



非精确Newton方法中线性迭代收敛判据研究

冯选燕 燕振国 朱华君 马燕凯 冯新龙

Study on the convergence criterion of linear iteration in inexact Newton methods

FENG Xuanyan, YAN Zhenguo, ZHU Huajun, MA Yankai, FENG Xinlong

引用本文:

冯选燕, 燕振国, 朱华君, 等. 非精确Newton方法中线性迭代收敛判据研究[J]. 空气动力学学报, 2023, 41(12): 28–36. DOI: 10.7638/kqdlxxb-2023.0001

FENG Xuanyan, YAN Zhenguo, ZHU Huajun, et al. Study on the convergence criterion of linear iteration in inexact Newton methods[J]. *Acta Aerodynamica Sinica*, 2023, 41(12): 28–36. DOI: 10.7638/kqdlxxb-2023.0001

在线阅读 View online: <https://doi.org/10.7638/kqdlxxb-2023.0001>

您可能感兴趣的其他文章

Articles you may be interested in

基于雅可比矩阵精确计算的GMRES隐式方法在间断Galerkin有限元中的应用

Applications of GMRES based on exact calculations of Jacobian matrix in discontinuous Galerkin methods

空气动力学学报. 2019, 37(1): 121–132 <https://doi.org/10.7638/kqdlxxb-2018.0189>

空化流动隐式求解方法研究

Investigation on implicit numerical method for cavitation flow

空气动力学学报. 2017, 35(1): 27–32 <https://doi.org/10.7638/kqdlxxb-2015.0027>

基于物面弱边界条件的并行全隐式流场求解方法

A parallel full implicit flow solver based on weakly imposed wall boundary condition

空气动力学学报. 2017, 35(1): 46–56 <https://doi.org/10.7638/kqdlxxb-2016.0121>

适用于非定常流模拟的分布式并行GMRES方法

A GMRES method on distributed parallel computers for unsteady flow simulation

空气动力学学报. 2019, 37(5): 834–843 <https://doi.org/10.7638/kqdlxxb-2017.0185>

高超声速边界层转捩预测中的关键科学问题——感受性、扰动演化及转捩判据研究进展

Progress in key scientific problems of hypersonic boundary-layer transition prediction: receptivity, evolution of disturbances and transition criterion

空气动力学学报. 2020, 38(2): 355–367 <https://doi.org/10.7638/kqdlxxb-2020.0056>

用隐式高精度间断伽辽金方法模拟可压层流和湍流

Implicit high order discontinuous Galerkin method for compressible laminar and turbulent flow simulation

空气动力学学报. 2018, 36(3): 535–541 <https://doi.org/10.7638/kqdlxxb-2017.0183>

地址: 四川省绵阳市二环路南段8号11信箱9分箱

电话: 0816-2463375

Email: kqdlxxb@163.com

网址: <http://kqdlxxb.xml-journal.net/>



关注微信公众号

获得更多资讯信息

文章编号: 0258-1825(2023)12-0028-09

非精确 Newton 方法中线性迭代收敛判据研究

冯选燕^{1,2}, 燕振国^{2,*}, 朱华君^{2,*}, 马燕凯², 冯新龙¹

(1. 新疆大学 数学与系统科学学院, 乌鲁木齐 830000;

2. 空天飞行空气动力学科学与技术全国重点实验室, 绵阳 621000)

摘要:在计算流体力学中, 采用隐式时间推进方法时通常需要采用 Newton 类迭代方法求解大型非线性离散系统。每步非线性迭代需求解由非线性系统 Jacobian 矩阵组成的大型线性方程组, 其中线性方程组求解误差会对非线性系统的收敛性产生显著影响, 然而对存在 Jacobian 矩阵误差情况下的线性迭代收敛判据缺乏深入的研究。本文针对上述问题, 首先给出了存在 Jacobian 矩阵误差和线性迭代误差情况下 Newton 迭代式的形式, 并通过数值测试验证了 Jacobian 矩阵误差对迭代产生较大影响的可能性; 其次对常见的两种不同类型的线性迭代收敛判据进行了数值测试, 重点研究了存在 Jacobian 矩阵误差情况下容易产生的过度求解问题; 最后, 结合上述两类判据的特点发展了一种新的线性迭代收敛判据, 结果表明: 新提出的迭代收敛判据能够有效缓解过度求解问题, 从而提高计算效率。

关键词:Newton 方法; 隐式时间推进; Jacobian 矩阵误差; 线性迭代; 收敛判据; 矩阵刚性; 计算流体力学

中图分类号: O242.2; V211.3 文献标识码: A doi: 10.7638/kqdlxxb-2023.0001

Study on the convergence criterion of linear iteration in inexact Newton methods

FENG Xuanyan^{1,2}, YAN Zhenguo^{2,*}, ZHU Huajun^{2,*}, MA Yankai², FENG Xinlong¹

(1. School of Mathematics and System Science, Xinjiang University, Urumqi 830000, China;

2. National Key Laboratory of Aerodynamic Science and Technology for Air and Space Flight, Mianyang 621000, China)

Abstract: Newton-like iteration methods are usually used for implicit time stepping to solve large scale nonlinear systems. Each of the nonlinear iterative step requires solving large linear equations composed of a Jacobian matrix of the nonlinear system, and the error of solving such linear equations can have a significant impact on the convergence of the nonlinear system. However, the convergence criterion of the linear iteration is less studied when the Jacobian matrix error exists in the Newton method. Aiming at the above problem, this work first presents the Newton iterative formula with both the Jacobian matrix error and the linear iteration error, and verifies the significant influence of the Jacobian matrix error on the iteration through numerical tests. Then, two different types of linear iterative convergence criteria are tested numerically, and the produced over-solving problem is investigated at the presence of Jacobian matrix error. Finally, a new convergence criterion of the linear iteration is developed based on two of the previous convergence criteria. Numerical tests suggest that the proposed method is effective to relieve over-solving problems, thus improves the computational efficiency.

Keywords: Newton method; implicit time stepping; Jacobian matrix error; linear iteration; convergence criterion; matrix stiffness; CFD

收稿日期: 2023-01-04; 修订日期: 2023-03-03; 录用日期: 2023-03-20; 网络出版时间: 2023-04-19

基金项目: 国家自然科学基金(11902344, 12172375); 国家重大项目(GJXM92579); 新疆维吾尔自治区自然科学基金重点项目(2022D01D32)

作者简介: 冯选燕(1995—), 重庆人, 硕士研究生, 研究方向: 计算数学. E-mail: fengxuanyan@stu.xju.edu.cn

通信作者: 燕振国*, 副研究员, 研究方向: 高阶精度计算方法及其应用. E-mail: yanzhg@mail.ustc.edu.cn

朱华君*, 副研究员, 研究方向: 高阶精度计算方法及其应用. E-mail: hjzhu@skla.carcn

引用格式: 冯选燕, 燕振国, 朱华君, 等. 非精确 Newton 方法中线性迭代收敛判据研究[J]. 空气动力学学报, 2023, 41(12): 28–36.

