



科学计算 科技创新的 第三种方法*

文/陈志明

中国科学院数学与系统科学研究院 北京 100190

【摘要】 科学计算利用先进的计算能力认识 and 解决复杂的科学工程问题,它融建模、算法、软件研制和计算模拟为一体,是计算机实现其在高科技领域应用的必不可少的纽带和工具。计算、理论和实验一起已成为当今世界科学技术创新的主要方式。本文简要回顾了计算数学和科学计算的发展历史,从数值代数、数值逼近、最优化方法和微分方程计算方法等4个方面简要介绍了计算数学的主要研究内容和最新进展,最后提出进一步发展我国科学计算的建议。

【关键词】 科学计算 计算数学 科技创新

DOI :10.3969/j.issn.1000-3045.2012.02.004

1 计算数学和科学计算的兴起和发展

1947年 Von Neumann 和 Goldstine 在《美国数学会通报》发表了题为 高阶矩阵的数值求逆 的著名论文,开启了现代计算数学的研究。计算数学研究为可在计算机上运行的数值算法的构造及其数学理论,包括算法的收敛性、精确性、稳定性和计算复杂性等。60多年来,伴随着计算机技术的进步,计算数学得到了蓬勃发展,逐渐成为一个独立和重要的学科。

上个世纪90年代,由于微电子技术的发

展和应用方面需求的推动,计算机得到了飞速发展,计算数学、应用数学、计算机科学以及应用领域结合在一起产生了科学计算这一新的交叉学科。科学计算利用先进的计算能力认识 and 解决复杂的科学工程问题,它融建模、算法、软件研制和计算模拟为一体,是计算机实现其在高科技领域应用的必不可少的纽带和工具。计算、理论和实验一起已成为当今世界科学技术创新的主要方式。

进入21世纪以来,高性能计算机发展迅速,已经进入了千万亿次时代。千万亿次科学计算应用将显著提升人类在气候与生态环境、航空航天、生命科学、材料科学、国家

* 收稿日期 2012年2月21日

安全等领域中的科技创新能力,产生重大科学理论和应用突破。当前,基于通用CPU和GPU异构的千万亿次计算机的总处理器核数已超过10万。不断膨胀的并行规模给并行算法研究及应用程序研制不断提出新的挑战。研制适应于千万亿次科学计算的高性能应用软件成为我国及世界科技发展所面临的一个重大问题。

高性能科学计算由于其在国家安全和科技创新方面的重要作用日益受到世界各国的重视。以美国为例,自1996年实施加速战略计算计划(ASCI计划)以来,在国防、气候与环境、能源、航空航天、工业仿真优化、生物制药等行业领域,针对流体力学、结构力学、材料科学、天体物理、等离子体物理、全球气候变化、生物计算等学科,发展并行算法库,研制并行应用程序,突破了高效使用数千至数万核开展大规模并行计算的若干关键瓶颈,在某些方面甚至达到了10多万核的并行规模,形成了一批有代表性的并行数值软件包和并行应用程序,如ETSc、Hypre、SAMRAI等。

欧盟投入巨资建设欧洲高性能计算生态系统,2007年包括德国、英国、法国、意大利等14个欧洲国家的代表宣布将合作开展欧洲先进计算合作(PRACE)计划,大力发展先进的计算方法、算法,研制大型实际应用软件,以加速提高欧盟的高性能计算水平。日本通过成功研制峰值计算速度达40多万亿次的超级计算机地球模拟器,大力推进高性能科学与工程计算的研究,在全球气候变化数值模拟研究方面取得了突出的成果,目前日本正在推进耗资10亿美元的万万亿次计算机计划,研究纳米和生命科学。

我国非常重视计算数学和科学计算的发展,在历次科学规划中都将计算数学和科学计算列为重点发展领域。早在1956年制定的《十二年科学技术发展规划》中,计算数学的发展和计算机在科学技术中的应用就已与计算机硬件的研制开发相并列。上个世纪90年代,科技部先后资助了两期攀登计划项目大规模科学与工程计算的方法和理

论,1999年起将大规模科学计算研究、高性能科学计算研究和适应于千万亿次科学计算的新型计算模式作为国家重点基础研究发展计划(973计划)项目予以连续支持,2011年国家自然科学基金委又启动了高性能科学计算的基础算法和可计算建模的重大研究计划。

