李群方法里的矩阵指数计算*

张林华[1],吴 永[2]
(1. 重庆师范大学 数学与计算机科学学院,重庆 400047;2. 重庆工学院 数理学院,重庆 400050)

摘 要:矩阵指数计算与力学计算中的动力学问题、最优控制的计算问题等密切相关,是数值代数里研究得最为广泛的课题之一。目前虽有以 PSSA 和 PIM 为代表的经典算法以及其最佳运算量的估计,但远未令人满意。近年来在国外流行的李群方法,由于具有重大的科学价值,在李群算法发展的需求下,李群、李代数里的指数计算,也成为了研究的热点。由于它要求逼近计算在李群与李代数里进行,一般不能直接使用经典的方法。因此,它比经典的指数逼近计算更为困难。本文系统地阐述了目前流行的几种主要逼近算法,对这些算法进行了详细的评估,并提出了一些有待深入研究的问题。

关键词:非线性度李群方法;李理论;矩阵指数;对称空间;李三系
中图分类号:O241.81;O152.5 文献标识码:A 文章编号:1672-6693(2008)03-0017-04

矩阵指数逼近计算是数值代数里最为古老且被研究最为广泛的课题之一,19 种令人犹豫不决的算法[1],以及许多针对具体问题的有效的算法,远远未令人满意,并且还留下许多公开的问题。由于在理论与应用上具有巨大价值,始终吸引不少研究者,如我国钟万勰院士,利用 2-幂逼近方法,发现了目前在动力学中广泛使用的精细积分算法。

本文主要就当前比较流行,并在控制论、机器人、材料科学、相对论、电动力学、分子动力学和量子力学等有重要应用的李群算法里矩阵指数逼近问题进行分析讨论。这里的逼近不同于经典的逼近方法,它要求矩阵属于某个李代数 $g\subset gl(n,C)$-全体 n 阶复矩阵李代数,而相应的指数、指数逼近属于李代数对应的李群 $G\subset GL(n,C)$-全体 n 阶可逆复矩阵群,所以理论上难度更大。

1 理论背景

先简要介绍李理论与微分流形的一些基本概念,详细内容可参考文献[2]。一般地,设任意 $F\in X(M)$ 是流形 M 上向量场,$F: M\to TM=\cup_{p\in M}T_pM$,这里 $F(p)\in T_pM$(p 点的切空间),F 在点 $p\in M$ 的积分曲线是指映射 $y:\mathbf{R}\to M$,$y(0)=p\in M$,并且 $\frac{\mathrm{d}}{\mathrm{d}t}y=F(y)$,$y(0)=p\in M$,等价于 $F(y)\in T_yM$,$y\in M$。

定义 1 流形 M 上的微分方程,是指满足下列条件的方程

$$y'=F(y),t\geqslant 0,y(0)=p\in M \tag{1}$$

其中 $F\in X(M)$,F 的流是指映射 $\psi_{F,t}:M\to M$,使得对 $\forall t\in\mathbf{R}$,$y(t)=\psi_{F,t}(y(0))=\psi_{F,t}(P)$ 是微分方程(1)保持在 M 上的解,并称微分方程(1)在 M 上变化,有 $\frac{\mathrm{d}\psi_{F,t}(y)}{\mathrm{d}t}|_{t=0}=F(y)$。

经典的数值方法直接应用到(1)式常常会产生漂移。

为了便于应用,很多时候需假设有下列流形上微分方程的一般表示方法[2]。

假设 设 $y'=F(t,y)$ 设是 M 上的方程(一般情形),存在李代数 g,函数 $f:\mathbf{R}\times M\to g$,李作用 $\lambda:g\times M\to M$,李代数同态映射 $\lambda_*:g\to X(M)$,使对 $y(t)\in M$ 有

$$y'=(\lambda_*f(t,y))(y),y(0)=p \tag{2}$$

而 $\lambda_*(X)(y)=\frac{\mathrm{d}}{\mathrm{d}t}\lambda(tX,y)|_{t=0}$。

设 G 是李群,g 是 G 的李代数,$\exp:g\to G$ 定义映射使 $\exp(a)=\sum_{i=0}^{\infty}\frac{a^i}{i!}\in G$,称为指数映射。关于李群、李代数的伴随映射 Ad 与 ad 有关系