FENG X Y, YAN Z G, ZHU H J, et al. Study on the convergence criterion of linear iteration in inexact Newton methods[J]. Acta Aerodynamica Sinica, 2023, 41(12): 28–36(in Chinese). doi: 10.7638/kqdlxxb-2023.0001

0 引言

非线性问题广泛存在于各类学科和实际问题中, 是科学研究中的难点之一^[1]。其高效数值求解方法的研究不仅在计算数学基础理论上具有重要意义, 而且在科学与工程计算中处于重要地位, 能够促进计算流体力学(Computational Fluid Dynamics, CFD)等科学研究和工程应用的发展。

具体到 CFD 领域, 其控制方程(Navier-Stokes equations, 即 N-S 方程组)具有高度非线性特征, 对于绝大部分实际流动问题, 往往只能对控制方程进行离散求解, 为了获得全离散方程, 需要同时进行空间离散和时间离散。通常的空间离散包括有限差分法、有限体积法和间断有限元(discontinuous Galerkin-finite element method, DG-FEM)等数值方法^[2-4], 时间离散则通常被划分为显式和隐式两类^[5]。其中隐式时间推进方法是解决 CFD 刚性流动模拟效率问题的重要手段^[6], 然而隐式时间离散后不可避免地需要求解非线性离散系统。发展这类大型非线性离散系统的基础理论和高效求解方法是提高 CFD 计算效率的关键手段^[7]。

Newton 法是求解非线性方程的常用方法之一, 其二阶的收敛效率在高效 CFD 计算中十分具有吸引力^[8]。然而传统的 Newton 法需要计算非线性残差方程的 Jacobian(雅可比)矩阵来构造线性迭代方程组, 并且要求对线性方程组进行精确求解, 上述两点在大规模 CFD 计算中都是较难实现的。为了避免上述困难, 通常会在上述两个环节引入近似。例如, CFD 中常用的 Jacobian-free Newton-Krylov(JFNK)方法就采用有限差分来近似 Jacobian 矩阵, 引入了 Jacobian 矩阵计算误差^[9-11]; 而线性化后的方程组系统通常采用迭代方法进行求解, 迭代方法会引入线性迭代误差^[12-13]。所以 CFD 等大型模拟中实际上采用的是非精确 Newton 方法^[14]。

对于非精确 Newton 方法, 国内外开展了近似方法的系列研究工作。Vanden 和 Orkwis^[15]讨论了精确和数值 Jacobian 矩阵的求解过程及其对计算的影响, 数值测试结果表明: 在这两种情况下获得了较一致的迭代收敛次数及相近的收敛残差, 并指出对于较为复杂的非线性系统, 数值 Jacobian 矩阵可能更具有优势。Ezertas^[16]等在此基础上进一步研究了数值 Jacobian 矩阵计算中误差的来源, 探究并给出了使得

有限差分近似误差最小化的有限差分扰动幅值。测试发现: 在该条件下, 采用数值 Jacobian 矩阵表现出了与精确 Jacobian 矩阵相当的收敛性能。基于此, 数值 Jacobian 矩阵具有实现简单等优势, 被广泛应用于 CFD 大规模数值模拟中。然而, Yang 等^[17]的数值测试结果表明: 采用无矩阵(Jacobian-free, JF)方法的数值 Jacobian 矩阵并不能获得与精确 Jacobian 矩阵相当的收敛特性, 有限差分误差的引入会对迭代的收敛过程产生明显的影响。

另外, 非精确 Newton 方法在每一步的内迭代中通过线性迭代收敛判据来控制求解线性方程组时解的精确程度, 在此过程中会产生迭代的收敛误差。对于该问题, 文献^[18]将其称为 Newton 迭代的强制项, 其中强制项控制系数 η_k 的取值大小对收敛性有很大的影响。一些学者^[19]结合理论分析和数值测试给出了该参数的不同选取方式, 并给出了强制项能够保证 Newton 方法二阶收敛速度的理论分析。然而, 在考虑 Jacobian 矩阵误差的情况下, 线性迭代收敛判据的选取方式尚未得到深入研究, 需要进一步考虑 Jacobian 矩阵误差的影响并展开对比研究。

因此, 本文重点对存在 Jacobian 矩阵误差情况下不同线性迭代收敛判据的表现进行了研究, 分析了不同线性迭代判据所引起的过度求解问题。结合两种类型的判据发展了一种新的线性迭代收敛判据, 并利用数值测试对新提出线性迭代收敛判据的优势进行了验证。

1 控制方程和数值方法

1.1 隐式时间推进方法

对于本文所关心的 CFD 问题, 其求解的 Euler(欧拉)和 N-S 方程, 可以描述为以下形式:

$$\frac{\partial \bar{Q}}{\partial t} + \nabla \cdot \mathbf{F}(\bar{Q}) = 0 \quad (1)$$

其中, $\bar{Q}$ 为守恒变量, $\mathbf{F}$ 为对应的守恒通量, 其具体形式参见文献^[5]。在确定空间离散格式之后, 控制方程可写为以下半离散形式:

$$\frac{\partial Q}{\partial t} + N(Q) = 0 \quad (2)$$

其中, Q 为 $\bar{Q}$ 的数值近似, $N(Q)$ 为空间残差向量。

对于定常流动的数值模拟, 可以忽略时间导数项, 此时离散方程简化为:

$$N(Q) = 0 \quad (3)$$

对于 Euler 或 N-S 方程, 式(3)是非线性方程组。采用 Newton 方法对该方程进行线性化后代求解, 迭代式为:

$$\begin{aligned} J(Q^n)\Delta Q^n &= -N(Q^n) \\ J(Q) &= \partial N(Q)/\partial Q \end{aligned} \quad (4)$$

其中, $J(Q)$ 是残差向量 $N(Q)$ 的 Jacobian 矩阵, $\Delta Q^n = Q^{n+1} - Q^n$ 。

为了增强计算的稳定性, 通常会在式(4)左端添加伪时间项, 此时不考虑时间项的计算精度, 时间离散通常采用实现简单、存储量小的 Euler 隐式格式, 此时迭代方程变为:

$$\left[\frac{1}{\Delta t} I + J(Q^n) \right] \Delta Q^n = -N(Q^n) \quad (5)$$

式(5)中时间步长 $\Delta t \rightarrow \infty$ 时, 退化为式(4)的形式。

1.2 Newton-Krylov 方法

Newton-Krylov 方法^[20]目前是 CFD 领域常用的大型非线性方程组求解算法, 基于 Krylov 空间, 其具体形式为:

$$K_N(J, r^{(k)}) \equiv \text{span}\{r^{(k)}, Jr^{(k)}, \dots, J^{M-1}r^{(k)}\} \quad (6)$$