2 计算数学的主要研究内容和一些新进展

随着高性能计算机技术的进步,计算机的计算能力和数据处理能力得到大幅度提升,当前科学与工程计算所需要解决的问题越来越复杂、越来越细致、越来越接近实际问题的模型。这对充分发挥计算机巨大能力,实际问题的高效计算方法提出了越来越迫切的需求。下面从分数值代数、数值逼近、最优化方法和微分方程计算方法4个方面简要介绍计算数学的主要研究内容和最新进展。

2.1 数值代数

数值代数主要研究代数方程组、代数特征值问题和最小二乘问题的数值求解方法,数值代数的算法在科学计算中具有基础作用。偏微分方程离散产生的大型稀疏代数方程组快速算法一直是科学计算的重要研究课题,过去30年取得了很大进展,其标志性的例子是椭圆离散问题的多重网格方法,它从上世纪70年代后期的几何多重网格方法开始,发展成通用的可扩展的代数多重网格方法AMG,以通用AMG程序为基础,研究离散偏微分方程的可扩展算法成为一个非常重要的研究方向。另一个重要进展是稀疏代数方程组的直接法,目前对于二维偏微分方程离散的代数方程组,多界面方法及其程序MUMPS的计算复杂性已降为 $O(N^{1.5})$,较好满足了许多二维偏微分方程计算的实际需求,如何降低多界面方法对三维离散偏微分方程的计算复杂性成为重要的研究课题。

大型稀疏代数问题的快速算法的另一个基本问题是特征值问题的计算方法,上世纪90年代发展成熟的带隐式重启动的Arnoldi算法已经能够求解大部分实际应用中出现的大规模线性特征值问

题,但对于一些特殊应用中出现的特殊要求的特征值问题还不能有效求解。关于特征向量非线性的非线性特征值问题,尤其是对于电子结构计算中的非线性特征值问题,现有的算法从收敛性、收敛速度、计算指定的特征值等方面均不能满足要求。

2.2 数值逼近

对数据、图像、函数等对象的逼近是计算方法中最基本的手段之一。众所周知,现代信息技术的基础是数据的表示和变换方法,有效的数据表示和数学变换为计算机处理大规模信息提供了可能。随着信息技术的高速发展和越来越复杂数据处理的需要,新问题不断涌现,逼近论已成为计算方法中最为活跃的一个部分。数值逼近的一个重要领域是计算几何,它是由函数逼近论、微分几何学以及计算方法等学科交叉形成的学科,主要研究几何形状的构造、计算机表示、分析和综合,是计算机辅助几何设计的数学基础。

逼近论的最新进展包括基函数、冗余框架或更一般的冗余词典的稀疏逼近,其基本思想是利用一个非线性逼近格式得到函数、数据或图像等对象的稀疏逼近。过去20多年小波理论及其应用的快速发展和最近压缩感知的快速发展为非线性稀疏逼近提供了重要的理论基础和计算方法,这些理论和方法与调和分析、小波理论、非线性逼近和优化方法等理论密切关联,它们突破了经典理论方法的局限,极大地提高了复杂数据信息处理的能力,成为图像科学、计算机图形学、数据挖掘、机器学习理论的重要工具。

2.3 最优化方法

最优化问题广泛见之于工程、国防、经济、管理等许多重要领域,在结构设计、化学反应设计、电力分配、石油开采等方面都有直接的应用。最优化计算方法还和计算数

学中的数值逼近、常微分方程中的变分原理、微分方程反演以及非线性代数方程组等分支和问题有交叉和应用。许多领域诸如压缩感知、数据挖掘、核磁共振、最优控制、图像处理、矩阵方程中的优化问题其规模往往很大,这对于最优化领域的计算方法的设计既是挑战又是机遇。

偏微分方程约束优化问题与无穷维优化问题在航天飞行器设计、卫星红外扫描仪、医疗设备形状设计、刀片涡轮片执行器控制、水库管理、天气预报、大气污染控制、石油勘探和开采等许多工业、医学和经济等领域中有着广泛应用,是这些领域中最具有挑战性的困难之一。这些优化问题的目标函数值和梯度值的计算往往需要求解极复杂的偏微分方程,与常见的优化问题相比,问题的规模常常很大,需要上千万甚至上亿自由度,有时目标函数是非光滑、非一致凸甚至是非凸的,往往需要求解一连串的优化问题及大规模非光滑全局优化问题,而且优化问题的鲁棒解和有效解需要很好地融合问题的结构、数值模拟和优化算法3个方面。

