等谱流问题的李群算法[*1)]

田益民
(中国科学院软件研究所并行计算研究室　北京 100080)

秦孟兆
(中科院计算数学研究所北京 2719 信箱　北京 100080)

THE LIE GROUP METHOD IN ISOSPECTRAL FLOW

Tian Yimin
(Parallel Computing Division, Institute of Software, Chinese Academy of Sciences, Beijing 100080, *China)*

Qin Mengzhao
(Institute of Computational Mathematics and Scientific Computing, Chinese Academy of Sciences, Beijing 100080, *China)*

Abstract

In this paper, we studied the application of Lie group method in isospectral problem and we made numerical experiments by explicit Euler and implicit Euler with simple iteration and Newton iteration respectively.

Key words: isospectral flow, Lie group method, Newton iteration

§1. 引　言

非均匀介质中声波传播的研究，有着重要的应用：一个是声波在地壳中的传播，可以用来研究地质结构；一个是在海洋中的传播，可以用来研究海洋中物体的情况．在声波传播的研究中，需要研究 Helmhotz 方程

$$\phi_{zz}+\phi_{rr}+n^2(z,r)\phi=0, \tag{1}$$

其中 ϕ 是声波的震幅， z 是深度， r 是范围， (n,r) 是介质的反射率．由 Helmhotz 方程可以得到

$$i\frac{\partial\phi}{\partial\alpha}=H\phi, \tag{2}$$

∗ 2003 年 8 月 25 日收到.
1) 中国科学院知识创新重大项目： KZCX1-SW-18 资助；中国科学院声学研究所声场声信息国家重点实验资助；国家重点基础研究项目 (G1999,032800).

其中

$$H=-\sqrt{n^2(n,r)-p^2},\quad p=-i\frac{\partial}{\partial z}.$$

在对声波传播的研究中，大角度问题的处理是比较困难的，人们想到去求相应矩阵的谱，并由此去解决问题. 在方程 (2) 的数值研究中，需要对等谱流问题进行数值研究.

等谱流问题是如下矩阵微分方程初值问题 [7]:

$$Y'=[B,Y],\qquad Y(0)=Y_0, \tag{3}$$

其中 $Y,B\in R^{d\times d}$. 初值 Y_0 是给定的 $d\times d$ 矩阵，矩阵函数 $B\equiv B(t,Y)$ 依赖于 Y 及时间 t. 方括号表示矩阵交换子的李括号， $[B,Y]=BY-YB$. B 是反对称的，一般情况下 Y 对称. 每一个流都被 B 所决定，例如， Toda 流，双括号流， KvM 流等 [2].

直到几年前，人们才在计算中得到结论 [7]: 以往的常微分方程的数值方法都不是等谱的，因此不能用于解决等谱流问题.

给定关于谱 $\sigma(Y_0)=\{\lambda_1,\lambda_2,\cdots,\lambda_d\}$ 矩阵 Y_0 的解析函数 f, 如果

$$B(Y)=f(Y)_+-f(Y)_-, \tag{4}$$

则我们称 (4) 为 QR 流，其中下标 $+,-$ 分别表示矩阵函数 $f(Y)$ 的上三角部分和下三角部分. 取 $f(x)=x$, 对三对角矩阵 Y_0, 我们得到 Toda 流， $Y(t)$ 形如

$$Y(t)=\begin{bmatrix}\beta_1 & \alpha_1 & 0 & \cdots & 0\\ \alpha_1 & \beta_2 & \alpha_2 & \ddots & \vdots\\ 0 & \alpha_2 & \ddots & \ddots & 0\\ \vdots & \ddots & \ddots & \ddots & \alpha_{d-1}\\ 0 & \cdots & 0 & \alpha_{d-1} & \beta_d\end{bmatrix}_{d\times d}. \tag{5}$$