该类方法在 Krylov 空间内构造正交的基向量, 并在该空间中寻找使得线性方程组误差最小的数值解, 因此有残差单调下降等优势。其收敛速度依赖于线性方程组系数矩阵的条件数(由矩阵的特征值分布决定), 当矩阵的条件数较小时, Krylov 方法可以在几次搜索内将残差降到很小的量级。这里为了解 Euler 和 N-S 方程产生的非对称系统, 采用广义最小残差法 (generalized minimum residual, GMRES)^[21-22], 具体如下:

```
for k = 1, 2, ..., N %N次重启
    r0 = f - Ju0 %初始化残差
    q1 = r0 / ||r0||2 %第一个正交基
    for i = 1, 2, ..., M %GMRES(M)迭代
        wi = Jqi %计算新基
        for j = 1, 2, ..., i %新基正交化
            hji = qjTwi
            wi = wi - hjiqj
        end
        hi+1,i = ||wi||2 %新正交基2范数
        qi+1 = wi / hi+1,i %新正交基单位化
        if (||ri|| < σ) %收敛判据
            end
    end
end
```

其中: M 代表最大内迭代次数, σ 为 GMRES 迭代的收敛判据。 Jq_i 的计算可采用有限差分近似得到, 具体

形式在本文 1.3 节中详细介绍。 N 代表 GMRES 的最大重启次数, 重启可以有效避免存储过多 Krylov 空间向量, 减少正交化操作, 是大规模计算中常用的技术。

而采用 Krylov 方法求解线性方程组时带来的线性迭代误差, 主要依赖于线性迭代收敛判据:

$$\|N(Q^k) + J(Q)\Delta Q^k\| \leq \sigma \quad (7)$$

其中, $\Delta Q^k = Q^{k+1} - Q^k$; σ 可控制线性迭代的求解精度, 当其取值过大时, 较大的线性迭代误差可能对 Newton 迭代的收敛造成较大影响, 而取值过小又容易导致过度求解, 造成计算浪费, 因此 σ 的选择对于发展高效的数值方法来说是十分重要的。

1.3 线性化方法及其误差

在每一步 Newton 迭代的过程中, 需要采取直接法或迭代法求解由 Jacobian 矩阵 J 组成的线性方程组。在标准的 Newton 方法中需要采用精确方法计算 J , 然而对于大规模 CFD 计算而言, 计算精确的雅可比矩阵所产生的计算量和存储量都是较大的。因此, 历史上提出了多种近似雅可比矩阵方法, 其中一种是对雅可比矩阵的表达式进行近似, 例如在经典的 LU-SGS 方法中对流通量和黏性通量的 Jacobian 矩阵分别采用一阶精度近似和黏性谱半径近似^[5]; 另外一种方法则是采用数值方法计算 Jacobian 矩阵, 例如 CFD 中常用的 JF 方法就是采用有限差分方法近似 Jacobian 矩阵; 此外, 还存在上述两种近似的混合方法^[11]。本文主要讨论采用 JF 方法的误差及其影响。

在构造 Krylov 子空间(式 6)过程中, 令 $q = J^{j-1}r^{(k)}$, 则 Jq 的计算可采用有限差分近似计算 Jacobian 矩阵与向量的乘积:

$$Jq \approx \frac{N(Q + \varepsilon q) - N(Q)}{\varepsilon} \quad (8)$$

式中, ε 为有限差分扰动幅值, 对方程求解的精度及稳定性有较大的影响, 关于它的一种计算方法^[23]为:

$$\varepsilon = \sqrt{\|Q\|_2 + 1} \sqrt{\varepsilon_{\text{machine}}} / \|q\|_2 \quad (9)$$

这里的 $\sqrt{\varepsilon_{\text{machine}}}$ 为所采用数据类型的机器零, 基于双精度条件下展开计算, 取 $\sqrt{\varepsilon_{\text{machine}}} = 1 \times 10^{-8}$ 。

2 误差影响研究

本文考虑的非精确 Newton 迭代法在求解大型非线性方程组时相对精确 Newton 方法存在两类误差: Jacobian 矩阵误差和线性迭代误差。这里首先给出存在两种误差情况下非精确 Newton 方法的迭代式, 其

次给出了两种误差对 Newton 迭代式的影响，最后分析了 Jacobian 矩阵误差来源验证及其影响。

2.1 误差及其分析

非精确 Newton 迭代的过程中，Jacobian 矩阵误差和线性迭代误差都对整个非线性方程组求解过程中的收敛性造成一定影响。对于经典 Newton 法，解的迭代式为：

$$\mathbf{Q}^{n+1} = \mathbf{Q}^n - [\mathbf{I}/\Delta t + \mathbf{J}(\mathbf{Q}^n)]^{-1} \mathbf{N}(\mathbf{Q}^n) \quad (10)$$

对于近似 Jacobian 矩阵 $\mathbf{J}_*$ ，本文定义 Jacobian 矩阵的误差矩阵 $\mathbf{E}_J$ ：

$$\mathbf{E}_J = \mathbf{J}(\mathbf{Q}^n)_* - \mathbf{J}(\mathbf{Q}^n) \quad (11)$$

此时解的迭代式变为：

$$\mathbf{Q}^{n+1} = \mathbf{Q}^n - [\mathbf{I}/\Delta t + \mathbf{J}(\mathbf{Q}^n) + \mathbf{E}_J]^{-1} \mathbf{N}(\mathbf{Q}^n) \quad (12)$$

在考虑线性系统迭代误差 $\mathbf{E}_d$ 的情况下，Newton 方法的迭代式为：

$$\mathbf{Q}^{n+1} = \mathbf{Q}^n - [\mathbf{I}/\Delta t + \mathbf{J}(\mathbf{Q}^n)]^{-1} [\mathbf{N}(\mathbf{Q}^n) + \mathbf{E}_d] \quad (13)$$

当同时存在 Jacobian 矩阵计算误差和迭代误差时，相应的非精确 Newton 方法的迭代式变为：

$$\mathbf{Q}^{n+1} = \mathbf{Q}^n - [\mathbf{I}/\Delta t + \mathbf{J}(\mathbf{Q}^n) + \mathbf{E}_J]^{-1} [\mathbf{N}(\mathbf{Q}^n) + \mathbf{E}_d] \quad (14)$$

作为示例，这里给出 Euler 方程采用 DG 方法离散后的具体离散形式和 Jacobian 矩阵的形式，详细矩阵系数和 N-S 方程的离散形式参考文献 [11]。具体残差形式为：

$$\mathbf{N}(\mathbf{Q}^n) = \mathbf{M}^{-1} \mathcal{L}(\mathbf{Q}^n) \quad (15)$$