* 收稿日期:2008-03-03
资助项目:重庆市教委基金项目(No. KJ080805);重庆市自然科学基金项目(No. CSTC2007BB2411);重庆师范大学基金项目(No. 07XLB036)
作者简介:张林华(1966-),男,副教授,博士,研究方向为代数群论和编码密码学。

$$\mathrm{ad}a(b)=[a,b],\mathrm{Adexp}(a)=\exp(\mathrm{ad}a)$$

$\mathrm{Ad}p(a)\in g$ 成立,$p\in G$,a、$b\in g$ 等等[2]。

定义 2　设 $a(t)\in g$,定义 $\mathrm{dexp}:g\times g\to g$,满足

$$\frac{\mathrm{d}}{\mathrm{d}t}\exp(a(t))=\mathrm{dexp}_{a(t)}a'(t)\cdot\exp(a(t))$$

称 d 为指数映射 exp 的微分,dexp 为指数映射 exp 的右平凡化,显然有

$$\mathrm{dexp}_a=\frac{\exp(ad_a)-I}{ad_a}$$

由 $\frac{e^x-1}{x}=1+\frac{1}{2!}x+\cdots+\frac{1}{(i+1)!}x^i+\cdots$ 得

$$\mathrm{dexp}_a(b)=b+\frac{1}{2!}[a,b]+\frac{1}{3!}[a,[a,b]]+\cdots+\frac{1}{(i+1)!}[a,[a,[\cdots[a,b]\cdots]]]+\cdots$$

显然 $\mathrm{dexp}_a^{-1}=\frac{ad_a}{\exp(ad_a)-I}$,设 $\frac{x}{e^x-1}=\sum_{i=1}^{\infty}\frac{B_i}{i!}x^i$,$B_i$ 是 Bernoulli 数,所以

$$\mathrm{dexp}_a^{-1}(b)=b-\frac{1}{2}[a,b]+\frac{1}{12}[a,[a,b]]+\cdots=\sum_{i=0}^{\infty}\frac{B_i}{i!}ad_a^i(b)$$

一般地,对于光滑函数 $\psi:g\to G$,定义 $d\psi:g\times g\to g$,满足

$$\frac{\mathrm{d}}{\mathrm{d}t}\psi(a(t))=\mathrm{d}\psi_{a(t)}(a'(t))\psi(a(t))$$

定理 1　设 $c(t)$ 使得

$$\exp c(t)=\exp(ta)*\exp(tb)$$

则　$c'(t)=\mathrm{dexp}_{c(t)}^{-1}(a)$,$0\leqslant t\leqslant 1$,$c(0)=b$

特别地,若

$$y'=A(t,y)y,t\geqslant 0,y(0)\in G=M$$

且 $A:\mathbb{R}_+\times G\to g$,设此方程的解为

$$y=\exp(\Omega(t))y(0),\forall t,\Omega(t)\in g$$

且满足

$$\Omega'(t)=A+\frac{1}{2}+\frac{1}{12}[A,\Omega],\Omega]-\frac{1}{720}[[[[A,\Omega],\Omega],\Omega],\Omega]+\cdots,\Omega(0)=0\quad(3)$$

则(3)式(也可截段求解)可用通常的数值方法在线性空间 g 里求解,且对任一近似解 $\Omega_n\in g$,有 $y=\exp(\Omega_n)y(0)\in G$,即将流形上的方程转化为线性空间里的方程,并再利用经典数值方法求解,这种方法就是一种李群方法。由此可见指数逼近计算是李群方法的关键。