B(Y) 按 (4) 将 $Y(t)$ 变换成

$$\begin{bmatrix}0 & \alpha_1 & 0 & \cdots & 0\\ -\alpha_1 & 0 & \alpha_2 & \ddots & \vdots\\ 0 & -\alpha_2 & \ddots & \ddots & 0\\ \vdots & \ddots & \ddots & \ddots & \alpha_{d-1}\\ 0 & \cdots & 0 & -\alpha_{d-1} & 0\end{bmatrix}. \tag{6}$$

特征值问题是求给定矩阵的特征值，而特征值反问题就是寻找具有给定的特征值和某种结构的矩阵. 这样的问题在振动理论，控制论，层面 X 射线照相术，系统认证，地球物理学等许多领域都有着广泛的应用. 等谱流对于构造特征值反问题的实用算法起着重要的作用. 我们寻求一具有给定的特征值 $\eta\in R^N$ 同时属于某一类 $\tau\in S_N$ 的矩阵， S_N 表示 $N\times N$ 实对称矩阵. 通常是设计一矩阵函数 B, 使得它的吸引不动点在 τ 中. 这样对于 $Y(0)=diag(\eta)$, 能得到特征值反问题的解.

例如，关于 Toeplize 矩阵的特征值反问题，此时 τ 是 $N\times N$ Toeplize 矩阵

$$X\in\tau,\Longleftrightarrow x_{k,l}=t_{|k-l|},\quad k,m,l=1,2,\cdots,N,$$

其中 $t_0,t_1,\cdots,t_N$ 是任意实数 [3]. 令

$$b_{k,l}(Y)=\begin{cases} y_{k,l-1}-y_{k+1,l}, & 1\le k<l\le N,\\ 0, & 1\le k=l\le N,\\ y_{k,l+1}-y_{k-1,l}, & 1\le l<k\le N\end{cases}$$

为一 Toeplize 零化子，这里 B 是反对称矩阵，且对 $Y\in\tau$ 有 $B(Y)=O$, 因此，解是不动点.

§2. 等谱流的李群方法及李群算法中的牛顿迭代

2.1 等谱流问题的李群方法

在 (3) 中，算子 B 乃反对称矩阵. 把 (3) 与下列微分方程联系起来

$$u'=B(Y)u,\quad u(0)=I. \tag{7}$$

由于 B 反对称，(7) 的解是正交流 (i.e. $uu^T=u^Tu=\mathrm{I}$), (3) 的解可以写成

$$Y(t)=u(t)Y_0u^T(t),\quad u(0)=I, \tag{8}$$

因此， $Y(t)$ 正交相似于初值 Y_0.

对正交矩阵组成的李群使用显式李群方法可以得到实用的数值求解正交流的等谱流方法.

在等谱流问题中，由于 u 是非线性的正交流，可以用 Runge Kutta 型李群方法 [6]. 此时可考虑相伴的非线性正交流

$$u\prime=B(t,u)u,\quad u_n=I,\quad t\in[t_n,t_n+h],$$

其中

$$B(t,u)=B(u(t)Y_nu(t)^T).$$

当我们得到 $u(t_{n+1})$ 的近似值 u_{n+1} 后，即可计算 $Y_{n+1}=u_{n+1}Y_nu_{n+1}^T$, Y_{n+1} 正交相似于 Y_n, 因此是等谱的. 在用隐式算法时，要用到牛顿迭代方法. 李群算法中的牛顿迭代与欧氏空间上的传统的牛顿迭代有很大的不同.

2.2 李群算法中的牛顿迭代

对方程

$$y'=g(y), \tag{9}$$

隐式欧拉法为

$$y^{k+1}=y^k+hg(y^{k+1}). \tag{10}$$

从李群 G 到其所对应的李代数 g 的微分方程

$$y' = y \cdot g(y), \quad y \in G \subset GL(n,R),\ g: G \to g \subset gl(n,R), \tag{11}$$

则 Crouch-Grossman 和 Munthe-Kaas 形式的欧拉方法为

$$y^{k+1} = y^k \cdot exp(hg(y^{k+1})). \tag{12}$$

一般情形下不能显式地解此方程，因此，计算序列 $(y_n^{k+1})_{n\geq 0}$，此序列收敛到方程 (12) 的解. 显而易见，不动点迭代是一种这样的序列

$$y_{n+1}^{k+1} = y^k \cdot exp(hg(y_n^{k+1})), \quad n = 0, 1, \cdots \tag{13}$$