其中临时变量 $\mathcal{L}(\mathbf{Q}^n)$ 为：

$$\mathcal{L}(\mathbf{Q}^n) = \sum_{j=1}^d \mathbf{B}^T \mathbf{D}_j^T \mathbf{A}(w\mathbf{J}) \mathbf{F}_j(\mathbf{Q}^n) - (\mathbf{B}^T \mathbf{M}_c)^T \mathbf{A}(w^\Gamma \mathbf{J}^\Gamma) \widehat{\mathbf{F}}^n(\mathbf{Q}^n) \quad (16)$$

这里 d 为空间维数， $\mathbf{M} = \mathbf{B}^T \mathbf{A}(w\mathbf{J}) \mathbf{B}$ 为质量矩阵， $\mathbf{A}$ 为对角矩阵， $\mathbf{D}_j$ 为第 j 个方向的微分矩阵， $\mathbf{B}$ 为从求解点或者模态系数到通量点的变换矩阵、 $\mathbf{B}^\Gamma$ 为到单元边界上的通量点的变换矩阵， w 和 $\mathbf{J}$ 分别代表数值积分权重和网格变换 Jacobian， $\mathbf{M}_c$ 为单元边界基函数 ϕ^Γ 和基函数 ϕ 之间的映射矩阵， $\widehat{\mathbf{F}}^n$ 为对流数值通量。

其中，空间残差项 $\mathbf{N}(\mathbf{Q})$ 关于守恒变量 $\mathbf{Q}$ 求导的 Jacobian 矩阵 $\mathbf{J}(\mathbf{Q}^n)$ 形式：

$$\mathbf{J}(\mathbf{Q}^n) = \mathbf{M}^{-1} \frac{\partial \mathcal{L}(\mathbf{Q}^n)}{\partial \mathbf{Q}} \quad (17)$$

$\partial \mathcal{L}(\mathbf{Q}^n)/\partial \mathbf{Q}$ 具有块结构，例如 $\partial \mathcal{L}_{e_1}(\mathbf{Q}^n)/\partial \mathbf{Q}_{e_2}$ 代表第 e_1 行和第 e_2 列的块 Jacobian 矩阵元素，其含义为第 e_1 个单元的 $\mathcal{L}$ 对第 e_2 个单元的 $\mathbf{Q}$ 的 Jacobian 矩阵块。其具体形式为：

$$\begin{aligned} \frac{\partial \mathcal{L}_{e_1}(\mathbf{Q}^n)}{\partial \mathbf{Q}_{e_2}} &= \sum_{j=1}^d \mathbf{B}_{e_1}^T \mathbf{D}_{j,e_1}^T \mathbf{A} \left[w_{e_1} \mathbf{J}_{e_1} \left(\frac{\partial \mathbf{F}_{j,e_1}}{\partial \mathbf{Q}_{e_1}} \right) \right] \mathbf{B}_{e_1} \delta_{e_1 e_2} + \\ &\quad (\mathbf{B}_{e_1}^T \mathbf{M}_{c,e_1})^T \mathbf{A} \left[w^\Gamma \mathbf{J}^\Gamma \left(\frac{\partial \widehat{\mathbf{F}}_{e_1}^n}{\partial \widehat{\mathbf{Q}}^{+/-}} \right) \right] \frac{\partial \widehat{\mathbf{Q}}^{+/-}}{\partial \mathbf{Q}_{e_2}} \mathbf{B}_{e_2} \end{aligned} \quad (18)$$

其中， $\partial \widehat{\mathbf{Q}}^{+/-}/\partial \mathbf{Q}_{e_2}$ 为 $\mathbf{Q}_{e_2}$ 到 $\widehat{\mathbf{Q}}^{+/-}$ 的插值矩阵， $\widehat{\mathbf{Q}}^+$ 和 $\widehat{\mathbf{Q}}^-$ 分别为单元边界内外两侧的变量值， $\delta_{e_1 e_2}$ 为克罗内克算子，具体形式可参考文献 [11]。

2.2 Jacobian 矩阵误差来源验证及其影响

本小节首先简单验证了采用有限差分近似的 Jacobian 矩阵的相对误差变化规律，验证了 Jacobian 矩阵误差主要来源于差分格式带来的截断误差和计算机的舍入误差 [16]，而且表明在特定的数据精度和计算格式下，都存在一个使得误差最小的 ε ，即最优 ε ，如图 1 所示。图中 FD 表示一阶前向差分，CD 表示二阶中心差分。

为进一步测试 Jacobian 矩阵误差对收敛速度的影响，这里利用周期计算域内的一维 Euler 系统所搭建的简单模型问题进行研究，此处选取 120 个网格点，通过控制网格拉伸比使得网格向计算域两侧汇聚，并通过控制计算网格的大小比值来控制问题的刚性，此处网格拉伸比为 1.05。空间离散采用一阶迎风格式，该模型问题 Jacobian 矩阵的条件数约为 1×10^8 。图 2 中，横坐标 (Non-Iter) 代表非线性迭代次数，结果验证了当数值 Jacobian 矩阵近似误差较大时会引起效率降低甚至无法收敛的问题 [16-17]，然而当 Jacobian 矩阵误差较小时则可能避免上述问题。

3 线性迭代收敛判据研究

对于 Newton 类非线性系统求解方法，线性迭代误差虽然不影响计算最终收敛后的精度，但是对非线性系统能否收敛以及收敛时的计算效率有较大的影响。其中，相关研究 [18] 给出的线性系统迭代收敛判据的选取准则要求，在保证 Newton 方法二阶收敛效率的前提下尽量避免线性系统的过度求解。所谓过度求解就是当线性系统的求解精度达到一定程度以后，继续提升线性系统的求解精度将无法显著提升非线性系统的收敛效率，此时额外增加的线性系统迭代步数就造成了线性系统的过度求解。从提高计算效