2 逼近计算方法

由于大多数李群方法需计算矩阵指数,而一般矩阵指数逼近本质上较简单,如数值分析里的有理逼近,Krylov 子空间法,Schur 分解法等等。但这里的逼近是指李代数 g 到它对应的李群 G 里的指数逼近,这样的限制,使得逼近不同于一般的矩阵指数逼近方法,人们发现虽然有的经典方法能够借助,但多数方法逼近不稳定。因此,这里的指数逼近难度更大一些。虽然关于低维代数,如二维、三维代数,矩阵指数计算也有精确的公式,但这些只是例外情形。下面主要就目前几种较为流行的方法进行分析讨论。

2.1 分裂方法

分裂方法,较早用于构造辛方法,如 1990 年,H. Yoshida[3],冯康等人 1995 年证明了李代数 $sl(N)$ 到李群 $SL(N)$ 的解析映射(零对应于单位阵)只能是指数映射[4]。这里的方法主要思想就是将矩阵分裂成其它便于计算的矩阵指数。设有 m 阶矩阵 $B\in g$,$B=\sum_{i=1}^{k}B_i$ 且有

1) $B_i\in g$,$i=1,2,\cdots,k$;

2) 每个 $\exp(tB_i)$ 易于计算,如低秩 B_i;

3) 上面指数乘积是廉价的。

如有一阶逼近

$$\exp(tB)\approx F(tB)=\exp(tB_1)\cdots\exp(tB_k)$$

若用　$F=\exp\left(\frac{1}{2}t\beta_k\right)\cdots\exp\left(\frac{1}{2}t\beta_1\right)\exp\left(\frac{1}{2}t\beta_2\right)\cdots\exp\left(\frac{1}{2}t\beta_k\right)$(Strang 方法)

合成算法

$$F(\gamma_r t)\cdots F(\gamma_1 t)\ F(\gamma_0 t)F(\gamma_1 t)\cdots F(\gamma_r t)$$

当 $r=2^{l-1}-1$ 时,人们猜想阶为 $2l$,但目前还未解决。若想得到高阶的逼近,一般得用多于 k 个的满足以上条件的指数积。对这类方法,人们利用了很多技术、技巧,如 Yoshida 技术等,得到了许多不同阶、不同计算成本与特性(如对称算法)的逼近,并且还在不断地发展。这类方法若不看它的几何优点,就计算成本来讲,它能与一些经典方法竞争。可以说这类方法是矩阵指数逼近中研究内容最为丰富,颇具竞争力的方法。并对其它方法的实现也产生重要作用。

2.2 第二类正则坐标法

这类方法首先考虑李代数 g 的固定的基 $e_1,e_2,$

$\cdots, e_d, g$ 对应于李群 G。设 U 是 g 里的 0 的邻域 $\exp U \subset G, v = \sum_{i=1}^{d} v_i e_i$,在 1998 年,Celledoni 和 Iserles 等人首先使用所谓的第二类标准坐标逼近指数[5],若选取所有 $v_i = \alpha_i(t)$ 适当,可得好的逼近,但如何构造 $\alpha_1(t), \cdots, \alpha_d(t)$ 是一个难题,目前即使对 $sl(2, \mathbf{R})$ 也未解决。其后,Owren 和 Marthinsen(1999 年)使用李代数表示理论[6],用可容许有序基得到了好的逼近算法,并发现 Chevalley 基是这样的基。但目前这类方法只在复半单李代数里用表示理论找到了这样的基,关于实李代数目前还是一个公开的问题,即使是考虑实斜对称 $n \times n$ 矩阵代数 B_l(或 D_l,),$n = 2l + 1$(或 $2l$),也是一个有趣的公开问题。

2.3 广义极坐标分解方法

在线性代数中,任何 $N \times N$ 矩阵 A 可分解为酉矩阵 A 和 Hermitian 半正定矩阵的积,若 A 可逆,则得到所谓的矩阵的极分解。在李理论里,利用微分几何中对称空间与李三系理论[7],人们找到了所谓的矩阵的广义极坐标分解方法。1994 年 Lawson[8] 证明了在李群里极分解的存在性和唯一性。

