

GMRES(m)算法在离散不适定问题中的应用

张海燕, 闵涛, 刘相国

西安理工大学理学院, 西安 710054

[摘要] 基于投影方法的规划算法——Krylov 子空间技术, 研究了离散不适定正则化和 Krylov 子空间广义极小残余算法(GMRES(m))的基本理论, 特别是残余向量与 Krylov 子空间的关系。利用离散不适定正则化方法, 将不适定问题转化为适定问题, 利用广义极小残余算法对此适定问题进行数值求解。数值结果表明该算法是可靠和有效的。

[关键词] GMRES(m)算法; 不适定; Arnoldi; 正则化

[中图分类号] O241.8

[文献标识码] A

[文章编号] 1000-7857(2007)13-0054-06

Application of GMRES(m) Algorithm for Discrete Ill-posed Problems

ZHANG Haiyan, MIN Tao, LIU Xiangguo

School of Sciences, Xi'an University of Technology, Xi'an 710054, China

Abstract: This paper studies the fundamental theory of the discrete ill-posed regularization and the generalized minimal residual algorithm (GMRES(m)) based on Krylov subspace methods and specially the relationship between residual vector and Krylov subspace. The techniques are based on the projection process. Discrete ill-posed regularization methods convert ill-posed problems into posed problems, so the generalized minimal residual algorithm can be used in this kind of ill-posed problems. Numerical results show the reliability and efficiency of the algorithm.

Key Words: GMRES(m) algorithm; ill-posed; Arnoldi; regularization

CLC Number: O241.8

Document Code: A

Article ID: 1000-7857(2007)13-0054-06

0 引言

随着科学和工程计算技术的发展, 越来越多的问题需要求解大规模离散不适定的线性方程组, 如数理方程反问题、有限差分及有限元数值求解, 对这几类方程的快速有效求解成为数值代数研究的热点之一。1986年, 由 Saad^[1-2]和 Schlitz 提出了基于 Galerkin 原理的 GMRES(m)算法, 此算法是一种求解非对称线性适定方程组的 Krylov 子空间投影法^[3], 该算法采用正交投影或斜投影在子空间产生迭代向量进行计算。该方法基于 Krylov 向量的完全正交化, 因而在 Krylov 子空间只求

出一个近似解。由于所需计算量和内存量较少等方面的优势, 目前广泛应用于各种工程应用领域。

本文在深入研究离散不适定问题基本理论、正则化方法及其 GMRES(m)算法的基础上, 提出了求解离散不适定问题的 GMRES(m)算法, 并编写程序, 进行了大量的数值试验。

1 离散不适定问题和正则化方法

1.1 离散不适定问题

对于线性方程组

收稿日期: 2007-05-13

基金项目: 国家自然科学基金项目(50579061), 教育部博士基金项目(106-220546)

作者简介: 张海燕, 陕西省西安市雁翔路 58 号西安理工大学理学院, 研究方向为科学工程计算与计算机模拟;

E-mail: zhanghaiyan0536@126.com

闵涛(通讯作者), 陕西省西安市雁翔路 58 号西安理工大学应用数学系, 教授, 研究方向为数理方程反问题;

E-mail: mintao@xaut.edu.cn

$$Az=u$$

的求解问题或最小二乘问题 $\min \|u-Az\|_2$, 若下面两个条件之一成立就称问题为离散不适定问题:

- 1) 矩阵 A 的条件数非常大,或者说矩阵 A 的最大奇异值与最小奇异值之比非常大;
- 2) 矩阵 A 的奇异值逐渐下降趋于零。

在求解上述离散不适定问题时,为了克服选择法和拟解法所受的局限,Tikhonov 于 1963 年在其开创性的工作中提出了求解此类不适定问题的正则化方法,这一方法为处理不适定问题和反问题奠定了坚实和广泛的理论基础。