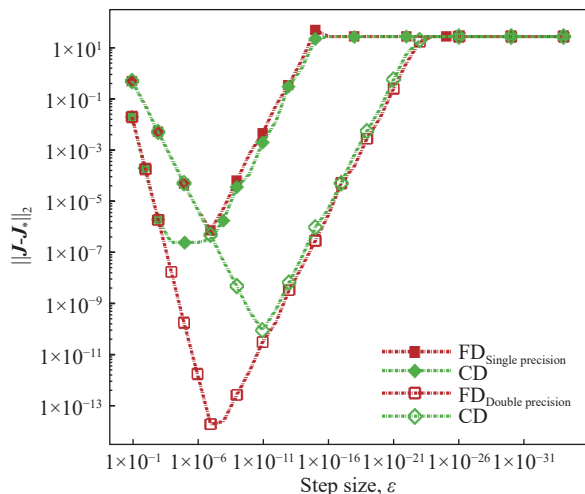


图 1 数值 Jacobian 矩阵的相对误差变化规律

Fig. 1 Relative error variations of the numerical Jacobian matrices

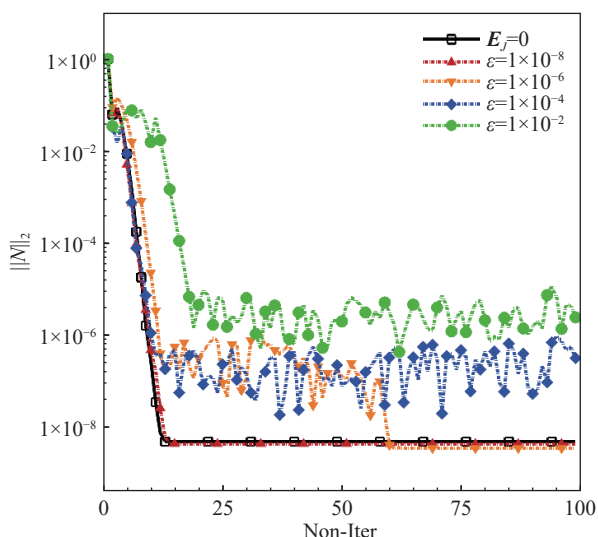


图 2 不同量级 Jacobian 矩阵误差对非精确 Newton 迭代的影响

Fig. 2 Influence of the Jacobian matrix error order on the inexact Newton iteration

率角度, 避免过度求解显然是必须考虑的问题。

常用的迭代误差收敛判据有两种, 一种是绝对量型收敛判据, 即选取某个绝对的数值作为迭代收敛的判据, 为了使得该策略中的参数更具有普适性, 本文采用如下的形式:

$$\|N(Q^k) + J\Delta Q^k\|_2 \leq \sigma = \zeta_k \|Q^0\|_2 \quad (19)$$

另外一种为相对型收敛判据, 即将线性系统的迭代收敛判据与非线性系统的残差进行关联, 其形式如下:

$$\|N(Q^k) + J\Delta Q^k\|_2 \leq \sigma = \eta_k \|N(Q^k)\|_2 \quad (20)$$

此处 η_k 为线性系统收敛判据的主要控制参数, 其含义为线性系统迭代误差要相对当前的非线性系统

残差小 η_k 倍以上。文献 [19] 对式 (20) 的线性迭代收敛判据进行了较系统的研究, 并从能够保持 Newton 方法二阶收敛效率的角度给出了相应的数学证明和数值测试; 然而并未对式 (19) 的绝对量型收敛判据和存在 Jacobian 矩阵误差情况下式 (20) 的迭代收敛误差的影响开展研究。

3.1 不同 η_k 的对比研究

这里采用风雷高精度软件平台 [24] 开展测试研究。控制方程采用 N-S 方程, 算例为黏性 NACA 0012 翼型流动问题, 来流 $Ma = 0.1$, 雷诺数取 100, 攻角为 0° , 翼型表面采用无滑移边界条件, 计算区域采用 20 倍翼型长度, 在外边界采用远场边界条件。采用 2842 个四边形网格对计算域进行离散, 如图 3 所示。在当前的定常计算中, 添加了伪时间项来控制线性系统的刚性以增强计算的稳定性, 计算结果如图 4 所示。本计算主要关注残差的收敛过程。

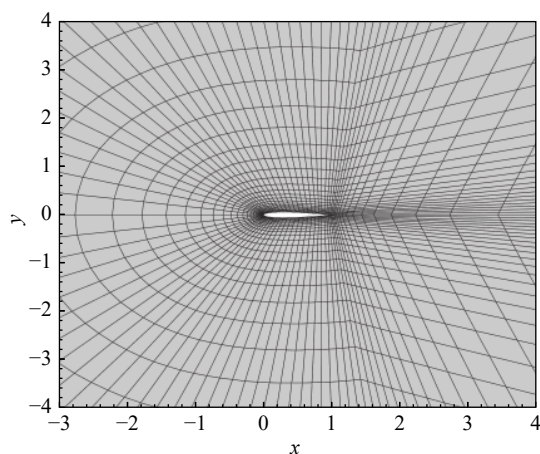


图 3 NACA 0012 算例的计算网格分布

Fig. 3 Computational grid distribution for the case of NACA 0012

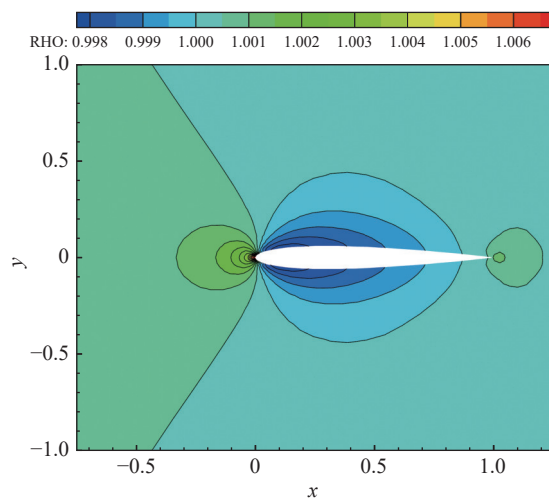


图 4 NACA 0012 算例计算结果密度分布

Fig. 4 Density distribution for the flow around NACA 0012

如图 5 所示, 首先对比测试了 η_k 分别取 1×10^{-3} 、 1×10^{-4} 和 1×10^{-6} 情况下采用式(8)的数值 Jacobian 矩阵(JF1)的迭代收敛情况, 作为对比, 本文同时给出了采用精确 Jacobian 的结果(JF0)。从线性迭代步数(Iter)上来讲, 显然采用更小的 η_k 所需的线性迭代步数显著增加, 线性方程组的求解精度相应提升。然而除了 $\eta_k=1 \times 10^{-3}$ 以外, 所有非线性系统残差基本完全重叠; 当 $\eta_k=1 \times 10^{-3}$ 时, 收敛曲线在收敛初段和最终阶段已经出现了较明显的偏差。上述结果表明, 当 $\eta_k \leq 1 \times 10^{-4}$ 时线性迭代已经足够精确, 继续降低 η_k 则会造成过度求解问题, 使计算效率降低。