设 $G \subset GL(\mathbf{R}, n)$ 是对应于李代数 g 的李群,$\sigma \neq id$ 是 G 的对合自同构,由对合自同构 $\sigma \neq id$,可诱导 g 上面的对合自同构 $d\sigma(Z) = \frac{\mathrm{d}}{\mathrm{d}t}|_{t=0} \sigma(\exp(tZ))$,则有 $g = p \oplus e$,这里 $e = \{Z \in g; d\sigma(Z) = Z\}$ 是 g 的子代数,$p = \{Z \in g; d\sigma(Z) = -Z\}$ 是一李三系。对 $\forall z \in G$临近 G 的单位元的元素,则所谓的 z 的广义极分解,是指对充分小的 t,有元素

$$z = \exp(tZ) \in G, Z = K + P$$

是 $g = p \oplus e$ 里的分解,并且有 $d\sigma(K) = K, d\sigma(P) = -P$,则有分解 $z = x \cdot y$,其中

$$x = \exp(X(t)) \in G_\sigma = \{w \in G; \sigma(w) = w^{-1}\}$$

G_σ 关于运算 $x \cdot y = xy^{-1}x$ 构成对称空间,

$$y = \exp(Y(t)) \in G^\sigma = \{w \in G; \sigma(w) = w\}$$

是 G 的子群。2000 年,H. Munthe-Kaas 等人用不同于 Lawson 的方法也证明了这样的分解的存在和局部唯一,但全局唯一性问题目前还没很好解决[9]。

设 $\sigma_1 = \sigma, \sigma_2, \cdots, \sigma_m$ 是 G 上的对合自同构,则 $g = p_1 \oplus e_1$,由于 e_1 是 g 的子代数,所以归纳有 $g = p_1 \oplus p_2 \oplus \cdots \oplus p_m \oplus e_m$,相应地有

$$\exp(tZ) \approx \exp(X^{[1]}(t)) \exp(Y^{[1]}(t)), \cdots$$

$$\exp(tZ) \approx$$

$$\exp(X^{[1]}(t)) \cdots \exp(X^{[m]}(t)) \exp(Y^{[m]}(t))$$

一般称这样的分解为广义极坐标分解,还可构造对称形式与镜像形式。2001 年,A. Zanna 对此类方法进行了详细的误差分析[10]。

关于算法实现问题,S. Krogstad 等人基于分裂技术提出了具有一般性的有效方法[11],这类方法,关于典型群计算复杂度是 $O(n^3)$。2005 年,A. Iserle 等人采用上 Hessenberg 型与去上(peel-up)技术结合对诸如正交群、辛群、Lorenz 群、迷向群和纯量群等较好地实现,但稳定性有待研究,对其它群也有待推广[12]。人们将矩阵指数的对称的 GPC 与非对称的 GPC 作用到向量 $v \in \mathbf{R}^n$ 得出对称方法有较好的结果,而作用到矩阵 $B \in C^{n \times n}$则非对称方法较廉价,但若方程关于时间对称,则人们偏好用对称逼近方法。当前线性代数中矩阵极分解技术及其它技术推广到一般李群情形是人们继续研究的课题[13-17]。

3 结论

本文分析讨论了在李群与李代数里的矩阵指数逼近计算,人们利用换位元技术也可得到高阶逼近,而由矩阵指数导出的平凡化切及它的逆的逼近也是重要的研究课题。从目前研究趋势看,选用线性代数技术是人们自然的选择,但是利用李理论的结构及表示理论研究矩阵指数逼近计算,尤其是对称空间理论与李三系理论应用到指数映射等的逼近,将是未来研究的重点[12-17]。

参考文献:

[1] MOLER C B, VANLOAD C F. Nineteen Dubious Ways to Compute the Exponential of a Matrix[J]. SIAM Review, 1978, 20:801-836.

[2] ISERLES A, MUNTHE-KAAS H, ZANNA A. Lie-group Methods[M]. Cambridge: Cambridge University Press, 2000.