1.2 正则化方法基本理论

正则化方法^[4]的基本思想:设法构造一个连续算子(正则算子)去逼近不连续算子 A^{-1} ,将不适定的问题化为一个近似的适定问题,从而得到原问题的近似解。

引进一个依赖于正则参数 α 的从空间 U 到空间 F 上的算子

$$R(\cdot, \alpha): U \rightarrow F$$

本文试图用 $R(u_\delta, \alpha)$ 作为 z_T 的近似,即算子 $R(\cdot, \alpha)$ 为 A^{-1} 的近似。为了克服逆算子 A^{-1} 的缺陷(不连续性),这个算子 $R(\cdot, \alpha)$ 对每个 $\alpha > 0$ 都应当是连续的,且当 $\alpha \rightarrow 0$ 时,这个算子应当收敛到 A^{-1} 。具体内容如下。

定义 1(正则算子) 如果算子 $R(\cdot, \alpha)$ 满足:

- 1) 对每个 $\alpha > 0$, 及任意的 $u \in U$, $R(u, \alpha)$ 关于 u 连续;
- 2) 对每个 $z \in F$, 有

$$\lim_{\alpha \rightarrow 0} R(u, \alpha) = z$$

则称 $R(\cdot, \alpha)$ 为方程 $Az=u$ 的正则算子, α 称为正则参数。

定义 2(正则解) 设 $\rho_U(u_\delta, u_T) \leq \delta$, 若存在这样的正则参数 α (依赖于 u_δ 与 α), 使得 $\lim_{\delta \rightarrow 0} \rho_F(z_\alpha, z_T) = 0$, 其中, $z_\alpha = R(u_\delta, \alpha)$, 则称 z_α 为正则解。

每一个满足定义 1 条件的正则算子 $R(u, \alpha)$, 连同决定正则参数 α 的不同原则和方法, 都定义了构造原问题的近似解的一个稳定算法。于是, 寻求原问题的稳定近似解的过程可归结如下。

1) 构造正则算子 $R(u, \alpha)$ 。Tikhonov 曾详细讨论过基于变分原理的构造方法借助于积分变换将其在象空间中实现的方法。

2) 选择适当的正则参数 α 。关于正则参数的选择, 通常有先验的和后验的两种策略。不少先验的策略有理论分析价值, 但在实际中常常难以验证其赖以施用的条件; 目前, 后验策略的研究颇为流行, 已发展了 Morozov 偏差原理、广义偏差原理、Arcangel 准则、广义 Arcangel 准则、拟最优准则和广义交叉检验准则等。

于是, 针对求解不适定问题, 本文先构造出正则算子, 选择正则参数, 然后利用 GMRES(m) 算法对此方程进行数值求解。

2 GMRES(m) 算法

2.1 GMRES(m) 算法理论

考虑方程

$$Az=u \quad (1)$$

式中, $A \in R^{n \times n}$ 为一非奇异矩阵, $u \in R^n$ 为一给定的向量。记 K_m 和 L_m 是 R^n 中两个 m 维子空间, 取 $K_m = \text{span}\{r_0, Ar_0, \dots, A^{m-1}r_0\}$, $L_m = AK_m$, 并称由 A 和 r_0 生成的空间为 Krylov 子空间^[5]。利用 Galerkin 原理, 取 $x_k \in K_m$, x_k 选取过程如下。

对于选定的初值 $z_0 \in R^n$, 第 m 步是求剩余范数

$$\varphi(x) = \|u - Az\|_2 \quad (2)$$

在 m 维超平面

$$z_0 + K(A, r_0, m) \quad (3)$$

上的极小点 z_m , 其中 $r_0 = u - Az_0$, 即

$$z_m = z_0 + x_m \quad (4)$$

x_m 是最小二乘问题

$$\min_{x \in K(A, r_0, m)} \|u - A(z_0 + x)\|_2 = \min_{x \in K(A, r_0, m)} \|r_0 - Ax\|_2 \quad (5)$$