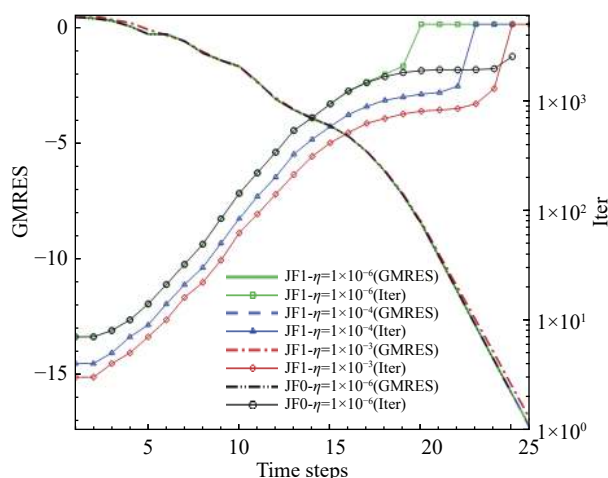


图 5 NACA 0012 算例在不同 Jacobian 矩阵下非线性迭代残差和 GMRES 迭代步数对比

Fig. 5 Comparison of the nonlinear iterative residual and GMRES iteration steps for different Jacobian matrices in the case of NACA 0012

另外需要强调的是, 在迭代接近收敛时, 基于 η_k 的迭代收敛判据都会造成过度求解问题。本文进一步通过增加 GMRES 的迭代重启次数来增加允许的最大线性系统迭代总步数, 从而进一步研究造成上述问题的原因, 结果如图 6 所示。在迭代的最终收敛阶段, 无论将允许的最大线性系统迭代步数增大到多大(在本文的测试范围内), 线性方程组均无法收敛到 $\eta_k=1 \times 10^{-6}$, 相应的线性方程迭代的总步数达到允许的最大值。与之相反, 采用精确 Jacobian 矩阵的计算结果, 虽然在收敛的最终阶段迭代步数同样增加, 但是并不存在上述现象, 说明上述问题是当所要求的线性系统迭代误差太小时, Jacobian 矩阵误差会造成线性系统无法收敛引起的。然而从非线性系统的残差角度来看, 各条曲线是基本重合的, 进一步说明采用近似 Jacobian 所增加的线性系统迭代是由于收敛判据过严引起的过度求解, 需要发展更加合理的迭代收敛判据避免上述问题。

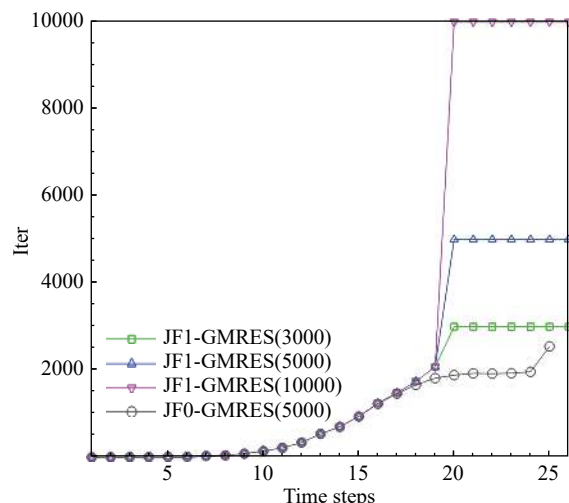


图 6 不同 GMRES 允许的最大迭代步数下的线性迭代次数

Fig. 6 Linear iterative number under different allowable maximum GMRES iteration steps

3.2 不同 ζ_k 的对比研究

类似 3.1 节, 这里对式(19)的收敛判据进行测试, 结果如图 7 所示。本文分别给出了 ζ_k 取 1×10^{-6} 、 1×10^{-10} 、 1×10^{-14} 的测试结果, 为了对比, 图中同时给出了 $\eta_k=1 \times 10^{-6}$ 情况下的收敛曲线。可以看到, 采用 ζ_k 类型的迭代收敛判据, 只有在 ζ_k 取很小的值时才能保证迭代收敛。但是, 当 ζ_k 取值很小时(如 1×10^{-14}), 在迭代的初始阶段会造成较严重的过度求解现象。而在迭代接近收敛的阶段, 采用 ζ_k 类型的迭代收敛判据却可以较好地避免过度求解的问题。

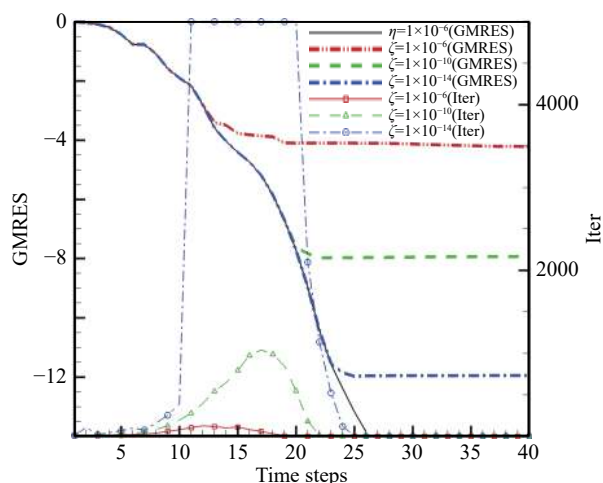


图 7 NACA 0012 算例中不同 ζ 取值情况下非线性迭代残差和 GMRES 迭代步数对比

Fig. 7 Comparison of the nonlinear iterative residual and GMRES iteration steps for different ζ in the case of NACA 0012

3.3 一种新的线性迭代收敛判据

基于以上研究结果, 本文发展了一种新的迭代收敛判据:

$$\sigma = \min(\max(\eta_k \|N(Q^k)\|_2, \zeta_k \|Q^0\|_2), \eta_{\min} \|N(Q^k)\|_2) \quad (21)$$

其中 $\max$ 函数内两项分别为式(19)和式(20)对应的绝对收敛判据和相对收敛判据,这里取两项中的较大值。在非线性迭代初期 $\|N(Q^k)\|_2$ 较大,因此 $\eta_k \|N(Q^k)\|_2$ 的量级会大于 $\zeta_k \|Q^0\|_2$,从而避免了后者在非线性迭代初期所引起的严重过度求解问题;随着迭代逐渐接近收敛, $\|N(Q^k)\|_2$ 逐渐减小,只要取合理的参数,可以使得 $\max$ 函数在接近收敛时取 $\zeta_k \|Q^0\|_2$,此时可以有效避免图5中的过度求解问题。在此基础上,本文还在 $\max$ 函数外添加了相应的 $\min$ 函数,其作用是保证每个线性迭代至少要下降一定量级(本文中取 $\eta_{\min} = 0.01$),从而避免类似图7中迭代不能完全收敛的问题。本文中 ζ_k 取 1×10^{-13} ,主要从计算的舍入误差角度进行选取,由于最终线性迭代误差会加到 Q^k 上, $\zeta_k = 1 \times 10^{-13}$ 的参数选择可以避免线性方程组收敛到比Newton迭代步本身舍入误差更小的量级,从而避免类似图5中的过度求解问题。

