

计量经济学中若干计算问题的现状与进展

何 耀

改革开放以来,国内在数量经济学、计量经济学以及与经济研究相关的动态最优化、博弈论等方面的研究和应用有了快速的发展。计算机在研究中起的帮助作用越来越大。与发达国家相比,我们的硬件装备是紧随潮流的,但从“软”的方面来讲,我们在研究方法、工具软件和网络的开发使用等方面还有很大的差距。从目前来看,计算机技术的发展对于计量经济学估计方法的促进作用愈显突出。计算技术的含义是广泛的,至少包括计算机硬件、计算机应用软件、网络技术、计算方法等多个方面的内容。80年代后期以来,以计算机为载体的计算技术渗透到人类生活的各个方面。计量经济学与“计算”有着与生俱来的联系,近年来二者之间的“互动”也有必要提及。文章仅就我们涉及到的资料加以归纳分类,并不全面,但我们的本意在于引起经济学特别是应用经济学领域工作者对近期计量经济学发展趋势的注意,更进一步我们希望在高校经济类课程的教学过程中,注意调整学生的知识结构。经济学研究手段的现代化必定寄希望于新一代的学人。从培养教育入手是唯一现实的道路。

以下分五个方面对现状与进展作简叙。

一、动态规划模型的估计

首先我们看看在动态规划模型(以下简称DP模型)估计方面的进展。80年代初以前,这方面还几乎没有任何实质性的工作出现,一方面是对这类模型的估计理论还没有,但更重要的是由于计算机的性能远远不能满足估计的计算需要。直到1984年,由米勒(R. Miller)和沃尔平(K. Wolpin)等人首先提出了对DP模型估计的理论框架。模型估计的基本构想是相对简单的:他们构造的优化算法是由“外部”和“内部”二部分构

成的。所谓“外部”优化算法是在 N 维空间(有限维)上搜索数据最佳拟合状态下DP模型的未知参数,这里所涉及的最佳拟合可以是极大似然估计意义上的;“内部”DP算法是针对每一个参数试算值来求解动态规划问题,“内部”DP算法常常涉及的是相当高维数的空间,比如在国外对退休问题的动态规划模型的处理中,问题被定式为在一个上百万维空间中压缩映射的不动点计算。这意味着DP问题的求解需要惊人数量的计算量和计算存储量,其中关于不确定性计算的变换矩阵就需要上百兆的存储空间。而上面提到的“外部”优化算法必须反复对最佳拟合参数向量进行成百上千次的搜索求解。十几年前,这种计算问题所需要的计算速度和存储容量还没有任何一种计算机可以满足,但在今天,一台个人使用的微机就能处理它。现在DP模型已经在经济研究领域得到广泛的应用,比如投资组合与储蓄的选择,人类生育控制,专利的更新,核电厂的监控管理,保险公司对于退休行为的分析,企业对资产折旧的抉择等等。

与DP模型估计相关的经济均衡模型估计方法的进展有必要在此提及。在一般情况下,建立计量经济模型时,往往考虑的是要估计包含时间和不确定性的单一动因模型(Single-agent model),这类模型可被视为“违背自然的对策”,因为模型是将自然运动的规律处理为外生的,而仅把问题中对研究目标的未来变动有决定性影响的因素作为内生变量来处理。但是近来出现了一些包含多种对整体和全局有重大影响因素的多动因模型,这些模型就是包括理性预期(rational expectations)和动态博弈理论模型(dynamic games-theory model)在内的一般经济均衡模型。但遗憾的是目前还几乎没有对此类模型进行有效估计的普适性算法。这将是今后一段时期内计量经济学研究的前沿问题。它所面临的理论问题和计算问题都是非常困难

的。目前用于估计和求解理性预期模型和动态博弈理论均衡模型的算法都是针对特殊情况而专门设计的特殊目标算法 (special purpose algorithms)。比如 Muth 针对价格变化作出理性预期的离散时间随机模型, Cagan 针对恶性通货膨胀作出的货币动态模型都是此类为人熟知的早期模型。总之, 目前还不存在象求解单一动因 DP 模型的“向后递归”(backward-recursion)方法一样的求解一般均衡问题的普适性可靠算法, 即使是具有普遍意义的求解不动点的算法——Scarf 算法(即单纯形算法), 也只是针对静态一般经济均衡模型的。但是无论如何, 这些针对特殊问题的特定目标算法毕竟在开拓通向普适算法的道路。人们也预期更高级的计算技术渗入到这一困难的研究领域中来。