的解。

记 $\beta = \|r_0\|$, 假定从 $v_1 = r_0/\beta$ 出发, 利用 Arnoldi 过程求出正交向量组 $v_1, v_2, \dots, v_{m+1}$ 和 h_{ij} , 则有

$$AV_m = V_{m+1}\bar{H}_m \quad (6)$$

其中, $V_m = [v_1, v_2, \dots, v_m]$ 是利用 Arnoldi 过程^[6]构造的子空间 K_m 的正交基矩阵, $\bar{H}_m \in R^{(m+1) \times m}$ 是一个 Hessenberg 矩阵, 即

$$\bar{H}_m = \begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1,m-1} & h_{1,m} \\ h_{21} & h_{22} & \cdots & h_{2,m-1} & h_{2,m} \\ 0 & h_{32} & \cdots & h_{3,m-1} & h_{3,m} \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & h_{m-1,m} & h_{m,m} \\ 0 & 0 & \cdots & 0 & h_{m+1,m} \end{bmatrix} \quad (7)$$

利用式(6)可得

$$\begin{aligned} \min_{x \in K(A, r_0, m)} \|r_0 - Ax\|_2 &= \min \{ \|r_0 - AV_m y\|_2, y \in R^m \} \\ &= \min \{ \|r_0 - V_{m+1} \bar{H}_m y\|_2, y \in R^m \} \\ &= \min \{ \| \|r_0\|_2 e_1 - \bar{H}_m y\|_2, y \in R^m \} \quad (8) \end{aligned}$$

其中, 在等式推导过程中用到了 $r_0 = \beta v_1$ 和 $V_{m+1}^T V_{m+1} = I_{m+1}$ 。由于 $\bar{H}_m$ 的特殊形式, 式(7)是容易求解的。

2.2 求解不适定问题的 GMRES(m) 算法

考虑方程

$$Az=u \quad (9)$$

当矩阵 A 的条件数很大时, 该方程是不适定的, 选取合

适的正则参数,利用正则化方法将式(9)转化为式(10)

$$(A^T A + \alpha I)z = A^T u \quad (10)$$

并记 $A_1 = A^T A + \alpha I$, $u_1 = A^T u$, 其中 A^T 是 A 的转置, 则上述方程变为

$$A_1 z = u_1 \quad (11)$$

下面利用 GMRES(m) 算法求解不适定问题式(11), 具体算法如下。

1) 初始化

输入 A_1 , b_1 , 初始向量 x_0 , 最大迭代次数 m 及精度要求 ε , 于是迭代的初始值为

$$r_0 = u_1 - A_1 z_0 \quad \beta = \|r_0\| \quad v_1 = r_0 / \beta \quad V_1 = \{v_1\}$$

2) Arnoldi 过程^[7-8]

① 正交化

$$h_{ij} = v_i^T A_1 v_j \quad i=1, 2, \dots, j$$

$$\hat{v}_{j+1} = A_1 v_j - \sum_{i=1}^j h_{ij} v_i \quad j=1, 2, \dots, m$$

$$h_{j+1,j} = \|\hat{v}_{j+1}\|_2$$

② 标准化

如果 $h_{j+1,j} < \varepsilon$, 则转到步骤 5); 否则, $v_{j+1} = \hat{v}_{j+1} / h_{j+1,j}$, $j=j+1$, 转到步骤 2) 中的正交化。

③ 求解新的 V_{j+1} 与 $\bar{H}_j$

$$V_{j+1} = (V_j, v_{j+1}) \quad \bar{H}_j = \begin{bmatrix} \bar{H}_{j-1} & \bar{h}_j \\ 0 & h_{j+1,j} \end{bmatrix}_{(j+1) \times j}$$

$$\bar{h}_j = (h_{1j}, h_{2j}, \dots, h_{jj})^T$$

3) 解最小二乘问题^[9]