在传统的隐式欧拉法中，上述类型的不动点迭代是不太可行的 [5]，因此就有必要选择其它形式的牛顿迭代. 在方程 (13) 中，对于接近 $y^{(k)}$ 的 y，有

$$y = y^{(k)} \cdot exp(v(y)), \quad v(y) \in g,$$

因此，需要解方程

$$f(y) = 0, \tag{14}$$

其中

$$\begin{aligned} f:\ & G \ \to\ g \\ & y \ \mapsto\ v(y) - hg(y). \end{aligned}$$

我们研究解 $f(y)=0$ 的算法，f 是从李群到其相对应的李代数的映射.

f 在点 $y \in G$ 的微分是 $f_y^{'}: TG|_y \mapsto g$ 定义为

$$f_y^{'}(\Delta(y)) = \frac{d}{dt}|_{t=0} f(y \cdot exp(tL_{y^{-1}}^{'})). \tag{15}$$

$f_y^{'}$ 关于切向量 Δy 的像可以被看作由左乘而得的元素 $v: v \in g$. 指数映射把由 t 作尺度变换后的 v 映射为李群中的一元素 z，则 $f_y^{'}(\Delta_y)$ 可以写为

$$f_y^{'}(\Delta(y)) = \lim_{t\to 0} \frac{f(y \cdot z) - f(y)}{t}. \tag{16}$$

$f_y^{'}$ 可以由函数 df_y 表示

$$df_y = (f \circ L_y)^{'} = f_y^{'} \circ L_y^{'}, \tag{17}$$

因此

$$df_y(u) = f_y^{'}(L_y^{'} u) = \frac{d}{dt}|_{dt=0} f(y \cdot exp(tu)). \tag{18}$$

应用 (15)，李群 G 上的牛顿法可以这样进行：对给定的 $y_0 \in G$，首先由 (15) 确定微分 $f_{y_0}^{'}$，然后寻找 $\Delta_{y_0} \in TG|_{y_0}$，使其满足

$$f_{y_0}^{'} + f(y_0) = 0, \tag{19}$$

最后以 $y_1 = y_0 \cdot exp(L_{y_0^{-1}}^{'})(\Delta_{y_0})$ 更换 y_0.

方法 1. 李群上的牛顿法

- 对给定的 $y_n \in G$, 由 (18) 确定 $df_{y_n}(u_n)$.
- 寻找 $u_n \in g$ 满足

$$df_{y_n}(u_n) + f(y_n) = 0. \tag{20}$$

- 计算 $y_{n+1} = y_n \cdot exp(u_n)$.

例 1. 给定 $z \in G, h > 0$, 方程

$$y = z \cdot exp(hg(y))$$

在 $z^{-1}y \in N(e)$ 时，与下述方程等价

$$f(y) \overset{def}{=} v(y) - hg(y) = 0. \tag{21}$$

这里

$$L_{z^{-1}}(y) = z^{-1}y = exp(v(y)). \tag{22}$$

$f(y_n)$ 为

$$f(y_n) = exp^{-1}(z^{-1} \cdot y_n) - hg(y_n).$$

微分 df_{y_n} 可由 v 和 g 的微分得到

$$df_{y_n} = dv_n - hdg_{y_n}.$$

由 (22) 我们可以得到

$$L'_{z^{-1}} = exp'_{v(y)} \circ v',$$

所以有

$$dexp_{v(y)}(dv_{y_n}) = L'_{exp(-v(y_n))} \circ v' \circ L'_{y_n} = L'_{exp(-v(y_n))} \cdot z^{-1} \cdot y_n = L'_e = Id_g.$$