2.4 微分方程计算方法

哈密尔顿系统辛几何算法是由我国著名数学家冯康院士于上世纪80年代开始系统创立的,1997年获得国家自然科学奖一等奖。无论在理论上还是在应用上,它的重要性均受到国际学术界的高度重视。特别是,它的长时间计算的优越性使得天体物理学、量子物理学、纳米材料和分子生物学等众多领域的科学家们进一步认识了科学计算这一研究方法的重要作用。哈密尔顿系统辛几何算法的基本思想是数值格式应该尽可能多的保持原系统的本质特征和内在对称性。根据这个思想,近年来国外学者先后建立了动力系统李群算法、无穷维哈密尔顿系统多辛几何算法、基于微分形式的有限元



中国科学院

方法和随机动力系统的保结构算法等。现已形成了包括哈密尔顿系统辛几何算法、切触系统的切触算法、保体积系统的保体积算法等在内的动力系统保结构算法的理论与应用体系。

偏微分方程描述许多构成实际物理过程的各个不同阶段的物理模型,其计算方法包括如何针对不同偏微分方程的问题设计合适的网格和离散格式,如何设计可扩展的并行算法及其并行实现技术,在离散网格上给出方程的近似解。偏微分方程的高效计算方法研究在过去30年中取得了巨大的进步,突出表现在网格的自适应构造方法及理论研究的兴起,非结构网格和混合网格上流体力学方程高精度高分辨率离散格式的进展。当前,在网格自适应方法的研究方面,主要有针对流体力学计算的结构网格自适应技术和针对椭圆抛物方程应用的基于后验误差估计的自适应有限元方法;在高精度高分辨率离散格式方面,流固耦合、多介质多相流问题的高效、皮实数值格式以及密切相关的网格构造的ALE方法,与ALE方法密切相关的网格自适应技术和移动网格技术、利用曲面造型技术产生的数据直接进行有限元结构分析的等几何方法、充分利用所研究问题的结构的外微分有限元方法等得到越来越多的计算数学工作者的关注。

3 科学计算进一步发展的建议

3.1 大力发展大规模科学计算的基础算法

科学计算能力包括计算机硬件设备和应用软件及支撑软件的算法的能力。2005年美国总统一信息技术咨询委员会报告指出:尽管处理器性能的显著增长广为人知,然而改进算法和程序库对于提高计算模拟能力的贡献是如此之大,如同在硬件上的改进一样。以在科学计算应用中广泛出现三维拉普拉斯方程计算求解为例,从上世纪50年代的高斯消去法到80年代的多重网格法,算法的改进使计算量从正比于网格数 N 的 $7/3$ 次方下降到最优的计算量正比于 N ,对于 N 等于100万,计算效率就改进1亿倍!2009年出版的美国世界技术评估

中心WTEC报告中对1998—2006年获著名超级计算Gorden Bell奖的应用程序进行了评估,指出尽管获奖程序的应用领域各不相同,但共同点是,算法(线性代数、图剖分、区域分裂、高阶离散)的进步使得获Gorden Bell奖应用程序对计算能力提高的贡献超过摩尔定律。

当前科学计算所要解决的数值模拟问题往往非常复杂,给数值方法研究带来了巨大的挑战。数值方法研究目前面临的突出共同难点表现为:高维数、计算规模大、多时空尺度、强非线性、不适定、长时间、奇异性、几何复杂、高度病态、精度要求高等,并非有了高性能计算机就可以解决这些难点。数值模拟的困难常常表现为规模大得难以承受或失去时效,算法不收敛或误差积累使结果面目全非;花费大量计算机时却得不到结果或只得到错误结果;由于问题的奇异性使计算非正常中止;问题太复杂使算法难以实现等。这些难点问题近年来受到广泛关注,已成为科学计算的研究热点。

3.2 大力发展高性能科学计算应用软件平台和框架

当前高性能计算机体系结构日趋复杂,CPU/GPU异构、数千计算结点、结点内多处理器、处理器内多核、核内多功能部件及多级存储是其显著特点,对应用软件的研制提出了挑战。当前,我国科学计算应用软件的研制面临两大瓶颈:第一,计算效率低,应用程序通常只能发挥数百处理器核峰值性能的10%以下;第二,研制周期长,与高性能计算机的快速发展不匹配。不突破这两个瓶颈,我国并行应用软件的研制就很难在总体上跟上高性能计算机的发展速度,无法将计算机技术的进步有效地用于科技创新,无法在国家重大应用领域发挥其应有的价值。

为了突破“计算效率低”和“研制周期长”两大瓶颈,近年来我国科学工作者根据并行应用程序的多层软件体系结构,提出了“集成共性、支撑个性”的新型并行软件研制方法,在此新思路下,研制成功三维并行结构自适应软件框架JASMIN和三维