$$\|r_m\| = \min_{y \in R^m} \|\beta e_1 - \bar{H}_m y_m\| \quad (12)$$

得 $y_m, x_m = x_0 + H_m y_m$ 。

4) 如果 $\|r_m\| = \|u_1 - A_1 z_m\| < \varepsilon$, 则输出 x_m 。

5) $z_0 = z_m, r_0 = r_m, v_1 = r_0 / \|r_0\|$, 转向步骤 2)。

3 数值试验

3.1 数值模拟 1

以 Hilbert 矩阵为系数矩阵的病态线性方程组为例进行数值试验, 考虑下述方程组

$$A_n x = b \quad A_n = (a_{ij})_{n \times n} \quad i, j = 1, 2, \dots, n$$

选取 $x_i(s) = s e^{-s}, 0 \leq s \leq 1$, 其离散形式是 $x_i(s_i) = s_i e^{-s_i}$, 其中 $s_i = ih, h = \frac{1}{n}$, 于是真实的右端项 b_i 可由 $A_{n \times n} x_i$ 得到, 对右端项加随机扰动后为 $b_i + \delta \text{rand}(b_i)$, 其中, rand 为随机函数。试验中, 取初始向量 $x_0 = [0, 0, \dots, 0]^T \in R^n$, 收敛判断参数 $\varepsilon = 10^{-9}$ 。

当 $n=2\ 000$ 时, δ 分别取 $10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}$, 自变量所在区间 $[0, 2]$ 等间距地取 25 个结点, 数值解也是相应地取 25 个结点, 利用求解不适定问题的 GMRES(m) 算法进行编程实现并进行试验结果分析, 其结果由表 1 和图 1 给出。假设该方程组的数值解为 x_i , 真解为 $\tilde{x}_i$, 均为 $n \times 1$ 维向量, 则相对误差 $e = \|\tilde{x}_i - x_i\|_2 / \|\tilde{x}_i\|_2$, 绝对误差为 $|\tilde{x}_i - x_i|$ 的最大值。

为了进一步探讨求解不适定问题 GMRES(m) 的有效性, 下面列出了当 $\delta = 10^{-4}$, 迭代次数 $m=60, n$ 分为 1 000, 2 000, 3 000, 4 000 的计算结果, 如表 2 所示。

可以看出, 即使对于严重的病态 Hilbert 矩阵 (当 $n=3\ 000$ 时, 其条件数为 $4.899\ 144\ 689\ 581\ 336 \times 10^{21}$), 通过求解离散不适定问题的 GMRES(m) 算法在不同的剖分数 n 下求出正则化参数, 从而得到不适定线性方程组的数值解, 最大误差和相对误差不大, 能够得到令人满意的结果。但是, 对于重新开始迭代步长的选择也是至关重要的, 合适的 m 能够加速 GMRES(m) 算法的收敛, m 过小可能导致迭代的收敛失败, 过大则导致计算量的成倍增加及收敛速度的减缓。另外, 迭代初始向量的选择对于加快算法收敛也具有重要作用。

3.2 数值模拟 2

积分方程

$$\int_0^1 k(s, t) f(t) dt = g(s) \quad 0 \leq s \leq 1 \quad (13)$$

其中, 它的核 $k(s, t) = (s^2 + t^2)^{\frac{1}{2}}, g(s) = \frac{1}{3} \left((1+s^2)^{\frac{3}{2}} - s^2 \right)$ 。对于上述方程左端采用数值积分离散化, 其积分算子离散为 $n \times n$ 矩阵, 用 A 表示, 表 3 所示为 $m=70, n$ 分别为

表 1 在右端项扰动 δ 的情况下最大误差和相对误差的计算结果

Table 1 Calculation results of the maximum error and the relative error with the right perturb δ

δ	迭代次数 m	正则化参数 α	最大误差	相对误差
10^{-2}	50	$2.206\ 313\ 401\ 082\ 070 \times 10^{-4}$	0.058 523 821 599 69	0.067 111 630 115 04
10^{-3}	60	$6.108\ 303\ 401\ 082\ 027 \times 10^{-4}$	0.043 281 101 466 34	0.060 243 383 360 41
10^{-4}	70	$1.206\ 303\ 400\ 082\ 090 \times 10^{-6}$	0.027 010 495 333 83	0.033 111 585 177 46
10^{-5}	80	$1.215\ 313\ 411\ 067\ 070 \times 10^{-7}$	0.018 187 002 313 11	0.022 631 504 569 66