自然地，方程 (20) 变为

$$dexp^{-1}_{v(y_n)} - hdg_{y_n}(u_n) + f(y_n) = 0. \tag{23}$$

不失一般性，我们假设存在一已知的 $z \in G$ 和某一 $v_n \in g$ 使得

$$y_n = z \cdot exp(v_n) = L_z \circ exp(v_n).$$

例如， z 为 $f(y) = 0$ 的解 y 的逼近，使得 $z^{-1}y_n \in N(e)$. 因此有

$$f(y_n) = f \circ L_z \circ exp(v_n) \overset{def}{=} \widetilde{f}(v_n).$$

函数 $\widetilde{f} = f \circ L_z \circ exp$ 是从 g 到 g 的映射. 特别地， $\widetilde{f}$ 在 v_n 的微分 $d\widetilde{f}_{u_n}$ 是由下式定义的线性映射

$$d\widetilde{f}_{v_n} = \frac{d}{dt}|_{t=0}\widetilde{f}(v_n + tu). \tag{24}$$

如果 g 为有限维，则牛顿法可以应用于方程 $\widetilde{f}(v) = 0$.

方法 2. 李代数上的牛顿法

- 对给定的 $y_n \in G$, $y_n = z \circ exp(v_n)$, 由 (24) 确定 $d\widetilde{f}_{v_n}$.
- 寻找 $u_n \in g$ 满足
$$d\widetilde{f}_{v_n}(u_n) + \widetilde{f}(v_n) = 0. \tag{25}$$
- 计算 $v_{n+1} = v_n + u_n, y_{n+1} = z \cdot exp(v_{n+1})$.

例 2. 用 (22) 例 1 中的方程 (21) 可以写为
$$d\widetilde{f}_v \stackrel{def}{=} f(z \cdot exp(v)) = v - hg(z \cdot exp(v)). \tag{26}$$
$\widetilde{f}$ 在 v_n 的微分为
$$\begin{aligned} d\widetilde{f}_{v_n} &= Id_g - hd(g \circ L_z \circ exp)_{v_n} \\ &= Id_g - hdg_{y_n} \circ L'_{y_n^{-1}} \circ exp'_{v_n} \\ &= Id_g - hdg_{y_n} \circ dexp_{v_n}, \end{aligned}$$
因此 (25) 变为
$$u_n - hdg_{y_n} \circ dexp_{v_n}(u_n) + \widetilde{f}(v_n) = 0. \tag{27}$$
因为
$$\widetilde{f}(v_n) = f(z \circ exp(v_n)) = f(y_n). \tag{28}$$
可以证明，第二种算法是第一种算法的线性化 [4].

§3. 数值实验

对矩阵 [1]
$$Y = \begin{pmatrix} d(1) & b & 0 & \cdots & 0 & c & 0 & 0 & \cdots & 0 \\ b & d(2) & b & \cdots & 0 & 0 & c & 0 & \cdots & 0 \\ 0 & b & d(3) & \cdots & 0 & 0 & 0 & c & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & 0 \\ 0 & 0 & 0 & \cdots & d(m) & 0 & 0 & 0 & \cdots & c \\ c & 0 & 0 & \cdots & 0 & d(m+1) & b & 0 & \cdots & 0 \\ 0 & c & 0 & \cdots & 0 & b & d(m+2) & b & \cdots & 0 \\ 0 & 0 & c & \cdots & 0 & 0 & b & d(m+3) & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & 0 & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & c & 0 & 0 & 0 & \cdots & d(n) \end{pmatrix},$$
用李群算法求其特征值, 其中 $n = 2m, b = -1, c = -1, 1 \le i \le m+1, v(i) = 1/(200^2), d(i) = v(i) + 4, m+2 \le i \le n, v(i) = 1/(600^2), d(i) = v(i) + 4$.

由矩阵 Y 所生成的李代数可以有不同的定义方式，若定义为 $Y^- - Y^{-T}$, 则对应 QR 流，定义为 Y^-, 则对应 LU 流，而定义为 $Y^- - Y^0/2$, 则对应 $Cholesky$ 流. 同样，也可以定义双括号流 $B[Y] = [Y,\ P(Y)]$.