并行自适应有限元软件平台 PHG, 这两个平台在基础性共性算法(例如网格自适应和数百处理器核上的求解器)层次上对用户屏蔽并行实现细节, 并较好地解决了自适应并行实现中的负载平衡难题, 在数 10 万亿次国产并行计算机上, 实现了高效并行计算。

在框架和平台的支撑下, 各专业领域的科学计算研究人员可以集中于物理模型和计算方法的创新研究, 无需了解并行计算的细节, 就可以将新的物理模型和计算方法快速融入到大规模并行计算中, 而计算机系统的研究人员, 则可以集中于更高速度和更大规模计算机系统的研制, 而无须顾及实现具体科学和工程计算的细节。

3.3 大力加强自主高性能计算科学软件研制

高性能科学计算应用软件的重要特点是多学科交叉, 它是数学、物理、力学等基础学科和相应应用学科及计算机软件技术相结合而形成的以算法为核心, 以计算机系统为支撑的知识密集型集成化信息产品, 领域专业性非常强, 只有建立高水平的多学科交叉研究队伍, 针对实际科学问题经过长时间的积累和沉淀, 才能研制成功。

高水平的计算科学软件需要先进的算法, 我们建议在国家重大科技项目执行中能特别重视高性能计算科学软件的自主研制, 鼓励更多的计算数学工作者深入科学计算中的瓶颈问题如材料计算、流体计算、电磁场计算、辐射流体力学计算、纳米计算和生

物计算中的算法研究、多尺度模型的分析与计算以及非平衡态的计算等, 算法的创新要结合适合于计算的模型进行研究。

致谢 本文得到北京应用物理与计算数学研究所江松、莫则尧研究员, 新加坡国立大学沈佐伟教授, 中山大学许跃生教授, 复旦大学苏仰锋教授, 中科院数学与系统科学院白中治、戴彧虹、洪佳林、徐国良、张林波研究员和许志强副研究员的大力帮助, 特此致谢。

主要参考文献

- 1 2005 年总统信息技术咨询委员会报告. 计算科学 确保美国竞争力. 北京应用物理与计算数学研究所信息中心译, 2006 年 4 月.
- 2 WTEC Panel Report. International assessment on research and development in simulation-based engineering and science. Washington: World Technology Evaluation Center, 2009. <http://wtec.org>.
- 3 中国科学院数学学部 高性能计算战略研究 咨询组. 加速发展我国高性能计算的若干建议. 科研信息化技术与应用, 2008, (3).
- 4 Bader D A. 面向千万亿次计算的算法与应用. 都志辉等译. 北京 清华大学出版社, 2008.
- 5 Grcar J F. John von neumann's analysis of gaussian elimination and the origins of modern numerical analysis. SIAM Review, 2011, 53: 607-682.
- 6 国家自然科学基金委. 2011-2020 年我国数学学科发展战略研究, 2010 年.

Scientific Computing: The Third Pillar in Science and Technology Innovations

Chen Zhiming

(Academy of Mathematics and Systems Science, CAS 100190 Beijing)

Abstract Scientific Computing understands and solves complex scientific and engineering problems using advanced computational capabilities. It integrates the studies of modeling, algorithm, software, and simulation and becomes the indispensable link and tool of the application of modern high performance computer in science and

technology. Computation, together with the theory and experiment, has become the main method in current worldwide science and technology innovations. In this paper, we first briefly review the history and development of computational mathematics and scientific computing. Then we introduce the main research directions and recent development of computational mathematics in four fields: numerical algebra, numerical approximation, optimization method, and numerical methods for differential equations. Finally we propose several suggestions on further developing scientific computing in China.

Keywords scientific computing, computational mathematics, science and technology innovations

陈志明 中科院数学与系统科学院研究员、计算数学与科学与工程计算所所长。1965年7月出生。1992年在德国Augsburg大学获自然科学博士学位。科学与工程计算国家重点实验室主任,国家973项目高性能科学计算研究和适应于千万亿次科学计算的新型计算模式首席科学家。主要从事数值分析与科学计算研究,近年研究方向为偏微分方程的自适应有限元方法、非均匀多孔介质流动问题的多尺度计算方法。2000年入选中科院百人计划,2000年获国家杰出青年基金,2001年获第四届冯康科学计算奖,2009年获国家自然科学基金二等奖。2006年在西班牙举行的国际数学家大会上做45分钟邀请报告。E-mail:zmchen@lsec.cc.ac.cn