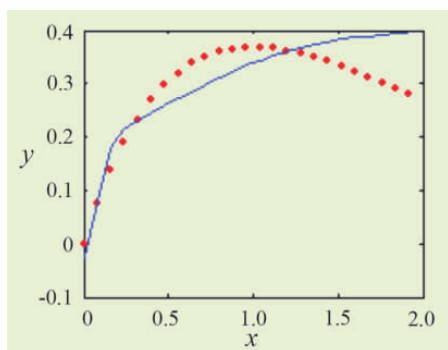
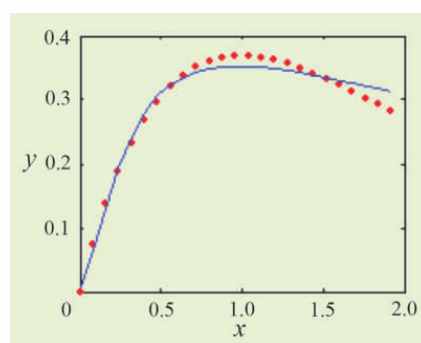
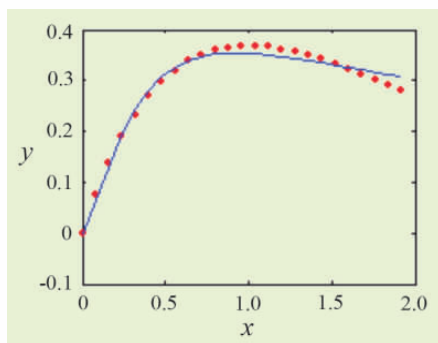
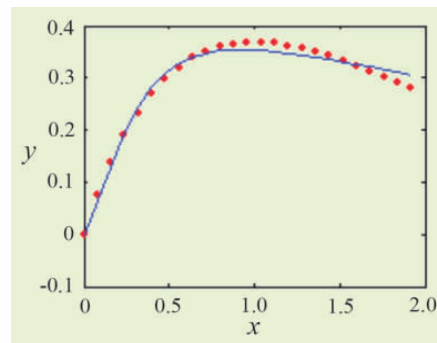
(a) $\delta=10^{-2}$ (b) $\delta=10^{-3}$ (c) $\delta=10^{-4}$ (d) $\delta=10^{-5}$

图 1 真解(实线)与数值解(点线)的对比

Fig. 1 Comparison between exact solution (solid line) and numerical solution (dotted line)

1 000, 2 000 时, 自变量所在区间 $[0, 1]$ 等间距地取 20 个结点, 数值解在 $[0, 1]$ 相应地取 20 个结点, 其数值模拟结果的偏差原理、Tikhonov 正则化方法和 GMRES(m) 3 种方法的最大误差比较, 其图示结果由图 2 和图 3 给出。

由图 3~图 6 可以看出, 当 n 为 1 000, 2 000 时, 偏差原理、Tikhonov 正则化方法和 GMRES(m) 3 种方法的求解中, GMRES(m) 算法在数值求解中效果最好, 且最大误差最小。

表 2 当 n 在不同数值下最大误差和相对误差的计算结果Table 2 Calculation results of the maximum error and the relative error with the different n

n	正则化参数 α	最大误差	相对误差
1 000	$1.108\ 303\ 401\ 082\ 070 \times 10^{-3}$	0.046 548 779 869 27	0.076 131 841 481 47
2 000	$6.108\ 303\ 401\ 082\ 070 \times 10^{-4}$	0.043 281 101 466 34	0.060 243 383 360 41
3 000	$1.108\ 303\ 401\ 082\ 070 \times 10^{-3}$	0.035 240 395 962 78	0.058 986 619 481 54
4 000	$1.108\ 303\ 401\ 082\ 070 \times 10^{-5}$	0.030 821 392 559 62	0.039 370 284 764 13