我们定义矩阵 Y 所生成的李代数为

$$B = Y^- - Y^+,$$

其中 Y^-, Y^+ 分别为矩阵 Y 的下三角部分所构成的下三角矩阵和上三角部分所构成的上三角矩阵. 然后, 分别用李群算法中的显式欧拉法和隐式欧拉法计算, 在隐式欧拉计算中, 分别用简单迭代和牛顿迭代.

我们将用以上三种方法进行的具体计算数据列于表 1.

表 1

		M1		M2		M3			
t	n	k	T	k	T	k	err	k1	T
0.3	4	111	33.3	112	33.6	130			39
0.3	10	185	55.5	186	55.8	183	0.1	22	61.5
0.3	20	566	169.8	567	170.1	546	0.1	39	175.5
0.3	40	4628	1388.4	4629	1388.7	4568	0.1	76	1393.2
0.4	4					101			40.4
5	4					15	0.1	6	76.8
5	10					19	0.1	22	101.6
5	20					42	0.1	39	221.7
5	40					282	0.1	76	1432.8
6	4					13	0.1	6	79.8
6	10					17	0.1	22	108.6
6	20					36	0.1	39	227.7
6	40					236	0.1	76	1438.8
100	4					5	0.1	6	501.8
100	10					6	0.1	22	606.6
100	20					8	0.07	64	819.2
100	40					21	0.01	410	2223
1000	4					4	0.1	6	4001.8
1000	10					4	0.1	22	4006.6
1000	20					5	0.08	42	5012.6
1000	40					7	0.05	156	7046.8

表 1 中的 $M1$, $M2$, $M3$, 分别表示显式欧拉方法、隐式欧拉简单迭代和隐式欧拉牛顿迭代, t 表示时间步长, n 表示矩阵阶数, k 表示时间方向上计算层数, T 表示在时间方向上的总积累量, err 表示用显格式计算初值的误差标准, $k1$ 表示用显格式计算初值在时间方向上的计算层数.

经过对分别用以上三种方法得到的计算数据进行对比分析, 可以看出, 在显式欧拉法中, 时间步长 h 只能取较小的值, 而在隐式欧拉牛顿迭代中, 时间步长 h 可以取较大一些的值. 例如, 当 $n = 4$ 等值时 h 甚至可以取到 1000,

用牛顿法时对初值有一定的要求, 当初值在解的某一邻域内时, 使用牛顿法可以迅速收敛到解. 在计算中, 矩阵对角线以外的元素最大允许误差为 10^{-10}, 当 $n = 4, 6$ 时用牛顿法

可以直接计算, 取 $n = 20, 40$ 时, 则需要先用显式欧拉计算, 当矩阵对角线以外的元素最大允许误差小于 ϵ 时, 再用牛顿法计算, 则可以用较大的步长计算.

该工作得到科学院"环渤海 (湾) 地区前新生代海相油田油气资源研究" 及声学所声场信息国家重点实验室资助.

参 考 文 献

[1] 罗明秋, 朱国同, 刘洪, 李幼铭, 适用于三维隐式叠前深度偏移中矩阵求逆的混合算法, 地球物理学报 (待刊).

[2] M.P. Calvo, A. Iserles and A. Zanna, *Numerical solution of isospectral flows*, DSAMP Techn. Rep. NA03, University of Cambridge, 1995, to appear in Math of Comp.

[3] M.T. Chu and K.R. Driessel, *Can real sysmmetric Toplitz matrices have arbitrary real spectral*, Technical report, Idaho State University.

[4] Arieh Iserles and S.P. Norsett, *On the solution of linear differential equations in Lie groups*, preprint.

[5] A. Iserles, S.P. Nϕrsett(1999),*On the solution of linear differentials in Lie groups*

[6] H.Munthe-Kaas, *Runge-Kutta methods on Lie Groups*, submitted to BIT.

[7] Antonella Zanna, *Lie-Group Methods for Isospection Flows*, Numerical Analysis Reports