[3] YOSHIDA H. Construction of Higher Order Symplectic Integrator s[J]. Phy lett A, 1990,150:262-268.

[4] FENG K, ZAI J. Volume-preserving Algorithms for Source-free Dynamical Systems[J]. Numer Math,1995, 71:451-463.

[5] CELLEDONI E, ISERLES A. Methods for the Approximation of the Matrix Exponential in a Lie-algebraic Settin-g[J]. IMA J Num Anal, 2001,21:463-488.

[6] OWREN B, MARTHINSEN A. Integration Methods Based on Canonocal Coordinates of the Second Kind[J]. Numer Mth, 2001,87(4):763-790.

[7] MUNTHE-KAAS H, QUISPEL G R, ZANNA A. Application of Symmetric Spaces and Lie Triple Systems in Numeri-

cal Analysis[EB/OL]. (2002-01-10)[2008-03-03]. http://www.focm.net/gi/gips/2002/002.pdf.
[8] LAWSON J D. Polar and O l 'shanskii Decoposition s[J]. J Reine Angew Math, 1994, 448: 191-219.
[9] MUNTHE-KAAS H, QUISPEL R, ZANNA A. The Polar Decomposition on Lie Groups with Involutive Autom- Orphisms[J]. Found Comp Math, 2001, 1: 297-324.
[10] ZANNA A. Error Analysis for Exponential Splitting Based on the Generalized Polar Decomposition I: Local and Global Bounds[EB/OL]. (2002-01-10)[2008-03-03]. http://www.focm.net/gi/gips/2002/001.pdf.
[11] KROGSTAD S, MUNTHE-KAAS H, ZANNA A. Generalized Polar Coordinates on Lie Groups and Numerical Integrators. (2003-06-19)[2008-03-03]. http://www.focm.net/gi/gips/2003/005.pdf.
[12] ISERLES A, ZANNA A. Efficient Computation of the Matrix Exponential by Generalized Polar Decompositions [J]. SIAM J Num Anal, 2005, 42: 2218-2256.
[13] CELLEDONI E, ISERLES A. Approxiating the Exponential from a Lie Algebra to a Lie Group[J]. Math of Comput, 2000, 69: 1457-1480.
[14] ZANNA A, MUNTHEKAAS H. Generalized Polar Decompositions for the Approximation of the Matrix Exponential [J]. SIAM J Matrix Anal, 2002, 23(3): 840-862.
[15] **孟道骥，史毅茜. 李三系分解的唯一性**[J]. **数学进展**, 2004, 33(1): 52-56.
[16] CELLEDONI E, ISERLES A, NORSETT S P, et al. Complexity Theory for Lie-group Solvers[J]. J Complexity, 2002, 18: 242-286.
[17] **吴永，李正良. 流形上微分方程的算法**[J]. **数学进展**, 2006, 35(4): 385-394.

Computation of Matrix Exponential in Lie Group Methods

ZHANG Lin-hua[1], *WU Yong*[2]

(1. College of Mathematics and Computer Science, Chongqing Normal University, Chongqing 400047;
2. School of Mathematics and Physics, Chongqing Institute of Technology, Chongqing 400050, China)

Abstract: The computation of matrix exponential which ties up structural mechanics and optimum control is one of the extensive topics in numerical mathematics. Though there are many algorithms about the matrix exponential such as Pade-scaling and squaring-approximation(PSSA), precise integration method(PIM) and some results about those optimal computational cost, it is far from satisfying to us. Currently Lie group methods are of huge value and they are popular abroad, and lead to the computation of matrix exponential to be a hot topic. Since numerical approximation must run in Lie groups or Lie algebras, some classic methods are no effect. It is more difficult than classic method. In this paper a few popular computation methods of matrix exponential are reviewed systematically and evaluated in detail, and some problems which need our thorough research are put forward.

Key words: Lie group method; Lie theory; Matrix exponential; Symmetric space; Lie triple system

(责任编辑 黄 颖)