表 3 偏差原理、Tikhonov 正则化方法 和 GMRES(m) 算法计算结果

Table 3 Calculation results of discrepancy principle, Tikhonov regularization method and GMRES(m) algorithm

n	正则化参数 α	偏差原理	Tikhonov	GMRES(m)
1 000	$1.034\ 432\ 704\ 661\ 631 \times 10^{-3}$	0.139 954 934 827 26	0.214 292 595 944 78	0.088 172 986 962 79
1 500	$7.661\ 174\ 339\ 330\ 052 \times 10^{-4}$	0.139 293 383 013 55	0.199 114 848 406 61	0.079 960 387 734 90
2 000	$4.661\ 187\ 596\ 412\ 336 \times 10^{-4}$	0.136 968 034 387 66	0.163 372 187 334 24	0.069 717 165 617 85

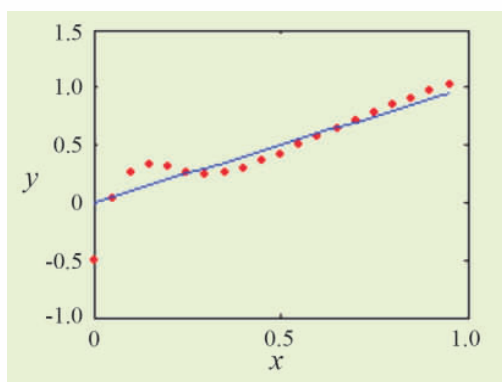
图 2 (a) $n=1\ 000$ 时真解(实线)与偏差原理数值解(点线)的对比

Fig. 2 (a) Comparison between exact solution (solid line) and discrepancy principle numerical solution (dotted line) with $n=1\ 000$

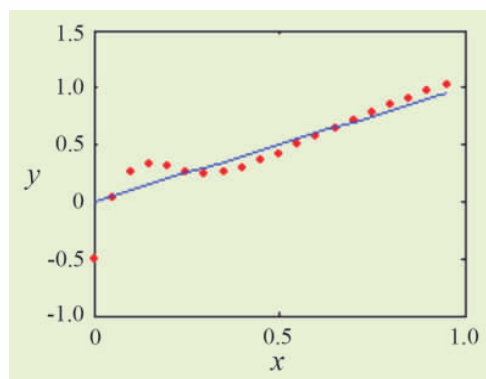
图 2 (b) $n=1\ 000$ 时真解(实线)与 Tikhonov 正则化方法数值解(点线)的对比

Fig. 2 (b) Comparison between exact solution (solid line) and Tikhonov regularization method solution (dotted line) with $n=1\ 000$

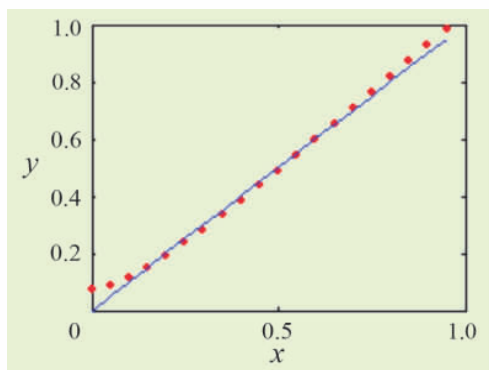
图 2 (c) $n=1\ 000$ 时真解(实线)与 GMRES(m)算法数值解(点线)的对比

Fig. 2 (c) Comparison between exact solution (solid line) and GMRES(m) algorithm numerical solution (dotted line) with $n=1\ 000$