图8同样采用NACA 0012算例对不同迭代收敛判据进行了对比,直至迭代第19步采用式(21)的新迭代收敛判据(NewTol)得到了与 $\eta_k = 1 \times 10^{-6}$ 完全相同的计算结果。此后采用 $\eta_k = 1 \times 10^{-6}$ 的线性迭代步数迅速增加到允许的最大值,线性迭代无法收敛,而采用式(21)的线性迭代步数则开始下降,说明如设计的那样,式(21)中 $\zeta_k \|Q^0\|_2$ 部分开始发挥作用从而大幅减少了线性迭代的计算量。直至第23步(残差收敛至 1×10^{-13})时采用新判据的非线性系统残差依然与 $\eta_k = 1 \times 10^{-6}$ 重叠,随后采用式(21)的线性迭代步数开始显著减小,非线性残差曲线不再重叠。表1给出了

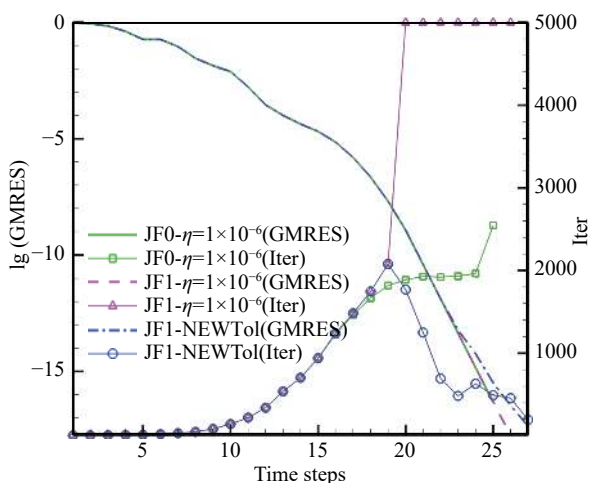


图8 非线性迭代收敛判据的残差和迭代步数

Fig. 8 Nonlinear iterative residual and GMRES iteration steps for the new linear iteration convergence criterion

残差收敛到不同量级时所需要的线性迭代总步数,该指标能够很好地反映出所消耗的总计算量。当残差收敛到 1×10^{-10} 以下时,JF0- $\eta_k = 1 \times 10^{-6}$ (精确 Jacobian、 $\eta_k = 1 \times 10^{-6}$)、JF1- $\eta_k = 1 \times 10^{-6}$ (近似 Jacobian、 $\eta_k = 1 \times 10^{-6}$)和JF1-NewTol(近似 Jacobian、新判据)所需线性迭代步数分别为13113、19670和12692步;当残差收敛到 1×10^{-13} 以下时,所需线性迭代步数分别为16964、29670和13870步;当残差收敛到 1×10^{-16} 以下时,所需线性迭代步数分别为21473、44670和15453步。可以看到,新发展的线性迭代收敛判据能够有效避免过度求解的问题,显著降低所需要的线性系统迭代步数。线性迭代步数最多降低3倍,从而提高计算效率。

表1 不同判据情况下迭代收敛所需线性迭代步数对比

Table 1 Comparison of the linear iteration steps under different convergence criteria

		非线性系统收敛残差		
		1×10^{-10}	1×10^{-13}	1×10^{-16}
精确Jacobian	$\eta_k = 1 \times 10^{-6}$	13113	16964	21473
近似Jacobian	$\eta_k = 1 \times 10^{-6}$	19670	29670	44670
	NewTol	12692	13870	15453

为了进一步验证上述迭代收敛策略在不同计算状态中的通用性,本文通过改变离散格式的计算精度来改变非线性系统的特性和规模。将DG格式的计算精度提高到三阶,此时总共求解的网格自由度变为17052个。图9给出了三阶精度的计算结果,可以看到,采用新的线性迭代收敛判据同样能够很好地避免接近收敛状态下的过度求解问题,从而显著提高计算效率。

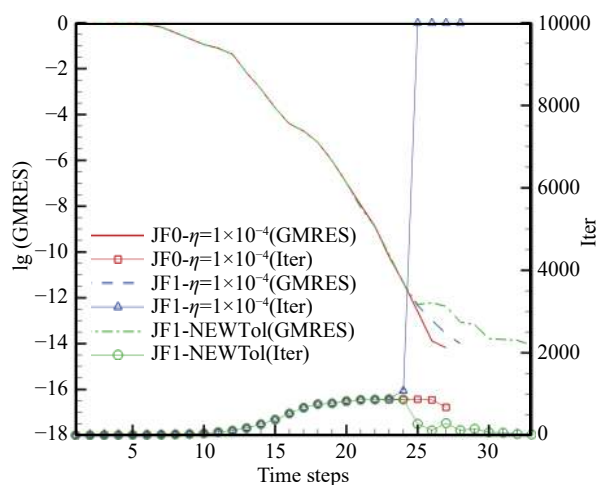


图9 三阶DG格式计算中新线性迭代收敛判据的残差和迭代步数

Fig. 9 Nonlinear iterative residual and GMRES iteration steps for the new linear iteration convergence criterion with third-order DG scheme

4 结 论

针对 CFD 中常见的基于 Newton 类方法的非线性系统求解问题, 本文通过数值测试, 对存在 Jacobian 矩阵误差情况下每个 Newton 迭代步中线性迭代收敛判据进行了研究, 最终结合两种类型的迭代收敛判据, 发展了一种新的判据。得到的主要结论如下:

1) 存在 Jacobian 矩阵误差情况下, 线性迭代收敛判据的影响会更加显著;

2) 基于绝对量的线性迭代收敛判据容易在非线性迭代初期引起过度求解问题; 相对非线性残差的线性迭代收敛判据容易在接近最终收敛时引起过度求解问题;

3) 新型线性迭代收敛判据能够有效地降低过度求解问题, 显著降低收敛所需要的线性迭代步数, 从而提高计算效率。

下一步, 我们将进一步对 Jacobian 矩阵误差对非线性系统迭代收敛效率的影响开展研究, 并将发展的线性迭代收敛误差应用到更加复杂的流动模拟中。

参 考 文 献:

- [1] CAI X, KEYES D E, VENKATKRISHNAN V. Newton-Krylov-Schwarz: An Implicit Solver for CFD[R]. Institute for Computer Applications in Science and Engineering (ICASE), 1995.
- [2] 邓小刚, 刘昕, 毛枚良, 等. 高精度加权紧致非线性格式的研究进展[J]. *力学进展*, 2007, 37(3): 417–427.
DENG X G, LIU X, MAO M L, et al. Advances in high-order accurate weighted compact nonlinear schemes[J]. *Advances in Mechanics*, 2007, 37(3): 417–427 (in Chinese).
doi: 10.3321/j.issn:1000-0992.2007.03.009
- [3] 雷国东, 李万爱, 任玉新. 求解可压缩流的高精度非结构网格WENO有限体积法[J]. *计算物理*, 2011, 28(5): 633–640.
LEI G D, LI W A, REN Y X. A high-order unstructured-grid WENO FVM for compressible flow computation[J]. *Chinese Journal of Computational Physics*, 2011, 28(5): 633–640 (in Chinese).
doi: 10.19596/j.cnki.1001-246x.2011.05.001
- [4] LUO H, ABSILLIS G, NOURGALIEV R. A moving discontinuous Galerkin finite element method with interface condition enforcement for compressible flows[J]. *Journal of Computational Physics*, 2021, 445: 110618.
doi: 10.1016/j.jcp.2021.110618
- [5] 燕振国. 高精度混合线性紧致格式的隐式时间推进方法研究[D]. 绵阳: 中国空气动力研究与发展中心, 2013.
YAN Z G. Investigation of implicit time integration methods with HDCS schemes[D]. Mianyang, China: China Aerodynamics Research and Development Center, 2013 (in Chinese).
- [6] PULLIAM T H. Development of implicit methods in CFD NASA Ames Research Center 1970s-1980s[J]. *Computers & Fluids*, 2011, 41(1): 65–71.
doi: 10.1016/j.compfluid.2010.09.016
- [7] 阎超, 屈峰, 赵雅甜, 等. 航空航天CFD物理模型和计算方法的述评与挑战[J]. *空气动力学学报*, 2020, 38(5): 829–857.
- YAN C, QU F, ZHAO Y T, et al. Review of development and challenges for physical modeling and numerical scheme of CFD in aeronautics and astronautics[J]. *Acta Aerodynamica Sinica*, 2020, 38(5): 829–857 (in Chinese).
doi: 10.7638/kqdlxxb-2020.0072
- [8] SHANNO D F, KETTLER P C. Optimal conditioning of quasi-Newton methods[J]. *Mathematics of Computation*, 1970, 24(111): 657–664.
doi: 10.1090/S0025-5718-1970-0274030-6
- [9] CHISHOLM T T, ZINGG D W. A Jacobian-free Newton-Krylov algorithm for compressible turbulent fluid flows[J]. *Journal of Computational Physics*, 2009, 228(9): 3490–3507.
doi: 10.1016/j.jcp.2009.02.004
- [10] LE D V, WHITE J, PERAIRE J, et al. An implicit immersed boundary method for three-dimensional fluid-membrane interactions[J]. *Journal of Computational Physics*, 2009, 228(22): 8427–8445.
doi: 10.1016/j.jcp.2009.08.018
- [11] YAN Z G, PAN Y, CASTIGLIONI G, et al. Nektar++: Design and implementation of an implicit, spectral/hp element, compressible flow solver using a Jacobian-free Newton Krylov approach [EB/OL]. 2020: arXiv: 2002.04222.
<https://arxiv.org/abs/2002.04222>
- [12] PERSSON P-O, PERAIRE J. Newton-GMRES preconditioning for discontinuous Galerkin discretizations of the navier-stokes equations[J]. *SIAM Journal on Scientific Computing*, 2008, 30(6): 2709–2733.
doi: 10.1137/070692108
- [13] ZHAO X Y, SUN D F, TOH K C. A Newton-CG augmented Lagrangian method for semidefinite programming[J]. *SIAM Journal on Optimization*, 2010, 20(4): 1737–1765.
doi: 10.1137/080718206
- [14] DEMBO R S, EISENSTAT S T, STEIHAUG T. Inexact Newton Methods[J]. *SIAM Journal on Numerical Analysis*, 1982, 19(2): 400–408.
doi: 10.1137/0719025
- [15] VANDEN K J, ORKWIS P D. Comparison of numerical and analytical Jacobians[J]. *AIAA Journal*, 1996, 34(6): 1125–1129.
doi: 10.2514/3.13202
- [16] EZERTAS A, EYI S. Performances of numerical and analytical Jacobians in flow and sensitivity analysis[C]// 19th AIAA Computational Fluid Dynamics, San Antonio, Texas, 2009.
- [17] YANG X Q, CHENG J, LUO H, et al. Robust implicit direct discontinuous Galerkin method for simulating the compressible turbulent flows[J]. *AIAA Journal*, 2019, 57(3): 1113–1132.
doi: 10.2514/1.J057172
- [18] EISENSTAT S C, WALKER H F. Choosing the forcing terms in an inexact Newton method[J]. *SIAM Journal on Scientific Computing*, 1996, 17(1): 16–32.
doi: 10.1137/0917003
- [19] AN H B, MO Z Y, LIU X P. A choice of forcing terms in inexact Newton method[J]. *Journal of Computational and Applied Mathematics*, 2007, 200(1): 47–60.
doi: 10.1016/j.cam.2005.12.030
- [20] KNOLL D A, MCHUGH P R, KEYES D E. Newton-Krylov methods for low-Mach-number compressible combustion[J]. *AIAA Journal*, 1996, 34(5): 961–967.
doi: 10.2514/3.13174
- [21] HAYAMI K, YIN J F, ITO T. GMRES methods for least squares problems[J]. *SIAM Journal on Matrix Analysis and Applications*, 2010, 31(5): 2400–2430.
doi: 10.1137/070696313

- [22] 燕振国, 刘化勇, 毛枚良, 等. 基于高阶耗散紧致格式的GMRES方法收敛特性研究[J]. *航空学报*, 2014, 35(5): 1181–1192.
YAN Z G, LIU H Y, MAO M L, et al. Convergence property investigation of GMRES method based on high-order dissipative compact scheme[J]. *Acta Aeronautica et Astronautica Sinica*, 2014, 35(5): 1181–1192 (in Chinese).
doi: [10.7527/S1000-6893.2013.0430](https://doi.org/10.7527/S1000-6893.2013.0430)
- [23] PARK H, NOURGALIEV R R, MARTINEAU R C, et al. On physics-based preconditioning of the Navier-Stokes equations[J]. *Journal of Computational Physics*, 2009, 228(24): 9131–9146.
doi: [10.1016/j.jcp.2009.09.015](https://doi.org/10.1016/j.jcp.2009.09.015)
- [24] 赵钟, 张来平, 何磊, 等. 适用于任意网格的大规模并行CFD计算框架PHengLEI[J]. *计算机学报*, 2019, 42(11): 2368–2383.
ZHAO Z, ZHANG L P, HE L, et al. PHengLEI: A large scale parallel CFD framework for arbitrary grids[J]. *Chinese Journal of Computers*, 2019, 42(11): 2368–2383 (in Chinese).
doi: [10.11897/SP.J.1016.2019.02368](https://doi.org/10.11897/SP.J.1016.2019.02368)

(本文责编: 周奇杰)