二、以模拟求估计

用计量经济方法对人们经济行为的分析也由于计算机的发展而得到极大的促进。用“模拟求估计”(Estimation by Simulation)的方法是计量经济学的突破性进展。把人的行为视为一个优化过程可以说是经济学处理行为分析的基本切入点。在十几年前, 典型的行为分析在计算上是将模型的优化解处理为闭型的, 以便于模型特定参数的计算达到所需的精度。这种处理隐含着以下的事实: 给定一个特定参数的设置, 观测数据的潜在特征就能在任何指定的精度上被计算出来, 但是其逆问题即从观测数据出发去解释参数值的合理性却得不到解决。到 90 年代初麦克费顿(D. McFadden)等人首次在理论上取得突破, 他们认为原先计量模型估计中难以处理的问题, 其关键在于用多重积分方法求得的模型估计解并不是闭型的, 而是非闭型解。对含有多元标准正态分布的多重积分的离散抉择概率单位模型进行估计, 其目的是为了计算似然函数的抉择概率, 而概率单位模型对于带有多于 2 个以上元素的抉择集合问题的处理是相当困难的。为此他们提出新的模拟方法以避免直接估计上述的多重积分, 代之以对不可观测的现象作蒙特—卡罗抽样, 进而产生无偏的模拟抉择, 即用模拟抉择的期望去代替由直接数值积分法所估计的抉择概率。这一方法在理论上有很大的优点, 但对于计算技术有强烈的依赖。这是因为标准的模拟方法所产生的统计目标函数——似然函数是典型的非连续不可微的, 这使计算似然函数的非线性爬山法的应用有困难, 因为爬山法中有计算数值微分的必要, 而且函数的不连续性导致对大量局部优化解的计算, 用以寻求整体优化参数的估计还有很大的困难。人们由此开始认识到更复杂的模型将会更加耗费机时。但随着计算机性能的提高, 凯纳(M. Keane)等人已在 90 年代初发展了这一算法, 使其变得更象一个“光滑的模拟”方

法, 并将其应用在就业与工资的时序—横截面数据集分析上。针对爬山法只是求出局部优化参数的缺点, 更多的研究者把注意力放在整体最优估计的算法上, 近年来出现的“模拟渐近冷却”算法(Simulated annealing algorithm)就是一个逐渐依概率收敛到整体最优的方法。最初不能推广它的原因并不是算法理论的障碍, 而恰恰是机器计算速度的限制, 但随着计算机性能的高速发展, 整体最优算法也正在扩展它的应用范围, 这也是当前计量经济方法研究的热点之一。

三、交互式的图形数据分析

基本数据的采集和分析, 调查数据的登录, 变量指标的设置、统计制表和数据清理、匹配等等是计量经济工作者公认的繁重事务, 不论在国内还是国外, 往往称其为“费力不讨好”的工作(dirty work), 研究者们也往往将此类工作留给研究生或年青的助手去做, 具有讽刺意味的是这些工作的好坏恰恰是计量经济经验分析成败的关键, 特别是近年来国外发达国家的经济实证研究中经常使用大规模的时序——横截面数据板块(panel data), 比如 NLS(美国的 National Longitudinal Survey of Labor Market Experience)和 PSD(密西根大学的 Panel Study of Income Dynamics)这二个常被提及的典型数据板块, 其中 PSD 是包括 6 000 个家庭、15 000 个人并从 1968 年开始跟踪调查, 直到目前还在不断扩充的时序——横截面统计数据集合, 对如此庞大的数据集进行整理、匹配于研究和分析工作的成功具有重要意义, 但这种工作的繁重是不言而喻的, 处理必须借助于计算机和高效率的软件。在 80 年代, 数据处理工作使用的还是非交互式的标准程序语言如 Fortran 和 Pascal, 且数据分析工作要在主机上做一系列的循环试算去定义不同的数据, 然后制作统计表格, 再将结果图形化, 以显示数据的特征。要实现以上目的, 工作者还需要作编辑——汇编——连接——输入的前期操作。到 90 年代中期后, 个人计算机的高性能和高效率的操作系统、数学应用软件使人们从繁杂的数据处理工作中解放出来。基于 Windows 和 Unix 平台上的交互式图形数据功能的软件包大量涌现, 如 Gauss, Mathematica, Matlab, MathCAD, Maple 等等, 除了数据登录需要键盘输入外, 数据编辑整理、制表、特征图形化等都可用简单的程序指令在屏幕上适时处理, 甚至连计量经济学教学中使用的软件 Eview(其前身是 MicroTSP)都可以做较简单的数据分析工作。现在所期望的下一步目标, 可能是语音输入技术的完善和不同数据格式间相互转换的软件以及网络上大容量信息传输的技术突破。这些计算技术的发展会使计量经济工作者更乐于去做关键性的数据分析工作。而更现实的目标是计量经济学的研究者和他们