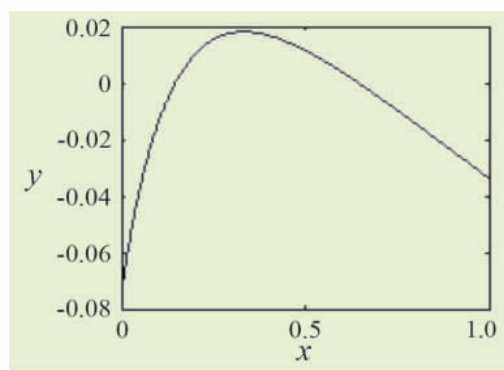
图 2 (d) $n=1\ 000$ 时 GMRES(m)数值解的误差

Fig. 2 (d) Error of GMRES(m) algorithm numerical solution with $n=1\ 000$

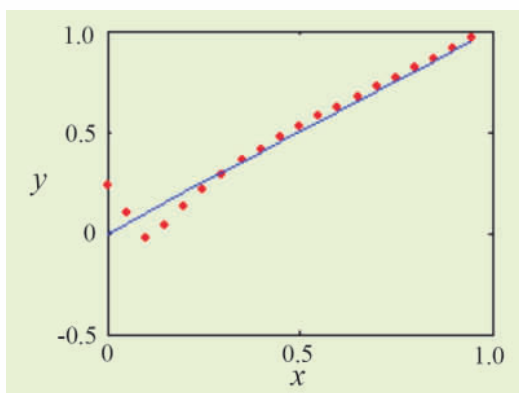
图 3 (a) $n=2\ 000$ 时真解(实线)与偏差原理数值解(点线)的对比

Fig. 3 (a) Comparison between exact solution (solid line) and discrepancy principle numerical solution (dotted line) with $n=2\ 000$

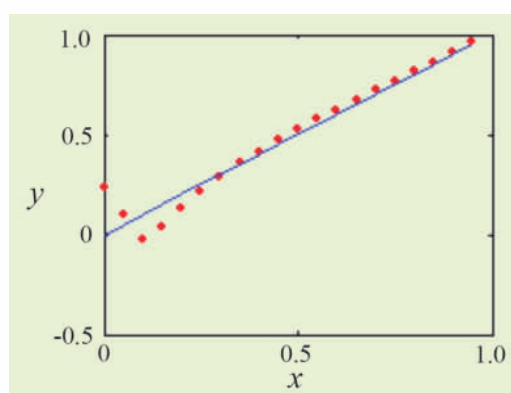
图 3 (b) $n=2\ 000$ 时真解(实线)与 Tikhonov 正则化方法数值解(点线)的对比

Fig. 3 (b) Comparison between exact solution (solid line) and Tikhonov regularization method solution (dotted line) with $n=2\ 000$

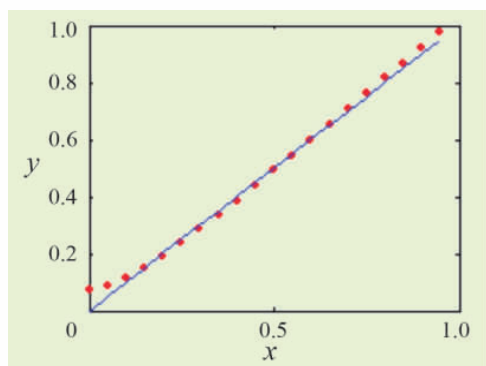


图 3 (c) $n=2\ 000$ 时真解(实线)与 GMRES(m) 算法数值解(点线)的对比

Fig. 3 (c) Comparison between exact solution (solid line) and GMRES(m) algorithm numerical solution (dotted line) with $n=2\ 000$

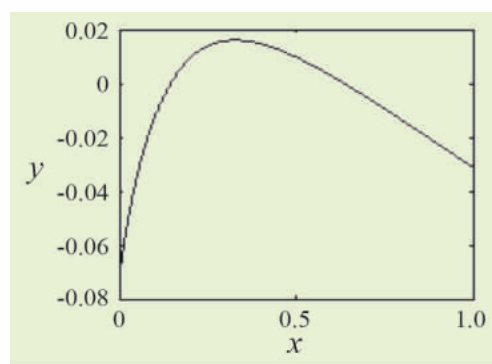


图 3 (d) $n=2\ 000$ 时 GMRES(m) 数值解的误差
Fig. 3 (d) Error of GMRES(m) algorithm numerical solution with $n=2\ 000$

