方程数值解新技术[M]. 北京: 科学出版社, 1997: 165-167.
- [22] 林群. 微分方程数值解法基础教程[M]. 第 2 版. 北京: 科学出版社, 2003: 87-111.
- [23] 李开泰, 黄艾香, 黄庆怀. 有限元方法及其应用[M]. 北京: 科学出版社, 2006: 20-27.
- [24] 朱伯芳. 有限单元法原理与应用[M]. 第 4 版. 北京: 中国水利水电出版社, 2018: 120-121.