的学生都应熟悉已有的数据分析和计量分析软件的使用方法。

四、应用于计量经济学的人工智能方法

在规范的经济分析中,总是把理论假设尽量考虑得复杂而合理,以便于把理论中的经济代理人行为优化趋真于实际的经济人行为。而一旦将此种行为优化问题定式为数学模型,人们马上发展它是不可解的,用计算机专业语言表达这种情况时,称之为“NP—完全”问题,即此类问题的规模只要有细微的增加,其计算量就会巨量膨胀,并轻易导致超级计算机过载停机,除非机器容量是无限大的。为了克服这种困难,一种新的思想是使用人工智能方法设计一些只受简单行为规则指导的“经济代理人”,并赋予它们以学习能力,“代理人”被置于一个模拟经济环境中,并考察它们的学习是否会使其行为变得与预先理论假设中的复杂行为相一致。1988年,伊斯利(D. Easley)和莱得亚得(J. Ledyard)使用上述思想对在纽约股票交易所和芝加哥交易所(CBT)实行多年的“双向竞价规则”(double auction institution)下的交易者行为进行了仿真。他们的工作表明,如果交易者是用相对简单的“经验法则”(rules of thumb)指导它们的交易行为,交易价格和数量的轨线将收敛到理论上的瓦尔拉斯均衡状态,即这一结果趋真于实际金融市场交易行为的结果。为了进一步证实上述人工智能思路的正确性,1990年帕尔默(R. Palmer)和拉斯特(J. Rust)等人发起了一个“双向竞价”仿真的比赛,有30位计量经济学家和计算机科学家提交了他们的仿真计算程序。程序中收入的交易法则五花八门,从最简单的“经验法则”到最复杂的交易算法如神经网络和细胞适应性曲线拟合法等等,不一而足。最终的胜利者所提出的方法恰恰是由最简单的非适应性“经验法则”所指导的,这一结果也使人联想到博弈论中囚徒困境(Prisoners dilemma)问题二人对局的纳什均衡——最简单的完全信息静态博弈结果。参赛者发现“机器人交易者”所营造的市场与人类交易者的真实市场非常相似,这进一步使人们相信可以用更简单的方法去构造经济环境中的行为趋真模型。计算技术日益成为经济学理论的创新所必需的重要资源之一。我们在理论经济学和应用经济学模型的研制中,有必要重视那些从数学、计算机科学、信息科学、物理学等自然科学中移植到经济学中的分析方法。

五、贝叶斯推断

贝叶斯推断方法的发展给应用经济学研究提供了不少便利。其与有限样本分布相关联的特性,使估计避免了渐近逼近

的方式。由于它更合理地运用了特征变量(latent variables),使得以前被传统经典观点视为不可解决的多步预测问题(multi-step forecasting)变得简单起来,而应用这类方法必须付出的主要代价是要估计复杂化的多维积分。但由于目前计算机与数学计算软件的进一步发展,这一方法也逐步在计量经济学的研究中得到更多的应用。传统的求积分方法已远远不足以应付高维数空间中的问题,可是以下三种快速发展的方法在许多应用问题中有良好的表现:

(1)由梯尔尼(L. Tierney)和卡丹(J. B. Kadane)在1986年提出的拉普拉斯展开法所需的逼近收敛速度大大快于传统的似然函数展开,机器上的计算时间也短得多。

(2)由格维克(J. Gewere)在1989年提出的带有重要性抽样的蒙特—卡罗积分法,目前已经被收入到不少商业计算软件中。它与上述的展开方法不同之处在于其对所需的多重积分计算的逼近精度可以任意好。缺点是计算时间更长。

(3)由盖尔方德(A. E. Gelfand)等人在1990年提出的吉布斯抽样(Gibbs Sampling)和数据扩张法(data augmentation),已经在统计应用中。此类方法发展的部分动力是来自图像处理的研究中。在计量经济学的应用中,此类方法的稳定性优于蒙特—卡罗积分法,计算时间也比上述二种方法短。

注释:

R. Miller (1984), "Job Matching and Occupational Choice", *Journal of Political Economy*, pp. 92

S. J. Turnovsky (1995), "Methods of Macroeconomic Dynamics", MIT Press

A. E. Gelfand and M. Smith (1990), "Sampling - Based Approaches to Calculating Marginal Densities", *Journal of the American Statistical Association*, pp. 85

W. H. Greene, "Econometric Analysis", Prentice Hall, 1997, Third Edition

这四个数学软件不仅具有交互式的图形功能,它们更强大的功能还在于数学的形式计算和数值计算方面。

D. Easley and J. Ledyard (1988), "Theories of Price Formation and Exchange in Double Oral Auctions", *Social Sciences Working Paper*, P. 661, California Institute of Technology

D. McFadden (1989), "A Method of Simulated Moments for Estimation of Discrete Response Models Without Numerical Integration", *Econometrica*, pp. 57, (No. 5, September).

(作者单位: 武汉大学经济学系 武汉 430072)

(责任编辑: 金 萍)