4 结论

本文基于 Galerkin 原理提出了求解离散不适定问题 GMRES(m) 算法, 为计算不适定问题提供了一种新的求解方法。数值试验表明, 该算法能很好地求解病态线性方程和积分方程, 特别是大型离散不适定方程, 克服了传统方法的不足, 大大提高了计算精度。

参考文献 (References)

- [1] SAAD Y, SCHULTZ H. GMRES. A generalized minimal residual algorithm for solving nonsymmetric linear systems[J]. SIAM J Sci Comput, 1986, 35(7): 856–869.
- [2] BROWN P N. A theoretical comparison of the Arnoldi and GMRES algorithms [J]. SIAM J Sci Stat Comput, 1991, 12 (1): 58–78.
- [3] 张可村, 赵英良. 数值计算的算法与分析[M]. 北京: 科学出版社, 2004.
ZHANG Kecun, ZHAO Yingliang. Algorithms and analysis for numerical methods[M]. Beijing: Science Press, 2004.
- [4] 肖庭延, 于慎根, 王彦飞. 反问题的数值解法 [M]. 北京: 科学出版社, 2003.
XIAO Tingyan, YU Shengen, WANG Yanfei. Numerical method for inverse problems [M]. Beijing: Science Press, 2003.
- [5] 蔡大用, 白峰杉. 现代科学计算[M]. 北京: 科学出版社, 2000.
CAI Dayong, BAI Fengshan. Modern scientific calculation [M]. Beijing: Science Press, 2000.
- [6] 贾仲孝, 陈桂芝. 解大规模非对称矩阵特征问题的精化 Arnoldi 方法的一种变形 [J]. 数值计算与计算机应用, 2003, 24(2): 101–110.
JIA Zhongxiao, CHEN Guizhi. A variant of the refined Arnoldi method for large scale matrix Eigen problems[J]. Journal of Numerical Methods and Computer Applications, 2003, 24(2): 101–110.
- [7] 钟宝江. 一种灵活的混合 GMRES 算法 [J]. 高等计算数学学报, 2001, 23(3): 261–272.
ZHONG Baojiang. A flexible hybrid GMRES algorithm[J]. Numerical Mathematics A Journal of Chinese Universities, 2001, 23(3): 261–272.
- [8] BAGLAMA J, CALVETTI D, GOLUB G H, *et al.* Adaptively preconditioned GMRES algorithm[J]. SIAM J Sci Comput, 1998, 20(4): 243–269.
- [9] HUANG Y, van der VORST H. Some observations on the convergence behavior of GMRES [R]. Technical report 89–90. Faculty of Technical Mathematics and Informatics. Delft, 1989.

(责任编辑 代丽)

·学术动态·

2007 年度中国科协精品科技期刊工程项目申报工作启动

2007 年度中国科协精品科技期刊工程项目申报工作日前已经启动。中国科协学会学术部负责对项目实施立项、组织评审、检查和追踪问效。本年度申报的范围和条件是: ① 2006 年度未获得资助的期刊出版单位; ② 2006 年度资助的非延续项目的期刊出版单位; ③ 2006 年度延续项目, 具

备上一类别项目条件的, 可申报上一等级项目; ④ 申报 A 类和 B 类的期刊须是中国科协所属全国学会主办学术期刊, 申报 C 类项目的期刊须是中国科协主管或省科协主管学术期刊。科普类期刊和内部刊物不在申报范围之列。2007 年, 中国科协将资助期刊 110 个, 包括 2006 年延续项目 78

个, 新申报项目 32 个(其中 A 类项目 1 个, B 类项目 10 个, C 类项目 21 个(含省级项目 2 个))。申报材料报送时间: 2007 年 7 月 16–25 日。详细信息请浏览中国科协网页 (<http://www.cast.org.cn>) “学会与学术”栏目。